9 Evidence-based policy at the end of the Cartesian dream

The case of mathematical modelling

Andrea Saltelli and Silvio Funtowicz

The end of the dream

Other authors in this volume have already discussed at length their interpretation of the end of the Cartesian dream. New reflections of their analyses were provoked by an article in *The Economist*, a weekly business magazine of largely neoclassical and positivistic views on economics. Commenting on the poor state of current scientific practices, the magazine proclaims 'How Science goes wrong' on its cover, and its first editorial reads (*The Economist* 2013a, 11):

> Science still commands enormous – if sometimes bemused – respect. But its privileged status is founded on the capacity to be right most of the time and to correct its mistakes when it gets things wrong. … The false trails laid down by shoddy research are an unforgivable barrier to understanding.

This attack on science's privilege reminded us of another quote, coming from Paul Feyerabend (2010, p. xviii), *enfant terrible* of modern epistemology and *bête noire* of all positivisms:

> Science must be protected by ideologies; and societies, especially democratic societies, must be protected from science. … The theoretical authority of science is much smaller than it is supposed to be. Its social authority, on the other hand, has now become so overpowering that political interference is necessary to restore a balanced development.

When *The Economist* and Feyerabend speak with one voice, a dream must be at its end.

What prompted *The Economist* to devote its cover page to an issue of science's governance? One of several reasons was the troubling wave of retractions affecting applied science. Laboratory experiments cannot be trusted without further, independent verification (Sanderson 2013) and 'bloggers put chemical reactions through the replication mill'. In another article, rules are proposed to spot 'suspected work [… in] the majority of preclinical cancer papers in top tier journals' (Begley 2013).

The Economist (2013b, 21–4) argues that technical shortcomings are among the main causes of trouble with scientific practice, including scientists' incapacity to balance false positives and false negatives[1] and poor refereeing. The truth is perhaps even more worrisome, as revealed by one of the sources quoted by the same magazine, Ioannides (2005), according to whom:

> In this framework, a research finding is less likely to be true when [… *a list of statistical limitations*]; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance.

In other words Ioannides hints at normative issues associated with scientific practice. The ethos of science is normally associated with the Mertonian principles known by the acronym of CUDOS (Merton 1942); one of which, under the name of Organized Scepticism, prescribes that 'All ideas must be tested and are subject to rigorous, structured community scrutiny'. These norms[2] must have had a powerful appeal to previous generations of scientists; so Richard Feynman (1974, 341):

> there is one feature … that we all hope you have learned in studying science in school … It's a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty – a kind of leaning over backwards. … Details that could throw doubt on your interpretation must be given, if you know them. … give all of the information to help others to judge the value of your contribution.

If this is not enough to appreciate the anti-climax of lost innocence, here is Danish writer Peter Høeg (1993, 19):

> That is what we meant by science. That both question and answer are tied up with uncertainty, and that they are painful. But that there is no way around them. And that you hide nothing; instead, everything is brought out into the open.

What separates Feynman and Høeg from the sloppy practitioners harshly criticised by *The Economist*? Could it be that a set of counter-norms, as described by Mitroff (1974, 592):

- solitariness (secrecy, miserism) often used to keep findings secret in order to be able to claim patent rights;
- dogmatism, because careers are built around the purported truth of a particular theory or hypothesis

are becoming the new norms, replacing the Mertonian principles?

It may appear that there is today a greater incentive to operate in the context of *pseudo-science*, here defined as 'where uncertainties in inputs must be suppressed lest outputs become indeterminate' (Funtowicz and Ravetz 1990). Not only is the concealment of uncertainty widespread, as suggested by *The Economist*, but also its opposite, its amplification, e.g. the fabrication of uncertainty, driven by policy agendas or industrial interests (Michaels 2005; Oreskes and Conway 2010).

A useful discussion on present-day practices in science and how these must appear to scientists faithful to the old traditions is Philip Mirowski's *Science-Mart: Privatizing American Science* (2011a). Mirowski argues that there is a crisis in the self-governance practices of science, and that the decline in the quality and character of science is linked to its commoditisation, driven by a combination of neoliberal credo and a close adherence to a neoclassic economics paradigm. Accordingly after the 1980s, neoliberal ideologies succeeded in decreasing state intervention in the funding of science, which became increasingly privatised.

Mirowski describes how in-house science laboratories of major corporations were closed, and research outsourced to universities which became more and more committed to the commercialisation of research findings. He then goes on to illustrate how research was further outsourced, this time to contract-based private organisations. As a result, knowledge as a monetised commodity has replaced knowledge as public good, producing among scientists the 'greater financial and other interest and prejudice' noted by Ioannides. In other words, there is a positive incentive to engage in pseudo-science.

A similarity can be detected between Mirowski's account of the neoclassic economic agenda as applied to research, recent critiques of Ricardian economics as applied to innovation (Reinert 2008; Mazzucato 2013) and the postmodern account of knowledge's legitimisation as formulated by Jean-François Lyotard in *La condition postmoderne* (1979).

Increased controversy is another visible characteristic of present scientific practices, particularly in innovation research or technoscience. From GMOs to climate, from bees and pesticides to shale gas fracking, from endocrine disruptors to refrigerant in Mercedes cars: an ever larger number of issues appear to become *wicked*, meaning that they are deeply entangled in a web of hardly separable facts, interests and values (Horst *et al.* 1973).

The media play an increasingly ambiguous role, opening an advertising channel to entrepreneurial scientists on one hand, and on the other, openly challenging trust in science with a language previously reserved to more mundane types of controversies. The manner for settling scientific disputes has evolved or degenerated, according to different perspectives. The media offer, for instance, headings such as 'Beware the rise of the government scientists turned lobbyists' (Monbiot 2013), and in the journal *Nature* an article proclaims that 'European bans on MON810 maize is the clear evidence of government interference with science' (Kuntz 2013).

Stringent standards for policy-relevant science and for the quality of the evidence are now insistently called for, even from the columns of *Nature*, where Ian Boyd (2013), speaking in its capacity of science adviser to DEFRA, the UK government department for environment, food and rural affairs, laments 'concern about unreliability in scientific literature' and 'systematic bias in research'.

Norms associated with scientific enterprise and scientific advice are under concerned scrutiny (see e.g. Pielke 2007; Jasanoff 2013; Gluckman 2014), and the media show a keen interest in the topics of science's governance and science–policy interaction. See, for instance, *The Economist* (2014a) taking good note of the creation of the Meta-Research Innovation Centre launched at Stanford (METRICS), involving the already cited I. Ioannidis, to combat *bad science*. According to Jasanoff (2013) 'a prime casualty in the age of information and informatics appears to be public confidence in the power of reason'. Perhaps the public is simply learning that science should not be trusted as faith, and that emerging scientific practices, so closely related to economics, policy and politics, should be democratically scrutinised.

Battling 'bad modelling'

How should we interpret the Cartesian dream in the context of mathematical modelling? A particularly explicit formulation of the dream was made by the French philosopher, mathematician and political scientist Marie Jean Antoine Nicolas de Caritat (1743–94), known as Marquis de Condorcet. 'Condorcet elaborated the utopia of a science-based society as one of welfare, equality, justice and happiness' (Rommetveit *et al.* 2013). Central to this vision was human's ability to calculate, to master mathematics, seen after Galileo as the language used by God to code the universe.

Fast forward to the present time and we read in the *Washington Post* that 'Based on mountains of data from 39 models and accurate within five years in either direction for any of the locations they studied … Washington DC climate will shift in 2047' (Bernstein 2013). *Prima facie* the dream of Condorcet has come true. We can predict nature and make the necessary arrangements to prevent problems ahead. Or can we? Some journalistic exaggeration needs to be taken into consideration. In the more sober scientific article at the source of the *Washington Post*'s piece (Mora 2013) the uncertainty is assessed at 14 years rather than five. Still it is legitimate to suspect that this is one of the many instances where the Knightian concept of uncertainty has been reduced to quantitative risk.[3] Should one be reassured by the fact that 39 models were used (or were deemed necessary) to arrive at the 2047 forecast? Or should we reflect about the forbiddingly complex nature of these inferences?

Another telling example is in Saltelli and d'Hombres (2010), discussing the so-called Stern Review, a cost benefit analysis of the merits of early intervention to mitigate climate change.[4] In this particular case the analysis extended two centuries beyond the present time and was equipped with a sensitivity analysis which was particularly unconvincing. A rich literature is by now available to criticise mathematical hubris, from Taleb's *Black Swan* (2007) to Pilkey and Pilkey-Jarvis's *Useless Arithmetic* (2007). Mathematical modelling paradox is best described by Naomi Oreskes (2000, 35), according to whom:

> In many cases, these [model-based] 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. … to be of value in theory testing, the predictions involved must be capable of refuting the theory that generated them … This is where predictions … become particularly sticky. … models are complex amalgam of theoretical and phenomenological laws (and the governing equations and algorithms that represent them), empirical input parameters, and a model conceptualisation. When a model generates a prediction, of what precisely is the prediction a test? The laws? The input data? The conceptualisation? Any part (or several parts) of the model might be in error, and there is no simple way to determine which one it is.

A different perspective from which to look at mathematical modelling is through the *ceteris paribus* assumption. According to Joseph Stiglitz (2011, 594): ‘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.’ The problem is when those things we leave out come back to haunt us. Said otherwise: *ceteris* are never *paribus*.

The issue is not new, and it is endemic in the parameters-rich models used in natural sciences, as well as the parsimonious models wanted in econometrics. Keynes alluded to it with his usual style in a dispute with Tinbergen, asking the rhetorical question (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?

In recent papers (Saltelli *et al.* 2013; Saltelli and Funtowicz 2014) a new set of specific criteria has been proposed for proper use of model-based inference in the policy process (sensitivity auditing). The rules, aimed at ensuring transparency and balance in the use of models, are:

1. Check against rhetoric use of mathematical modelling.
2. Adopt an ‘assumption hunting’ attitude.
3. Detect pseudo-science.
4. Find sensitive assumptions before these find you.
5. Aim for transparency.
6. Do the right sums.
7. Focus the analysis on the key question answered by the model, exploring holistically the entire space of the assumptions.

It may be interesting to compare these rules with a suggestion from Ian Boyd (2013, in the *Nature* article already cited):

> We need an international audited standard that grades studies, or perhaps journals. It would evaluate how research was commissioned, designed, conducted and reported. This audit procedure would assess many of the fundamental components of scientific studies, such as appropriate statistical power; precision and accuracy of measurements; and validation data for assays and models. It would also consider conflicts of interest, actual or implied, and more challenging issues about the extent to which the conclusions follow from the data. Any research paper or journal that does not present all the information needed for audit would automatically attract a low grade. Such a system would provide policy officials and others with a reliable way of assessing evidence quality, and it would drive up standards in scientific research to reverse the worrying trends that suggest underlying bias.

Though Boyd's proposed international standards are independent from our rules, the similarity of context and intents is evident.

An important caveat is in order before introducing the rules in detail. The purpose of the rules is not to discourage the use of mathematical modelling in policy-related science. On the contrary, we do believe that modelling has a role to play, provided it is not used rhetorically or inappropriately. We distinguish between policy simulations, when e.g. macro-economic models are used to explore the effects of different shocks on economic variables, from policy justification, when the same models are used to justify policy interventions.

In 2010, the Hearing Charter of the House Committee on Science and Technology received sworn testimony by economists Sidney Winter, Scott Page, Robert Solow, David Colander and V.V. Chari on why the financial and economic crisis was not foreseen by existing modelling tools, and in particular, from the dynamic stochastic general equilibrium models (DSGE; Mirowski 2011b). The chairman of the committee made precisely this point in remarking:

> DSGE 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 last question sounds rhetorical, though we appreciate the distinction between a *theorist tool*, what we would call a policy simulation tool, and a platform for policy advice, which we would call a policy justification tool. It is somewhat implicit in this formulation that policy simulation and policy justification perform quite different functions, though it must be extremely tempting, not to say an automatic reflex of the analysts, to assume that the former can be deployed for the latter.

The seven rules

The point of departure for the development of the rules is the consideration that good practices for sensitivity analysis, enshrined in existing guidelines for mathematical modelling, are insufficient to ensure quality in the treatment of uncertainty in the contested arena of science for policy. In an adversarial context, not only the nature of the evidence, but also the degree of certainty and uncertainty associated with the evidence will be the subject of heated debate by all the relevant parties.

The problem is succinctly illustrated in the following coastal zone oil drilling example in the Norwegian islands of Lofoten:

> When there is low uncertainty, it is often because a topic is not interesting. But as soon as the stakes rise, uncertainty becomes important. … uncertainty is the result of three things: incomplete science, bad science and corrupted science. In this latter case, corrupted science is produced purposefully to create debate or even confusion. … Uncertainty can be seen as a tool that is used to prevent or support action. In the case of Lofoten, uncertainty is part of the political game, and is used by decision-makers, industry actors, the local population, environmentalists and NGOs. (Blanchard 2013)

It is in this type of context, that of post-normal science, where *facts are uncertain, values in dispute, stakes high and decisions urgent* (Funtowicz and Ravetz 1993), that the rules find their justification. The rules presuppose a participatory style of decision-making, one where knowledge is co-produced, where a hybridisation of science and politics takes place and where a new public, capable of bringing fresh insight in the solution of a problem, is created (Lane *et al.* 2011; Feyerabend 1975, 262[5]). In such a situation, the rules facilitate the work of mediation between the abstract rules of mathematical modelling and the policy issues at stake.

In the case of the deployment of mathematical models for impact assessments, the rules of the checklist could be introduced as a set of potentially adversarial questions to be *anticipated* by practitioners, including the following:

- X was treated as a constant when we know it is uncertain by at least 30 per cent.
- A 5 per cent error in X would be sufficient to make your statement about Z fragile.
- The model is but one of the plausible models – model uncertainty has been neglected.
- The level of confidence in a desired result has been artificially inflated by minimizing the inputs' uncertainty.
- Uncertainty in the input has been inflated in order to invalidate an undesired inference.
- The model is a black box – why should we trust your results?
- The framing of the analysis is not socially robust (a class of stakeholders has been neglected).
- The question which was answered is a question nobody asked.

Sensitivity auditing can also be related to NUSAP (Funtowicz and Ravetz 1990; van der Sluijs *et al.* 2005), a system for the quality assessment of quantitative information. NUSAP also belongs to the tradition of post-normal science, and has been used e.g. in the field of climate science (Kloprogge and Van der Sluijs 2006). When using NUSAP, a relevant number (N) comes with its units (U), its standard error (S), as well as with an assessment (A) of the process leading to the formulation of the number. Finally, relevant numbers (e.g. those which may feed into a policy decision) must have pedigrees, which may describe the track record of the team proposing the number, or the available history of related or similar number predictions. Both assessment and pedigree can be in the form of checklists (see also www.nusap.net).

As mentioned above we can relate the checklist to the NUSAP tradition. In this case the sensitivity auditing checklist could be seen as part of a model-assessment or pedigree, answering questions such as:

1 Is the model redundant?
2 Are there important implicit assumptions?
3 Is uncertainty instrumentally amplified or compressed?
4 Was a sensitivity analysis performed prior to publication of the inference?
5 Is the model transparent?
6 Does the model address the right question?
7 Was sensitivity analysis performed holistically?

We'll now introduce the checklist, illustrating the rules in detail.

Rule 1. Check against rhetorical use of mathematical modelling

This rule should be rather evident to the reader at this point of our discussion. We term rhetorical, a model use which aims to confirm (at times with a disproportionate use of mathematics and computer time) an already taken decision, based on considerations of power or interest. The larger the model, the easier it is to fiddle with its parameters to obtain whatever result one might wish (Hornberger and Spear 1981). As noted by Stiglitz (2010, 161) – discussing the case of the mathematical tools used to price collateralised debt obligations leading to the financial crisis – perverse incentives generate flawed models.

The issue was popularised by Douglas Adams in his book series *Dirk Gently, The Holistic Detective*:

> 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.
> (Adams 1987, 69)

Rule 2. Adopt an 'assumption hunting' attitude

We refer here to our discussion on the *ceteris paribus* assumption. The rule could thus be read as: which *ceteris* were assumed to be *paribus*? What was assumed out (which effect or process was not included)? What was assumed in (which parameters were fixed by the developers and on which basis). It is frequently easy to deconstruct the model by reconstructing the series of assumptions which went into its construction.

Rule 3. Detect pseudo-science

Pseudo-science or Garbage In Garbage Out (GIGO) was defined by Funtowicz and Ravetz (1990) as a situation in which 'uncertainties in inputs must be suppressed lest outputs become indeterminate'. The modeller in violation of this rule fiddles with the uncertainty present in the input, in order to ensure that the output, the inference, is not so vague as to be practically useless (e.g. a policy's payoff bracketed between a big loss and a large gain). Similar prescriptions in econometrics recommend a thorough exploration of the space of the input assumptions (Kennedy 2007). As noted above, this rule can be played in reverse, with a party inflating uncertainty instead of minimising it, with the objective, for instance, of resisting a regulation by overestimating the uncertainty in a class of health effects (see examples in Saltelli *et al.* 2013).

Rule 4. Find sensitive assumptions before these find you

This rule reminds model developers, and *a fortiori* those building the case for a policy, to be clear about the limits of their analysis before going public with the findings. In an adversarial context an opposing party could otherwise apply rule 2 to invalidate the case. Doing such an analysis *a posteriori*, to fend off a received criticism, usually results in protracted and costly arguments. In the case of the Stern Review mentioned above, sensitivity analysis was performed by the team led by Nicholas Stern after its main findings had been criticised by an expert in cost benefit analysis. As discussed in Saltelli and d'Hombres (2010), Stern's position would have been stronger if he had performed the analysis before going public with his results.

Rule 5. Aim for transparency

This rule recommends that proponents of a policy present their evidence in a way that the relevant audiences, including the opponents, can understand. In other words, black box models, or proprietary models, owned by a third party, which cannot be consequently explored, are generally interpreted as an attempt to hide more than to show. At the time of writing the present chapter, a piece of legislation is under discussion in the US. The bill,[6] named 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"' (Wilkey 2014). The consequences of this draft bill are clearly ambiguous; a positive outcome might entail making a mathematical model fully available to all parties so it can be used as a policy simulation tool, with its assumptions made transparent.

Rule 6. Do the right sums

As the saying goes, doing the right sums is more important than doing the sums right, in line with Keynes's famous remark that it is better to be roughly right than precisely wrong. In the context of a policy study this would imply asking the relevant questions in order to resolve the problem that is salient and pertinent to the relevant stakeholders. As an example we can take a current and popular wicked issue: the case of genetically modified organisms (GMOs) used for crops and foods. Proponents of GMOs observe that citizens' hostility to these products is at odds with the evidence that GMOs do not have negative health effects. According to the results of an EU-funded study (Marris *et al.* 2001), food safety is not prominent in the list of citizens' concerns on GMOs. A list of concerns registered by Marris *et al.* includes:

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

For a recent illustration of this case, if we believe in the findings from the report cited above, we would consider this rule as violated by articles lambasting the US state of Vermont for its recently introduced GMO labelling law on the basis that scientific evidence proves GMO food safe for consumption.

> Montpelier is America's only McDonald's-free state capital. A fitting place, then, for a law designed to satisfy the unfounded fears of foodies.
>
> (*The Economist*, 10 May 2014)

> Just ask about genetically modified crops, declared safe by the scientific establishment, but reviled as Frankenfoods by the Subarus-and-sandals set.
>
> (*The Economist*, 10 May 2014)

While the GMO example does not refer to a particular mathematical model, there is an entire class of models which may fall under the watch of rule 1. These are all the cost benefit analysis or risk analysis performed to demonstrate the safety of a new technology after the technology has been introduced. As cogently noted by Langdon Winner (1986, 138–63), ecologists should not be led into the trap of arguing about the 'safety' of a technology after the technology has been introduced. They should instead question the broader power, policy and profit implications of that introduction.[7]

Rule 7. Focus the analysis on the key question answered by the model, exploring holistically the entire space of the assumptions

This rule, more technical, is a summary of good practices belonging to the discipline of sensitivity analysis (Saltelli *et al.* 2012). In a model-based study for impact assessment it is important that the sensitivity of the input assumption is directly related to what is being assessed, and not to some intermediate model result. At the same time, the space of the input assumptions should be explored thoroughly. The most popular sensitivity analysis practice found in the literature is that of *one-factor-at-a-time* (OAT; Saltelli and Annoni 2010). This consists of analysing the effect of varying one model input factor at a time while keeping all others fixed. The shortcomings of OAT are known from the statistical literature, but its use among modellers is still widespread.

Where do we go from here?

There is still a strong movement of scientists in favour of performing analyses of the cost of climate change. So, for instance, Revesz *et al*. (2014), writing in *Nature*:

> Costs of carbon emissions are being underestimated, but current estimates are still valuable for setting mitigation policy.
>
> … These [Those from climate change] are real risks that need to be accounted for in planning for adaptation and mitigation. Pricing the risks with integrated models of physics and economics lets their costs be compared to those of limiting climate change or investing in greater resilience.
>
> Yet the social-cost benchmark is under fire. Industry groups, politicians – including leaders of the energy and commerce committee of the US House of Representatives – and some academics say that uncertainties render the estimate useless.
>
> As legal, climate-science and economics experts, we believe that the current estimate for the social cost of carbon is useful for policy-making, notwithstanding the significant uncertainties.

Here we find all the ingredients of a science–policy mix: the normative stance of the embattled authors, together with the acknowledgment of the pervasive uncertainties, and the belief that costs of damage and costs of remedial actions can be compared. It is evident that even a weak application of the rules of the checklist would put these analyses into serious methodological difficulties, as the case of the Stern Review discussed above has shown. Ultimately we agree with Brian Wynne that 'science can be led to overreach itself in arbitrating public facts, meanings and norms' (Wynne 2010), and with Pilkey and Pilkey-Jarvis (2007, 86) that progress would be achieved if

> … the global change modelling community would firmly and publicly recognise that its efforts to truly quantify the future are an academic exercise and that existing field data on atmospheric temperatures, melting glaciers, … and other evidence should be relied on to a much greater degree to convince politicians that we have a problem.

In conclusion we believe that current modelling practices, in their development and use, are a significant threat to the legitimacy and the utility of science in contested policy settings, and that organised forms of quality control are needed. Transparency and parsimony seem to be important elements of quality control, which will encourage modellers to focus on the truly relevant assumptions and mechanisms.

The conditionality of model predictions must be a constant concern for those operating models in support of policy. This will result in greater credibility for models and greater clarity about what can be adjudicated by quantitative model-based quantification, and what should be deferred instead to democratic political institutions.

Acknowledgements

Helpful suggestions and corrections were offered by Bruna De Marchi, University of Bergen (NO), Centre for the Study of the Sciences and the Humanities (SVT). The opinions of the author cannot in any circumstance be attributed to the European Commission.

Notes

1 'In medical testing, and more generally in binary classification, a false positive is when a test result indicates that a condition – such as a disease – is present (the result is positive), but it is not in fact present (the result is false), while a false negative is when a test result indicates that a condition is not present (the result is negative), but it is in fact present (the result is false)' (Wikipedia). According to Ioannidis (2005) false positives and false negatives are poorly accounted for in the appraisal of the results of ongoing medical research.

2 The CUDOS set of norms runs as follows: Communalism – the common ownership of scientific discoveries, according to which scientists give up intellectual property rights in exchange for recognition and esteem … Universalism – according to which claims to truth are evaluated in terms of universal or impersonal criteria, and not on the basis of race, class, gender, religion, or nationality; Disinterestedness – according to

which scientists are rewarded for acting in ways that outwardly appear to be selfless; Organised Scepticism – all ideas must be tested and are subject to rigorous, structured community scrutiny.

3 In *Risk, Uncertainty, and Profit* F.H. Knight distinguishes between risk that can be computed and uncertainty which cannot. Knight's prescriptions are largely ignored in the modelling community. According to John Kay, a British economist, the issue was felt as crucial by Maynard Keynes: 'For Keynes, probability was about believability, not frequency. He denied that our thinking could be described by a probability distribution over all possible future events, … In the 1920s he became engaged in an intellectual battle on this issue, in which the leading protagonists on one side were Keynes and the Chicago economist Frank Knight, opposed by a Cambridge philosopher, Frank Ramsey, and later by Jimmie Savage, another Chicagoan. Keynes and Knight lost that debate, and Ramsey and Savage won, and the probabilistic approach has maintained academic primacy ever since. A principal reason was Ramsey's demonstration that anyone who did not follow his precepts – anyone who did not act on the basis of a subjective assessment of probabilities of future events – would be "Dutch booked". I used to tell students who queried the premise of "rational" behaviour in financial markets – where rational means are based on Bayesian subjective probabilities – that people had to behave in this way because if they did not, people would devise schemes that made money at their expense. I now believe that observation is correct but does not have the implication I sought. People do not behave in line with this theory, with the result that others in financial markets do devise schemes that make money at their expense.'

4 A cost benefit analysis extending till 2200 of a socio-economic-ecological system at the planetary scale seems to us an illustration of George Soros's Postulate of 'radical fallibility': 'Whenever we acquire some useful knowledge, we tend to extend it to areas where it is no longer applicable' (2009).

5 '… 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' (Feyerabend, 1975, 262).

6 See http://science.house.gov/sites/republicans.science.house.gov/files/documents/HR4012%20.pdf (last accessed April 2014). The peremptory wording of the bill is interesting: 'To prohibit the Environmental Protection Agency from proposing, finalizing, or disseminating regulations or assessments based upon science that is not transparent or reproducible.'

7 '... the risk debate is one that certain kinds of social interests can expect to lose by the very act of entering. In our times, under most circumstances in which the matter is likely to come up, deliberations about risk are bound to have a strongly conservative drift. The conservatism to which I refer is one that upholds the status quo of production and consumption in our industrial, market oriented society, a status quo supported by a long history of economic development in which countless new technological applications were introduced with scant regard to the possibility that they might cause harm' (Winner 1986).

References

Adams, D. 1987. *Dirk Gently's Holistic Detective Agency*, New York: Pocket Books.

Begley, C.G. 2013. Reproducibility: Six red flags for suspect work, *Nature*, 497: 433–4.

Bernstein, L. 2013. DC climate will shift in 2047 researchers say; tropics will feel unprecedented change first, *Washington Post*, 9 Oct.

Blanchard, A. 2013. Interview in Sverre Ole Drønen, *Oil and Uncertainty,* 11 Nov. (http://wwwuibno/en/news/45772/oil-and-uncertainty-watch-new-video-messages-lofoten) accessed April 2014.

Boyd, I. 2013. A standard for policy-relevant science: Ian Boyd calls for an auditing process to help policy-makers to navigate research bias, *Nature Comment,* 501 (12 Sept.): 160.

The Economist. 2013a. How science goes wrong, 19 Oct.

The Economist. 2013b. Trouble at the lab, 19 Oct.

The Economist. 2014a. Metaphysicians (combating bad science), 15 March.

The Economist. 2014b. Genetically modified food: The little state that could kneecap the biotech industry, 10 May.

Feyerabend, P. 1975/2010. *Against Method*, London: Verso.

Feynman, R. 1974. Cargo cult science, Caltech commencement address; also in *Surely You're Joking Mr Feynman!*, New York: W.W. Norton & Co., 1997.

Funtowicz, S.O., and Ravetz, J.R. 1986. Policy-related research: A notational scheme for the expression of quantitative technical information, *Journal of the Operational Research Society,* 37: 1–5.

Funtowicz, S., and Ravetz, J. 1990. *Uncertainty and Quality in Science for Policy,* Dordrecht: Kluwer Academic.

Funtowicz, S.O., and Ravetz, J.R. 1993. Science for the post-normal age, *Futures,* 25(7): 739–55.

Gluckman, P. 2014. The art of science advice to government, *Nature,* 507: 163–5.

Høeg, P. 1993. *Borderliners*, Toronto: Delta Publishing.

Hornberger, G.M., and Spear, R.C. 1981. An approach to the preliminary analysis of environmental systems, *Journal of Environmental Management,* 12(1): 7–18.

Horst, W.J., Rittel, Melvin, and Webber, M. 1973. Dilemmas in a general theory of planning*, Policy Sciences,* 4: 155–69.

Ioannidis, J.P.A. 2005. Why most published research findings are false, *PLoS Medicine*, 2(8): 696–701.

Jasanoff, S. 2013. The science of science advice, in Robert Doubleday and James Wilsdon (eds), *Future Directions for Scientific Advice in Whitehall* (http://www.csap.cam.ac.uk/events/future-directions-scientific-advice-whitehall) accessed Aug. 2014.

Kennedy, P. 2007. *A Guide to Econometrics,* 5th edn, Oxford: Blackwell Publishing.

Keynes, J.M. 1940. On a method of statistical business-cycle research: A comment, *Economic Journal,* 50(197): 154–6.

Kloprogge, P., and van der Sluijs, J. 2006. The inclusion of stakeholder knowledge and perspectives in integrated assessment of climate change, *Climatic Change*, 75(3): 359–89.

Knight, F.H. 1921. *Risk Uncertainty and Profit*, Ithaca, NY: Cornell University Library.

Kuntz, M., Davison, J., and Ricroch, A.E. 2013. What the French ban of Bt MON810 maize means for science-based risk assessment, Correspondence, *Nature Biotechnology*, 31(6): 498–9.

Lane, S.N., Odoni, N., Landström, C., Whatmore, S.J., Ward, N., and Bradley, S. 2011. Doing flood risk science differently: An experiment in radical scientific method, *Transactions of the Institute of British Geographers*, 36: 15–36.

Marris, C., Wynne, B., Simmons, P., and Weldon, Sue. 2001. *Final Report of the PABE Research Project Funded by the Commission of European Communities,* Contract number: FAIR CT98-3844 (DG12-SSMI) Dec, Lancaster: University of Lancaster.

Mazzucato, M. 2013. *The Entrepreneurial State: Debunking Public vs Private Sector Myths*, London: Anthem Press.

Merton, R.K. 1942. The normative structure of science, in R.K. Merton (ed.), *The Sociology of Science: Theoretical and Empirical Investigations,* Chicago, IL: University of Chicago Press.

Michaels, D. 2005. Doubt is their product, *Scientific American*, 292(6): 96–101.

Mirowski, P. 2011a. *Science-Mart: Privatizing American Science*, Cambridge, MA: Harvard University Press.

Mirowski, P. 2011b. The seekers, or how mainstream economists have defended their discipline since 2008, Part IV (http://wwwnakedcapitalismcom/2011/12/philip-mirowski-the-seekers-or-how-mainstream-economists-have-defended-their-discipline-since-2008-%E2%80%93%C2%A0part-ivhtml) accessed April 2014.

Mitroff, I.I. 1974. Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists, *American Sociological Review*, 39: 579–95.

Monbiot, G. 2013. Beware the rise of the government scientists turned lobbyists, *Guardian*, 29 April.

Mora. C., *et al.* 2013. The projected timing of climate departure from recent variability, *Nature,* 502: 183–7.

Oreskes, N. 2000. Why predict? Historical perspectives on prediction in Earth science, in D. Sarewitz, R.A. Pielke, Jr and R. Byerly, Jr (eds), *Prediction: Science Decision Making and the Future of Nature*, Washington, DC: Island Press.

Oreskes, N., and Conway, E.M. 2010 *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*, New York: Bloomsbury Press.

Pielke, R.A., Jr. 2007. *The Honest Broker: Making Sense of Science in Policy and Politics*, Cambridge: Cambridge University Press.

Pilkey, O.H., and Pilkey-Jarvis, L. 2007. *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*, New York: Columbia University Press.

Reinert, E.S. 2008. *How Rich Countries Got Rich and Why Poor Countries Stay Poor*, New York: Public Affairs.

Revesz, R.L. *et al*. 2014. Global warming: Improve economic models of climate change, *Nature Comment,* 508: 173–5.

Rommetveit, K., Strand, R., Fjelland, R., and Funtowicz, S. 2013. What can history teach us about the prospects of a European Research Area? Study procured by the Joint Research Centre, EUR report 2612 (http://wwwuibno/sites/w3uibno/files/attachments/histera_final_report_25_2pdf) accessed April 2014.

Saltelli, A., and Annoni, P. 2010. How to avoid a perfunctory sensitivity analysis, *Environmental Modeling and Software*, 25: 1508–17.

Saltelli, A., and d'Hombres, B. 2010. Sensitivity analysis didn't help: A practitioner's critique of the Stern review, *Global Environmental Change*, 20: 298–302.

Saltelli, A., and Funtowicz, S. 2014. 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): 79–85.

Saltelli, A., Ratto, M., Tarantola, S., and Campolongo, F. 2012. Sensitivity analysis for chemical models, *Chemical Reviews*, 112(5): PR1–PR21 (Perennial Review of the 2005 version).

Saltelli, A., Guimarães Pereira, A., van der Sluijs, J.P., and Funtowicz, S. 2013. What do I make of your Latinorum? Sensitivity auditing of mathematical modelling, *International Journal of Foresight and Innovation Policy,* 9(2–4): 213–34.

Sanderson, K. 2013. Bloggers put chemical reactions through the replication mill, *Nature*, 21 Jan.

Soros, G. 2009. *The Crash of 2008 and What it Means: The New Paradigm for Financial Markets,* New York: PublicAffairs.

Stiglitz, J. 2010. *Freefall: Free Markets and the Sinking of the Global Economy*, London: Penguin.

Stiglitz, J.E. 2011. Rethinking macroeconomics: What failed and how to repair it, *Journal of the European Economic Association,* 9(4): 591–645.

Taleb, N.N. 2007. *The Black Swan: The Impact of the Highly Improbable*, London: Random House.

van der Sluijs, J., *et al.* 2005. Experiences with the NUSAP system for multidimensional uncertainty assessment, *Water Science and Technology,* 52(6): 133–44.

Wilkey, R. 2014. House Republicans aim to limit power of environmental protection agency, *Huffington Post* (http://wwwhuffingtonpostcom/2014/02/07/secret-science-reform-act_n_4748024html) accessed April 2014.

Winner, L. 1986. *The Whale and the Reactor: A Search for Limits in an Age of High Technology,* Chicago, IL: University of Chicago Press.

Wynne, B. 2010. When doubt becomes a weapon, *Nature*, 466: 441–2.