Policy-related mathematics is infected by fallacious statistical inferences, patently unfalsifiable simulation models, spuriously quantified indicators of quality, and sorcerer’s-apprentice algorithms.  How can our scientific culture be weaned from its teddy-bear numbers?

This theme has two parts. The pejorative descriptions of the ‘four horsemen’ may be shocking to some, but they have mainly been extensively studied, and they are not controversial to those who are familiar with these pathologies of modern scientific life. The second sentence is quite deliberately paradoxical and challenging, for I believe that the syndrome is both at the root of these problems in policy -related mathematics, and also lies very, very deep in our modern scientific culture. My experience of trying to engage people with this problem has been rather like getting fish to appreciate hydrodynamics.

First, as to the “four horsemen”, we have a very precious document, namely the essay by Prof Dorothy Bishop about her experience with statistics (posted in Resources). What I find most significant there is not just that there has been widespread incompetence on some basic techniques, but rather that even scientists of her calibre were ignorant of this incompetence for a long tine. Now, in statistics as distinct from some of the other fields I have mentioned, the really does exist a tradition of high quality, self-critical craftsmanship, producing results as reliable as the techniques permit. And in spite of this, the quality assurance mechanisms in this crucial field of modern science were seriously defective. It was not merely a case of ignorance of good procedure; even when the errors were pointed out, prominent scientists and journal editors fought vigorously to prevent critics from having any effect. That story is told considerable detail by the rather neglected book by McCloskey and Ziliac,The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives.

There is a good explanation of this extraordinary severe lapse in quality assurance. In scientific inquiry statistics is only a tool, and is not itself the subject of scientific enquiry. Hence the discipline based on the refutation of false assumptions, imperfect as it is in the research context, is totally lacking in the case of statistics. It is only natural for people who are doing slipshod or shoddy science to misuse or even abuse their tools. But in statistics there seems to have been a sort of Gresham’s Law operating, in that the bad practice seems to have completely driven out the good, even to the point where people with who would have wished to practice competently had no idea of how to do it, or even that it existed.

In this way the problem that is commonly seen as ‘irreproducibility’ is a very deep one in the life of the scientific community. Some have already raised the question of just how much of the body of published science has any relation to the realities of the external world. Exploration of this issue would entail an overhaul of some centuries of epistemology, since the possibility of vacuous science previously was never imagined. It also has urgent ethical and practical implications, since if science comes to be correctly understood (by practitioners and publics) as untrustworthy, then the foundations of our modern technological and social systems are vulnerable. The movements for restoration of quality and reform in science, which are now focused on statistics, are therefore as important as any other, for the survival of our civilisation.

Moving on to models, we rely here on the work of Andrea Saltelli. Almost single-handed, he has striven to bring some discipline into the practice of computer simulation models. This vast field of practice never even had standards of good practice available to be neglected. Saltelli has attempted to create such standards, with some signs of success. Again, models raise epistemological issues. The most profound statement about them is the well known, “All models are wrong; some are useful.” Accepting this, we ask how the model-based sciences relate to Karl Popper’s philosophy of science. They are certainly neither inductive or deductive in the classical senses; nor are they refutable. What then? If those sciences rely on confirmations, then are they any different from those denounced by Popper (Marxism and Freudian psychology) as no better than astrology? What then for all the policiesd, public and institutional, which have come to depend so completely on computer simulation models?

Moving on to indexes and indicators, I must confess that just now I do not have the empirical material that I would like to have, namely examples of illegitimate manipulations with quasi-numerical (or psuedo-numerical) data. It is all too easy to assign numbers to qualitative information, such as rating scales, and then to manipulate those numbers as if they were genuine quantitative measures, that is getting a mean and variance of their distribution, as well as performing all arithmetical operations on them. Perhaps, on the analogy of statistics, this claim of illegitimacy is a surprise and shock to some scientists here. For me, the unawares misuse and abuse of quantities, even deeper than in the case of statistics, is the essential contradiction of science conceived in the tradition of Galileo and Descartes. All I have in evidence are the numerous cases, some quite ludicrous, of pseudo-precision in tabular information. (I have posted my presentation ‘Arithmetical Languages for Policy’ in Resources) Such expressions are actually mendacious; if I write 5.6789 then I am implicitly claiming that the quantity is more than 5.6788 and less than 5.6790. If I know that quantity to within, say, 10%, then those extra digits are literally pernicious nonsense. I am well aware that in advertising such thoughts I am passing from the status of scientific rebel to that of scientific crank. But here I stand; I can do no other.

Finally, we have sorcerer’s-apprentice algorithms, on which I am comfortably in with the consensus. There is just one observation, which seems to me not to have been made already. This is whether the technology of conformity, which in China has been developed to become quite hegemonic, will actually produce a lethal contradiction. Unless the Chinese state can produce some truly extraordinary social engineering, it is difficult to imagine how creativity, which is essentially disruptive even in science and technology, can survive in an atmosphere of stifling conformity. The trajectory is that within a few decades it would all become stale, as nonconformists are squeezed out and forgotten. It could be stable, like Tokugawa Japan, but would eventually become brittle against disruption.

Now I must defend my second proposition, about ‘teddy-bear numbers’. I owe this insight to Rafael Ramirez, who expressed it as ‘transitional objects’ as developed in the theory of Donald Winnicott. I had originally used the term ‘Linus’s blanket’, but the reference might by now be obsolete. Traces of this can be found in the literature about management, in connection with mathematical tools such as spreadsheets. The insight, for which I cannot claim any scientific evidence, is that in our culture, people who must make decisions find emotional support in numbers. This can explain the widespread, even universal, abuse of numbers for which there is an ample literature. (For further reading, I have posted my essay on ‘Faith and Reason in the mathematics of the credit crunch’)

Just as in the case of bad statistics, we can ask, how does it persist? What is there in the culture that makes the faith in numbers so deep that noone notices it, and that critics like myself are, effectively, cranks? To be very brief, I can only repeat what I said previously, that it all goes back to Descartes and Galileo, and their paradigm for knowledge, based on their reductionist-atomist metaphysics. It could be argued that ever since the 1920’s, with the work of Heisenberg, Schroedinger and Bohr, this metaphysics has been obsolete at best. But mathematics has not yet caught up with physics.

I have such a commitment to this issue because I do believe in an objective reality out there, and in truth and integrity in our expressions of it. I also believe that if we continue to live in lies about our quantitative representations, then we will experience intellectual corruption, incompetence and ultimately failures of many sorts. For a complementary perspective on this challenge, I have posted the important essay by Urmie Ray.

Postings: Dorothy Bishop on statistics, Faith & Reason, Chapter 17 extract, Urmie Ray.