Data and ideology of science

Sumit Bhaduri

A recent editorial in *Nature*[1] urged scientists to rise above politics and restate their value to society. It is a suggestion that may work if what is of value could be agreed upon by scientists, their paymasters and ordinary citizens. Generating data or evidence-based knowledge is what scientists are supposed to do. That probably is the enduring value of science. However, to the two biggest paymasters of the scientific profession – Government and industry – that kind of value means little. Science that delivers technologies for war or profit, preferably both, is of value to them. To complicate matters, to an average citizen all data-based knowledge has now become a domain for 'experts'. Consequently, when bad or bogus knowledge is touted as 'scientific' or 'expert advice', the value of all knowledge is undermined. Unfortunately, in the post-truth world 'expert advice' is often nothing but lies cloaked in bad or fraudulent statistical data. A closer examination of the ideology that drives this rampant misuse of data in the name of science is therefore necessary.

Data and data analytics basically mean information and its analysis using statistical techniques. The value of the information carried by data depends on its accuracy and rigour of the analysis. Although probability and statistics came to play a pivotal role in physics rather late, their importance in the natural sciences cannot be overemphasized. It was only in the later part of the 19th century that theories based on probability found a firm footing in experimental physics. That specific area of physics dealt with the interactions between matter and energy, and was called thermodynamics. Today it is called statistical thermodynamics because when tested against data, the statistics-based theory unlike conventional physics gave the perfect fit. Max Planck, the father of quantum mechanics, had to adopt a statistical approach to come up with the equation that was consistent with all the data. His adoption of the statistical approach, in his own words, 'was an act of despair ... I was ready to sacrifice any of my previous convictions about physics'.

The spectacular success of 19th and early 20th century science spurred the adoption of its methods, rather indiscriminately in many subjects. The behaviour and thinking of (wo)man, both as an individual and as a social being received much attention. Measurements of mental attributes under controlled conditions were attempted. Galtonian psychometry and the eugenics of the mid-19th century were the first steps towards social Darwinism, an ideology based on the belief that the differences among individuals, including the social and economic hierarchy in a society, can all be explained and justified in terms of genetic differences. However, starting with Cyril Burt, a British psychologist who did most of his research in the first half of the 20th century, much of the statistical data of psychometry, IQ measurements, social psychology, etc. turned out to be cooked up to fit a theory[2].

Statistical theory came into the discipline of economics, famously called the 'dismal science' by Thomas Carlyle, rather late. Irving Fischer, the first Ph D in economics from Yale University, USA, and the father of American neoliberal economic thoughts, had two faculty advisors – a sociologist and a physicist. Willard Gibbs, a pioneer of statistical thermodynamics was one of them. Gibbs' contributions to science and technology are immense and enduring. Without his theory of chemical equilibrium, today there would be no industrial manufacture of ammonia and half of the world's population would starve[3]. Based on the analogy between chemical equilibrium and the supply–demand equilibrium of an economy, Fischer built mathematical models for value and price. In the present context it is interesting to note as an aside that both Burt and Fischer were eugenics enthusiasts.

In experimental physics or chemistry theories play second fiddle to data. In economics though, model-building came before data. Questions such as how long it takes for a thermodynamic equilibrium to be established are exceedingly important because the answers may have serious practical consequences. A wrong answer could lead to what are called runaway reactions accompanied by violent explosions and disastrous accidents. They could be answered with confidence only if there are enough reliable data and a theoretical model that agrees with the data. In contrast, in much of economic model-building, such inconvenient questions are left unanswered. The so-called law of the market, viz. that supply creates its own demand, is just an assumption and a questionable one. The failure of demand to catch up with supply can and has had disastrous consequences for society. Economic policies that focused on the supply side and left economy and societies to the mercy of the market, contributed substantially to the two World Wars and the major social upheavals of the last century.

Much of current economic theorizing continues to assume this fundamentally faulty, ideal, supply versus demand equilibrium. It also assumes a perfect market, the availability of complete information and the absence of irrational agents. It is interesting to note that so far as the data for testing economic theories are concerned, growth and gross domestic product (GDP) as measures of a nation's economic activity came into existence only in the late 1930s. However, even today there is no single universally accepted method for calculating GDP, let alone for verifying the accuracy of such data.

As the 'Grexit' drama of this decade showed, what numbers to include and what to leave out in calculating the measurable part of a country's economy and its debt, depend basically on ideology and power equations between the nations. Greece had joined the European Union in 2002, but after the spectacular financial crash of 2008, under a new Government and a new chief statistician, was forced to adopt a revised method for data analysis. The data had to be revised, as overnight much of the finance capital turned out to be fictitious but legal. The lenders – the European banks, International Monetary Fund (IMF), etc. – were in no mood to revise their definition of capital, and the debtor eventually had to agree with them. Greece's deficit jumped up by about two and a half times, and severe austerity measures were imposed.

The austerity measures insisted upon by the lenders got the backing from a paper written by two high-profile

economists. This paper was described as one with 'more immediate influence on public debate than any previous paper in the history of economics'[4]. However, it turned out, not surprisingly as many would say, that the arguments were based on selective choice of statistical data. Once it became public knowledge, the authors had no choice but to acknowledge their mistake. Meanwhile in Greece, after several changes of governments, the country's IMF-trained chief statistician abruptly resigned in 2015. In 2018, he was convicted by the Supreme Court of Greece for 'breach of duty' with two years' suspended sentence.

The Cyril Burt episode may be the first documented incident of the use of fraudulent statistics in psychology research, but certainly not the last. Early on in this decade Diederick Staple, a Dutch professor of social psychology, was found to have used cooked-up statistics in as many as 55 papers. This incident prompted much soul searching among the experimental psychologists and others. It brought into focus the deliberate abuse and misuse of statistical techniques in many other areas. The havoc that false or mindless application of the P factor can inflict was pointed out by many[5]. What is especially relevant is that Staple himself had spelt out his intentions in an award speech long before his fraud became public knowledge[6].

The motivations behind the deliberate use of bad or fraudulent data do not come from mere ignorance. They come from an ideology where knowledge has a price but no intrinsic value. It is an ideology that has successfully commodified almost all of science and much of social science. The mythical all-knowing free market, the pillar on which the ideology rests, does not differentiate between good data, bad data and fraudulent data. It cares only about profit maximization and pays lip service to long-term innovations that improve the quality of life[7]. When 'data-based knowledge' is sold in such a market, there are many 'rational agents' who profit by the deliberate tweaking and falsification of data. It is not surprising that many of them, including scientists and other academics, turn out to be 'experts'.

Globally and locally the signals coming from society do not bode well for mankind or science. It is important to note that the chaotic global economy, geo-political manoeuvring, and the identity politics of the present have many similarities with the socio-economic settings of the early 20th century. At that time too, the belief that the market is autonomous and operates outside society, had allowed the commodification of labour, land and money – entities over which society till that time had a say. The tension between society and market had finally reached a breaking point and social disintegration in the form of two World Wars, fascism and revolutions followed[8]. What commodification and deliberate promotion of bad data will do to science, or for that matter the whole of mankind, only the future will tell but, as the editorial in *Nature*[9] succinctly puts it, 'we ignore our past at our peril'[9]. The first step for rising above politics, is to recognize and acknowledge the shortcomings of the ideology that haunts today's world.

1. https://www.nature.com/articles/d41586-019-02379-w?
2. https://www.nature.com/articles/d41586-018-06784-5 and https://www.nature.com/articles/d41586-018-00578-5. Also see Lewontin, R., Rose, S. and Kamin, L. J., *Not in our genes: Biology, Ideology and Human Nature*, Penguin Books, London, 1984, ISBN 0-394-72888-2.
3. Smil, V., *Nature*, 1999, **400**(6743), 415.
4. *The Economist*, 2013, p. 70; also see https://www.newyorker.com/news/john-cassidy/the-reinhart-and-rogoff-controversy-a-summing-up
5. https://www.nature.com/articles/d41586-017-07522-z and https://www.nature.com/articles/d41586-019-00857-9
6. https://www.academia.edu/6759974/Social_psychology_before_and_after_the_Stapel_fraud_case
7. https://www.currentscience.ac.in/Volumes/113/01/0018.pdf
8. Polanyi, K., *The Great Transformation: The Political and Economic Origins of our Time*, 1944. Reprinted with a forward by Joseph Stiglitz and introduction by Fred Block, Beacon Press, Boston, USA, 2001.
9. Editorial, *Nature*, 2019, **573**, 464.

Sumit Bhaduri lives at 562 Adenwala Road, Rustom Mansion, Mumbai 400 019, India.
e-mail: bhaduri.sumit@gmail.com