New School Economic Review, Volume **1**(1), 2004, 3-8

SEARCH FOR GOOD SCIENCE: A PERSONAL MEMOIR

Scott Moss

I have been arguing for some years that all of the economics with which I am acquainted, whether orthodox or heterodox, is bad science. Rather than to look for an independent definition of bad science, a more constructive approach is positively to define good science and then to define bad science as its complement.

All good science is inspiration constrained by evidence and observation. If this statement is itself good science, then it too must be inspired and it must be constrained by evidence and observation.

In my experience, inspiration comes from many sources, including, to my great fortune, my teachers and my colleagues. The four teachers who were instrumental in my thinking about what is good science were, from my time at the New School, Edward Nell, Robert Heilbroner and Adolph Lowe, and, from my time in Cambridge, Nicholas Kaldor. Nell and his sometime colleague Martin Hollis [9] argued that the conditions under which a theory can be tested should always be specified by the theory, and a failure of neoclassical economics was that those conditions were not specified. Nell (though not, I think, with the support of Hollis) argued that the *ceteris paribus* conditions of neoclassical economics constitute the relevant conditions of testing. Heilbroner was less concerned with the conditions of testing, which draw on the philosophical literature on truth theory, arguing instead that the problem with neoclassical economics was its ahistoricity—in effect, that the conditions under which a theory might be relevant are the historical conditions that it describes. Adolph Lowe and Nicholas Kaldor argued in different ways that theory should be pragmatic. Kaldor in particular argued that economic theory had no point unless it informs economic policy.

These three lines of argument—which are not mutually exclusive—led me to the view that good science consists of theories which specify explicit conditions of *application*, where the worthwhile applications are those that serve policy formation.

A second line of thinking about what would constitute good economic science resulted from my work on computer simulation of economic and social issues, reported first in my New School PhD thesis and then in two books (Moss 1981, 1984). When developing social simulation models, the conditions of their application were foremost in my mind. I wanted to be able to assess whether my agents were acting in ways that conformed to the behavior of real actors. For this reason I adopted a decision-making representation based on a feature of computer science called *endorsements* (Cohen 1985). Its function was to enable a user to interrogate the agents about the reasons for their actions. For example, an agent might report that it reduced simulated water consumption because there was a drought and because other agents reduced their water consumption; that it was influenced by those other agents because it shared some specified common interests with them (Downing, Moss and Wostl 2000). The virtues of this approach are not only that we have a better understanding of the reasons for model outputs, but we also have a means of comparing the behavior of agents with the behavior of individuals those agents are intended to represent. That is, we can *validate* our models with respect to the qualitative accounts actors give of their own behavior and that of others with whom they interact. Generally, we get this kind of qualitative output at a micro-level and numerical outputs at macro- (system- or model-) level, enabling us to *cross-validate* our models.

Note the difference between this approach to social simulation and economics (or, indeed, other social sciences). We do not start with any theoretical propositions. We build our models from the bottom up to describe how individuals interact with others and how they choose other individuals with whom to interact in various ways. It is generally the norm in the social sciences either to start with some prior theory or to devise a set of propositions or hypotheses (*cf.* any issue of *Management Science* or the *American Journal of Sociology*) that are then tested by some statistical means.

In seeking to develop descriptive models, we constrain them by observation and evidence. We build models with the participation of stakeholders who use the models to analyze the behaviors of themselves or their competitors, customers, suppliers, collaborators, regulators or other individuals and agencies whose actions affect them. Such models do not stand alone. They are tools for helping decision makers form their expectations, identify opportunities and threats, and formulate strategies and policies.

What about inspiration? Theory is always one source of inspiration. However, that is not a source we have found to be useful. Again, permit me to offer the account of the development of my own views on this issue.

While working at Cambridge with some of the leading heterodox economists of the day (c. 1970), I came to the conclusion that the notion of long run equilibrium was unhelpful. Informed by my discussions and arguments with these incredibly gifted economists, I set out to produce for my PhD thesis a macroeconomic theory that captured the lessons of the 1960s Cambridge controversies in capital theory (won by the above great and lost by the MIT giants led by Robert Solow and Paul Samuelson), but that did not contain any notion of equilibrium.

In my thesis and subsequent book (largely unread and probably unreadable), I was able to develop a formal accounting framework for my theoretical insights but never a closed-form analytical model. That is why I turned to simulation modeling.

As my models became increasingly empirical I began to obtain some curious results. The numerical outputs from my models and those of my colleagues showed clusters of volatile change in the aggregate values of variables summed over individual actions. This is the sort of phenomenon usually noticed in financial markets and in such macroeconomic variables as employment, GDP and inflation, yet I found the same result in models of other types of markets. I obtained data of sales values and volumes for a wide range of fast-moving consumer goods, and found the same statistical signature in the data of every good I investigated, including alcoholic drinks, shaving preparations, tea, biscuits and shampoo.

This sort of clustered volatility is reflected in heteroskedasticity, autocorrelation and leptokurtosis (thin-peakedness or fat tails) of the frequency distribution of both levels and differences, requiring ARCH, GARCH and GMM methods of econometric analysis (reviewed in (Bollerslev 2001). The same result is found in the statistical mechanics literature on self-organized criticality (Bak 1997). Jensen (1998) has pointed out that the result is associated with models in which individuals only respond to significant events (as opposed to optimizing over continuous functions), they interact strongly with other individuals (which would be considered an 'externality' in economics), they are influenced by but do not imitate one another, and their world is not in turmoil. It turned out that my colleagues and I had been building models based on these assumptions because the assumptions describe the behavior we were observing. Leptokurtosis, autocorrelation and heteroskedasticity emerge at the macro-level from this descriptively well-validated specification of micro-level behavior.

These specifications were inspired by work done in social psychology, experimental cognitive science (Anderson 1997), computational cognitive science (Laird, Newell and Rosenbloom 1987), business history (Chandler 1962, Porter 1972), artificial intelligence and, of course, the evidence I was trying to capture. I find it remarkable that, starting from a commitment to high theory and a disdain for practical, applied economics, I have become inspired by history and experiment and have a deep regard for the importance of being constrained by evidence.

I believed I was doing good science, but I felt the need for some external assessment of that belief. So I turned to the history of science. Consider a few undisputable giants of good science:

Brahe:	Collected very precise measurements of the positions of stars and planets
Kepler:	Built on Brahe's observations and advanced the heliocentric theory
Galileo:	Pioneered telescopic observation
Lavoisier:	Devised oxygen hypothesis to explain increased weight of burnt objects
Darwin:	Collected data for more than 20 years before publishing *Origin of Species*
Maxwell:	Deduced the electromagnetic nature of light from Faraday's experiments
Watson & Crick:	Relied on Rosalind Franklin's x-ray crystallographic data to demonstrate the existence of DNA

It would be hard to imagine anyone more deeply involved in theory and less engaged in experiment and direct observation than Einstein. Yet here are two quotations from his own account of the development of relativity theory (Einstein 1961):

"Even though classical mechanics does not supply us with a sufficiently broad basis for the theoretical presentation of all physical phenomena, we must grant it a considerable measure of 'truth,' since it supplies us with the actual motions of the heavenly bodies with a delicacy of detail little short of wonderful. The principle of relativity must therefore apply with great accuracy in the domain of *mechanics*. But that a principle of such broad generality should hold with such exactness in one domain of phenomena, and yet should be invalid for another, is *a priori* not very probable."

"If the principle of relativity were not valid we should…expect that the direction of motion [relative to the sun] of the earth at any moment would enter into the laws of nature, and also that physical systems in their behaviour would be dependent of the orientation in space with respect to the earth…. However, the most careful observations have never revealed such anisotropic properties in terrestrial space, *i.e.*, a physical non-equivalence of different directions."

It is clear to me from these quotations that, at least in retrospect, Einstein gave pride of place to the constraints of observation and evidence. Any history of the science of the 19th and 20th centuries will confirm the importance of evidence, observation and experiment in the development of modern physical, chemical and biological sciences. Compare the attitude of the unambiguously successful sciences with the record of economics. Here are some approaches in mainstream economics whose putative goal was to increase significantly our understanding of actual economic activity:

- general equilibrium theory (1880s)
- marginal productivity theory (1890s)
- imperfect and monopolistic competition theory (1930s)
- econometric forecasting (1930s)
- game theory (1940s)
- neoclassical growth theory (1950s)
- post-war Keynesian macroeconomics (1950s)
- fixed-point general equilibrium theory (1950s)
- monetarism (1960s)
- supergames (1970s)
- endogenous neoclassical growth theory (1980s)
- computable general equilibrium theory (1990s)
- replicator dynamics (1990s)

I have tried and failed to find a single instance where one of these theoretical developments has ever led to a correct forecast of a volatile episode—for example, a turning point in a trade cycle or stock exchange price. Since it is also the case that none of these theories has been validated against observed behavior at the micro-level, I must conclude that this brand of economics is far removed from any reality of which I have ever known.

I suspect that the same accusation that I have made against mainstream economics can be made with equal legitimacy against any of the economic heterodoxies, whether Marxist or Post-Keynesian, whether based on cellular automata or “routines” (in the sense of Nelson and Winter). For, while heterodox economists reject the orthodoxy for being unrealistic, they too fail to begin their analyses with the evidence. Instead, they start with some analytical or programming technique or some theoretical speculation and then devise verbal arguments or formal models constrained by that technique or that speculation rather than by the evidence. Comparison with the evidence, if it happens at all, comes at a very late stage in the argument.

This is not the place to enter into a detailed discussion of leptokurtosis. Let me conclude simply by noting that leptokurtosis is inconsistent with the law of large numbers and has been known since the 1960s (Fama 1963, Mandelbrot 1963) to be incompatible with statistical forecasting based on any assumptions of the existence of a population distribution with a finite variance. This means that if the behavior we observe is correctly described by the agents we program, then the formal basis of econometrics is incompatible with the data. The incompatibility can be hidden by aggregating the data over time (leptokurtosis in daily data will be hidden by taking monthly or annual data which, by virtue of the central limit theorem, will appear normal). This would perhaps explain the failure of economics to meet its own central claim of being good science – that it leads to correct predictions. In any case where it is not possible to “refine” the econometric models in light of preliminary results (see, for example Mayer (1975)), economics does not generate correct predictions.

Experience shows that economic theory of whatever stripe does not inspire useful science. Since its assumptions are unconstrained by evidence and observation and it fails to satisfy its own criterion of prediction, I must conclude that economics is not good science. However, there is good science to be done — if one abandons failed theory and starts with the evidence and the problem at hand.

BIBILIOGRAPHY

Anderson, J.R., *Rules of the Mind*. 1993, Hillsdale NJ: Lawrence Erlbaum Associates

Bak, P., 1997, *How Nature Works: The Science of Self Organized Criticality*, Oxford: Oxford University Press

Bollerslev, T., 2001, "Financial Econometrics: Past Developments and Future Challenges". *Journal of Econometrics* **100** (1), pp. 41-51

Chandler, A.D.J., 1962, *Strategy and Structure: Chapters in the History of the American Industrial Enterprise*,Cambridge MA: MIT Press

Cohen, P.R., 1985, *Heuristic Reasoning: An Artificial Intelligence Approach*, Boston: Pitman Advanced Publishing Program

Downing, T.E., S. Moss, and C. Pahl Wostl, 2000, "Understanding Climate Policy Using Participatory Agent Based Social Simulation" in S. Moss and P. Davidsson (eds.), 2000, *Multi Agent Based Social Simulation*, pp. 198-213, Springer Verlag: Berlin

Einstein, A., 1961, *Relativity: The Special and General Theory*, New York: Crown Publishers Inc.

Fama, E.F., 1963, "Mandelbrot and the Stable Paretian Hypothesis" *Journal of Business* **36** (4), pp. 420-429

Hollis, M. and E.J. Nell, 1975, *Rational Economic Man: A Philosophical Critique of Neo-Classical Economics*, Cambridge: Cambridge University Press

Jensen, H., 1998, *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*, Cambridge: Cambridge University Press

Laird, J.E., A. Newell, and P.S. Rosenbloom, 1987, "Soar: An architecture for general intelligence", *Artificial Intelligence* 1987 **33** (1): pp. 1-64.

Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices" *Journal of Business* **36** (4): pp. 394-419

Mayer, T., 1975, "Selecting Economic Hypotheses by Goodness of Fit" *The Economic Journal* **85** (340): pp. 877-883

Moss, S., 1981, *An economic theory of business strategy : an essay in dynamics without equilibrium*, Oxford: Martin Robertson

Moss, S., 1984. *Markets and macroeconomics : macroeconomic implications of rational individual behaviour,* Oxford, OX ; New York, NY, USA: B. Blackwell.

Porter, G. and H.C. Livesay, 1972, *Merchants and Manufacturers: Studies in the Changing of Nineteenth Century Marketing*, Baltimore: Johns Hopkins University Press.