Towards Good Social Science

Scott Moss and *Bruce Edmonds*
Centre for Policy Modelling
http://cfpm.org

1. Introduction

This paper is about good and bad science. It is not about good and bad scientists though we do consider conditions in which good scientists do bad science. In particular, we take it for granted that scientists who seek to explain observed events by adhering carefully to the best standards available from their training and the traditions of their discipline have the personal qualities of the good scientist. If, however, the best available standards lead to bad science, then we would say that the good scientists are doing bad science.

The purpose of this paper is to explore the ways in which agent-based simulation could support good social science. We will argue that, where the social sciences are concerned, simulations based on software agents could support good science provided that the design of the agents is itself based on good science[1].

Clearly, the pursuit of this argument first requires an account of what constitutes good science and a defense of this. This will be done, not so much by engaging in a general philosophical or methodological debate but rather by drawing on examples of scientific developments that have transformed our understanding of the world. The examples will draw on developments in the natural – physical and biological – sciences and the lessons learned therefrom will be applied to the social and computer sciences. The discussion of the natural sciences will constitute section 2 followed by a brief discussion of how these will differ from the social sciences in section 3. The failure of large sections of key social sciences – economics and sociology – will be explored in section 4. In section 6 we take a step back from the argument about good and bad science in order to consider the purpose of doing social science and how that purpose conditions what we mean by good social science. Finally, in section 5, we consider the role of software agents in the development of good social science and, in particular, the development of social policy.

2. What is good science?

Good science enables us to understand what we observe. Different sciences have different criteria of what it means to do this. In physics, the depth of understanding is judged on the prediction of specific events and phenomena or distributions of numerical measures. Evolutionary biologists do not predict the emergence of previously described species but they do provide an *explanation* of speciation that was developed by Darwin to cohere with the fossil record and has subsequently cohered with statistical and molecular genetics[2]. “Positivist” economists following Friedman (Friedman 1953) claim to test the goodness of their theories by prediction

[1] This is, of course, a necessary rather than a sufficient condition.

[2] They do claim to predict some more abstract properties, e.g. the rate of change in junk DNA.

independently of the (lack of) "realism" of the assumptions on which their models are predicated.

The major milestones in the history of natural science have all helped explain the occurrence of observed events. Following Cartwright (Cartwright 1983), we charaterise explanation as follows: "*Explanation (at least the high level explanations of theoretical science …) organize, briefly and efficiently, the unwieldy and perhaps unlearnable, mass of highly detailed knowledge that we have of the phenomena.* (p. 87)" Whilst explanation is itself satisfying, it is frequently also of practical import. High level explanatory power is essential for applications of physics to engineering, because it leads to prediction and hence guides design. But, as Cartwright also points out, engineers do not *directly* use physical theory to determine whether a bridge will stand the stresses of a particular traffic load. They have rules of thumb that are informed by physical laws even though those laws do not *strictly* apply in any particular case[3]. Cartwright takes it for granted, as Hollis and Nell (Hollis and Nell 1975) argued before her, that the *ceteris paribus* conditions of any theory or law constitute its conditions of application. And the *ceteris paribus* conditions in physics, biology, economics or probably any other field are rarely, if ever, satisfied in applications[4].

While Cartwright bases her arguments on what modern physicists actually do, there is no doubt that explanation as organising principle has been the defining characteristic of the most important milestones in natural science. For example, at the time Copernicus' heliocentric cosmology[5] was far less accurate in predicting the movements of celestial bodies across the sky than was Ptolemaic geocentric cosmology unless it was complicated by as many epicycles and equants. But it did, at coarse grain, provide an easily understood account of the retrograde movements of the planets. And that, after all, was the point[6].

Tycho Brahe rejected the heliocentric cosmology though the alternative geocentric formulation he developed is in fact a straight (if not entirely obvious) mapping from Copernicus'. Of course, Tycho's main contribution to the development of cosmology and science more generally was his careful measurement of the positions and motions of the main navigational stars, planets and the sun. It was these measurements, based on his development of highly accurate observational instruments, that enabled the development of new navigational tables based on Tychonian observation and Copernican cosmology and those new tables, the first in several centuries, made Copernican cosmology predominant of Ptolomaic. As is well known, Tycho's observations were used first by Kepler to formulate his laws of planetary motion and then by Galileo. Newton would not publish his further developments until he had evidence – obtained virtually under duress from John Flamsteed, the first Astronomer Royal – to confirm his laws of mechanics (White 1997).

---

[3] In order to be made to apply, a whole series of simplifications, approximations and assumptions need to be used, which can not be completely justified by reference to the laws.

[4] It takes a great deal of effort and skill to make the conditions hold, which is why designing convincing experiments (which to a large extent do this) is so difficult.

[5] The points made about Copernicus are from (Kuhn 1992).

[6] Later on, with the availability of good data, it could be shown that given any finite set of data that the Copernicum account using elliptical orbits would forcast future planetary positions more accurately than one using Ptolemaic epi-cycles.

The development of the theory of electromagnetism and, eventually, Einstein's special and then general theories of relativity and quantum mechanics were driven by experiment and observation of natural phenomena. Faraday identified a wide range of electricity-related phenomena in a connected series of brilliant experiments showing *inter alia* that electricity and magnetism were closely related. However, while Faraday was able to formulate a qualitative version of the law of electromagnetic induction, he lacked the mathematical competence to produce any quantifiable physical theory of electricity. It is important to note here that Faraday experimented for many years to demonstrate the generality of his law. For he knew from both his own observation and the reports of experiments by others that electricity could be produced chemically, from magnets, by friction, by maintaining different materials and different temperatures and by animals such as eels and torpedo fish. "He asked the question which may seem obvious to us now, but which illustrates his deep insight at the time – are these different forms of electricity the same? In 1832, he performed an elegant series of experiments in which he showed that he could produce similar … effects, no matter what the source of the electricity might be…. Although the law of induction began to emerge at an early stage, it took Faraday many years of to complete all the necessary experimental work to demonstrate the general validity of the law…." (Longair 2003)

Faraday's experimental results along with those of Coulomb, Ampère, Volta and others were brought together by James Clerk Maxwell who built explicitly on Faraday's "lines of force" suggested by the pattern of iron filings on a sheet of paper close to a magnet. Maxwell was explicitly looking for an analogy with some physical phenomena that were readily represented mathematically. His first such analogy was with fluid flows so that the higher the density of flux (the fluid analogue) at a point, the stronger at that point would be the electrical or magnetic force. Based on this analogy, Clerk Maxwell mathematically formalised Faraday's "lines of force" as a new and mathematically impeccable concept: the field.[7]

For Einstein (Einstein and Infeld 1938), Maxwell's formulation of the field equations "is the most important event in physics since Newton's time … because they form the pattern for a new type of law…. The characteristic features of Maxwell's equations, appearing in all other equations of modern physics, are summarized in one sentence. Maxwell's equations are laws representing the *structure* of the field." (p. 143) This contribution underlay the future development of both quantum physics and relativity theory.

The history of biology from Darwin to the Human Genome Project is no less concerned with the explanation of observed phenomena than physics has been. Darwin was a leading geologist of his day, very familiar with the geological records of fossils. His explanation of gradual change in geological formations was his first great success and formed one empirical leg of his eventual theory of evolution. The other, of course, was the observation of differences among species across separated islands such as the Galapagos. One element of Darwin's preparation for the publication of *The Origin of Species* was his 20 year study of barnacles drawing on specimens collected by leading marine biologists from around the world. While Darwin assumed there to be random variation of mixed inherited traits, the statistical genetics invented by Grigor Mendel to explain heredity and the subsequent molecular

[7] This account is from (Mahon 2003), pp. 61-63.

biological identification by Watson and Crick of the structure of DNA to explain the chemical mechanism of heredity and random variation were all based on exhaustive observational and, in particular, experimental evidence. The evidence was of very different kinds: fossil evidence, Darwin's exhaustive study of barnacles, Mendel's beans and other plants, Rosalind Franklin's x-ray crystollography.

Although we have concentrated on a small number of exceptionally important instances of unarguably good science, wherever we look in the natural sciences we find developments great and small that are driven and constrained by empirical observation. That is to say that evidence and observation have *priority* over theory there – (in the end) when evidence and theory disagree the theory is changed.

Of course, dividing everything into theory or evidence is a considerable simplification of the situation in any science. Typically there is a whole range of entities involved from abstract conceptual frameworks and organising laws down to concrete data models and descriptions (Suppes 1962). In between there are whole families of theories, analogies, procedures, bridging rules and models (Giere 1988). To take a widely known example: the conceptual framework of using vectors to represent forces, velocity, acceleration, momentum etc. concerning discrete entities (treated as particles) facilitates and allows, using Newton's Laws, the construction of explanations of the movement of observed mechanics and models which, in restricted cases, allows for the prediction of sets of measurements. The moral of this paper is that the more abstract entities, require a huge amount of supporting scaffolding all the way down to the most concrete entities. One is unlikely to get to the thin pinnacle without the laborious work of also building up a broad base that is more directly related to observation and evidence – there is no 'magic' short cut (such as assuming people act collectively *as if* they were rational plus a random element when no evidence supports this) to obtain useful abstract social theory. Before we get to the pinnacle (if we ever do) we will not know *which* of the details in our concrete descriptions and models that are the important ones. Thus, until that time, our laws will not have the elegance of some models and laws in physics – they will be messy (or "ugly" as our title has it).

While some explanations are better than others, the entire class of explanations constituting good science explain phenomena we observe without distorting or contradicting the conditions in which we observe them. Where an explanation does not apply directly to a set of conditions because, following Hollis and Nell and Cartwright, its *ceteris paribus* conditions are not satisfied, then it may be the case that the explanation informs the development of special formulations that serve the purpose – for example, to enable engineers to calculate load factors for bridges. A good explanation will also guide investigations into failure. When the NASA Mars orbiter dropped into the Martian atmosphere in September, 1999, the NASA investigative team did not have to consider whether the science and engineering principles were wrong – only whether there were either extraordinary conditions that they had not anticipated (a failure of the *ceteris paribus* conditions) or there had been a programming or calculating error. The embarassing truth, of course, was that one team was calculating the required thrust for a course correction using metric units while another team was implementing that calculation using imperial units. But the underlying science based on Newton's laws and many years of engineering experience were not in question and so usefully restricted the scope of the investigation to possibilities consistent with this knowledge.

3. Difference between the social and natural sciences

The discussion in section 2 described some of the most influential episodes of scientific development in terms of theoretical development driven and constrained by observation (including experimentation). These theoretical developments, following Cartwright (Cartwright 1983) were taken to constitute explanations of observed and observable phenomena under specified conditions or to inform the development of special explanations when the conditions of application of the theory are violated in specific ways. Natural scientists assume that the underlying processes that generate observed phenomena are themselves unchanging over time and space. Experience has shown this assumption to be productive in the sense that explanations based more or less closely on fundamental laws prove to be descriptively accurate regardless of time or place provided, of course, that their conditions of application are not violated. There is no such experience in the social sciences. Indeed, one might characterise social processes by their ability to periodically undergo fundamental (i.e. structural) change.

A second important difference between the natural and the social sciences turns on what can be directly observed and how. The laws of thermodynamics were reformulated by Maxwell as statistical laws and, more generally, the whole of quantum mechanics is statistical. Many of the fundamental components specified by physics that are necessary to explain observed phenomena are not themselves directly observable. Indeed they often require very long chains of inference, where the devices that do the manipulation and measurement are themselves constructed on the basis of validated laws etc. These long chains are possible due to the relative stability and reliability of these laws supported by the scientific and engineering practices developed over the years, but are also the result of huge effort by engineers and scientists. It is a characteristic of the natural sciences that a great emphasis is given to developing new methods of measurement (as compared to the development of abstract theory).

In the social sciences techniques of measurement/data collection are relatively poorly developed (Chattoe 2002), but a key difference is that often social phenomena are not so much objectively measured but subjectively interpreted by the human mind. The human mind is endowed with an encultured understanding of many of the social phenomena that concern us as a result of its socialisation. This has its pros and cons: we already understand, in an informal but richly meaningful way, many of the underlying causes of social phenomena (e.g. everybody left the party early because of the open hostility between the host couple), but this pre-theoretic understanding may be misleading – it is certainly not 100% reliable. In the past this has led some social scientists to reject all accounts by the participants of social situations as “anecdotal”, leading to an attempted “objectification” of social science. As we will argue below we reject the exclusion of the anecdotal, arguing that, as evidence, it still has *priority* over theory, despite its difficulties.

These two differences suggest that good social science will be in some respects different from good natural science. But that is no reason to abandon the fundamental proposition that good science of whatever stripe is driven and constrained by observation and that the purpose of science is to *usefully* explain phenomena. An appropriate meaning of “usefully” in this context will be explored in section 6. Drawing on the natural sciences, we will say that *a necessary condition for social science to be good science is that it coheres with directly observable evidence in as* *many ways as possible*. When there are competing or alternative theories, analytical procedures or explanatory relationships, those that cohere better with more observed phenomena and with more kinds of observed phenomena will be deemed to be better. If the coherence relationship is nested in the sense that one framework or procedure coheres with a proper subset of the observed phenomena with which a second framework or procedure coheres, then the second is *in that dimension* better science than the first.

The foregoing is really no more than an attempt to state in a reasonably explicit way that good science coheres with observation and better science coheres in some sense better and with more observations. Of course, it may be that this is just one criterion and others will prove to be more important in the determination of a *useful* basis for social science. As a first step in the investigation of that issue, we require some examples of social scientific analyses that patently do not cohere with observation. We can then consider how further to investigate their goodness (or badness).

4. Examples of bad social science

The papers considered in this section are examples of work which, by virtue of their appearance in the leading mainstream journals of their fields, can be taken as good examples of leading edge analysis. These papers will be assessed in relation to their coherence with observed phenomena. We will argue that they are examples of bad science strictly in relation to that empirical coherence. This argument can be couched entirely in relation to the papers in question. Whether altenrative approaches provide the basis for a better science on some clear sense will be taken up in section 5 hwere we consider how social science can be useful.

Economics

We consider a paper (Etro 2004) chosen at random from among the theoretical apers in the most recent (at the time of writing) issue of *The Economic Journal*. This is a strictly theoretical paper concerning patent races.

Etro studies “a patent race where the patentholder has the opportunity to make a strategic precommitment to a level of investment in R&D. This may happen through a specific investment in laboratories and related equipment for R&D, by hiring researchers or in a number of other ways. In the case of ‘contractual costs’ of R&D, that is, when a fixed initial investment determines the arrival rate of the innovation, the interpretation of a strategic precommitment for the incumbent monopolist is very standard. The leader can choose to invest before the other firms and, since the leader is by definition the firm who has discovered the latest technology, it is reasonable to assume that such a discovery was associated with a first mover advantage in the following patent race. More generally, *our strategic assumption seems a natural one* since the patentholder can be easily seen in a different perspective from all other entrants in the patent race. Moreover the first mover advantage could be a consequence of an even small technological advantage, so that our arguments should be seen as complementary to those based on technological advantages.” [p. 282, emphasis added.] The analysis is based on “[t]he fact that patentholders do not invest in R&D [which] implies a continuous leapfrogging and no persistence of monopolistic positions between one innovation and another.” Etro asserts that this “is a quite counterintuitive picture of what is going on in the real world” [p.282] but justifies this "fact" by an appeal to an empirical finding by (Blundell, Griffith et al. 1999) "witnesses a positivbe rleationship between market power and innovation activity that is consistent with" the claimed fact.

The Etro paper is purely theoretical. It has no direct reference to any empirical observation. The only explicit empirical reference in the paper is to Blundell, *et al.* Blundell and his co-authors themselves set out their statistical analysis on the basis of equilibrium theoretical precedents. They are clear that the key empirical relationship of concern in their paper "between innovation and market share could be artificial for at least three reasons" which they enumerate. So they test the robustness of their results "by examining an alternative measure of innovative output (patent counts) and examining a particular industry in greater deail (pharmaceuticals)" – though this detailed study is also statistical.

Neither Etro nor Blundell and his colleagues are doing anything that is, in the context of mainstream economics, in any way ideosyncratic. Etro takes directly into consideration no evidence whatsoever. His sole objective is to elaborate and extend previous theoretical results that themselves were based on no direct observation of innovating firms. The only evidence Etro does consider is from Blundell, *et al.* who themselves are concerned to extend previously published regression techniques and who do not appear to be influenced and do not cite any studies of actual innovating firms.

As in physics, there is no guidance from economic theory regarding the means of analysing any particular case that does not have all of its *ceteris paribus* conditions satisfied. A difference between physics and economics in this regard is that physical theory is well validated observationally, frequently experimentally, while no economy has ever been observed to satisfy the strong equilibrium conditions of economic theory. So one way of describing the difference between physical and economic theory is that applications in physics start from a well validated theoretical base while economic applications cannot although both resort to pragmatic processes to bridge the gap between the conditions for which the theories are defined and the conditions in which they are to be applied.

The Etro and the Blundell, *et al.* papers – both in leading, core economic journals – are representative of theoretical and applied economic analysis and neither those papers nor any of the works they cite seek to validate their analysis by any means other than statistically even though non-statistical means are available. One frequently hears economists dismissing case studies as anecdotal and therefore untrustworthy evidence. The commitment to statistical analysis and rejection of case study evidence amounts to a refusal to validate theories and even procedures for bridging to empirical conditions in any way but one. This refusal to cross validate (Moss and Edmonds 2003) is what distinguishes economics (and, we shall see, sociology) from the natural sciences. If validation is the hallmark of good science, then the Etro and Blundell, *et al.* papers are representatives of bad science in economics.

We do want to be clear that we are not accusing the authors of being bad scientists. Blundell, *et al.* stated carefully the limits and the weaknesses of their analysis. Etro was dealing with equilibrium systems. He suggests several times that particular assumptions are "natural" to make but this is in the context of the theory and is not stated to be accurate as a description of any observed state of affairs. So, insofar as good scientists are careful not to claim more for their results than can be justified and are clear about the bases of their analyses, the authors of the papers considered here are good scientists. However, because they are operating in an intellectual environment in which validation is either not undertaken – as in the Etro paper – or is restricted to one inconclusive approach to validation – as in the Blundell, *et al.* paper – we would have to say that these authors are good scientists doing bad science.

Sociology

The example papers chosen from sociology started with the random choice of an article from the most recent (at the time of writing) issue of the American Sociology Society's *Sociological Theory*. This paper (Jerolmack and Porpora 2004) is a critique of a rational choice theory of religion due to Rodney Stark (Stark 1999).

Stark's argument began "with the assumption that people make religious choices in the same way they make other choices, by weighing the costs against the benefits." [p. 265] Religion "is the only plausible source of certain rewards…. Since the realization of some of the most valuable of these rewards is deferred to the afterlife or to other nonempirical contexts, religion entails a high level of risk of nonfulfillment." Stark adopts the position that, though individuals seek to maximise the value of rewards over costs, they do so without precision and are constrained by whatever information and understanding they have. A number of propositions are offered in the paper which, taken together, amount to the axioms of subjective expected utility theory applied to religious experience. The form of these propositions collectively constitute a bridging process from the subjective expected utility theory of economics to the sociology of religion.

Just like economists typified by Etro and Blundell *et al.*, Stark does not refer directly or indirectly to a single case of individual religious belief or experience. He cites some antropologists' descriptions of religious observances and observations about the nature of various religions. But he does not produce any evidence whatsoever about the reasons why any one individual has performed any action indicated by or predicated on religious belief.

This is in marked contrast to the Jerolmack-Porpora critique. For example, Jerolmack and Porpora argue at one point against Stark's assertion that all behaviour is selfish by appealing to the case of St Ignatius. "Ascriptions of motives are notoriously difficult to make. Certainly, Stark … identifies motives on which Ignatius might have acted. That hardly establishes that Ignatius did in fact act on them. In particular, it is very unlikely that Ignatius was motivated by … fame … among future Christian generations [since] Christians still expected an imminent end to the world…." [p. 150] They also argued that martyrdom in particular is not just a subject of personal disutility. "In addition to enhancing the credibility of a religion, rigorous religious practices like martyrdom may also create the changed emotional energy of group identity …." [p.151] The claim that religiosity is driven in part by interaction within a group is not supported by evidence drawn from personal accounts whether published or obtained for the purpose. So far, we are left with two argued assertions based on different assumptions and neither founded clearly in evidence about individual behaviour or accounts of specific social interctions. Nonetheless, though it is not a major part of their paper, Jerolmack and Porpora later an (p. 155) report secondary evidence including a statement by a member of the clergy clearly based on personal experience rather than anything accounted for by rational choice theory.

5. The role of social simulation in good social science

The articles and form of argumentation reported in section 4 are typical of papers from the core journals of two social sciences: economics and sociology. With one exception, towards the end of the Jerolmack-Porpora paper, all of the argumentation is drawn from unvalidated theory with *ad hoc* bridging procedures from the theories to their respective applications. Such *ad hoc* bridging appears to be common to all science. But the difference between the natural and the social sciences is that, in the natural sciences, one anchor of the theory is predominantly theory that is well validated – frequently by many scientists in many different ways – in contexts where their *ceteris paribus* conditions are thought not to have been violated. We need to be very careful here since we know of no way of determining that all relevant *ceteris paribus* conditions have been identified by the scientific community much less that they have all been satisfied. Even so, in the social sciences there is typically little or no attempt to validate theory. The balance between theory generation and theory validation is very different in the social as compared to natural sciences. Attempts to validate or invalidate theory by experimental economists have shown in the most influential and seminal cases that the theory – especially rational choice theory – is invalid.[8] More generally, accounts of individual behaviour of episodes of social interaction are dismissed as anecdotal evidence.

The problem with this situation is not just that it leaves social scientists with unvalidated explanations of social phenomena – though it certainly has that effect. As Cartwright has noted, natural scientific theories are also not validated in most applications even though they are very well validated in some circumstances where the *ceteris paribus* conditions are satisfied. But the difference is that natural scientists can use the validated theory as constraints in the design of their bridging procedures. At the very least, the validated theories can indicate a starting point for such designs. The essentially pragmatic procedures drawing on the theory can then be validated in application with widespread validation of these procedures giving us confidence in the application of physical science to, for example, civil engineering or the control of spacecraft sent to the moon and other planets.

The position in the social sciences could hardly be more different. As in the natural sciences, bridging procedures are typically constrained by theory but the theory is itself wholly unvalidated. A particularly good example is to be found in the econometrics of financial economics. Asset prices on any stock exchange or organised commodity market are characterised by clustered volatility. Moreover, there are no reported correct real-time forecasts of the volatile clusters or the post-cluster levels in financial market indices or macroeconomic trade cycles (Moss 2002). Econometricians have developed estimating techniques based on the assumption of time varying parameters of a population distribution from which the observed data is taken to be a sample. Specifically, the observed time series is assumed to be drawn from a normally distributed population with constant mean and varying variance. The variance at the time of any observation is itself a function of previously observed errors so that standard regression techniques can be used to model the time series of variances. To capture the clustering of extreme events, the variance of a time series

[8] The two most important such experimental studies were undoubtedly Allais and Kahneman and Tverski – both the rational for Nobel prizes in economics (i.e. Swedish Bank Prizes in memory of Alfred Nobel).

for any observation is treated as a function of previous error terms. Recent, large error terms tend to generate large variances. Because the observation is then drawn from a population with a larger variance, the probability of the observation being relatively far from the mean is greater than when the population variance is smaller.[9]

The motivations offered for particular methods of estimating time varying parameters are invariably related to rational expectations, the mean-variance representation of risk and risk aversion or some similar equilibrium notion from economic theory. There are, however, no microeconomic equilibrium models that generate either analytically or by means of simulation the sort of clustered volatility observed in the time series data. Experimental data shows that the mean-variance representation of risk is invalid and, in relation to rational expectations equilibrium models, we know of no case in which the assumptions that individuals have been shown all to have the same (or, indeed, any) utility functions and share the same – much less, the *correct* – model of the economy or any relevant financial market. The standard, naïve response to this sort of point follows Friedman's (Friedman 1953) classic claim that the descriptive accuracy of assumptions is irrelevant and all that counts is predictive accurace. Since there has been no demonstration of the predictive accuracy of the estimation of time varying parameters, even this argument fails. The fact is that the theory has never been validated since its *ceteris paribus* conditions fly in the face of common observation, common sense and experimental evidence and none of its bridging procedures (such as time varying parameter estimation) have ever been validated by their predictive accuracy.

In the absense of validated theories and validated bridging procedures, the social sciences are in a similar position to that of (say) physicists investigating electricity in the hundred years from the mid 18th century. In the second half of the 18th century, Benjamin Franklin, Joseph Priestly and Augustin Coulomb demonstrated experimentally a number of properties of static electricity. In 1812, Simeon-Denis Poisson (formulator of the Poisson distribution) published his demonstration that observations in electrostatics could be rationalised mathematically by defining the new concept of electrostatic potential. The identification and development of current electricity enabled the conduct of a wide range of experiments by Galvani, Volta, Biôt, Ørsted and Savart leading to the recognition of similarities and relationships between electricity and magnetism. Faraday built on this experimental base and, on the basis of his experimental results developed the concept of lines of electromagnetic force subsequently elaborated experimentally and formulated mathematically by Clerk Maxwell.

There are many such cases in the natural sciences where observation and experimentation lead to conceptualisation. We know of no such cases in the core of mainstream economics or sociology, where the conceptualisation has tended to come first. Although there were observation-based theoretical developments particularly in industrial economics (Wilson and Andrews 1951), these never became mainstream and have now been largely forgotten.

We speculate – perhaps too generously – that one reason for social scientists to avoid detailed evidence about social processes and behaviour has been the lack of any equivalent of the experimental and observational (*e.g.*, the telescope) techniques

---

[9] Bollerslev (Bollerslev 2001) identifies the core econometric processes of relevance here to be the ARCH process (Engle 1982), the GMM process (Hansen 1982) and GARCH (Bollerslev 1986).

developed by the natural scientists. These frequently both demonstrated prior scientific understanding (for example, optics and the telescope) and made possible further observations used to modify or extend prior understanding. Our proposal is that agent-based social simulation can serve in the social sciences some of the functions of the experimental and observational apparatus invented and employed in the development of electromagnetics and field theory. There is, however, an essential caveat here. The design of the agents must not be constrained by any prior unvalidated theory[10]. The essential feature of software agents devised for purposes of social simulation is that they should be validated as good descriptions of the behaviour and social interaction of real individuals or collections of individuals.

Whilst it is clear that mathematics turned out to provide a good basis for describing observed and inferring the existence of as yet unobserved phenomena in the physical sciences, there is no such experience with mathematics or any other analytic systems in the social sciences. The behaviour we observe is typically best described qualitatively and the reasons individuals give for their behaviour is almost invariably qualitative as well. Therefore, validation of software agents as good representations of real individuals is facilitated by having the agents perceive events specified by qualitative descriptions, maintain the qualitative terms in processing those perceptions and then to act in ways that can be described qualitatively. Insofar as possible, the qualitative terms should be those used by the individuals they describe. A natural way to maintain this qualitative link between the language of actors and the language of the agents is to use production systems whereby the rule conditions describe the perceptions by the agents, processing is governed by some inference engine and the actions are specified by the consequents of the rules. We might find that algorithms that best support model validation are consistent with some class of formal logics, but that is not an issue of direct concern here. What is of concern is that no formal system or algorithm (including logics) should be chosen and imposed on any class of models without some independent evidence that that system does not constrain model design in ways that distort the descriptions by relevant individuals of the reasoning that leads to specific behaviours.

We do not insist that individuals always fully understand the reasons for their own behaviour. To the extent that they do not, it would be wrong to assert that models designed to capture the rationales given by the individuals being modelled are in any wider sense good descriptions of the social processes being modelled. In any event, the modeller does not usually have access to all of the relevant stakeholders in any social process. In the absense of such access, there are obvious limitations on the scope for the qualitative validation of models at micro level. Such limitations on validation will also limit the uses to which the models can be put with confidence.

6. Implications for social science research

We do not argue that all good social science is agent-based. Nor do we argue that all agent-based social science is good science. Far from it! We do argue that good

[10] Of course, it is impossible to completely avoid all prior unvalidated theory in the construction of agents, but this should not be a constraint. That is to say that one should be flexible about the way agents are constructed, paying particular care to ensure consistency with any available accounts, and be willing to change the design when and where it comes into conflict with evidence.

social science can be facilitated by agent-based social simulation because of the wider possibilities for validation it facilitates.

So far, we have been arguing in general terms about the characteristics of good science. We have previously addressed issues of abstraction and application. One of us (Moss 2002) has produced evidence demonstrating that supermarket sales values and volumes for fast moving consumer goods have the same statistical signature of clustered volatility as is found in financial markets. A range of agent-based social simulation models produce the same statistical signature at system level. These models are characterised by agents that are resistant to changing their behaviour unless some significant stimuli are present, they interact with other agents, agents influence but do not commonly imitate one another and the social processes captured by the model dominate the behaviour and macro level statistics produced by the model. While we have not produced counterexamples, we do see the need for further investigation of: the generality of the empirical finding of clustered volatility that *cannot* be forecast in real time; of how widely the agent and mechanism (interaction) design *is* a descriptively accurate representation of actual behaviour and social interaction; and the generality of the finding that unpredictable clustered volatility *follows from* agent and mechanism designs of the type indicated.

Another area we have investigated relates to social policy in complex conditions. Areas of application include integrated water resource management, the social costs of carbon emissions and, incipiently, the social determinants and consequences of land use change.

In policy applications, the purpose of the social simulation models is to capture the behaviour and social interaction of stakeholders as described by them (that is also consistent with known data models). In general, we would not expect stakeholders with conflicting goals to agree on either the reasons for the behaviour of other stakeholders or the consequences of their behaviour. Conflicting goals can result from different understandings of the social impacts of individual behaviour or, possibly, the different understandings can result from different goals. Goals and beliefs about the nature and impacts of social processes might well develop together. Consequently, we do not assert that different beliefs are necessarily held in order for antagonists to justify conflicting goals. We subscribe to the adage that one should never ascribe to malice or cupidity what can be adequately explained by stupidity, ignorance or narrow-mindedness. The important point here is that qualitatively specified social simulation models can be implemented in collaboration with stakeholders to explore the relationships between their own goals and their understandings of their social and, where relevant, their natural environments.

Where there is goal conflict, models can be implemented to capture the different understandings of the social and natural processes involved. Each set of stakeholders can explore the consistency between their prior beliefs and the simulation results from the models designed to capture those beliefs. Sometimes this can lead stakeholders to clarify, refine or modify their views concerning the nature of relevant social processes. We imagine, though with limited experience to date have not actually found, that in some situations the process of designing and validating the simulation models can lead to a new understanding of the social and natural environment that will facility the resolution of conflict. We have been much influenced in this view by the pioneering work in Senegal of colleagues from CIRAD (Barreteau, Bousquet et al. 2001).

We do have experience of building models to assist in the development of scenarios when there is no definitive scientific result. Where climate change is concerned, for example, global circulation models are either too simple for their *ceteris paribus* conditions to be satisfied or they are composed of several submodels – for example models of ocean and of atmospheric circulation – such that each violates the *ceteris paribus* conditions of the other. In these cases, the natural science models are no more definitive than the social models. There are some well validated natural science theories to form one end of a bridging process to models of climate change but there is insufficient experience to allow for any meaningful validation of these bridging processes. The social sciences provide neither validated theory nor validated bridging processes.

The best we know how to do in these circumstances is to integrate rough and ready models of the natural processes with social models designed and validated with stakeholder participation. The natural process models must allow for a wide range of natural phenomena both as a reslt of interaction among natural processes and interaction between natural and social processes.

Prediction is no longer an issue in applications to such complex problems. Exploration of the problem space with more formal and therefore more precise (if not more accurate) analytical apparatus is, for now and the foreseeable future, the best we can do.

7. Conclusion

This paper can be interpreted as propounding a neo-positivist position – that the social sciences should be more like the natural sciences. This is accurate in some respects but not in others. We argue that social science would be more successful if it learned from the more fundermental causes of the success of the natural sciences, but specifically argue against the shallow borrowing of the style of natural science models and the unthinking applications of techniques from the natural sciences.

In particular we *do* argue for the following:

- The fundermental *priority* of observation and evidence over models and theory
- The importance of the multiple validation of theory
- Not relying on theory that is not sufficiently validated (and certainly not as a justification for further theory generation)
- That evidence should not be excluded if at all possible, including anecdotal evidence from stakeholders
- That much more effort should be expended towards developing new techniques for the observation of social phenomena
- That much descriptive modelling at a low, concrete level will probably be necessary before successful and useful more general theory can be developed
- That some of the modelling and description needs to be of a formal (but probably computational and not analytic) nature so that we know what we are comparing and talking about

- That agent-based simulation can help facilitate the above

And we argue *against* these:

- The unthinking and inappropriate use of analytic and statistical techniques
- The unjustified assumption that there is always a single social reality “out there” to be represented and understood rather than alternatives that are continually being constructed
- That there will necessarily be unchanging and/or universal social laws underlying social phenomena
- That any particular technique (including agent-based simulation) will always be appropriate for all modelling tasks, rather the domain should guide the choice of technique from a large palette of possibilities
- That one can objectify social science, and hence exclude subjective and meaningful entities as reported by the people concerned
- That all aspects of social phenomena are amenable (even in principle) to scientific approaches

References

Barreteau, O., F. Bousquet, et al. (2001). "Role-playing games for opening the black box of multi-agent systems: method and lessons of its application to Senegal River Valley irrigated systems." Journal of Artificial Societies and Social Simulation 4(2): <http://www.soc.surrey.ac.uk/JASSS/4/2/5.html>.

Blundell, R., R. Griffith, et al. (1999). "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms." The Review of Economic Studies 66(3): 529-554.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." Journal of Econometrics 31: 307-327.

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

Cartwright, N. (1983). How the Laws of Physics Lie. Oxford, Oxford University Press.

Chattoe, E. (2002). Building Empirically Plausible Multi-Agent Systems: A Case Study of Innovation Diffusion. Socially Intelligent Agents - creating relationships with computers and robots. K. Dautenhahn, A. Bond, D. Canamero and B. Edmonds. Dordrecht, Kluwer.

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

Einstein, A. and L. Infeld (1938). The Evolution of Physics. New York, Simon & Schuster.

Engle, R. F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." Econometrica 50: 987-1007.

Etro, F. (2004). "Innovation by Leaders." The Economic Journal 114(495): 281-303.

Friedman, M. (1953). The Economics of Positive Methodology. Essays on Positive Economics. Chicago, University of Chicago Press.

Giere, R. N. (1988). Explaining Science: A Cognitive Approach. Chicago, University of Chicago Press.

Hansen, L. P. (1982). "Large Sample Properties of Generalized Method of Moments Estimators." Econometrica 50: 1029-1054.

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

Jerolmack, C. and D. Porpora (2004). "Religion, Rationality, and Experience: A Response to the New Rational Choice Theory of Religion." Sociological Theory 22(1): 140-160.

Kuhn, t. S. (1992). The Copernican Revolution: Planetary Astronomy in the Development of Western Thought. Cambridge MA, Harvard university Press.

Longair, M. S. (2003). Theoretical Concepts in Physics. Cambridge, Cambridge University Press.

Mahon, B. (2003). The Man who Changed Everything: The Life of James Clerk Maxwell. Chicester, John Wiley & Sons Ltd.

Moss, S. (2002). "Policy Analysis from First Principles." Proceedings of the US National Academy of Sciences 99(Suppl. 3): 7267-7274.

Moss, S. and B. Edmonds (2003). Sociology and Simulation: Statistical and Qualitative Cross-Validation. Manchester, Centre for Policy Modelling.

Stark, R. (1999). "Micro Foundations of Religion: A Revised Theory." Sociological Theory 17(3): 264-289.

Suppes, P. (1962). Models of Data. Logic Methodology and the Philosophy of Science: Proceedings of the 1960 International Congress. E. Nagel, P. Suppes and A. Tarski. Palo Alto, CA, Stanford University Press: 252-261.

White, M. (1997). Isaac Newton: the Last Sorcerer. London, Fourth Estate.

Wilson, T. and P. W. S. Andrews, Eds. (1951). Oxford Studies in the Price Mechanism. Oxford, Oxford University Press.