Economics and Philosophy, **18** (2002) 5–16

SYMPOSIUM ON *MARSHALL'S TENDENCIES*: 1 HOW MODELS HELP ECONOMISTS TO KNOW

Mary S. Morgan
University of Amsterdam and London School of Economics

Over the last 40 years or so, economics has become a modelling science: a science in which models have become one of the main epistemological tools both for theoretical and applied work. But providing an account of how models work and what they do for the economist is not easy. For the philosopher of economics like me, struggling with this question, John Sutton's views on the nature and design of economic models and how they work is indeed thought provoking. Because of my own interests, my review of Sutton's book: *Marshall's Tendencies: What Can Economists Know?* will focus on three related issues that I found especially intriguing in his treatment of the role of models in modern economics. The first is the way that Sutton's account fits with my own reading of the history of twentieth-century economics, namely that the focus of economic explanation has moved during the last century from 'theories' and 'laws' to 'models' and 'mechanisms'.[1] The second is to understand the epistemological connotations of Sutton's 'class of models' view. The third is to explore what Sutton means when he says a model 'works'. These three questions roughly coincide with the material presented in Chapters 1, 3 and 2 respectively of the book. As we shall see, Sutton gives us a practitioner account of applied economics which can fit within the

This commentary on John Sutton's book draws also from my attendance at his graduate-level course on the economics of industry during 2000–2001 which related the ideas in his book to his applied work in greater depth.

[1] Although I might dispute some of the details of Sutton's history, his general story fits easily into the broad claims I express here.

https://doi.org/10.1017/S0266267102001025 Published online by Cambridge University Press

standard terms used by philosophers of economics on theory-testing, but which reveals a number of novel elements for philosophical analysis.

THE STANDARD PARADIGM: FROM LAWS TO THEORIES TO MODELS TO MECHANISMS?

Sutton begins his work with the subject indicated in his title: Marshall's Tendencies, and spends the first chapter giving an account of the changing views in the twentieth century about his subtitle: What Can Economists Know? Although it is something of an anachronism to depict Marshall and Edgeworth as arguing over the methodology of *modelling*, it is a useful device for Sutton. His history of the past century, armed with concepts from the present, enables him to make a succinct account of how economists have thought about the problem of finding out about their world. This takes him on to his own position without too many detours. An unintended consequence of this retrospective review is that it reveals just how far the picture has changed during the twentieth century: from Marshallian talk of laws and theories, economists like Sutton now think in terms of models and mechanisms.

Sutton contrasts Marshall's view that economic laws are 'tendency' laws, capturing the main 'mechanisms' (p. 4) of economic events, and therefore predictive of outcomes despite the messy world in which they occur, with the criticism from Edgeworth. For Edgeworth, such laws or descriptions might work for some cases, but not for others, where other 'relevant' factors (which might be unknown, immeasurable or unobservable, etc.) would likely swamp the tendencies, making outcomes indeterminate. Sutton offers two interpretations of Edgeworth's objection: either it leads us to a 'class of models' or to a 'supermodel' view. The former describes the idea that we should aim for a set of models compatible with the evidence, but this concept is fluid in the early pages; it is only developed and gains analytical power in Chapter 3. The latter supposes a more general model which encompasses the others, an idea which we are familiar with from recent econometrics literature, though perhaps the connotations here are more that of a 'meta-model'.

According to Sutton's history, the 'standard paradigm' which emerged in the mid-twentieth century was a more sophisticated version of the early econometricians' interpretation of the Marshallian view. I would associate this most clearly with Tinbergen's work, for it was Tinbergen who adopted the terminology of models, and in a practical way, worked out many of the early ideas about their role. Sutton counts this view as one which supposes that there are 'true models', though these are difficult to pin down with measurements for various reasons. Haavelmo added the probabilistic interpretation which, as Sutton suggests, can be understood as softening the commitment to 'true models' by supposing that both models and the probabilistic approach were epistemological devices, and by adding clear notions of structure for both the economic 'true' model and the statistical noise. The standard (mid-century) paradigm for applied economists was, therefore, partly instrumental – in its commitment to models and to the probability approach – but retained a strong realist-empiricist stance that one could use evidence to learn about the underlying structures in the world.

Sutton's own later-century view retains some of the commitments of the standard paradigm. He, too, believes that economists need to explain as well as predict, that empirical work is essential to know something about the economy, and that models provide the instrument for this.[2] But there is also a considerable difference: whereas the standard paradigm hoped for 'complete' models with the aim of uncovering the structures that underlay economic phenomena, Sutton uses models to gain knowledge of mechanisms. This difference requires some explication.

One of the points raised but, I suggest, under-utilized by Sutton, was that this mid-century econometric version, or the 'standard paradigm' way of fitting theories to the world via models, hinged on the possibility of isolating – identifying, locating and measuring – the autonomous relations from out of the confluent relations. Confluent relations were reduced form relations, or the kind of relations fitted to time-series samples of data, whereas economists of the day sought the autonomous or structural relations between the variables which created those time patterns or reduced-form relations. For Haavelmo and his contemporaries, autonomous relations were essential to the working of the economy; they were the ones which would work more or less regardless of whether all the elements of the economy were going on the same. Haavelmo used, as an illustration, the performance of a car: the relation between the accelerator pedal and the speed of the car was a 'functional' relation with little autonomy with respect to the operation of a car, while the general laws of thermodynamics held a high degree of autonomy in this case.[3] (The use of 'functional' here seems to imply both fulfilling a role and being mathematically describable by a relation.) The general laws governing the 'functional relation' had greater autonomy, for they provided an accurate description of 'some parts of the mechanism [the automobile] *irrespective* of what happens in some *other* parts' (p. 28, his italics). But though these autonomous relations might be fearfully complicated, or not easily knowable or derivable from observations of

[2] Both Sutton and the followers of the 'standard paradigm' therefore diverge from the more thorough-going instrumentalist view of models associated with Friedman.

[3] Haavelmo also relates his definition of autonomy to a 'class of models' approach. Sutton footnotes the definition, but it would have been interesting to explore the relation between his own 'class of models' and Haavelmo's terminology and concept here.

the performance of cars, they were the relations econometricians should aim to find.

In the standard paradigm, models were devices to get at the autonomous (probably hidden) laws or governing relations that make the mechanisms work. For Sutton, models are the epistemological instruments to access mechanisms in the economy and it is the mechanisms that are responsible for the regularities in economic phenomena which we observe in fields such as industrial economics:

> That such regularities appear in data sets that span different industries, which vary hugely in many characteristics that we cannot hope to control for, suggests that some very strong and robust mechanisms are at work here. A major goal of research in this area is to uncover the natural description of what is driving such apparent regularities, in the hope of thereby obtaining indirect evidence on the workings of the competitive mechanisms that are melding these patterns in the data. (Sutton, pp. 65–6).

In Sutton's terms, we build models to represent and capture the mechanisms of economic life; in Haavelmo's paradigm, they built models to capture the laws which governed the mechanisms.

Though Sutton does not really define 'mechanisms', they have recently been discussed seriously in the philosophy of science literature by Machamer, Darden and Craver (2000). They defined them as follows: 'Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions.' (p. 3). Their work, in the context of explanations in biology, moves the agenda away from the purely mechanical notions associated with the terminology while at the same time appealing to a combination of functionalist and causal thinking, to a commitment to the importance of both the entities and the process or activities, and to the observation of local regularities which might not be readily describable as law-governed. All this seems compatible with Sutton's view. And, both they and Sutton might agree that the relation between the accelerator pedal and the speed of the car represents a mechanism: it fits with Machamer *et al*'s definition; it provides reliable and useful knowledge of a regularity; and it holds over a range of circumstances. In contrast, for Haavelmo, this regularity was a confluent relation with little degree of autonomy. Nevertheless, like Haavelmo's autonomous relations, Sutton requires that his mechanisms have a degree of stability and independence over many circumstances and these features are important for his applied work.

Another element which features in Sutton's account of the standard paradigm, but which again might do useful work in his later account of the 'class of models' is the issue of variability. Sutton observes that, in the case of econometrics, we need variability in the factors we suppose to be explanatory factors, but that we need very low variability in background conditions: that is we need homogenous time periods or cross sections. Variability in the former allows us to assess their influence; variability in the latter would prevent such an assessment. But variability appears to have a different quality and another role in the case of industrial economics, the case that figures in Sutton's 'class of models' account. There, the economic world is presented as looking less like a physical world operating according to some general theories (for which there are 'true models') and more like a natural history world in which there are many types of cases. In the world of firms and industries, there is very considerable natural variability – so that trying to obtain a 'general' model, in the sense of a model which will cover all cases, does not get us very far in describing this variation. For example, it is a contradiction of what we know to model all industries as being either in a perfectly competitive or monopolistic state. A 'complete' model, which might account for the differences we observe by the addition of many other factors leads to the 'supermodel' approach, one which retains a commitment to a 'true' model. But not only is this an impractical recipe for research, it does not even fit the material for it is not a question of making corrections or additions to the pure, idealized cases. Sutton takes the challenge to be how to build and use models to capture some of this natural variability without striving for either unobtainable completeness or explanatorily weak generality.

JOHN SUTTON'S 'CLASS OF MODELS' VIEW

Abandoning both the hope of easily or often finding 'true models' (or models that 'work', as the Black–Scholes case, discussed later) and the search for complete structural models, and yet embracing Haavelmo's (1944, p. 3) dictum that our models are 'our own artificial inventions' not 'hidden truths to be discovered', leads Sutton to his 'class of models' approach (in Chapter 3). It is not easy to gather from his earlier comments in Chapter 1 just what this might mean, but one of Sutton's talents is a gift for picking up and using analogies. He used Marshall's own analogy of the tides to explore the limitations of Marshall's view. For his own work, he uses the analogy of Carnot's investigations of steam engine efficiency. I take it that understanding this analogical example is essential for understanding what Sutton has in mind for the 'class of models' approach and, thus, how we should interpret his own methodology of applied economics.

Carnot's *diagram* of a steam engine was somewhat abstract but at the same time somewhat particular. It was particular in the sense that he represented a particular *kind* of steam engine (or mechanism) in various states of activity (see Sutton's Figure 3.1) but the diagram abstracted the main elements rather than giving a fully detailed drawing of the engine.

https://doi.org/10.1017/S0266267102001025 Published online by Cambridge University Press

Carnot derived, from this first diagram, a *graph* showing pressure related to volume (see Sutton's Figure 3.2) which embodied an outline representation of the work done during the operation of the machine shown in the first diagram. There are a number of important points Sutton made about this example which bear repeating and expanding. First, Carnot was working with a theory of heat which was soon invalidated, but as Sutton tells us, the representation of work done in the graph relating pressure to volume remained valid as a description, as a tool to compare the performance of different kinds of engines, and possibly (I imagine) even as a tool for designing better engines. Second, both his relatively abstract machine diagram and his graph of work done could be compared with much more detailed and realistic drawings of *specific* engines available at that time. Each of those more specific representations enabled the engineer to describe the behaviour of, and elicit the best performance from, only that one specific machine design. These drawings were no use for comparing different kinds of steam engines or giving access to a general description of the work done by such machines. In contrast, Carnot's somewhat more abstract diagram of a machine could be applied to all engines of that same kind or type, and his graph of efficiency, it later appeared, could be plotted for all kinds of engines. Thirdly, we might note that the graph used information on measurable elements (volume and pressure) to provide a picture and an indirect measure of a less measurable attribute, namely efficiency.

Sutton's 'class of models' idea operates in the same in-between level: between a specific model for each specific case and a general model which tries to capture all cases. The former, as Sutton himself drew my attention to, was the situation in industrial organization around 1990 when it appeared that theoretical models proliferated, and each empirical case came to be described by a purpose-built model. Game theory models in industrial organization had proved extremely fragile, and attempts to provide some general model at the level of application had proved problematic even though the field had some well-recognized empirical regularities. This lack of generalizing power in the theory and the narrow applied scope of models meant that models could not be predictive, and their explanatory power was extremely limited. 'Exemplifying theory', as Fisher (1989) called it, might be compared with an experiment which works once: interesting, but not the source of more generally applicable knowledge. While, as Sutton suggests here, game theory may have proved very important in moving the field of industrial economics onto a new trajectory of research, its importance has not been in providing good general descriptions, but only in changing the nature of the models used. Sutton's class of models approach retains game theory but can be seen as his response to the impasse caused by its lack of generalizing power.

The 'class of models' approach Sutton applies in industrial economics follows the same modelling strategy used by Carnot in his steam engine case. That, recall, has two elements: a schematic representation of a type of engine (the diagram) and a related description of characteristics of the performance of the engine (the graph). We can call these a representational model and a derived graph of measurable elements. In Sutton's case, industries are represented by a particular type or kind of model, namely a multi-stage game-theoretic model. I presume that this is what he calls a 'class of models', a group of models which share, like kin, a set of characteristic features. The generic model of the class is not supposed to offer a complete representation of one industry and its market, or be general for all industries. It is designed to lie between these: Sutton aims to capture in some part, some of the details of the reality for a class (possibly a broad one) of industrial structures and decision making. Certain design features embedded in this type of model, and thus common to the whole class of such models, enable Sutton to abstract certain characteristics of the industry which he terms 'viability' and 'stability'. Using this model design, he proposes some theorems about how these abstracted characteristics relate, via mechanisms (e.g., the mechanism relating R&D expenditure to sales revenue), to certain particular outcomes which can be observed. That is, these abstract characteristics derived from the economic model can be related to particular outcomes in the counterpart graphs of concentration against market size, rather as the 'work' done by a steam engine is graphed in terms of its measurable volume and pressure. The efficiency of the steam engine graphed out depended on the particular characteristics of the machine (mechanism) represented in the diagrammatic model; similarly, the viability and stability of industries, located in Sutton's graphs of market size and concentration ratios, depend on the particular mechanisms of industrial activity represented in the economic model.[4]

We might reasonably ask how this class of models idea gets us around the well-known joint problem of theory testing and model selection? This was the problem that Sutton had earlier argued bedevils work in the standard paradigm and forced him to look for a way around the standard approach. Here, I believe, we can draw a stronger conclusion than does Sutton himself (in his last chapter) from the implications of his class of models approach.

My own interpretation of testing in relation to his approach goes as follows. What gets tested here is not the representational model of the

[4] Both the Carnot case and Sutton's own use of models fits well into the 'models as mediating instruments' account that Morrison and I (see Morgan and Morrison, 1999, Chapter 2) provided to understand how models are used as investigative tools in theoretical and applied science. The in-between character of their sort of model is particularly interesting and deserves further analysis.

industry, and to the extent that economic theory is bound up with that representation, it remains untested. (Nor is there an estimation procedure as conventionally found in the standard paradigm of econometrics.) The 'test' of the model lies in the empirical work which compares the patterns predicted from such a class of models with the measurable outcomes obtained in data on industrial structure, industry by industry. This is a kind of characteristics test which we have seen much of in the new classical macroeconomics. In that domain, testing has often proved disappointing, because there is no inference link back into the theory or model from the test (see Kim *et al*, 1995), and, as John Sutton remarks, we easily fall into the problem of there being lots of models compatible with some observed data regularities or patterns. Others have interpreted such tests as being concerned with measurement functions not testing (see, Boumans, 2002). But there is an important difference in Sutton's work, for the characteristics tests here can be used to define the difference between kinds of industries represented in the class of models. The approach offers a way of pulling apart, and making sense of, data which appeared irregular or ambiguous, or were otherwise uninformative of behaviour under the older more general model measurement and testing regimes which looked at industry averages and aggregates. Characteristics testing here can be informative because it enables the economist to distinguish between types of industries in ways which allow inference back to the mechanisms (entities and their activities) in the model, rather than being passively accommodating to the model's predicted correlations, as has often been the case in characteristics testing in macroeconomics.

IN WHAT SENSE DO MODELS 'WORK'?

Sutton is a modest man as we can see in his repudiation of the 'true model' ambitions of the standard paradigm. He takes this stance because he finds it unhelpful for applied work where we often face the joint problems of theory testing and model selection. On the other hand, unlike those swayed by extreme instrumentalism or extreme relativism, he believes we can know something about the economy, and like Tinbergen, he puts immense energy into modelling as a way of finding out things. For example, Sutton writes: 'the bulk of empirical work in economics is not concerned with theory testing as such. Rather such work is investigative in nature' (p. 92). For Sutton, successful models are described not in terms of theory testing or model selection, but as models that 'work'.

But what is meant by saying 'models that work'? This is a classic practitioners' term, and we need to pay attention to the colloquialism here. Economists say that a 'model works' just as a scientist in another

https://doi.org/10.1017/S0266267102001025 Published online by Cambridge University Press

field might say an 'experiment works'. What are the connotations? For an experimentalist, it suggests that the experiment had been set up and run in such a way that results were forthcoming, that something was learned from the experiment. Philosophers of science find it as difficult to characterize learning from experiments as they do to describe learning from models, but I think it would be a mistake to assume in either case that such a claim is necessarily concerned with theory testing, for then the colloquial claim would surely be something about a theory or hypothesis being proved 'correct', or a hunch about something being proved 'right'. The notion that a model or experiment 'works' has quite other connotations. Philosophers of science, in recent thinking about experiment, have focused on the creative, explorative role of experiments, in which getting something to 'work' refers to the ability to successfully manipulate something or to produce new substances (as in biotechnology and chemistry). In the science studies literature on experiment that followed Hacking's (1983) and Franklin's (1986) seminal works on experiment, 'theory testing' is some way down the list of immediate experimental aims. I suggest that in several respects, getting models that work may be much like getting experiments to work (see, Guala, 2002 and Morgan, 2002).

Sutton does not give a general account of what he means by a model that 'works', but from the two examples he discusses in detail in Chapter 2, it seems he has in mind something like the following: A model which 'works' is one which succeeds in providing an intelligible, plausible, succinct description or explanation of a case (or type of case) at hand, which is also precise enough for successful empirical work or for policy design. Sutton claims that models that work in economics are rare, and so it is worth exploring the circumstances under which he suggests such models do 'work'. Readers may be surprised to learn from the two examples that Sutton gives: that of futures trading (Bachelier to Black–Scholes) and of auctions for oil-field tracts, that these models work because we know what they are already. We thus remain in the world of 'the true model' because scientist and scientific subject (the 'agent') share knowledge of the true model. It appears that this joint knowledge is a circumstance that attends these two examples, not necessarily a condition for models to work, and the first response to this characterization of the circumstances is that if these are the conditions for a model to work, they are going to be extremely rare – a sentiment that Sutton expresses himself.

The second response to this view, that models will work where both economist and economic actors know the model, is that it makes the modelling exercise pointless. But this is not so: in Sutton's first example, the futures market model is useful because we need to estimate certain parameters in it. In his second case, we have a game-theoretic argument based on the rules of the game and the situation, but we need to test the power of the reasoning by checking the model-predicted outcomes with some empirical work. These models can work as a vehicle for measurement or for theory testing because, knowing what the structure of the model is, we do not have the model selection problem. That is, knowing the model does not mean we know everything about the case or situation, there are still some things we need to find out. It is because we are able to do this, to gain more knowledge about the economy from using the model, that we are able to say 'the model works'.

But surely, the reader now wants to argue: Sutton does think that his models of industrial organization 'work' in some sense, although he does not know the true model, and nor, probably, do the agents he is modelling. There are cases where the standard paradigm works: '. . . happy situations in which outcomes are driven by a single market mechanism, whose operation is robust enough to override all secondary influences, [where] we may find a clear and sharp pattern emerging in the data whose interpretation is uncontroversial' (p. 23). Where it does not work, the economist has to design models to create, in the model-experiment, the same kinds of circumstances as in those happy cases. The industrial economics work in Chapter 3 shows how John Sutton sets about this. An important element of the answer here relates to the ability to choose or design a class of models which give derivable and distinguishable implications at the level of measurables and observables. Only models which enable empirical work to isolate something about the mechanisms of economics have the potential to work, for only these have the potential for empirical measurement or testing in ways which allow inferences back to the model (e.g., via characteristics tests). This condition might seem both self-evident and banal, as well as lacking in ambition, but though the ambition may be more attainable than the supermodel aim, it does not make applied economics any easier to do.

Sutton is a fund of analogies – it is one of the pleasures of reading his book. So let me use one here to emphasize this conclusion. I suggest that a useful analogy for the status and role of economic modelling with respect to understanding the behaviour of industries, is to compare it with the medical profession's understanding of the workings of joints. Over the last forty years or so, the medical profession has learnt to understand the hip joint sufficiently well to create artificial versions which provide highly effective working replacements. But knees remain problematic and feet too difficult to handle. The general principles of mechanics and of medicine only take one so far in dealing with the joints of the body. We need to know a lot of local particular detail and we need a rather simple case to get artificial joints to work. Similarly, the general principles of economics: rationality, competition, equilibrium etc., only take one so far in understanding how parts of the economy work.

Nevertheless, there are some bits which we can treat in relative isolation, and which we seem to have a really good grip on, where our models work well enough to use them for measurement and policy design. For, just like medics and engineers who use models to seek understanding of the way joints work, economists seek to understand their world by building models of the mechanisms, and hoping that in due time they will come to the point where some of the models work.

Sutton portrays the process of gaining economic knowledge as one without a rigid methodological recipe, but with a variety of flexible tools in the form of models. He works in an economic world full of particular cases, a few of which yield to straightforward analysis by a skilled and inventive economist, but most of which do not. Of those that do not, the most we can expect (at the moment) is to isolate mechanisms which operate across a number of the particular cases, across a class of cases.

CONCLUSION

Economists, both in their practice of applied economics and in their writings about methodology, rarely move beyond the accepted paradigm of their day. There are exceptions, among whom, Tinbergen and Haavelmo – the two economists who established the 'standard paradigm' in applied economics – present an interesting comparison. Tinbergen was the greatest applied econometrician of the early generation, and can be held responsible, through his practical work, for establishing the concept and importance of modelling in econometrics, though his writings on methodology were commonplace and his defence of his approach weak. Haavelmo's methodological tract of 1944 for econometrics is brilliant, but he hardly did any applied econometric work and that work certainly did not shine. Sutton may count as another exception, but one who puts methodological change into practice in his applied work and writes about those changes with insight. He is deservedly admired, by both theoreticians and applied economists, for his work in industrial economics. For a philosopher of science, the pleasure of reading this book is that Sutton, by writing with great intuition about what he does well, has shown how we might move around the standard paradigm that Tinbergen and Haavelmo set up. Sutton holds to their ideals as a distant dream, but meanwhile fashions a less ambitious programme which at least promises to give us some local knowledge, local knowledge gained from models that work to describe mechanisms, and that might in time prove to add up to something bigger or wider.

REFERENCES:

Boumans, M. 2002. 'Calibration of models in experiments'. In *Model-Based Reasoning: Science, Technology, Values*, pp. 75–93. L. Magnani and N. J. Nersessian (eds.). Kluwer

Fisher, F. M. 1989. 'Games economics play: a noncooperative view' *RAND Journal of Economics*, 20, 113–24

Franklin, A. 1986. *The Neglect of Experiment*. Cambridge University Press

Guala, F. 2002. 'Models, simulations, and experiments'. In *Model-Based Reasoning: Science, Technology, Values*, pp. 59–74. L. Magnani and N. J. Nersessian (eds.). Kluwer

Haavelmo, T. 1944. 'The probability approach in econometrics'. Supplement to *Econometrica*: 12

Hacking, I. 1983. *Representing and Intervening*. Cambridge University Press

Kim, J., N. De Marchi and M. S. Morgan. 1995. 'Empirical model particularities and belief in the natural rate hypothesis'. *Journal of Econometrics*, 67:81–102

Machamer, P., L. Darden and C. F. Craver. 2000. 'Thinking about mechanisms'. *Philosophy of Science*, 67:1–25

Morgan, M. S. 2001. 'Model experiments and models in experiments'. In *Model-Based Reasoning: Science, Technology, Values*, pp. 41–58. L. Magnani and N. J. Nersessian (eds.). Kluwer.

Morgan, M. S. and M. Morrison. 1999. *Models as Mediators*. Cambridge University Press