

SYMPOSIUM ON MARSHALL'S TENDENCIES: 5 SUTTON'S CRITIQUE OF ECONOMETRICS

KEVIN D. HOOVER

University of California

Through most of the history of economics, the most influential commentators on methodology were also eminent practitioners of economics. And even not so long ago, it was so. Milton Friedman, Paul Samuelson, Trygve Haavelmo, and Tjalling Koopmans were awarded Nobel prizes for their substantive contributions to economics, and were each important contributors to methodological thought. But the fashion has changed. Specialization has increased. Not only has methodology become its own field, but many practitioners have come to agree with Frank Hahn's (1992) view that methodology is a distraction to the practitioner, best left to the professional methodologists and philosophers, and of little practical import even when delivered from their pens. John Sutton's lectures, *Marshall's Tendencies: What Economists Can Know*, is a welcome return to the older fashion, for Sutton is an eminent practitioner of game theory and industrial organization. One of the main themes of these rich and nuanced lectures – and the one on which I shall focus – is the relationship of economic theory to econometric evidence. Sutton's reflections on econometrics appear to arise from the darker recesses of his practitioner's soul. While he affects a sunny disposition and ends on a hopeful note, his analysis articulates the lurking fear that econometrics is a hopeless project and that economics has little to learn from the interaction of theory and econometrics. Sutton's book is like a play in which virtue triumphs, but the villain gets all the good lines.

1. SUTTON'S NIGHTMARE

Sutton understands the economic world to be messy and complex. Human minds are limited to relatively simple representations or models.

The big question is how do we obtain adequate representations, knowing that any model is much less complicated than the world that it represents? Sutton never considers the genesis of theories or models, but instead concentrates on the questions: How should we choose among given theories? Or, how should we test a given theory? His focus is on Hans Reichenbach's context of justification rather than the context of discovery.

How can a simple model represent a complicated and messy economy? One answer, which Sutton ascribes to Alfred Marshall, is that theories or models represent the main tendencies of the economy. There are systematic and unsystematic factors, which may also be weak or strong. With data enough and time, statistical estimates of the systematic factors will become more and more precise, and even weak systematic factors will be isolated. Marshall analogized the tendencies of the economy to the tides, which could be seen as governed principally by the systematic forces of gravity and the rotation of the earth, as well as some weaker systematic forces, and the unsystematic forces of transient weather. One of Sutton's central points is that, while the theory of the tides has improved hand in hand with empirical observations, so that there is now a highly complete model, empirical relationships in economics are unstable and not well anchored to economic theory and have, since Marshall, made relatively little progress.

Sutton considers two responses to the shakiness of empirical economic relations. The optimistic one suggests that shifting empirical relationships are truly embedded in an as-yet-unarticulated *supermodel* or, equivalently, that there are unobserved factors that would *systematically* account for their shifts. The problem with empirical progress in economics is not that the tendency view is wrong. Rather it is just practically hard to implement. In his darker moments, Sutton is more pessimistic. His nightmare is that large and unsystematic factors – unknown and unknowable, untameable through statistical devices such as errors modelled by probability distributions – account for the shifting empirical relationships in economics. On this view, enough data and time are not helpful, because there are no stable attractors towards which empirical economic relationships should tend.

Fortunately, Sutton spends a good deal of his time awake. So, while he spends a lot of time attacking the foundations of Marshall's tendency view, which he often refers to as the *standard paradigm*, he also points to several cases in which it seems to succeed. He suggests that, even when it strictly fails, there may still be features of the economic world that are generalizable, that might distinguish a broad class of models from competing classes even if they are too imprecise to permit discrimination among the tightly specified models within the consistent class. If the characterization of the class is rich enough, such knowledge may be

highly useful. Still, Sutton fears the dark, and a vein of pessimism both for theory and for econometrics runs through the lectures.

On the tendency view, theory and econometrics are two sides of the same coin. Haavelmo's econometric methodology starts with a theory that identifies the structural relationships among potential economic factors. If the theory is specified fully and correctly, any residual error will be unsystematic and statistically well-behaved – for example, serially uncorrelated, homoscedastic, and independently distributed. Estimation assigns strengths to those factors and may, when the best estimate of the strength is zero, indicate which potential factors are unrealized. An alternative approach works in the other direction. It starts with a set of potential factors, possibly without a strong theoretical argument for them, and eliminates those which appear statistically insignificant or which do not deliver residual errors with the nice properties which indicate that they could have been generated by a correctly specified model. Any acceptable theory should be consistent with such estimates.¹ Sutton's attack on the tendency view undermines both approaches; for, even though they start from different ends, both assume that theory and data stand in a Marshallian relationship to one another.

2. UNIVERSAL AND LOCAL KNOWLEDGE

Sutton revives an old methodological debate. He acknowledges that Lionel Robbins had similar reservations about empirical studies of business cycles and that John Maynard Keynes had similar reservations about the econometric models of Jan Tinbergen. More recently, the same issues arise in the debate between calibrators and estimators of business cycle models (see, Hoover, 1995 and Hartley, Hoover and Salyer, 1997, 1998). Importantly, John Stuart Mill articulated the tendency view well before Marshall, although he was vastly more pessimistic than Marshall about the possibility of relating observable empirical outcomes back to the fundamental tendencies that generated them (Hausman, 1992).

A certain kind of physics enthralled many economists and philosophers. Economics aspires to the complete, universal laws of mechanics (Newtonian, relativistic, quantum) and is compared unfavorably with their empirically progressive experimental practice. Alex Rosenberg (1992) argues that economics fails as a science for failing to make notable empirical progress. While Mill (1848/1911, 1851) had a more generous view of the success of economics, his diagnosis of the limitations of economic science is grounded in a similar view of what science should

¹ This approach is well articulated by Hendry (1993, 1995). Sutton (p. 97) cites a good applied example of the approach (Davidson *et al.* 1978), but refers to it as if it were principally a theoretical approach.

be. The problem of economics for Mill, Robbins, and other anti-empirical economists is that the economy is too complex.

The first problem is just raw complexity: there are so many economic agents, governed by so many forces, and there is little possibility for controlled experiment. It is not widely recognized that the same is true for many areas of physical science. Practitioner's of the most admired areas of physics – such as particle physics – typically believe that all physical phenomena are derivable from a small set of universal laws. Yet, this is of no practical import to solid-state physicists, or to students of hydraulics or meteorology (to take examples within physics) or geology, chemistry, or biology (to take examples outside physics). 'Basic' physics may be helpful in these fields; yet it must be supplemented by quite specific experimental or observational knowledge not easily linked to the supposedly more basic science.

The second problem is that economics, unlike physical sciences, deals with agents who hold beliefs, desires, attitudes and other intentional states and who make choices. Intentionality sometimes introduces nonuniqueness. Sutton cites the example of the nonunique set of feasible exchanges in a two-by-two Edgeworth box. He provides examples of nonunique solutions to economic games. Mill subscribes to the view (shared in varied forms by Jevons and Marshall as well) that people have hierarchies of utilities or interests. Economic values and motivations may be powerful, but other – higher – values and motivations may dominate in particular cases and mask the economic tendencies.

Mill's solution, shared most clearly by Robbins and the Austrians, is to regard economics as an a priori discipline. For Mill, economics is a deductive science that works out the consequences of wealth maximization *ceteris paribus*. But, because other things are never really equal, the best that can be said is that there are tendencies. The countervailing factors are so rich, that these tendencies cannot be practically falsified. Robbins (1937) famously generalized Mill's account, defining economics as the science that studies the relationship between ends and scarce means with alternative uses.

Most economists accept Mill's and Robbins's view of the discipline. They differ in how much apriorism they believe it entails. Marshall was a moderate, who genuinely believed in the interplay of economic theory and real world experience. Unfortunately, he gave little advice on how to use empirical evidence to develop and enrich economic theory. Mill and Robbins profess moderation – they profess some use for empirical data. In fact, they are extremists – no inductive evidence could affect the core theory in their views. Economics, for Mill and Robbins, is not empty because they believe that that direct awareness of our own motivations is enough to provide premises from which logical economic deductions

follow. Violation of *ceteris paribus* accounts for any failure of the facts to confirm these deductions. They have faith in their understanding of human motivation and economic logic. Even an econometrician such as Haavelmo accepts some degree of apriorism, though he has sufficient faith in economic factors being strong enough to be revealed as a Marshallian tendency rather than hidden as a Millian tendency.

Sutton professes a Marshall-like moderation, yet he forcefully highlights Mill's and Robbins's deep fear – namely, that unsystematic changes in relevant factors will generate predictive failure. When Sutton, like Rosenberg, laments the lack of empirical progress in economics, he has predictive failure in mind. There are two types of prediction. The first is the prediction of a fact. This is what people want when they gaze into a crystal ball or when they ask an economist to tell them what the price of a stock or an interest rate will do over the coming week or month or year. The second is prediction of a relationship. The difference can be illustrated through a central idea from the theory of finance: stock prices should follow a random walk (see, Sutton, Chapter 2). The prediction in the second (relational) sense is that prediction in the first (crystal-ball) sense is impossible. Every failure to predict a particular stock price is evidence in favor of the prediction that stock prices follow a random walk.

The two senses of prediction are important because, on the one hand, failures of crystal-ball predictions are less important as measures of the success of a science in making sense of the world than failures of relational predictions and, on the other hand, some of Sutton's examples highlight crystal-ball failures. One such example is the supposed failure of standard macro models in the early 1970s (Sutton, pp. 33ff., 90). The key failure was the inability of the Phillips curve analysis to predict stagflation.² But this was a crystal-ball failure. A look at Phillips's (1958) original paper or at Samuelson and Solow's (1960) paper, which popularized the Phillips curve for policy analysis, reveals that all three authors understood that the curve would shift with a sustained increase in the rate of inflation. The great inflation of the 1970s and 1980s was unprecedented – such high rates of inflation had been previously experienced in relatively brief and otherwise highly unstable episodes. The breakdown of the Phillips curve can then be seen as a case of what Engle, Hendry, and Richard (1983, p. 285) refer to as 'nonexcitation' (Sutton, p. 19) in which the current specification could not have been

² Sutton refers to the failures of the IS/LM approach to macroeconomics and its inability to resolve differences between the monetarists and the Keynesians in the face of high inflation and high unemployment. He recognizes that it is not really Keynes's theory which is at stake, but what some Keynesians believe. I believe that it comes closer to the heart of the controversy, to which he refers, to cast it principally in terms of the Phillips curve of which Keynes was, of course, wholly innocent.

estimated in the 1960s, because some factors important to it had not shown sufficient variance at that time. Yet, the current specification may be consistent with the data of the 1960s looking backwards – the fundamental relationship did not change.³ A crystal-ball failure is consistent with relational confirmation.

The response of standard economists to the complexity of the economy and the instability of empirical relationships is to insist on theory. Economic relationships are useful when they are generalizable, when they apply to new contexts – that is, when they can be treated counterfactually to conduct thought experiments. Along with many other economists, Sutton demands the assurance of a theoretical explanation before he feels justified in generalizing an empirical observation outside its immediate context. For example, despite the fact that the inverse relationship between the vacancy rate and the unemployment rate (the Beveridge curve) is qualitatively robust, Sutton is loath to use it counterfactually for lack of a good theoretical model. His reluctance stands clearly in the Mill/Robbins tradition. Yet the class-of-models view shows that Sutton is more aware than they were that theory may sometimes itself be of little use.

I wish to suggest that the high opinion that economists hold of theory as the essential warrant of counterfactual inference comes in large part from their thrall to the ideals of 'basic' physics. We often use empirical regularities counterfactually, even absent any theoretical account. Sutton's example of option pricing, which he cites as a successful application of the standard paradigm, illustrates this fact. Bachelier's or Black and Scholes's option-price model is just a shell. They are only complete when empirical estimates of the volatility of stock prices are fed into them. These estimates are not themselves theoretically grounded – they cannot be derived from first principles – and they are assumed to hold counterfactually whenever the options-price formulae are used to make predictions. Such cases recur throughout economics and science. Despite the high standing of theory, empirical regularities are routinely used counterfactually to support inference outside their contexts. Often such regularities are merely local, so that, in practical problems, measurement and recalibration are essential to inference; but they appear to be stable enough to be helpful. I suggest that there is a seamless continuum running from pure empirical regularities to fundamental theory. In economics, even the most basic premises such as utility maximization are warranted as an empirical inference: even Mill's or

³ The point is not that a model can be fitted across two periods, but that a model that could not be fitted to an early period because of nonexcitation of key variables might be fitted separately to the later period and yet be consistent with the estimates of the earlier period. Robert Gordon (1998) *inter alia*, has demonstrated how, once one accounts for new excitations (e.g., supply shocks), the Phillips curve tells a consistent story through time.

Robbins's starting premises, known through direct introspection (through *verstehen* in the Austrian jargon) can be seen as based on the empirical observation of a very small sample – ourselves.

Once we are ready to grant counterfactual status to local regularity, then the nightmare of only classes of models with weak predictive capacities enjoying empirical support seems less dark. The thralldom to 'basic' physics is reflected in Sutton's approving recounting of Carnot's analysis of the efficiency of steam engines – a paradigm class-of-models account. The major use of the Carnot cycle is in theoretical thermodynamics – where it is extremely important. It is of more comfort for the theoretical physicist than for the practical engineer. The engineer must account for the properties of materials and the idiosyncrasies of particular designs through lots of experimentally derived regularities, which he must, of course, generalize to new contexts. In that respect, the engineer is more like Carnot's benighted rivals than like Carnot himself. Economics is no different. Practical economic analysis requires detailed empirical knowledge of facts and relationships, often highly local in place and time. Practical economics is not so different from practical physics in that regard.

3. TWO VISIONS OF ECONOMETRICS

My insistence on the importance of highly local regularities and their counterfactual use is similar to the claims of Sutton's colleague at the London School of Economics, the philosopher of science Nancy Cartwright (1983, 1989, 1999). She shares with Sutton and me the view that the world is a messy, complicated place in which universal laws are few and far between. She is less enthralled by 'basic' physics than most economists and philosophers. Like Sutton, however, she finds common econometric practices wanting. (I should note that Sutton's analysis stresses how misleading the Marshallian approach to econometrics can be. Yet, he writes: 'In most situations that we encounter in economics, the standard paradigm provides our most useful investigative framework' (Sutton, p. 33). Sutton's tone is always moderate. Like Marshall, he advocates the middle ground between theory and empirics; and – like Marshall – he gives us little guidance on how to stand in this middle ground. For Sutton, the standard paradigm works when it works, and does not work when it does not.)

Cartwright's (1999, especially Chapter 7) criticism of econometrics is based on the analogy between regressions and experiments. She views probabilities, not as ubiquitous properties of the world, but as characteristics that emerge only from highly structured experimental set-ups. She talks of *nomological machines* – constructions from which lawlike (including probabilistic lawlike) behavior emerges. Theory provides

blueprints for nomological machines. Economics has few opportunities for controlled experiments. Systems of regression equations (in Haavelmo's or the Cowles Commission's paradigm) aim to provide the missing controls statistically. Standard econometric practices can work on this view only when they follow the blueprint for the nomological machine, when they are highly structured according to the economic theory *and* when that theory accounts for the key features of the world. In practice, most econometrics does not meet this exacting standard. So, like Sutton, she is brought to question whether econometrics is up to the job and can deliver the goods.

There is an irony in Cartwright's view. For her, the laws of physics work in real cases only when they are applied with a 'not-too-literal' mind by people familiar with the empirical properties of real materials and with experience of when they work and when they do not (Cartwright, 1989, p. 8). That is, practical physics requires knowledge of local regularities, often without a deep or consistent theoretical base. In contrast, since she sees the 'laws of economics' emerging from systems of regressions in much the same way as the 'laws of physics' emerge from experimental set-ups, she fails to note the importance of local regularities in economics or to entertain the thought that they might be established econometrically (see, Hoover 2001, 2002). Acknowledging that universal regularities are genuinely hard to find, Cartwright sometimes ignores the fact that local regularities are ubiquitous – in physics and economics. It would be difficult to make our way through the world if that were not true.

Sutton's radical critique of econometrics is similar to Cartwright's: the standard paradigm should work only if econometrics recapitulates the true structure of the economy. But since he doubts that any theoretical model could capture that true structure (and sometimes writes as if the notion of a true structure were itself problematic), he doubts that econometrics can *generally* succeed. Unlike Cartwright, Sutton hints at a more moderate approach. The *weak* or *sophisticated* interpretation of the standard paradigm takes econometric evidence as 'diagnostic' rather than as representational of a true structure (Sutton, pp. 21–2, 100). The persistence of asymmetrical residuals in a linear regression, for instance, might suggest a nonlinear relationship. Such evidence points towards Sutton's preferred class of market-concentration models over the tightly specified Cournot model (Sutton, Chapter 3). Still, it falls far short of Cartwright's required standards of econometric specification.

Unfortunately – perhaps because he is himself principally a theorist – Sutton spends too little time in the lectures developing the notion of econometrics as a diagnostic instrument. Elsewhere (Hoover, 1994), I have drawn a similar distinction between 'econometrics as observation'

(Sutton's diagnostic tool) and 'econometrics as measurement' (Cartwright's completely articulated specification). The point is that econometrics that falls short of quantifying a universal theory can nevertheless help to isolate local regularities that shed light on what a theory must be. Sutton talks of theory as posing restrictions on the space of empirical outcomes. The fear seems to be that without such restrictions, unstructured empirical investigation would throw up too many 'regularities' that prove to be of little value, as they cannot be generalized beyond their immediate context. Another way to understand what I have called Sutton's nightmare is to see the warranted class of models as posing too few restrictions to justify any useful counterfactual analysis. But 'econometrics as observation' or 'econometrics as diagnostic tool' suggests a two-way street between theory and empirical evidence. Observed regularities place restrictions on the space of admissible theories *and* theories place restrictions on the space of admissible observations. This is a Marshallian suggestion, but one that neither Marshall nor Sutton nor anyone I know has developed into a systematic econometric methodology.

4. AN UNDOGMATIC CONCLUSION

As I have already observed, there is a mismatch between Sutton's radical analysis of the utility of econometrics and his general tone. The lectures articulate deep fears for the prospects for empirical economics. These fears are summarized at the end through two pessimistic propositions (Sutton, pp. 100ff.): first, econometrics does not effectively discriminate among competing theories; and, second, theory does not provide interesting restrictions on empirical observations. In this worst case, economic theory could perhaps provide a framework for cataloging economic outcomes, but not strong predictions. And, if that is all there is to it, then 'economic theory would be a poor kind of thing' (Sutton, p. 105). Sutton rejects these two pessimistic positions with the observation: 'we can hope for more. In some (perhaps many) situations, it is possible to find theories that work' (p. 105). Nonetheless, although he chips away at the pessimistic propositions throughout the book whenever he acknowledges the successes of theory and empirical economics, they remain the most forcefully argued positions in the book. The book ends on a hopeful note, but it does not quite convince.

Nevertheless, I share Sutton's hope. Perhaps the wisest methodological principle of the book is the observation that sometimes it is 'more useful' to take a looser approach to empirical evidence than the standard paradigm enjoins: 'There is no recipe for research' (Sutton, p. 33). And there are no guarantees. The philosopher and scientist C. S. Peirce once wrote: '. . . when we discuss a vexed question, we *hope* that there is some

ascertainable truth about it, and that the discussion is not to go on forever and to no purpose'. Sutton's lectures should be seen as cautionary: we may enjoy our success in understanding the economy; but we cannot expect too much. To which perhaps we might end with Peirce in mind: keep hope alive.

REFERENCES

- Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Clarendon Press
- Cartwright, N. 1989. *Nature's Capacities and Their Measurement*. Clarendon Press
- Cartwright, Nancy. 1999. *The Dappled World*. Cambridge University Press
- Davidson, J. E. H., D. F. Hendry, F. Srba and S. Yeo. 1978. 'Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom'. *Economic Journal*, 88:661–92
- Engle, R. F., D. F. Hendry and J.-F. Richard. 1983. 'Exogeneity'. *Econometrica*, 51:277–304
- Gordon, Robert J. 1998. 'Foundations of the Goldilocks economy: supply shocks and the time-varying NAIRU'. *Brookings Papers on Economic Activity*, No. 2, pp. 297–333
- Hahn, F. 1992. 'Reflections'. *Royal Economic Society Newsletter*, No. 77
- Hartley, J. E., K. D. Hoover and K. D. Salyer. 1997. 'The limits of business cycle research: assessing the real business cycle model'. *Oxford Review of Economic Policy*, 13:34–54
- Hartley, J. E., K. D. Hoover and K. D. Salyer (eds.). 1998. *Real Business Cycles: A Reader*. Routledge
- Hausman, D. M. 1992. *The Inexact and Separate Science of Economics*. Cambridge University Press
- Hendry, David F. 1993. *Econometrics: Alchemy or Science? Essays in Econometric Methodology*. Blackwell
- Hendry, David F. 1995. *Dynamic Econometrics*. Oxford University Press
- Hoover, Kevin D. 1994. 'Econometrics as observation: the Lucas critique and the nature of econometric inference', *Journal of Economic Methodology* 1:65–80
- Hoover, Kevin D. 2001. *The Methodology of Empirical Macroeconomics*. Cambridge University Press
- Hoover, Kevin D. 2002. 'Econometrics and reality'. In *Fact and Fiction in Economics: Models, Realism and Social Construction*, Uskali Mäki (ed.). Cambridge University Press
- Mill, J. S. [1848] 1911. *Principles of Political Economy with Some Applications to Social Philosophy*. Longman's and Green
- Mill, John Stuart. 1851. *A System of Logic, Ratiocinative and Deductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, 3rd edn. Vol. I. John W. Parker
- Phillips, A. W. 1958. 'The relation between unemployment and the rate of change of money wages in the United Kingdom 1961–1957'. *Economica*, NS 25(4):283–99
- Robbins, L. 1937. *An Essay on the Nature and Significance of Economic Science*. Macmillan
- Rosenberg, A. 1992. *Economics: Mathematical Politics or Science of Diminishing Returns?* Chicago University Press
- Samuelson, Paul A. and Robert M. Solow. 1960. 'Analytical aspects of anti-inflation policy'. *American Economic Review*, 50:177–94