Markets re-bound:
Sutton's Models as Methodology

Tom Cunningham*

J ohn Sutton has devised a model for explaining the patterns in market structure in different industries, especially the relationship between the concentration of an industry and that industry's use of advertising or investment in technology. His theory seems to do its job well but Sutton also proposes it as a prototype of a new kind of economic theory; one which avoids some of the problems of previous theories, especially the difficulty of empirical testing. In this essay I try to test his new method against two papers written in criticism of other types of economic theory: those of Milton Friedman and of Nancy Cartwright. Before explaining Sutton's theory I will very quickly describe three standard types of economic theory, and the problems applying to each. Firstly, reasoning about economic issues can be done with general principles, where each of the principles are derived independently. The principles could be derived from observations, or from axioms of maximising behaviour (e.g. "that people tend to copy the investments of others", or "that monopolists will not price at a point where demand is inelastic"). This is probably the most natural way of doing economics and is the type of reasoning associated with Keynes.

The problem is that the principles may not be robust. Firstly, empirical regularities are always liable to lapse. For this reason economists always prefer principles which are based on rational behaviour, because they seem to have a justification. Further, even principles which are derived simply from assumptions about maximising behaviour may, when composed, become invalid. Most interesting principles need to include "all other things equal" in their antecedent clause, but adding another principle will generally upset this assumption.

A second approach is simply to find statistical relationships within sets of historical data, and compile a matrix of correlations between each of the variables. This

*Tom studied Philosophy in Otago University, New Zealand. He is currently undertaking the MSc in Philosophy and Economics at the LSE and his current research interests are logical fallacies, anti-social choice and "anything that decreases the gross confusion and increases the net confusion of mankind."

matrix will then serve as a model of that economic system allowing prediction of the magnitude of the effects of changing one of the variables.

There are many unsatisfactory aspects to this way of doing economics. First, it might predict but it cannot explain economic relationships. Second, it still requires a theoretical understanding to know how to make predictions: e.g. if you want to estimate the effect of introducing a completely novel type of tax it will be unclear how to simulate that tax by manipulating the model's variables which had previously recorded only levels of conventional taxes. A third problem, described by Lucas (1976), is that these statistical models cannot easily incorporate peoples' expectations, and that the expectations are typically very important influences on economic outcomes.

A third approach to economic theory is to use a complete optimising model. If an entire system were to be derived from axioms about maximising behaviour this would avoid the problem of inconsistent assumptions. This would also naturally incorporate expectations. So, for a given situation we could construct a model which represents each of the actors and the actions open to each of them. The outcome of the model should be a consistent and robust prediction.

The main problem is the difficulty of collecting assumptions that are each justified, but collectively sufficient to make non-trivial predictions. Standard neoclassical assumptions can be used in many situations, but will often give the wrong prediction: the assumptions of infinite producers, infinite divisions, and perfect homogeneity, are too far removed from real life to give accurate numerical predictions. Alternatively, a game-theoretic structure can be built which incorporates more of the subtlety of the situation; the problem however, being that they typically require radical simplifications and are often sensitive to the particular assumptions used.

Sutton's Market Structure theory

Sutton's theory is submitted as an explanation of the relationship between the concentration of a market (the degree to which the market is dominated by a few firms), and the particular characteristics of that industry. His *Sunk Costs and Market Structure* (1991) suggested a model with which to predict the relationship between the use of advertising in an industry and the market's concentration. *Technology and Market Structure* (1998) uses a similar model to test the relationship between technology and market concentration.

Sutton gives two assumptions which produce most of his results: (1) Viability (the survivor principle): No company will continue to make losses. And (2) Stability (the arbitrage principle): No gaps exist in the markets which could be profitably filled. (1998, p.8)

These assumptions are unusual for economics because the agents are not as-

sumed to be behaving optimally. A situation which satisfies these two assumptions Sutton calls an "equilibrium configuration."

Sutton uses these spare premises to derive a variety of results. The general shape of the argument is to say that, in some space of industry characteristics, certain regions cannot be equilibria because they violate either the viability principle or the stability principle. However nothing is implied within the space of possible equilibria.

Sutton's principal hypothesis in *Technology and Market Structure* is that, among the many influences on the concentration of an industry, the R&D intensity of that industry sets a *lower bound* to the possible levels of concentration. In particular, an industry with high R&D-expenditure cannot be in a stable state if the market is shared between many small firms (low concentration). The reason is that R&D expenditure shows increasing returns to scale: R&D costs the same whether you produce a small or large amount, so it is more cost-effective to produce a large amount. Then by the arbitrage principle, if the market has low concentration it will be profitable for a large company to enter.

Sutton tests the lower-bound hypothesis against data on 44 US industries in 1977, and rejects, with 99% confidence, the hypothesis that homogeneity does not serve as a bound on market concentration in high R&D industries(1998, p.109).

Friedman & Truth of Assumptions
Milton Friedman, in "The Methodology of Positive Economics" (1953), argues that a theory should be judged just on its predictions, not on the truth of its assumptions. His paper is written as a defence of neoclassical economics, especially a defence of the simplifying assumption that all agents maximise: businesses maximise profits, and consumers maximise their utility.

There are many strands of his argument, but I think the two principal ones are these: (a) that in science it is irrelevant whether a theory's assumptions are true because only the quality of the predictions matter, and (b) that there is independent reason to believe that the neoclassical maximising assumptions are approximately true.[1] However, giving these two reasons is giving one too many: if (a) is true then (b) is irrelevant. If it doesn't matter whether the neoclassical assumptions are true then why argue that they are?

Friedman's paper has many imprecise statements, which could be resolved one

---

[1] "It is well known from the game-theoretic literature that, even when such models have a unique equilibrium outcome, the predictions of the models are delicately dependent on the institutional features of the bargaining environment … To control for the influence of such non-observables is in practice infeasible." (Sutton 1998, p343)

way or the other. For example, whether the 'predictions' of a theory should include its assumptions; whether "as if" statements should be interpreted as literal and false, or analogical and so true when stated with proper qualifications, and whether a theory has "assumptions" at all, or just a range of applications.

Let the maximisation hypothesis be this: 'businesses act in ways which maximise their profits; and consumers act in ways which maximise their expected discounted consumption preferences.' Now this hypothesis is a useful simplification because (a) there are independent sociological reasons to think it is true in many circumstances; and (b) models which use this assumption have fared well, eg. in predicting the effect of a new tax. I think that both Friedman's practice and arguments fit this way of putting things; though his bold conclusions, including "the more significant the theory, the more unrealistic the assumptions" (p.14), are not so striking.

Alan Musgrave argues that Friedman's particular fault was ambiguity in his use of the word "assumption", failing to distinguish between three different types (negligibility, domain, heuristic) and that none of these types of assumptions support Friedman's above claim.

Sutton is very careful to formulate his assumptions as weakly as possible. For example, in chapter ten of *Technology and Market Structure* he mentions that a standard assumption is Gibrat's law: "a firm's proportionate rate of growth [is] given by a random variable whose mean was independent of the firm's current size." (1998, p.241) When Sutton constructs a model with this type of assumption he says "we avoid introducing any specific form of relationship linking the firm's size to its rate of growth", and formulates his hypothesis as: "The probability that the next market opportunity is filled by any currently active firm is nondecreasing in the size of that firm". (1998, p.246)

Also in the particular case of the "maximization of returns hypothesis", Sutton finds a way to move from Friedman's justification for the hypothesis to conclusions drawn from the hypothesis, without committing himself to the hypothesis itself. Friedman says that "confidence in the maximization-of-returns hypothesis is justified by … [the fact that] unless the behaviour of businessmen in some way or other approximated behaviour consistent with the maximization of returns, it seems unlikely that they would remain in business for long". (1953, p.22) Instead of moving from this reasoning to the hypothesis, Sutton simply includes the underlying reasoning in his model, for he says "no firm has a portfolio of product varieties (locations) on which it makes an overall loss" (2001, p.72), and from this much weaker assumption he is able to deduce testable predictions.

In fact, the predictions are likely to be superior, given the greater realism of the hypothesis. Sutton does not say this directly, he puts it like this:

> ... instead of posing the intractable question, Do agents really maximize? It may be more fruitful to ask, to what extent can we deduce, by reference to observed actions, whether the agents were maximizing? In a setting where we do not know, and cannot identify, a true model that can be assumed to be common knowledge to agents, the most we can hope to do is to identify some subset of actions as being inconsistent with maximization within any admissible model. What we are led to, along this route, is a form of 'bounded rationality'. (2001, p.77)

I think this is the wrong way of putting things. Sutton is confusing imperfect rationality, and imperfect observation of rationality. If the problem were just imperfect observation of peoples' decision problems we could keep the rationality assumption, and simply make weaker assumptions about the games that were being played in the usual way. But Sutton does weaken the rationality assumption, and this is one of the sources of predicted indeterminacy above a lower bound. In *Technology and Market Structure,* Sutton discusses some case histories, illustrating the complex determinants of market structure above the lower bound, and he does not assume that each action which affects the outcome is a rational maximisation.

Speaking generally, Sutton says: "Sometimes, the most interesting alternative to a highly structured classical model is a model of a more primitive kind. It may be that the apparently successful prediction emanating from a highly structured model would follow from a less highly structured model that embodied some basic features." (2002, p.98) He gives two examples: Becker deriving downward-sloping demand curves from random behaviour, so long as it is constrained by a budget; and Hildenbrand's proof that the 'law of demand' is likely to hold at the aggregate level, without holding at the individual level . These examples suggest that the predictive success of the optimising hypothesis may not be such strong evidence for that hypothesis, since many of the predictions can be derived from weaker and more intuitively realistic assumptions.

In summary we can say that it was Friedman's purpose to defend the use of the 'rationality' hypothesis in economics, despite its unrealism, by appeal to its predictive success. Sutton argues that in industrial organization (arguably a case in which rational behaviour is most likely) using weaker assumptions actually produces superior predictions. However, Sutton still requires some weakened rationality of agents to generate suitably determinate predictions.

Cartwright & Auxiliary Assumptions

In "The Vanity of Rigour in Economics" (1999) Nancy Cartwright argues that the modern model-building methodology in economics is flawed. She takes two exam-

ples: Lucas' 1981 model of money illusion, and Pissarides' 1992 model of skill loss and unemployment. Both models find relationships between exogenous and endogenous variables through predicting the behaviour of a set of identical individuals maximising their expected returns in a very abstract situation. Cartwright says that the purpose of this type of model must be, if anything, to provide "tendency laws", which give relationships between variables. In the abstract models the principles derived are constant and exact (e.g. increasing skill-loss by 1% increases unemployment by 1.2%). No-one expects the observed economic relationships to be so clear-cut, but we might expect constant "tendencies" where the variables can be related in a particular way outside of the inevitable interference.

Cartwright's criticism is that some of the assumptions used are too strong to justify deriving genuine tendencies. Some of the assumptions simply set interferences to nothing, which is fair, but other assumptions used to make the logic tractable may bias the results.[2] For example, Lucas assumes many things about his consumers (identical, 2-period-lived, ignorant of government policy) which may be material to the conclusion. In other words, it is not clear whether the tendency derived might disappear or reverse if these assumptions were modified in uninteresting ways.

Cartwright criticises Pissarides and Lucas, as representative modern economists, for not coming up with stable tendency laws. However, in this paper she does not give reason for believing that stable tendency laws are indeed possible in economics. In fact, Lucas' rational expectations arguments could be redirected to argue against the existence of stable tendency laws. Lucas says that well-confirmed empirical relationships can predict quite incorrectly if they miss the underlying motivations of agents. Likewise a useful tendency law (an example of Cartwright's: "in any situation skill loss tends to produce persistence in unemployment shocks") could become simply false, and not merely hidden by interference, if underlying conditions change.

Depending on the particular definition of terms this reasoning could either be interpreted as an argument that relevant tendency laws will not be reliable, or that though tendency laws could be reliable, they will not exist. Either way, doubt is cast on the usefulness of the ideal of tendency laws for economics.

Sutton's theory bears an interesting relationship to Cartwright's ideal. First, Sutton's theory does not deliver tendency laws with the same structure as Cartwright's: they are not of the kind "more of A gets more of B, all other things being equal". It

---

[2] The distinction: "those that eliminate confounding factors and those that do not eliminate confounding factors but rather provide a simple enough structure to make a deductive study possible." (p.22)

might seem at first that the new structure is essentially similar: "more of A gets a higher lower bound of B". But another interesting difference is that Sutton can afford to drop the "all other things being equal" clause: the model is built to apply even when other causes operate in unknown ways.

The particularly interesting thing about the principles derived from Sutton's theory is that they can be composed without the fear of interactions. If we have one principle predicting a lower bound, and another principle predicting an upper bound, and if both principles are agnostic with respect to other interferences, they can be combined to predict an intermediate region.

Conclusion

I have argued the following two points. Firstly, Friedman's justification for the maximising hypothesis, based on a sacrifice of realism for the sake of predictions, is undercut by Sutton's demonstration of a theory which has more realistic assumptions but also better predictions. Secondly, Sutton's bounds approach does not involve as much "ceteris paribus" reasoning as other theories, and so principles derived from his models are likely to be more stable.

To investigate these ideas more thoroughly the next step would be to lay out some of Sutton's theorems in a completely rigorous manner, to show what role the maximisation assumptions play (especially, what different predictions will be made if full rationality is assumed), and to show the effect of interference. Also a natural suspicion will be that a "bounds" theory could be rewritten as a "classical" theory, or vice versa. This would be worth attempting.

Bibliography

Cartwright, N. (1999) "The Vanity of Rigour in Economics", CPNSS Discuss. Paper.

Friedman, M. (1953) "The Methodology of Positive Economics" in *Essays in Positive Economics*. Chicago: Chicago University Press.

Ingrao, B. (1989) "The Hidden Epistemology of Rational Expectations and the Paradoxes of Rationality", *Richerche Economiche* Vol. XLII.

Lucas, R. (1976) "Econometric Policy Evaluation: A Critique", in Lucas, R. *Studies in Business Cycle Theory*. Oxford: Basil Blackwell.

Musgrave. A. (1981) "'Unreal Assumptions' in Economic Theory: the F-Twist Untwisted", *Kyklos*, Vol. 34.

Sutton, J. (1998) *Technology and Market Structure*. Cambridge, MA: MIT Press.

Sutton, J. (2001) *Marshall's Tendencies*. Leuven & Cambridge, MA: Leuven University Press & MIT Press.