

MSc Economics & Philosophy Dissertation, 2006

“Models & Supermodels, Bound & Unbound”

Candidate number: 13820

PART ONE: BACKGROUND	3
PART TWO: INTERPRETATION OF A BOUNDED MODEL	8
PART THREE: AN ALTERNATIVE INTERPRETATION	16
PART FOUR: RELATED ISSUES	18
CONCLUSION	25
BIBLIOGRAPHY	27

Introduction

John Sutton has a new way of testing the standard economic theory of industrial structure – a theory which had come to be regarded as virtually un-testable.

Sutton and his critics bring up a lot of different concepts in trying to explain the difference between the orthodox method and this new method of testing: Sutton begins by making a distinction between a model and a class of models, other concepts which then come up are latent variables, inequalities, unobservables, systematic and unsystematic influences, large and small influences, multiple equilibria, and bounded rationality. In this essay I show that Sutton's original distinction is unsatisfactory, and because of this Sutton ends up making inconsistent statements about how his method works.

I suggest a new distinction: between tests which work on the assumption of unobservable influences being uncorrelated with observable influences, versus tests which work on the assumption of limits to the size of the influence made by the unobservables. Normally models are tested using the first assumption (*uncorrelated unobservables*), and Sutton's principal novelty is to use the second assumption (*bounded unobservables*).

Both distinctions, Sutton's and mine, are sufficient to distinguish his theory from the classical sort. The reason that my distinction is to be preferred is that only this distinction explains why the theory is able to solve the problem of testing in Industrial Organisation.

Having explained my distinction I discuss again the inconsistencies in Sutton's writing, and resolve them using the new distinction. I also follow up some ends which have come loose in examining Sutton's theory.

PART ONE: BACKGROUND

History of IO

Modelling the number and size of firms in an industry is one of the cleanest and most used parts of economics. The model can be explained very simply, it has realistic assumptions (relative to most other economic models), it is used a great deal in explanation and understanding, and the theory has been used for decades in justifying government regulation of competition.

An economic model can be built to represent an industry and predict precisely the sets of firms and products which would constitute an equilibrium in that industry, where an *equilibrium* means that each of the parties involved is making the best decision given their information, and given the decisions of the other parties. Constructing such a textbook model requires specifying just these characteristics of the industry: the amount that consumers would demand at each price (the demand schedule), the production technologies available to firms, the type of price competition which operates (e.g., Bertrand or Cournot), and any barriers to new firms entering. This is the theory of oligopoly, i.e. treating markets with only a few firms, with monopoly and perfect competition being extremes at either end.

Although the model seems to clarify and explain very well, and although the model has been used as a basis for estimating relationships and predicting outcomes, there have also naturally been various attempts to test the core assumptions, or at least to find out how close an approximation we can expect from it. And to test this model has turned out to be a very difficult task.

In the 1950s Joe Bain attempted to test an economic hypothesis about industry concentration with a dataset of American industries. The concentration of an industry is a measure of how equally it is shared among firms; an industry dominated by a few firms is 'concentrated,' one which is shared among many firms is 'fragmented'. The combined market share of the top three firms is often used as a measure of concentration. Bain looked for, and found, a positive relationship between industry concentration and industry profitability¹.

However Bain's results and his methodology are not now taken seriously. This is partly because in other studies the relationship appears only weakly if at all, but principally because a more thorough economic model of concentration and profitability regards both concentration and profitability as functions of deeper variables (i.e., they are both 'endogenous' variables), and this more thorough model is compatible with either a

¹ Reference in Schmalensee.

negative correlation, a positive correlation, or no correlation at all, depending on the values of those deeper variables.

So to test the model we need to measure the underlying or 'exogenous' variables, which are, in the simplest version of the model, the demand curve, the production function, the type of price competition, and the barriers to entry. To use industry data to estimate the effect of any one of these variables on an outcome (e.g., on profits or on concentration) we either need to collect data on all of the other variables, or to find one variable which varies independently of the others. This last option justifies the use of an instrumental variable: some outside factor which causes variation in one of the exogenous variables, as when weather is assumed to affect agricultural supply but not agricultural demand.

In a chapter in *The Handbook of Industrial Organisation*, published in 1998, Richard Schmalensee says that IO economists have come to consider cross-industry studies to be not capable of testing the theory, and are only useful for describing empirical regularities². On the one hand it is too difficult to make reasonable estimates of *all* the underlying variables over a range of industries. On the other hand there are very few instrumental variables of which it is reasonable to assume that they vary independently of the other determinants. This is partly because theorists have come up with many new elaborations on the basic model in which more and more factors can have an influence on the outcome³.

The conclusion then seems gloomy: because the model contains unobservables, it cannot be tested.

Professor John Sutton says that this problem of unobservables caused economists working on the economics of industry in the 1980s and 1990s to work on detailed studies of single industries, instead of comparisons across a range of industries.

Sutton's Solution

Despite these problems, Sutton has undertaken two large cross-industry studies, and despite Schmalensee's warnings, Sutton has attempted to divine from these studies deep facts about the nature of industrial organisation. The first study, published in 1991 as *Sunk Costs and Market Structure*, attempts to estimate the effect of advertising expenditure on concentration; the second, published in 1998 as *Technology and Market Structure*, estimates the effect of R&D expenditure on concentration.

Sutton has a quite simple general theory about how sunk costs can affect concentration. In some industries a certain level of fixed investment can increase the profitability of all

² Schmalensee, p953.

³ Schmalensee, p954.

your products, e.g. a higher level of advertising can increase the price that consumers will pay for each product, or a higher level of R&D can decrease the cost of producing each product. As the demand in a market increases, a higher level of fixed investment becomes profitable because of increased economies of scale. A higher level of fixed investment throughout the industry will in turn increase concentration, i.e. decrease the number of active firms, because a larger output is required to offset the fixed costs. The result is that an increase in market size can have two effects: the usual concentration-decreasing effect, but also a concentration-increasing effect through making new fixed investments worthwhile.

In sum, in industries where R&D or advertising are important, concentration can remain the same or even increase as the market size increases. In industries with little scope for fixed investment concentration will decrease with market size, as is normally predicted.

In *Sunk Costs and Market Structure* Sutton builds a conventional model to show how this can occur, using a traditional set of strong assumptions: Cournot competition, Nash equilibrium, and a simple functional form for the returns to investment (Sutton, 1991, p52). He shows that in this model of advertising-intensive industries concentration will first fall and then rise as market size increases, and as the market becomes indefinitely large concentration converges to a fixed level.

Conventionally the next step would be to run a regression of concentration on market size to test the theory. This brings us directly to the problem of unobservables, and the reason why cross-industry studies have not been much attempted recently. The nature of the problem can be put in two ways, and is put in two ways by Sutton: either that we do not know the structure of the different models which are true of each market⁴, or that we do not know the different values of the parameters in the same model as it applies to different markets⁵. The difference between these two constructions becomes important in the argument of this essay.

In any case, this is the point at which Sutton departs from the orthodox practice. He states two assumptions which are true of the model he has formulated, and which he claims are true of most orthodox models which could be formulated to predict concentration in advertising-intensive industries⁶. These two assumptions are not sufficient to derive a particular level of concentration, but they are sufficient to prove that for any industry which uses advertising there exists some *minimum* level below which concentration cannot fall, no matter how large the industry becomes, and no matter the value of the other influences. Sutton calls this result his 'non-convergence

⁴ Sutton, 2002, p70.

⁵ Sutton, 2002, p23.

⁶ Sutton, 1991, p69.

theorem,' in contrast to the classical result that concentration will converge to zero as market size increases.

This theorem, says Sutton, allows the general theory to be tested without having to know the true model for each industry, or having to know the values of many unobserved influences, and for this reason he calls his method the 'class of models approach'.⁷

The proof contrasts levels of fixed investment with levels of gross profit per consumer. Gross profit per consumer is the difference between the average revenue received from each consumer and the average cost to supply each consumer's goods. Each firm's fixed investment in R&D or advertising is assumed to increase their gross profit levels, either by increasing the price paid by consumers, or by stealing market share from other firms, or by increasing the total number of products sold. An increase in fixed investment is worthwhile for a firm if and only if it causes an equivalent or greater increase in their gross profits.

The conclusion of the proof is that there exists some minimal *proportion* of a market's total gross profits, below which the share of the largest firm will never fall. Concentration is generally measured in *revenue*, not gross profits, and indeed Sutton's own tests use revenue, but it requires only a weak assumption to show that a lower bound to the largest share of gross profits entails a lower bound to the largest share of revenue⁸.

Assume first that all firms' levels of gross profit per consumer are determined by levels of fixed investment, independently of the number of consumers. An increase in the number of consumers will then increase gross profit exactly proportionately. This assumption seems plausible regarding investment in R&D, but perhaps unusual for spending on advertising; this point is discussed later. This assumption also entails constant marginal costs.

Further assume that *no matter* other levels of investment, a new entrant who invests some particular multiple of the currently biggest-spending firm's level of investment will get at least some particular share of the market *total* gross profit per consumer. For example, say that in some market an entrant spending twice the highest existing level of investment will at least earn gross profit equivalent to 30% of the previous total gross profit, no matter what the distribution of investment among other firms is.

From these assumptions it follows that there exists some threshold level of investment as a proportion of total gross profits, such that if the biggest spender spends less than this amount then it would be profitable for a new entrant to enter with a higher level of

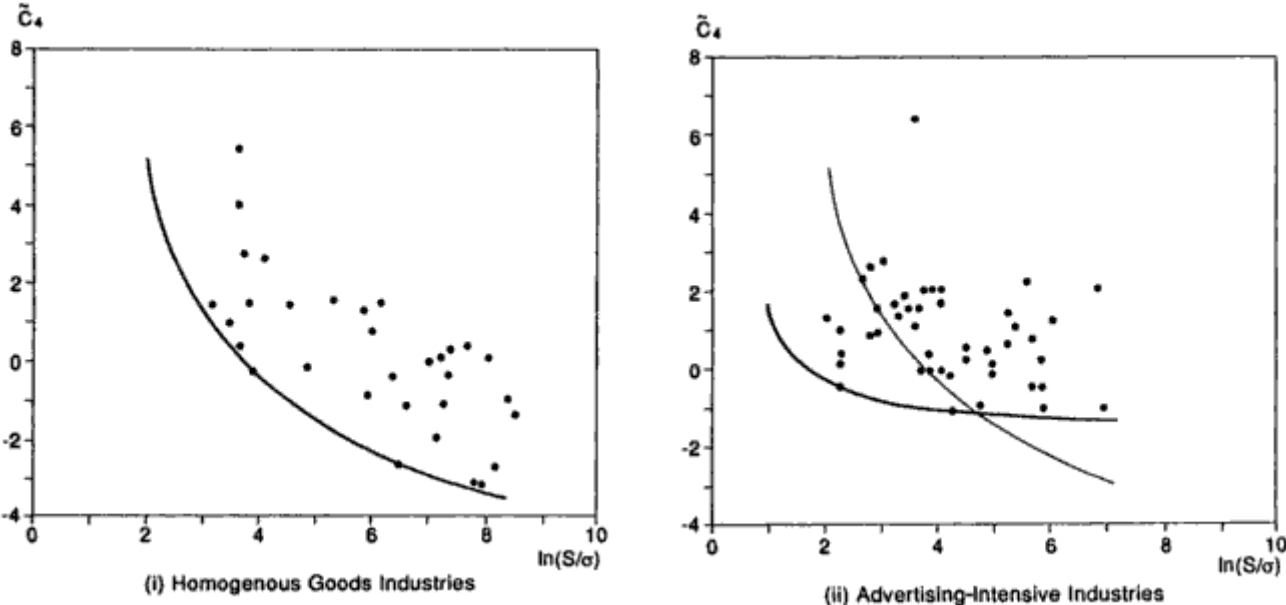
⁷ Sutton, 2000, p70 and passim.

⁸ See Sutton, 1991, p73.

fixed investment. For the example just quoted the threshold would be at 15% of total gross profits: if the investment of the biggest spender was only 10% of total gross profits the situation would not be stable, because an entrant who invested twice as much (equivalent to 20% of prior total gross profits) would earn at least 30% of the prior total gross profits, and so earn *net* profit. When investment is at the threshold level of 15% of total gross profits it is at least *possible* that no entrant would wish to enter, so any concentration level at or above 15% cannot be shown to be unstable. An entrant who doubled spending at 15% would spend the equivalent of 30% of prior gross profits, and in some circumstances receive only the equivalent of 30% of prior gross profits.

The principal test that Sutton makes of his theory consists of dividing the industries into those with low advertising spending⁹, and those with high spending, and then for either set of industries estimating an asymptotic lower bound on concentration with respect to market size, where one parameter of the lower bound specifies the difference between the asymptote and zero. The asymptote of the bound for the concentration of the low-advertising industries was found to be not significantly different from zero, but for the high-advertising industries it is significantly different from zero. Specifically, he estimated that the market share of the four largest firms (i.e. the C_4 ratio) will never go below 14% for a high-advertising industry, no matter how large the market becomes¹⁰.

Figure 1: Market Size and Concentration for European & US Food Industries around 1986¹¹



⁹ Advertising spending measured as a percentage of total revenue.

¹⁰ The charts are from Sutton 1991, p118.

¹¹ From Sutton, 1991, p118. $\hat{C}_4 = \ln[C_4 / (1 - C_4)]$, a logit transformation to make the values unbounded.

Sutton's two assumptions are not sufficient to justify any particular shape for the lower bound, they prove only that there does exist some minimum level below which concentration never falls. The fact that neither the level nor the shape of the bound can be predicted will be discussed again towards the end of this essay, but it is not crucial for the main point. The main point is the argument over what distinction explains why Sutton's method is able to test the standard theory of oligopoly, where the orthodox method has been unable to.

PART TWO: INTERPRETATION OF A BOUNDED MODEL

Sutton's 2002 book *Marshall's Tendencies* is the revised text of a set of lectures that he gave in 1997 on the testing of economic models. One of the chapters gives an interpretation of the difference between his 'class of models' approach and the standard approach. In commentaries on the book, published in an *Economics and Philosophy* symposium in 2002, Carl F. Christ and Eric Renault both criticise this interpretation; Sutton also published a reply in the same symposium issue.

Sutton and a "Class of Models"

The book considers three cases: areas in which the standard paradigm has been a success (auctions, option pricing); where it has been a failure (macroeconomics); and finally an area which Sutton annexes for economic modelling by using an amended version of the standard paradigm, his 'class of models'.

Sutton describes the history of the 'standard approach,' quoting particularly Marshall, Samuelson, and Haavelmo. The approach is described not as a framework which is perfectly comprehensive, but as a prototypical case of an economic model which can be successfully tested. This ideal case Sutton thinks has sometimes misled economists on how to apply their theories. In the standard paradigm:¹²

(1) the true model captures a "complete" set of factors that exert large and systematic influences, (2) all remaining influences can be treated as a noise component that can be modelled as a draw from some probability distribution, and (3) the model determines a unique equilibrium.

The main problem for the standard model is the influence of unobservable variables which cannot be controlled for¹³.

¹² Sutton, 2002, p20.

¹³ Sutton, 2000, pp. 32, 57; Sutton, 2002, p55.

In contrast Sutton's 'class of models', takes, as suggested by the name, an entire class of models when none of the individual models can be ruled out on the basis of the evidence. As described earlier, in Sutton's treatment of sunk costs he just uses a few abstract assumptions, and asserts that these assumptions would be satisfied by the entire class of reasonable models, then shows that an empirically testable prediction can be derived from the assumptions. I think that this is a fair representation of the essence of his idea, but I cannot here give a precise statement of the 'class of models' method because I do not think that the idea is consistent in itself. The best I can do is mention the most important statements that Sutton makes about this method:

1. *A class of models is a solution to the problem of unobservable factors. (2000:57)*
2. *A class of models is used when it cannot be determined which of a set of complete models is the correct one. (2000:8, 2000:70).*
3. *A class of models lessens the distinction between large and small influences made in the standard approach. (2000:20).*
4. *A class of models has less demanding rationality assumptions than the standard approach (2000:76).*
5. *The class of models is not an alternative, but a complement, to the standard approach (2000:85,2002:59)*
6. *A class of models can itself be considered a model only if the differences between the models within the class can be parameterized. (2002:58)*
7. *A class of models need not have a "bounds" structure (2000:86, 2002:57).*

Two examples are given of the new paradigm: his own 'bounded' theory of concentration in industrial organisation, and Sadi Carnot's abstract theory of the operation of a steam engine, which predicts a limit to the efficiency of all actual steam engines of a certain class.

Inconsistencies

What follows is a discussion of some of the inconsistencies which occur in Sutton's discussion of the class-of-models. I do not second-guess Sutton's substantive judgements about which models work and which do not, but I think I do show that he appears to be saying inconsistent things.

Supermodels and Parameterizability

An obvious issue that comes up is whether a “class of models” is itself a model, only different from a standard model in being more general. Sutton seems to say exactly this in passing, early in *Marshall's Tendencies*. He is mentioning that Edgeworth's criticism of Marshallian economics is fundamentally the same as his own criticism¹⁴:

there are two ways of looking at Edgeworth's objection ... Different detailed models may be equally plausible a priori, which would lead us to different outcomes (the “class of models” view). Another way of putting things is to imagine that there exists some supermodel that embodies all the particular models ... These two versions of Edgeworth's objection are equivalent ...

Carl Christ puts this more forcefully, in his essay commenting on Sutton's book¹⁵:

Using [the Cowles Commission] terminology, a class of models is itself a model. Thus Sutton's proposed retreat to a class-of-models is the perfectly natural choice of a less specific model when one is not able to settle on a more specific model.

But in reply Sutton insists that his “class of models” can not be interpreted as a supermodel. The reason is that some of the differences between included models cannot be parameterized¹⁶:

[Christ] likes to think of the class of models as being a kind of supermodel within which we embed all the constituent models. Now this is almost true; it would be exactly true if we could reduce all the differences between models within the set to some unobserved parameters, so that we could move across the models by simply shifting the values of these unobserved parameters. Now this is indeed sometimes possible, though in the market structure examples I considered in Chapter 3, for instance, it is not practicable to do this. The two unobservables that matter in these problems are the form of the price competition and the nature of the entry process. For the former, it is indeed possible to find a suitable parameterization which bring us from Bertrand competition at one end, to joint profit maximization at the other (Symeonides, 2001), but there is no way of classifying, ordering or ranking the huge variety of entry models that we might consider, so this strategy just does not work here. The only way to handle things is by formulating the theory in terms of a set of constraints that must be satisfied by any equilibrium of any model of our class.

Parameterization

In the paragraph just quoted Sutton seems to have made a clear division among unobservables on the basis of whether or not their effect can be reduced to a set of parameters, which in turn determines whether a class-of-models can be regarded as a

¹⁴ Sutton, 2000, p8.

¹⁵ Christ, 2002, p23.

¹⁶ Sutton, 2002, p58.

supermodel; all this notwithstanding Sutton's earlier statement that the two formulations are equivalent.

However I do not think that Sutton's parameterization distinction is fit to distinguish his theory from a classical model. First, in his own model price competition is *not* parameterized as a single variable, instead it is a *function* which maps concentration to price, and which captures the effects of a great variety of underlying models¹⁷. The following quotation shows that the theory reduce the effects of a wide class of qualitatively different models on price competition to a single function:

Why use this *function* as a primitive in the analysis? ... The results [of dynamic price competition] are delicately dependent on many unobservable features of the model, and within any particular model specification, multiple equilibria are common (the Bertrand, Cournot, and joint profit maximisation examples considered above are only three of a very wide class of models).

Further, in Sutton's model the nature of the entry process is parameterized in much the same way as the nature of the price competition. The following quote clearly defines a class of models, but it also clearly defines a set of parameters with which to distinguish each model in that class:

Rather than specify any particular form of entry process, we begin with a general class of multistage games, consisting of $T+1$ stages. Each firm is assigned some stage, labelled t , that defines its date of arrival in the market. At any stage from t to T inclusive, the firm may take actions. An action involves making one or more discrete and irreversible investment. An investment results in the entry of a product at some location in an abstract space of locations. While this setup clearly includes location games ... it is sufficiently abstract to encompass many other kinds of model that have been used in the literature on market structure. This framework allows a wide menu of entry games: these range from cases where all firms are assigned the same arrival date ... to cases in which some firms that arrive late will condition their actions on the actions already taken by an early entrant.¹⁸

So, on the one hand, the unobservable factor which Sutton says can be parameterized with a variable (the price competition) is actually parameterized with a function in his model; and the unobservable which Sutton says cannot be parameterized (the entry process) is parameterized with a set of variables. I discuss further the parameterization of qualitatively different models in a later section.

Latent Variables

The same trouble of comparing a supermodel to a class of models comes up in Eric Renault's comment on Sutton's book. Renault, like Christ, prefers to talk about a single more general model instead of a class of models. In particular Renault introduces the idea of 'latent variables,' which are, like the classical error term, unobservable, but

¹⁷ Sutton, 1991, p35.

¹⁸ Sutton, 2000, p71.

unlike the error term, are incorporated directly into the equilibrium model, and do not operate independently. Renault says:

Sutton neatly explains that Edgeworth's objections can be accounted for within two equivalent paradigms: one is the 'class of models' approach, while the other one leads to the adoption of a 'more complete' supermodel that should incorporate additional latent variables.¹⁹

Renault goes on to argue that a latent-variable approach is different and superior (apparently, as Sutton did, contradicting himself in saying the two approaches are "equivalent"). The superiority lies in the greater specificity of a latent-variable model. Regarding a model of option pricing Renault says:

The state-variable [or latent-variable] approach described above appears to be much more productive than the strategy of downplaying Marshall's tendencies with the view that observed 'pricing errors' are simply mispricing that is 'indeterminate within a certain region'²⁰.

Here Renault means the "indeterminacy within a certain region" to apply to Sutton's class of models. Sutton responds to Renault's arguments separately from Christ's, despite them both apparently driving at the same point. In his response he again says that the two approaches are equivalent, and yet one approach is fundamentally different from the other:

the bounds approach to market structure, and the latent variable approach favoured by Professor Renault, are – as we both remark – essentially equivalent 'ways out' of Edgeworth's problem of 'indeterminacy'. The key issue is this: by accounting for observed outcomes with reference to latent variables, we might in principle allow ourselves enough leeway to reconcile any set of observations with a preferred underlying theory.²¹

But it is not clear why a latent-variable theory should be considered unfalsifiable and a class-of-models not. Sutton himself provides a reformulation of his bounds model in terms of latent variables which seems to show that at least one latent variable model is falsifiable²². And on the other hand it hardly seems difficult to construct a class of models which is consistent with any set of observations, i.e. unfalsifiable. I return to this point later.

Completeness of a Model

Related to the questions about supermodels is Sutton's labelling of some models as "complete" or "fully specified" models. He says at one point:²³

¹⁹ Renault, 2002, p38.

²⁰ Renault, 2002, p38.

²¹ Sutton, 2002, p61.

²² Sutton, 2002, p61.

²³ Sutton, 2002, p58.

What views am I arguing against? ... [that] 'the only "proper" kind of model is a fully specified model of the classical kind.

However Sutton never gives a definition of completeness. Discussion of what might make a model complete is brought up again towards the end of this essay.

Existence of a True Model

In *Marshall's Tendencies* Sutton distinguishes the weak and strong interpretation of the standard paradigm. Under the weak interpretation the true model is just a useful heuristic device; under the strong interpretation there really exists a true model. Sutton's final judgement is that the weak and strong interpretations are suitable in different circumstances²⁴.

This scepticism about the existence of a true model naturally raises a difficult question: if we do not believe that any true model represents the situation, then why should we expect a class of such models to be useful? It is as if, to learn about a particular goose, we note everything that is common between ducks. This is generally unclear in the discussion of classes of models, one example of this is in the discussion of rationality.

Rationality

Sutton says that one virtue of his market structure theory, and the class-of-models approach more generally, is that it assumes a weakened form of rationality, 'bounded rationality'. For his proof of the existence of a lower bound on industrial concentration he requires, instead of the usual principle of profit maximisation, two principles which are weaker: the principle of 'viability' (no firm makes a loss), and 'stability' (no profitable niche in the marketplace remains unfilled).

Sutton claims this as a general virtue of the class-of-models approach:

once we follow a class-of-models approach we automatically abandon the need to insist upon strict profit-maximizing behavior. The reason is as follows: if we want to judge, on the basis of the data, whether an agent has taken a maximizing decision, we need to know the "true model" within which the agent was working. [...] So instead of posing the intractable question, Do agents really maximise? 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".²⁵

²⁴ Sutton, 2000, p102.

²⁵ Sutton, 2000, p76.

This claim for the class-of-models approach seems strange because a class of models has been said to be simply a generalisation over a set of models of the standard type²⁶, and Sutton says that standard models assume Nash equilibria, and so complete rationality²⁷, so how can a class of models not require an assumption which each of its member models requires? There are two ways of taking this, the class of models being defined either bottom-up, finding the commonalities between models, or top-down, imposing the weakest assumptions necessary to generate testable predictions.

The first idea is that once you begin generalising over models, and your predictions are becoming progressively weaker, at some point you notice that certain assumptions of rationality have become redundant and can be removed, as if gingerly removing a stick from a pile of pick-up-sticks. In the case of market structure it may be that once assumptions about entry and price competition are removed the testable predictions are identical whether you assume a full Nash equilibrium, or you assume just viability and stability. It is not clear whether this is actually the case. It is essentially a technical issue, and there is no definite claim either way in Sutton's writings. It is perhaps *near* to being the case, because in the body of Sutton's 1991 book all the proofs assume Nash equilibrium and only in the afterword does he mention the possibility of substituting weaker assumptions which could deliver the same conclusions.

If this is what Sutton had in mind: that, although all markets are in fact entirely in equilibrium, the class-of-models predictions happen to require slightly weaker assumptions, then the fact is barely worth commenting on; there is no reason to leave out an assumption which you anyway think is true.

The second way of taking Sutton's claim is to see a class-of-models as built up from very weak assumptions just sufficient to generate testable predictions. This point may be reached without including all the assumptions that are common to the standard models, e.g. we may impose viability and stability without full rationality, just because the latter is not needed to generate predictions.

This boils down to an important question: does Sutton think it reasonable to expect that some market structures are not Nash equilibria?

The answer from his books must be "yes". He says in his 1991 book that the requirement of rationality can seem implausible²⁸:

²⁶ Sutton, 2002, p59, 85.

²⁷ Sutton, 2002, p70.

²⁸ Sutton, 1991, p320.

One important objection to many current game-theoretic models is that they rest on the notion that firms carry out quite subtle calculations that require an extremely detailed knowledge of the environment they face.

Sutton also dwells on how much less demanding his assumptions are than those of full Nash equilibria²⁹:

in imposing the viability and stability conditions, all that we require of *all* agents is that they do not violate the viability condition. For the stability condition to be satisfied, what we require is that ... there is always one smart agent available who will fill the gap in the market.

But if actual markets are not in Nash equilibrium, then the success of Sutton's theory is not due to it containing a class of models of the standard type: it only works because it happens to include models *not* of the standard type.

Finally, note that Sutton's use of the term "bounded rationality" is very different from that generally used. Sutton says that his theory involves bounded rationality because it has weaker assumptions than full rationality³⁰:

instead of posing the intractable question, Do agents really maximise? It may be more fruitful to ask, to what extent can we deduce, by reference to observed actions, whether the agents were maximizing?

However, "bounded rationality" is normally used to mean something substantive about actual behaviour or decision making, especially that people do not obey some rule of rationality, i.e. it involves assumptions inconsistent with full rationality, instead of weaker than full rationality³¹.

Summary

Using the "class of models" distinction leads Sutton to make various inconsistent statements. He both asserts and denies that a class is equivalent to a supermodel. He makes two further distinctions – parameterizability and falsifiability – which both fail to distinguish a class from a model. He also appears to claim that a class of models is used just to generalize over models of the classical sort, yet that it has the significant benefit of not making assumptions common to all these classical models, for example that of a Nash equilibrium.

²⁹ Sutton, 2000, p77.

³⁰ Sutton, 2000,p76.

³¹ See, e.g., Gigerenzer & Todd, 1999, p12.

PART THREE: AN ALTERNATIVE INTERPRETATION

I think that the fundamental departure of Sutton's method can be seen in a quite different way from that proposed by himself, Christ and Renault. I suggest this: classically, when testing a model against observational data the test is only valid under the assumption that the observed and unobserved elements are uncorrelated; and this is a restriction with which econometricians are always struggling. However, when some information is known about the distribution of the unobservable influences, the uncorrelatedness assumption may not be required to perform a test. Sutton shows that if his model is correct then the influence of the unobservable elements forms a one-sided distribution, which means that the model can be tested without making any assumptions about the correlation between observable and unobservable parameters.

Let us say we have a model of the effect of exogenous variables on an endogenous variable, of the form $y = f(\mathbf{x})$ and suppose that all the exogenous variables in the set \mathbf{x} are observable. We wish to test this model against the data, but we know that the dependent variable is also influenced by other, unobservable variables \mathbf{u} :

$$y = f(\mathbf{x}) + g(\mathbf{u})$$

The problem then is that if we both know nothing about the distribution of the unobservable variables \mathbf{u} , and nothing about the form of the function $g(\cdot)$, then we cannot test the model because any distribution of observable outcomes is consistent with the model $f(\cdot)$, given some distribution of the unobservable influences, $g(\mathbf{u})$. For a test to be valid we must impose some assumptions on the second term in the equation.

The workhorse assumption of econometrics is that the unobserved influences are uncorrelated with the observed influences, i.e. that

$$E[\mathbf{u}|\mathbf{x}] = E[\mathbf{u}]$$

which allows us, using a central limit theorem, to make judgements about the probability distribution of a sample statistic, and so allows the standard practice of rejecting a hypothesis when the magnitude of some sample statistic is unlikely given the hypothesis.

The uncorrelatedness of the outside influences is therefore very important, and much of the effort of testing a model is spent on justifying the assumption of no correlation, generally by "adjusting for" or "controlling for" outside influences, i.e. by including within the model other variables which influence the outcome, and are correlated, and are observable. Once all the variables which can be included have been included, no test is available which can demonstrate the lack of other correlated influences. Instead a

judgement must be made about how plausible it is that there should be no other influences. This is why Sutton keeps repeating throughout *Marshall's Tendencies* his motivating phrase, the problem for theories of market structure being: “unobservable factors that we cannot measure, proxy, or control for”³².

One way to entirely solve the problem is to use a randomised experiment: if you randomly choose cases in which to manipulate \mathbf{x} then the randomness guarantees that \mathbf{u} will be uncorrelated, and so tests of the model can proceed. In industrial organisation generally economists can not of course afford to experiment by manipulating an industry's demand, or its responsiveness to advertising, or its production technology.

The general problem of testing can be summarised in this way: trying to find a test statistic, $F(\cdot)$, which has a threshold value $k(\alpha)$, where under the null hypothesis the probability of it exceeding that threshold is less than α . So in general, we need to prove a theorem of the form:

$$P[F(\mathbf{X}) \geq k(\alpha)] \leq \alpha$$

With an uncorrelatedness assumption we are able to get a version of this theorem (this is assuming a linear model and unknown error variance):

$$P[\sum(y - \mathbf{x}\beta)^2/\sigma \geq F_{\alpha,m}] = \alpha$$

There does exist an alternative assumption which allows testing, the use of which, I believe, constitutes the best distinction between Sutton's bounded models and orthodox economic models. Any assumption which limits the range of the function $g(\cdot)$ restricts the range of possible outcomes, allowing a test. For example if we assume that $g(\cdot) \geq 0$, then we know that $f(\mathbf{x}) \geq y$, which is a testable restriction on outcomes: if we ever find a data point where $f(\mathbf{x}) < y$ then one of our assumptions must be incorrect, i.e. either the theory is wrong or our testing assumptions are wrong. The new testing theorem can then be written as:

$$P[f(\mathbf{x}) - y < 0] = 0 \leq \alpha$$

Although this new assumption, of bounded unobservables, helps with constructing a test of a model, it remains much less generally useful than the uncorrelatedness assumption, particularly because the uncorrelatedness assumption can allow full identification of the parameters in a model. The bounded assumption does allow some identification, see Manski (1995).

³² Sutton, 2000, p8, 23, 32, 58, 70, 85, 92.

This substitution of assumptions to allow a test is almost what Sutton does when testing his model: he confirms that no industry has a concentration lower than that predicted by his lower bound.

Almost, but not quite. The actual way that Sutton tests his bound is slightly inconsistent with this characterization. I discuss this in a later section.

PART FOUR: RELATED ISSUES

Argument over Supermodels

I think that Sutton's class of models is itself model, as Sutton, Christ, and Renault each assert at points, but which Sutton and Renault also deny at other points.

As I have already mentioned Sutton argued that a "supermodel" is not equivalent to a "class of models," because the differences among a class of models cannot always be parameterized. The solution to this problem is already used by Sutton himself: reducing the complicated differences between two models to a function. For example, take a pair of models with qualitatively different choice problems, where the relationship between agents' choices and their expected payoffs are represented with the functions³³:

$$f_1(x,y,z)$$
$$f_2(x,y,a,b,c)$$

How do you create a supermodel, meant to understand the operation of x and y , which can encompass both of these models? If the supermodel will be used in conjunction with the assumption of a Nash equilibrium then the problem is easily solved, because the values of the idiosyncratic variables (z , a , b and c) can simply be assumed to be whatever maximises the payoff; i.e. these functions can be used:

$$g_1(x,y) = \max_z f_1(x,y,z)$$
$$g_2(x,y) = \max_{a,b,c} f_2(x,y,a,b,c)$$

This is entirely standard practice in applying game theory: for example a firm can produce anywhere within the production possibility frontier, allowing many possible combinations of inputs and outputs, but their behaviour is normally represented with just one choice variable: the level of output, and all other choice variables are assumed

³³ The functions are assumed to include the actions of all the other agents so that an equilibrium can be calculated, Nash or otherwise. To be explicit this could be written as:
 $f_{1,i}(x,y,z|\mathbf{f}_{1,-i})$.

to be at their profit-maximising level with respect to that variable. This is exactly how Sutton handles the parameterization of price competition and entry, as noted earlier.

Sutton also has an argument against the 'latent variable' model³⁴:

The key issue is this: by accounting for observed outcomes with reference to latent variables, we might in principle allow ourselves enough leeway to reconcile any set of observations with a preferred underlying theory.

Consider the one situation in which this objection would have force: an observation which is inconsistent with some class of models, but is consistent with some latent variable model – meaning that there is some value of the latent variables which allows that observation. Why, then, is the class-of-models inconsistent with the observation? Because it has a set of assumptions which rule it out. But if those assumptions are reasonable enough to put in a class-of-models then they should likewise be reasonable enough to put into a latent-variable model, and so would likewise rule out that observation. Is there something about latent-variable models which makes it impossible to add assumptions which restrict the range of outcomes? Clearly not, because Sutton himself suggests a reformulation of his own bounds theory in terms of latent variables (2002:61), and emphasises that it is still possible to reject the theory in this form.

So this disposes of Sutton's two arguments against considering his class-of-models as a model itself: his parameterization argument, and his falsifiability argument. Before going on to the next subject I will discuss two other ways that a class of models could be distinguished from a model: the realism of the assumptions, and the existence of a Nash equilibrium.

At first glance, there seems to be a contrast in the realism of the assumptions used in a standard model, and those used in Sutton's "class". In developing his theory in *Sunk Costs and Market Structure* Sutton gives a fully-specified model of the standard kind before moving on to his "class of models," of which this complete model is one example. The complete model has a number of unrealistic assumptions including that the consumers all choose according to identical utility functions, that the consumers all have the same income, and that the firms all face identical production conditions. These assumptions are clearly quite false of virtually all actual markets, and so as is common with economic models, some predictions can be drawn from the model which are flatly wrong. For example, this model predicts that all firms active in an industry will have identical market shares, which in Sutton's data is quite untrue. So the model can only be used with a sort of care over which predictions should be taken seriously.

³⁴ Sutton, 2002, p61.

In contrast to this model, done in the orthodox style, Sutton's 'class of models' uses assumptions which appear less unrealistic. Nothing is assumed about the utility or the income of consumers: consumer demand only enters the model in terms of how market share reacts to fixed investments. Similarly, firms are not required to choose their profit-maximising strategy, weaker assumptions are substituted – as discussed earlier.

However this distinction is also not sufficient, because the 'class of models' is not able to shake all unrealistic assumptions. The proof of the nonconvergence theorem still assumes the following propositions, false of most markets: (1) constant marginal costs³⁵; and (2) the effect of advertising on willingness to pay is independent of market size³⁶. The second assumption is discussed further in a later section.

The last distinction to be tried out is that of whether or not a theory can be shown to have a Nash equilibrium in pure strategies. In *Sunk Costs and Market Structure* Sutton mentions that in some of the games which his class includes a pure-strategy equilibrium will not exist³⁷. Sutton says he avoids the common problem in modelling of having to impose assumptions to ensure that a pure-strategy equilibrium exists: "we keep the class of models as broad as possible and seek only to *characterize* equilibria, whenever they exist."

So a distinction between a model and a class of models could be the existence of pure strategy equilibria. But this would be a very weak distinction, because, first, all finite games do have equilibria, though some of them may be mixed³⁸; and second, is not uncommon for models in Industrial Organisation to have only mixed equilibria.

The Two Distinctions

Despite these considerations it is possible that a consistent and sensible way could be found to distinguish a class of models from a model. Then Sutton's IO theory would be distinguished both by its 'bounded unobservables', and its being a 'class of models'. The question would become which of these distinctions explains the theory's ability to deal with the problem of testing which Sutton poses in *Marshall's Tendencies*. This can be answered by considering cases where the two distinctions diverge.

First, a theory which has bounded unobservables, but which is not a class of models, does solve the problem of testing in the presence of unobservables, as explained in the previous section on the basis of model testing.

³⁵ Sutton, 1991, p69; Sutton, 2002, p67.

³⁶ Sutton, 1991, p72; Sutton, 2002, p74.

³⁷ Sutton, 1991, p75.

³⁸ As proved by Nash.

Second, consider a theory which uses a class of models, but is not bounded. Sutton suggests that such a theory may still solve the problem of unobservables³⁹:

[the theory of sunk costs] leaves us with a nice statistical description of the data, in terms of a “one-sided error distribution.” In most cases, the unobservables for which we cannot control will not be so easily handled. This may not rule out the use of a class-of-models approach – but it will probably make it harder to implement successfully.

However there is not further discussion of how a class-of-models can solve the problem without a one-sided error distribution (i.e., bounded unobservables), and Sutton’s only other example of a class of models is Carnot’s theory of steam engines, which also has a one-sided error distribution⁴⁰.

Finally I should note that Carnot’s theory of steam engines appears not to be a very good example of a class of models itself: it appears to be fully-specified and simplified, as opposed to Sutton’s ideal of a class as partially-specified and not too simplified.

Sutton says of the model:

Instead of building a “realistic” mathematical representation of some typical engine, [Carnot] began with a different question. He asked: is there anything that can be said, independently of the details of design, about the factors that must ultimately limit the efficiency achievable by *any* engine? [...] Carnot reduced the description of the engine to a simple abstract representation, in an attempt to isolate some basic features common to a large class of engines.

Note that Carnot’s representation is common to a “class of engines”, not a “class of models”. In fact Carnot’s simplifying assumptions were strictly true neither of the engines of his day, nor of others’ models.

Fitting the Bound

Much of the discussion so far has been about testing whether a bound on the data is violated. However, although Sutton predicts a bound, his model cannot actually predict where the bound will be or even what shape it has, so he cannot directly test whether it has been violated. Instead he first assumes a functional form for the bound, then estimates the parameters of that function, and then statistically tests whether the estimated bound converges to zero or not, for both of the two sets of industries: those with low advertising and those with high advertising. He finds, as expected, that the estimated point of convergence of the bound on low-advertising industries is not significantly different from zero, but the point of convergence on the high-advertising industries is significantly different from zero⁴¹.

³⁹ Sutton, 2000, p85.

⁴⁰ Sutton, 2000, p62.

⁴¹ Sutton, 1991, ch5.

Clearly this test sacrifices many of the advantages of the 'bounded' method of testing models.

The distribution of unobservables may distort the measurement of the bound: e.g., the true bound could be much lower than his estimated bound (which is 14% for the C_4 of high-advertising industries), if in the data-set some unobservable influence was correlated with market size. The bound estimated would then be a false bottom, and vulnerable to a change in the distribution of the unobservable influences.

The problem of having an estimated bound, instead of a predicted bound, is of course that the estimation process is vulnerable to the problem of correlated unobservables, and so we are back to where we started. Again the unobservable influences must be assumed to be uncorrelated. Sutton's many comments about his method being motivated by the problem of unobservables which we cannot "measure, proxy, or control for" seem wasted.

However, the test Sutton uses can be justified by a slightly weaker assumption than full uncorrelatedness. For a regular regression we assume that the unobservable influences are uncorrelated with the observable influences. For this regression we only require that the minimum value of the uncorrelated influences are uncorrelated; or that a certain proportion of the conditional distribution of the influence of unobservables is within a certain distance from 0. From the certain proportion and the certain distance you could calculate upper and lower limits for the existence of the bound. Sutton does not discuss anything similar to this alternative weaker assumption, but it is the only possible justification which I can see for his practice.

Finally, though I think that this is an unsatisfactory part of Sutton's theory, I do not think that it makes any material difference in the battle between Sutton's "class of models" and my "bounded unobservables" interpretations of his theory. The problem in IO remains: unobservables; the solution remains: predicting bounds on the data. Because Sutton cannot derive from theory any quantitative predictions about his bound, the application relies – partly – on the old assumption of uncorrelated unobservables.

In light of this I think that Sutton's statements about testing of his theory are a little misleading, for example⁴²:

We can ... find a bound, relative to which the effects of our unobservables all operate in the same direction. This is the key to retaining a testable theory: we can reject the theory if we can show that the restrictions imposed by the properties of these bounds are violated

⁴² Sutton, 2002, p61.

Whereas, first: no observation in his data-set could have contradicted his bound, because there was no bound predicted. And second, although Sutton predicts that the high-advertising bound does not converge to zero, he cannot say where it should converge to, so his test simply rejects the null hypothesis that the point of convergence is zero. No test of the same type is capable of rejecting his alternate hypothesis, of a non-zero point of convergence, because an infinitesimally small point of convergence is consistent with his theory. So indeed any set of observations is consistent with his theory.

Economic Objections

There are two points about the construction of Sutton's theory that I would like to mention, which are relevant to judging its success in solving the problem of unobservables.

Elasticity. Sutton's non-convergence conclusion depends on an assumption about the shape of the advertising response function, which measures the effect of advertising expenditure on consumers' willingness-to-pay. The assumption is described as if it is very natural, but in fact is a strong assumption on the shape of the function, it is never directly justified, and there exist plausible functional forms which do not satisfy this assumption.

Sutton discusses what it is justified to assume about the advertising response function⁴³:

Two empirical features regarding advertising response functions are well established: first, threshold effects may exist (in the sense that advertising below a certain threshold level has little impact). Second, apart from the possible existence of such thresholds, the effectiveness of advertising outlays is subject to diminishing returns.

Then when constructing his class of models he assumes that this function has a finite point of maximum elasticity (where u is willingness-to-pay and F is expenditure, so $F(u)$ measures the cost of improvements in willingness-to-pay)⁴⁴:

CONDITION 2: The function $F(u)$ is continuous and increasing on $u \geq 1$; and its elasticity is bounded above, viz., for some beta,

$$\frac{u}{F} \frac{dF}{du} < \beta, \text{ for all } u \geq 1.$$

This assumption ensures that the cost of increasing willingness-to-pay does not increase too quickly; when the elasticity is increasing then the standard concentration-decreasing effect may outweigh Sutton's concentration-increasing effect, but if elasticity has a

⁴³ Sutton, 1991, p51.

⁴⁴ Sutton, 1991, p72.

maximum value then concentration will have a minimum value. However, consider the function $F(u) = a^u$, for which elasticity is unboundedly increasing, which means that concentration could indefinitely fall even in an advertising-intensive industry, contradicting the theory. So this assumption is both substantive and crucial.

It seems to me that Sutton misrepresents the situation when discussing the assumption⁴⁵:

A failure of this property, then, would arise in a situation where there is some *absolute* level of advertising outlays beyond which further increases become ineffective. What matters in practice is whether such a saturation level is achieved across the empirically relevant range of parameter values. What happens if such a saturation level exists? In terms of the theory, we might model this by letting the $A(u)$ schedule become vertical beyond some point (u_∞, A_∞) . If this case holds, then that [sic] increases in market size may lead to an indefinite fragmentation of the market, as an indefinitely large number of entrants each undertake an advertising level A_∞ .

Sutton implies that a function violating his assumption must have some *finite* point of infinite elasticity, which does sound unlikely, but in fact this is not necessary. An exponential function has unboundedly increasing elasticity, and so violates the crucial assumption, but at any finite point an exponential function has a finite elasticity. The mention of the “empirically relevant range” further confuses things, because Sutton’s theory is about what value concentration converges to in the *limit*; further, the exponential function has no “saturation level” in the empirically relevant range, but still contradicts the theory because it has increasing elasticity, and so allows decreasing market concentration.

Advertising. One of Sutton’s assumptions turns out to be strange on closer examination: that advertising is a fixed cost, i.e. the effect of money spent on advertising on raising willingness-to-pay is assumed to be independent of the level of production and independent of the market size. This means that if £100,000 of advertising is sufficient to double the price paid for your product in a small market, it is also sufficient to double the price in a large market, which seems implausible. There is an ambiguity in whether a cost is “fixed”: an investment in R&D is fixed with respect to the number of units sold, and to the market size, whereas it seems to me that spending on advertising is fixed with respect to the first, but not the second. Some advertising costs are fixed in both senses, for example the money spent on the creative work of an advertising agency; however I do not think that this cost dominates, and indeed Sutton measures the advertising levels of industries by simply counting the number of advertisements, and then estimating the amount of money which they would have cost⁴⁶.

⁴⁵ Sutton, 1991, p78. Note that $A(u)$ and $F(u)$ can be treated as equivalent for the purposes of this discussion, see Sutton, 1991, p69.

⁴⁶ Sutton, 1991, p101.

If we were to treat advertising as fixed with respect to quantity sold, but varying with respect to the size of the market, then a simple classical model can be produced which displays nonconvergence. Let the common level of sunk costs in an advertising intensive industry be proportional to market size:

$$\sigma = \sigma_0 + S \cdot \gamma$$

Where σ_0 is the minimum efficient scale, S is the market size, and γ represents a level advertising necessary to survive in the market. Then concentration, in a symmetric Cournot model, is⁴⁷:

$$C_1 = [S / (\sigma_0 + S \cdot \gamma)]^{-1/2}$$

Which does not converge to zero as S goes to infinity:

$$\lim_{S \rightarrow \infty} C_1 = \gamma^{1/2}$$

So that an industry in which each firm must spend 2% of total market revenue on advertising will never have less than 15% C_1 concentration. In a way this model just restates Sutton's central intuition: that advertising expenditure increases with market size, and may increase so fast that it counteracts the tendency of concentration to decrease. However there is a considerable difference between my formulation and Sutton's: this model does *not* belong to Sutton's class, because it violates the assumption that willingness-to-pay depends on fixed costs irrespective of market size.

CONCLUSION

Sutton's 'bounded' theory of market structure predicts the existence of a minimum level of concentration, determined by observable factors (market size and the level of advertising), and which can never be broken by unobservable influences (e.g., entry conditions, price competition). Because of this robustness to unobservable influences Sutton appears to be able to use cross-section industry data to test his theory, whereas normal theories cannot be tested with cross-section data because of the problem of unobservables being correlated with the observable independent variables.

Sutton explains his bounded theory, and its success, as differing from the standard approach in its covering a *class* of models instead of only a single model, and so predicting outcomes which would hold for any model in that class. In contrast, I think that the success of the theory is due to its treatment of the unobservable influences:

⁴⁷ Using the formula for the number of firms in Sutton, 1991, p31.

because Sutton's bounded theory puts a limit on the influence of the unobserved factors, his theory can be tested without the assumption that is usually so difficult to justify: that the unobserved factors vary independently of the observed factors.

In fact both distinctions are sufficient to differentiate Sutton's theory from a standard orthodox theory: his theory is both bounded, on the one hand, and includes a class of orthodox models on the other hand. However it is its boundedness, not its being a class, that allows the theory to solve the problem of unobservables. Trying to explain the success of the model in solving the problem of unobservables using the 'class of models' distinction gets Sutton into trouble because it is a pragmatic distinction, i.e. only defined with respect to some purpose. Fundamentally, as Christ teaches, a class of models is itself a model.

The issue becomes more complicated if, as I have mentioned, there are some signs that Sutton's theory in fact does not solve the problem of unobservables. Because the shape of the bound, and the height of its asymptote, are not predicted by the theory the parameters must be estimated from the data, and this estimation requires at least a weak version of the standard assumption of uncorrelatedness. This also means that no test can be constructed capable of rejecting the theory.

Despite these problems we do observe an absence of industries roughly where absences were predicted: particularly an absence of industries with a large market size, high levels of advertising, and C_4 concentration less than 14%. However, because the theory only predicts that there is some minimum level of concentration which is not broken, it does not predict what that level is, and it would be consistent with the theory if the lower bound were 0.5%. So any practical applications of from the theory requires assuming not just that the elasticity of cost is bounded, but that it has a quite low bound⁴⁸.

This last fact does most of the practical work, but the reason that this fact holds – the nature of advertising which keeps costs from increasing too fast as market size increases – is not discussed. It is simply regarded as a brute fact of nature. This makes the claimed robustness of the bound doubtful⁴⁹: if we do not know why cost elasticity is generally low, we do not know what changes are likely to raise the elasticity.

The problem is illustrated in a recent study of EU industrial concentration using Sutton's method: whereas Sutton estimated a bound at 14%, Lyons, Matraves and Moffatt (2001) estimated a bound at 7% for UK advertising-intensive industries, and at only 2% for EU-wide advertising-intensive industries. Lyons et. al's sample of industries is different, as are some points of their methodology, but their results make clear that the position of

⁴⁸ Strictly: $\max_k \alpha (k)/k^\beta \geq 0.14$

⁴⁹ Sutton, 1991, p11 and *passim*; Sutton, 2000, p85 and *passim*.

the bound must depend on some factors beyond just market size and advertising intensity, and without knowledge of how those factors work, the bound is much less useful than it could be.

BIBLIOGRAPHY

- Christ, C. (2002). "Sutton on Marshall's Tendencies: A Comment." *Economics and Philosophy*, 18 (2002) pp. 21-27.
- Friedman, M. (1953). "Methodology of Positive Economics" in *Essays in Positive Economics*.
- Gigerenzer, K., and Todd, P. M., (1991). *Simple Heuristics That Make Us Smart*, Oxford: Oxford University Press.
- Lyons, B., Matraives, C., Moffatt, P., (2001). "Industrial Concentration and Market Integration in the European Union." *Economica*, 68, pp. 1-26.
- Manski, C. F. (1995). *Problems of Identification in the Social Sciences*. Cambridge, MA: Harvard University Press.
- Renault, E. (2002). "Comments on Marshall's Tendencies." *Economics and Philosophy*, 18 (2002) pp. 29-44.
- Schmalensee, R. (1989). "Inter-Industry Differences of Structure and Performance," in Richard Schmalensee and Robert Willig (eds.), *Handbook of Industrial Organisation*, Amsterdam: North Holland.
- Sutton, J. (1991). *Sunk Costs and Market Structure*. Cambridge, MA: MIT Press.
- Sutton, J. (1998). *Technology and Market Structure*. Cambridge, MA: MIT Press.
- Sutton, J. (2000). *Marshall's Tendencies*. Cambridge, MA: MIT Press.

Sutton, J. (2002). "Marshall's Tendencies: A Reply". *Economics and Philosophy*, 18 (2002) pp. 55-62.

Sutton, J. (2002B). "Rich Trades, Scarce Capabilities: Industrial Development Revisited". *The Economic and Social Review*, Vol. 33, No. 1, pp. 1-22.