

REFUTING GROUPS OF THEORIES

Peter Bowbrick¹

ABSTRACT

There is an infinite number of possible wrong economic theories. It is impossible to refute them one by one. Either they must be left unchallenged, or a method of refuting groups of theories must be used. One approach is set out here.

"The ingenuity of these nineteenth century writers knew no bounds when it came to giving reasons for ignoring apparent refutations of an economic prediction, but no grounds, empirical or otherwise, were ever stated in terms of which one might reject a particular theory" Mark Blaug, *The Methodology of Economics*, (1980, p55).

INTRODUCTION

There is an infinite number of possible economic theories based on different combinations of assumptions and logic, and virtually every paper in the literature purports to present a new model, a new variant of existing theory or new theory. Unless the discipline has a method of identifying and rejecting 'bad' theories, the few which are 'good' will be snowed under. It is not possible to avoid this by refuting individual papers or variants of papers, unless most of the papers published are successful refutations. Instead, groups of theories must be refuted. This paper shows an application of an approach to refuting groups of theories.

Over the last thirty years I have found several groups of theory on the economics of quality which were rigorous and practically applicable, but the subject is dominated by another set of theories, those based on rational economic man (REM) making optimal choices with perfect knowledge about his own preferences, the objective characteristics of all products, and the availability and price of all products. The seminal works are Lancaster (1966, 1971, 1979)², Rosen (1974), Houthakker (1954), Thiel (1952) and Brems (1948,1957).

This paper shows an approach to refuting the groups of REM theories that are based on these seminal works, by identifying failures in assumptions, logic, prediction and testability. As the weight given to these failures by economists varies according to their epistemological standpoint, the analysis is presented first, then its force is examined in the light of five epistemological standpoints. The detailed economic analysis of the weaknesses is not presented here.

FUNDAMENTAL ASSUMPTIONS AND LOGIC

There are more than 10,000 and probably up to 20,000 papers in the REM literature on the economics of quality. Virtually all new papers present a new variant of a theory or of a situation-specific model. It is not feasible to attempt to refute them one by one, and even if one successfully refuted the first 9000, it would be no indication that the next thousand were wrong.

Accordingly, four areas were identified which are fundamental to the research programme at so basic a level that the theories could not reach their first paradigm case (the optimal purchase of the individual) without them. The four are: 1) assumptions on individual consumers' preferences, 2) supply assumptions, 3) the logic of characteristics space, and 4) the logic of subjective and objective quality. This paper is concerned with papers that depend on at least one of these, but most depend on all of them. The impact of the criticisms of

¹ Copyright © Peter Bowbrick, peter@bowbrick.eu 07772746759. The right of Peter Bowbrick to be identified as the Author of the Work has been asserted by him in accordance with the Copyright, Designs and Patents Act.

² Becker (1965), Muth (1966), Gorman (1956) and Ironmonger (1972) produced similar theory, but it lacked the rigour of analysis of fundamental assumptions and logic, and had little impact, so cannot be considered seminal.

assumptions and logic which will be set out is greatest not on the papers which set out the fundamentals and which are cited here, but on the papers being written today which are derived from the same fundamentals through a long chain of logic: 'for the want of a nail a kingdom was lost'.

Grounds for rejecting a theory on the basis of its assumptions, from some epistemological positions, are:

- a. The assumptions conflict with observed reality. (Simplifications of observed reality are acceptable and desirable).
- b. The assumptions, explicit or implicit, contradict each other, so the theory is logically impossible. In practice this may occur when 'relaxing initial assumptions' involves accepting all the conclusions of previous analysis but changing the assumptions on which they were based for subsequent analysis. It is common for writers to accept the conclusions of previous writers, and then proceeding with the analysis using a different set of assumptions.
- c. The assumptions rule out all possible real situations.
- d. The assumptions rule out so many possible real situations that the theory is trivial.
- e. The assumptions cannot be tested against reality, so it is not possible to say whether they apply in a particular case. Any failure of predictions can be explained away as 'the assumptions cannot have applied here' rather than 'the theory is wrong'. The theory is of the form 'This theory predicts accurately if the consumer has an even number of guardian angels'. This makes the theory untestable or 'unscientific' in Popper's sense.

Assumptions on Consumer Preferences

The theories rely on the fundamental assumptions that each individual consumer values each characteristic positively and that the indifference curve for two characteristics is similar in shape to the textbook indifference curve for two goods³ (Figure 1). This implies transitivity, completeness, continuity, strict convexity, non-satiation and all characteristics positively desired. (See Lancaster 1971 for analysis of these fundamental assumptions).

An error here has been to use an indifference analysis developed for two different **goods** which are bought separately and consumed separately. When considering the **characteristics** of a good the theory should recognize that for most goods the characteristics are necessarily bought together and consumed together.

Figure 2 plots indifference curves for two characteristics which the consumer values separately, making the normal assumption of increasing then declining marginal utility. These curves bear no relation to those assumed by REM theories of quality. Figure 3 shows a bullseye where the consumer likes a medium-sweet, medium-acid wine, but finds a wine that is too sweet as unpalatable as one that is too dry. Similar curves will be the norm for most pairs of characteristics.⁴

For nearly every good there is the possibility of contamination or infection, implying another, very different indifference curve with all the points on the axes, or on a product possibility curve (Figure 4). Bowbrick (1992) shows that the main indifference curves for a range of everyday goods, and a few of the indifference

³ In REM theories, a good is usually defined as a unique mix of objective characteristics, so that one mix makes a Washington Navel orange, while oranges in general are a group of goods.

⁴ In the early work on indifference curves it was recognized that such indifference curves were the norm even for two goods (e.g. Boulding, 1965), but if, in addition to this preference assumption, a supply assumption was made that more consumption of a good meant more outlay, a rational consumer would buy on only one arc of the circle, so only the part of the circle shown in Figure 1 is relevant. With the characteristics of a good, acid in wine, for instance, this supply assumption is absurd. The REM theorists have been presenting the results of some supply and preference assumptions as though they were purely a preference assumption. The error is taking the results of previous theory, but changing the assumptions on which they were based for the rest of the analysis.

curves for virtually all goods are very different to those of REM theory. It is unlikely in the extreme that, for any good, **all** individual consumers will have indifference curves like those in Figure 1 for **all** pairs of characteristics, but this is what the theory demands.

Fundamental assumptions on supply

The fundamental assumption on supply is that the buyer will have to pay more to get a product with a higher level of any characteristic. This means that wines get more expensive, the more acid they are, that cars get more expensive the softer their seats are, and that a computer gets more expensive the bigger it is and the heavier it is. Clearly, with most goods there is no such simple relationship between the level of the *objective* characteristics of a good and its price, and it can be shown that it is particularly unlikely with the price making markets which REM theory implies. Without the assumption that this is universal, however, it is not possible to reach the first paradigm case of the optimal choice of the individual. The Figure 1 solution of optimizing where the indifference curve is tangential to the indifference curve does not hold.

Virtually all REM theory of quality depends crucially on this assumption.

The Fundamental Logic of Characteristics Space

REM theories of quality are analyses of supply and demand within ‘characteristics space’, a space defined by axes [level of] characteristic A and [level of] characteristics B as axes, rather than Good 1 and Good 2, as in the traditional theory of Figure 1.

There is a logical error here. There is not just a single characteristics space, as assumed in the theory. Any change in the definition of the axes necessarily changes the scale and the shape of curves plotted within the curves, while at the same time the indifference curves, budget lines etc. necessarily have a different meaning. The logic that leads to the first paradigm case with one definition of the characteristics of axes need not do so with another definition.

One direction of change in characteristics space may be shown by plotting indifference curves for the quality of sugar and curry powder a) in a mouthful of ice cream, b) a meal, c) one’s diet and e) total consumption of all goods and services. While REM theories usually start by assuming (e) it is difficult to argue that people have a concept of curry powder in total consumption, and the analysis slips imperceptibly into another characteristics space. Lancaster, for example, uses at least a dozen different characteristics spaces:

1. Total amount of characteristic in total consumption. This requires the assumptions of linearity and additivity. It appears to be the characteristics space used for the basic paradigm case.
2. Total amount of characteristic in the diet (1971, p17).
3. Total amount of characteristic in a single unit of a good. This is the space used for the automobile example (1971, pp.157-174).
4. One axis being ‘Cleaning power per dollar’ for goods in the product group detergent (1966, p.153). This conflates two characteristics and introduces concepts like value for money. It does not appear in Lancaster (1971).
5. Level of characteristic obtained from one or more goods in one product group. This appears to be the characteristics space used for most of the analysis, including that which at first sight uses the paradigm case (1971, pp.125-9).
6. Characteristics per unit of a good (1979, p28).
7. A space with a ‘normalized’ efficiency frontier, implying some kind of ‘normalized’ definition of characteristics (1971). This is used for his second paradigm case. In fact different ‘normalized’ spaces may be created starting from any of the six previous spaces and be related to total consumption, to an automobile etc. so there are many more than seven spaces used.⁵

In spite of these major changes in the definitions of the space used and the axes, none of the curves change shape when they are plotted in a different space! I am not aware of anyone working in REM theory who

⁵ Lancaster’s ‘normalized’ curves are presented in a ‘characteristics in total consumption’ space, but most of his followers present the identical diagrams in a different space ‘[Normalized] level of characteristic [in total consumption] per dollar’, without explanation.

notes the switch in characteristics space, much less who adjusts the shape accordingly. This is similar in effect to altering one's fundamental assumptions in the middle of an analysis.

The Fundamental Logic of Subjective and Objective Quality

A major attraction of REM theories of quality is that they claim to be able to operate with objective characteristics, ignoring perceptions, psychology and so on.⁶ They work on the assumptions that a) goods have objective characteristics on which individuals make their choices, b) all individuals see the same characteristics and perceive them identically, but c) they may value them differently.

These assumptions conflict with those of much of mainstream economics, including the theory of monopolistic competition, the economics of markets, the economics of information and the economics of advertising. They are shown to unrealistic by the whole of marketing. They are, therefore, another set of unrealistic fundamental assumptions.

At the same time they introduce logical errors. REM theory requires that it is possible to plot all individuals' indifference curves in the same space and that the axes are specified in terms of objective characteristics. If it is not possible to plot **all** individual's preferences on the same graph with the same axes it is not possible to reach the second paradigm case used for comparison of the purchase decisions of two individuals. It is not possible to plot all individuals' preferences in the same characteristics space if, for example, some people perceive non-existent characteristics like lucky numbers, ignore 'objectively important' characteristics like BSE in beef, or just perceive different characteristics or subjective attributes as important.

It is not enough that individuals should perceive the same characteristics: they should perceive them in the same way. It is possible to plot a car's power in objective characteristics space if **everybody** sees it in terms of BHP, but not if everybody is perfectly informed, but some see it in terms of acceleration, some in terms of top speed and some in terms of ability to pull a caravan.

In practice, theorists drop the assumption of objectivity the moment they start talking of real products, but usually without noticing that they are doing so, or mentioning it (e.g. Lancaster's 1971 automobile example). It is evidently not possible to talk about real products without introducing subjectivity.

If an individual, John Smith cannot perceive objective characteristics, the most he can do is plot his indifference surfaces against the axes 'My perception of the level of characteristic A per kilogramme' for instance, but Mary Jones cannot plot her indifference surfaces within this space, because her perceptions may be quite different. There would also be problems with two detergents occupying completely different positions in subjective space, even when their objective properties were identical. If an outside economist were to try to plot the two in a common space, it would be one determined by the outsider's perceptions, again subjective. In effect, each point (that is to say each brand, each grade, each quality) would then be identified by the researcher, then plotted in his or her own space. Inevitably, shapes would change: a totally rational curve in Mary Jones' subjective space (even one of the shape of Figure 1) could change shape and appear totally irrational in the researcher's subjective space.

The moment this subjectivity is accepted, therefore, it becomes impossible to reach the second paradigm case where the preferences of individuals are compared in the same space.

This is not to say that economics with subjectivity is impossible. One version of hedonic theory goes 'The subjective attributes which individuals ascribe to this group of goods has been identified, and so have the levels of attribute that they ascribe to it. The hypothesis is that there will be the greatest demand for those

⁶ The characteristics which appear in the analysis are assumed to be objectively quantifiable, as well as objectively identifiable, even though they are important characteristics (colour for example) that do not fit this specification. Although colour can be objectively defined by primary colour composition and degree of saturation, colour differences cannot be put on a simple scale like size or horsepower or vitamin C content so that everyone agreed that good A has twice as much per pound as good B. Lancaster 1979, p.18

goods which have the highest levels of A, B and C, the attributes which most consumers value most. A secondary hypothesis is that if the marginal producer increases the level of these attributes for his or her product, whether by advertising or by a change in specifications, he or she will get a higher price or sell more.' This formulation does not take objective characteristics into account at all.

THE BOUNDARY ASSUMPTIONS

The boundary assumptions set out the domain within which a theory is expected to work. There are several levels of boundary, from the broad divisions between Lancaster's approach, Rosen's approach, etc., to the boundary assumptions of each variant within an approach. Any criticism of a set of boundary assumptions applies only to theories within those assumptions, to a single variant of a theory or a group of theories. It is important to note that theory does not develop conveniently from trunk, to branch, to twig with each new variant adding a progressively smaller and less important assumption. Variant 2001 may be variant 2000 with a new or changed fundamental assumption, meaning that all the conclusions and predictions are different. Variant 2050 may be in the 'characteristics approach' while variant 2051 takes this on board, and adds in some assumptions or logic from Rosen. Even a single new assumption may completely change predictions and conclusions. It is not always immediately apparent which group of theories would be affected by an attack on boundary assumptions.

The only attacks of boundary assumptions that I am aware of in the literature are attacks on Lancaster's approach by Hendler (1975), Ladd and Zober (1977) and Lucas (1975) who show that they are extremely restrictive. Lancaster's theory only works for instance when the satisfaction gained from a characteristic is independent of the good in which it is consumed, so one gram of chilli powder gives the same utility whether it is in a stew or an ice cream. While I find many of their criticisms persuasive, they have had little impact: the approach they attacked remains the dominant approach, and only 1.5% of people citing Lancaster in recent years cite these critics. This is partly because they do not attack the fundamental assumptions and logic identified above and stay very much within the paradigm they are attacking, tinkering with boundary assumptions and logic, rather than saying 'This theory has such limited application that it is trivial: let us abandon it and work with an entirely different theory.'

Ad hoc Assumptions

The research programme is seriously, even fatally, damaged by the number of implicit and explicit *ad hoc* assumptions running through it. These are not realistic assumptions made to turn a theory into a situation specific model. They are arbitrary assumptions made because the analysis cannot proceed any further with just the fundamental and boundary assumptions (Popper 1972 pp15-16; 1976 pp40, 42). Often they are introduced purely to make mathematical analysis possible or to exclude inconvenient complications.

The dominant approach in the literature follows Lancaster and so incorporates his *ad hoc* assumptions. In *Consumer Demand* (1971) there are 63 explicit assumptions, at least 40 of which are *ad hoc*, for example:

'The Cobb Douglas functional form is assumed.' (Lancaster 1971 p73).

'Uniform distribution is assumed so that average income is constant ...' (Lancaster, 1971, p79).

'Goods are completely separable, sharing no characteristics' (Lancaster, 1971 p126).

All other goods may be treated as identical, all being equally close or distant substitutes for this group. (Lancaster 1971, pp128-9).

A characteristic may be treated as irrelevant if it does not appear in the preferences of a large proportion of the consumer population (Lancaster 1971, p146), which implies that we can ignore the fact that 20% of the population hates garlic.

'The most heroic assumption is the uniformity assumption on the nature and distribution of preferences ... In geometric terms it implies that the transformed indifference curves in specification-quantity space are all of identical shape and are tangent to the [Product

Differentiation Curve] at the specification corresponding to the most preferred good' (Lancaster 1979, p47).

When one is dealing with a group of closely related goods, all other goods may be treated as equally close substitutes for this group (Lancaster 1971 pp128-9). [He uses 'good' in the sense of a single product line].

There is a uniform distribution of income so that average income is constant over preferences and there is a rectangular distribution of preferences, with constant density taken to be unity (Lancaster 1971, p79).

The consumption technology is linear, after ignoring invariant characteristics, and a characteristic is irrelevant if there is a linear dependence in the technology (Lancaster 1971 p142). 'In many cases it will be appropriate to assume that characteristics technically related in this way are also related in the view of the consumer so that he reacts to any one the related characteristics not to each of them separately.' (Lancaster 1971, p144).

There are in addition many *ceteris paribus* assumptions. These are quite unexceptional if they are dropped at a later stage of the analysis. Since they are not, they are *ad hoc* assumptions in disguise.

The long list of *ad hoc* assumptions effectively rules out all real life situations, as the theory demands that they apply to **all** individuals and **all** goods. It is probable that it introduces conflicting assumptions. It is clearly impossible to determine whether or not the assumptions hold in any real situation so the theory is not testable.

While Lancaster was scrupulous in attempting to make all his assumptions explicit, his followers and competitors are not. Most make *ad hoc* assumptions and most of these are implicit. While Lancaster's work is among the most cited in economics, I have never seen anyone working in this tradition saying which of Lancaster's *ad hoc* assumptions are taken on board and which are rejected, or indeed recognizing that any *ad hoc* assumptions were made. The probability, therefore, is that every new bit of theory will introduce *ad hoc* assumptions which conflict with those made in the previous theory which it relies on, or which, taken with previous *ad hoc* assumptions effectively rule out all reality. By the time a dozen workers have built on previous work it is impossible to know which *ad hoc* assumptions have been made, so the most modern theory is particularly suspect. Any one of the new assumptions could change the conclusions and predictions of the theory.

It is common for one worker to take the conclusions of a previous worker, and then to change the assumptions they are based on so that the analysis can proceed, producing a logical fallacy. Sometimes a worker's later work uses the conclusions of earlier work, but continues with a different set of assumptions.⁷ This is common when a writer 'relaxes initial assumptions'. The error is unavoidable when there is a large number of forgotten or implicit *ad hoc* assumptions.

Formally, contradictory or all exclusive assumptions are sufficient to refute the theory. If one can identify such *ad hoc* assumptions in the seminal works, the effect is to invalidate the later parts of those works (the assumptions are often introduced after the initial paradigm case) and any other theories that rely on the later part of these works. In practice it is a long and painstaking task to identify the implicit and explicit *ad hoc* assumptions used in each variant. Without this, the clearest possible refutation of one variant has no obvious impact on other variants derived from it. In practice, perhaps, it will be impossible to convince an audience which habitually ignores the fact that it is introducing new assumptions into a theory that contradictions in these assumptions are important.

THE THEORIES HAVE NOT BEEN TESTED

⁷ Lancaster 1979, for instance, starts by accepting the conclusions of Lancaster 1971, then changes the assumptions for the subsequent analysis.

Can theories be tested?

Economic theories do not pretend to make predictions about the real world. Predictions are made by models which typically combine several theories and realistic assumptions with data (which is a form of realistic assumption). It is not possible to test theories directly by the accuracy of their predictions.

The possibility remains that it may be possible to test theory indirectly: if models using one theory are 'generally' good predictors, and if models using another are 'less good', the first may be preferred, though it is not possible to compare the performance of theories which have a different domain. The following problems with indirect testing are widely accepted.

Under most epistemologies it is recognized that no theory can be expected to apply everywhere, so a theory can only be expected to apply within the domain of its boundary assumptions. This means that a theory can only be tested if it is possible to say unequivocally that its assumptions apply in this case. If not, poor predictions may be taken as showing not that the theory is wrong, but that the assumptions cannot have applied in this case. With REM theory it is not possible to show that the fundamental, boundary or *ad hoc* assumptions of REM quality theory hold, though it may be possible to show that they do not. It is formally impossible, for instance to plot an individual's multi-dimensional indifference surface, let alone to confirm that the many *ad hoc* assumptions apply. This means that the theory cannot be tested.

Economists are agreed that it is very difficult to test even a single model by its predictions. There are a great many recognized and legitimate excuses for poor predictions such as inaccurate data, sampling error, changes in exchange rates and the fact that the prediction was presented as a probability not as an exact figure. It is seldom that a model would be abandoned merely because of a few bad predictions.⁸ "The ingenuity of these nineteenth century writers knew no bounds when it came to giving reasons for ignoring apparent refutations of an economic prediction, but no grounds, empirical or otherwise, were ever stated in terms of which one might reject a particular theory" Blaug (1980, p55).⁹ Depending on the complexity of the model and its purpose, one might instead have repeated tests of a progressively refined and adjusted model.

Another problem is that any model has elements of several theories incorporated in it, as well as realistic assumptions and data. The failure of a prediction cannot be attributed to any one of them. Furthermore half a dozen totally 'correct' theories can be put together to give a model which predicts badly. No predictive failure, therefore, invalidates any one theory.

The theories cited in papers are often not the ones actually used in the analysis. Many papers claiming to be based on the REM theories of Lancaster or Rosen are in fact based on the hedonic theory of Waugh (1928). Lancaster (1971 p113-4) complains of this, but it can be seen in many recent papers for instance Larue (1991), Williams (1991), McDaniels, Kamlet and Fischer (1992), Ortono and Scacciati (1992), Thomas (1993), Berliant and Raa (1991), Thomas (1993), Johnson and Fornell (1987), Heffernan (1990). No number of tests based on such theories can add to or subtract from our confidence in REM theory.

Each of the main variants of REM theory has had hundreds or even thousands of variants produced, with new variants being published each year. Variants sharing most of the same assumptions and logic may give very different predictions, because of differences in other key assumptions and logic, because of different *ad hoc* assumptions, and because of the way they fit into situation specific models for instance. It is conceivable that some variants will produce good predictors, others not. It is probable that some will be good predictors in some situations over some ranges of variables when used with some other variables, with errors in different theories cancelling out, or by lucky chance. A few successful predictions in a narrow range of situations gives no corroboration to a theory.

⁸ These problems are generally recognized by practical economists, as well as by methodologists including the Victorians,

⁹ Hutchinson (1938), Machlup (1963) and "sophisticated falsificationists" (including Popper). "The ingenuity of these nineteenth century writers knew no bounds when it came to giving reasons for ignoring apparent refutations of an economic prediction, but no grounds, empirical or otherwise, were ever stated in terms of which one might reject a particular theory" Blaug (1980, p55).

It is not possible to test groups of theories by their predictions. The tests must be of models derived from a single variant. Each variant is presented as an attempt to produce a theory which avoids the problems of previous theories and is therefore a better predictor. The fact that 9000 variants are bad predictors is no indication that other variants have not solved their problems. This is in contrast to refuting groups of theories because they share the same conflicting assumptions or incorrect logic.

In order to say that models based on one variant of one theory generally predict well it would be necessary that a range of different types of model had been constructed and applied to a range of situations, and that there had been repeated tests of each model, refining it and improving it, in order to overcome the recognized problems described above. This would require a very carefully designed experimental protocol with some hundreds of replications, for each variant.

In fact, there has been no attempt at such a formal test of any variant of REM quality theory. There have been a few papers trying models based trying different theoretical approaches on a single set of data, but no conclusion could be drawn from such an exercise. Very little of the practical work based on REM theory has been designed as a test: rather the theory has been used in a model to determine for instance which characteristics of a product should appeal to customers. Since no one is likely to launch a product which scores badly on such tests, there is no way if knowing that the predictions that these products would fail were wrong. Real world economics does not provide evidence from repeated tests: most economic decisions are made on economic models used between one and ten times, not enough to test them - in practice decision makers act on them because the assumptions are realistic and the logic appears to be sound. For this and other reasons a trawl through the literature will give no indication whether any variant of a theory or group of theories is generally a good predictor.

It is sometimes argued that a) Lancaster and Rosen provided the theoretical justification for hedonic theory; b) hedonic theory is generally a good predictor; c) therefore Lancaster and Rosen's (rather different) REM theory must be correct c) therefore any theory in this tradition is a good predictor. This is of the same form as a) flat earth theory suggests that plane geometry is appropriate for measuring the earth's surface; b) plane geometry gives very accurate results in measuring football pitches or mapping towns; c) therefore the earth is flat; d) therefore any theory assuming a flat earth will be a good predictor. In fact, hedonic theory was introduced by Waugh (1928) and it would still be used whether or not REM theorists had attempted to justify it. Lancaster, writing nearly 40 years later, makes no attempt to justify it¹⁰. The prices used by Rosen and others are not in fact the prices used in hedonic theory, so they are arguing about different things.¹¹ Many of the variants of hedonic theory in common use do not overlap with REM theory: the example was given above of an analysis based purely on subjective attributes. REM theory gives the same support for hedonic analysis that the central pier gives to Sydney harbour bridge: it is an imaginary construct supplying support to something that does not need it.

WEAKNESSES IDENTIFIED

The weaknesses that have been identified above in the REM approaches to quality may be summarized as follows:

¹⁰ The only part of his seminal works in which he uses something vaguely resembling it is Chapter 10 of *Consumer Demand*, in which the structure of logic and assumptions so painstakingly built up in previous chapters is abandoned.

¹¹ Hedonic prices are defined as the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amount of characteristics associated with them. Econometrically, implicit prices are estimated by the first-step regression analysis (product price regressed on characteristics) in the construction of hedonic price indexes'. (Rosen, 1974, p.34). The REM analysis is confusing: Rosen (1974) appears to assume that the set of prices facing buyers is at the same time:

- a market clearing price
- an average equilibrium price at the end of a day's trading
- the price facing each buyer and each seller at all periods through the day

There is also a confusion between the 'price lists' required for REM theory and the hedonic prices obtained from regressions.

Fundamental Assumptions

The fundamental assumptions on consumer preferences conflict with observed reality. They are not simplifications of reality. It is most unlikely that they will hold for any product group bought by any individual, but the theory requires that they hold for all individuals and all product groups.

The fundamental assumptions on price conflict with reality in most cases, and as they are usually tacit assumptions, there is confusion as to what they are and what prices should be used. In general, the price assumptions are most likely to hold in price-making markets where other REM assumptions do not apply.

Fundamental logic

There are errors in the logic of characteristics space, a logic which is fundamental to the whole theory.

Most variants and all the main branches make fundamental errors in the logic of subjective and objective quality factors. This is in addition to unrealistic fundamental assumptions on the subject.

Boundary assumptions

The boundary assumptions limit each group of the theory to a very small domain.

***Ad hoc* assumptions**

Formally, the *ad hoc* assumptions invalidate most of what has been written, by introducing conflicting assumptions or ruling out most of reality. It is doubtful whether any variant has *ad hoc* assumptions which apply for any one individual for one product, let alone for all individuals for all products, as the theory demands.

Do the assumptions apply?

It may be possible to say that some assumptions do not apply in a particular instance, but it is never possible to say that all assumptions do apply. Any predictive failure, therefore, can be taken as showing that the assumptions cannot have applied and the theory was operating outside its domain, rather than that the theory is wrong.

Has the theory been tested?

None of the many variants or groups of theories has been subjected to a programme of crucial tests by prediction. It is either impossible or impracticable to test theories by their predictions.

THE IMPACT OF THESE CONCLUSIONS

Those theoreticians who created REM theories of quality think that its value derives from correct logic based on carefully stated assumptions. They would accept that the theories must be rejected for logical errors, contradictory assumptions or assumptions that rule out all of reality,¹² but they show little or no interest in testing by predictions. Practical economists know that it is easy to test assumptions that are meant to be realistic for realism and relatively easy to check the logic of a model, but it is very difficult in practice to check predictions. For them realistic assumptions are essential: they would lose their jobs immediately if they presented a model, however good a predictor it was, based on the assumption that cheese grows on apple trees. They prefer to work from what we know to what we do not yet know, rather than from what we already know to be unrealistic assumptions. Some economists believe that it is acceptable or better to work with unrealistic assumptions (rather than simplified assumptions) so that we must accept all (published?) variants of theories until they have been rejected because of their poor predictions.

The epistemological stances underlying these positions is seldom made explicit and there is a great deal of confusion. Some ways of interpreting these stances are as follows:

Theory as logic

Most practitioners take theory to be a string of logic. Strings of logic taken from different economic theories are 'cut and pasted' and applied to assumptions and data about a specific situation to create a situation-

¹² Lancaster (1979) for instance accepts that his 1975 paper must be rejected because it rules out all possible markets.

specific model. These economists require that a theory is logically correct and is based on assumptions reasonably close to those used for this specific model, preferably being simplifications. There is no suggestion that any theory represents 'the truth' or that it applies everywhere, but the model is intended to represent the truth.

This epistemological stance rejects REM theory on the grounds of assumptions contrary to reality and logical errors. It rejects the theories where their highly restrictive fundamental, boundary and *ad hoc* assumptions do not apply. Since it is never possible to state that all the assumptions do apply in a particular instance, the Popperians would reject the theory as 'unscientific'.

This stance makes it possible to reject **all** REM theories, in so far as they share the weaknesses identified. It is possible to reject **all** the theories in some branches or groups of theory. It is not necessary to refute them one by one.

Models developed from situation-specific assumptions plus bits of theory can, in principle, be tested by their assumptions, their logic, or their predictions. Groups of models can be rejected if they share the same wrong logic or assumptions. It is possible in principle, though difficult in practice, to test a single model by its predictions, but it is difficult to argue that one model or a dozen similar models have in fact failed, and, if so, that all 'similar' models fail. Still less could it be argued from this that a variant of a theory or a group of variants fail. The fact that dozens of aeroplanes fall out of the air each year does not imply that the theory of aerodynamics is wrong or that all aeroplanes of the same model will also crash.

Theory as truth

Another stance is that theories in economics mirror the truth in much the same way that theories in physical science do. General laws can be derived by observation. Since the theories correspond to the truth, it follows that the logic must be correct and the assumptions realistic, which would rule out REM theory.

Again models are based on several theories, and can be tested on their assumptions, their logic or their predictions.

Again, it is possible to reject a group of theories or a group of models on the basis of assumptions that conflict with reality or logical errors, but it is somewhere between impractical and impossible to reject groups of theories on the basis of predictions (except, perhaps where the predictions are totally ridiculous).

Theory as probability

The theory as probability stance has it that economic theory is truth but only when its assumptions apply - no theory is expected to apply everywhere. It is normal, usual, common, probable or possible that the assumptions apply in a 'randomly selected' situation.

This epistemology can be used in the same way as Theory as Logic. Realistic assumptions are identified, appropriate theory is selected, and models are constructed which may be tested by their assumptions, logic or predictions.

Alternatively, the theory may be applied to all situations, without first checking to see that the assumptions are realistic. It is believed that normally, usually, commonly, probably or possibly the assumptions will apply and the theory will be the truth for that situation. The option of checking its assumptions (or its logic, in some formulations) has been abandoned - in spite of the general perception of economists that it is easy to check logic and assumptions that are meant to be realistic, but difficult to test predictions.

In a variant of this epistemology the theory is expected to work on some occasions and not on others in a given situation (implying that the assumptions apply on some occasions and not on others in a given situation). In this case there must be repeated tests of predictions over time to give any indication of the probability that the theory will apply.

For this alternative approach to have any practical value, it must be possible to assess the probability that the assumptions of a variant of the theory will apply in a certain class of situation. This could be done by checking the assumptions in a class of situation and determining in what percentage of situations they

conformed with the assumptions of the theory. If this was feasible, however, the Theory as Truth or Theory as Logic approach would be used instead.

The alternative is to build a standard model for a type of situation and test a sample of situations in a class of situations by their predictions. This requires replicated tests with stochastic sampling from a carefully defined parent population of situations (the buying behaviour of middle class customers in high street chemists?). There would have to be repeated tests in each sample situation because of the generally accepted difficulty of testing predictions. This would have to be repeated for every class of situation in which the variant of the theory might be used (working class customers buying vegetables in street markets?). At the end it would not be clear whether the predictive failures were because the model was logically wrong, because the assumptions did not apply in these instances, or because the model or testing procedure was inadequate. It is doubtful whether it would ever be possible to convince economists of this persuasion that they should reject even a single variant of a theory in this way. More important perhaps, there is no possibility that the attempt will be made, because of the enormous cost. If this is so of a single variant, then it is far truer of the infinite number of theories and variants that could be generated.

Even this number of replications is inadequate if all the theories used in a model are 'theory as probability', each of which may or may not apply in that situation at that time.

In the first use of this epistemology, where it is used like Theory as Logic, it can reject large groups of theories or models on assumptions and logic. In the second use, where assumptions are not checked against reality, it is impossible to reject groups of theories or models, and virtually impossible to reject even a single model.

'As if' theory

One epistemological standpoint is that we can use a theory based on unrealistic assumptions, and even incorrect logic 'as if' it was Theory as Truth. This follows from the belief 'We know that individuals do not behave like Rational Economic Man, but it is possible to derive credible demand curves from the assumption that they do. If we assume Rational Economic Man for all economic theory, we can expect similar success.'¹³ Presented like this, the logical foundation is similar to that for astrology. It begs the question 'why do we not just assume demand curves instead?' It is also difficult to understand the belief that because it is possible to get an elementary, two-dimensional, demand curve from REM assumptions, it will be equally possible to argue from a complicated set of assumptions about quality preferences and supply, involving multi-dimensional indifference surfaces, through a very complex aggregation procedure to get multi-dimensional market effects and market demand for quality. While the original REM assumptions applied to basic economic theory might be said to be grossly simplified generalizations of reality, the REM assumptions of quality theory contradict reality.

Since the realism of the assumptions and the correctness of the logic is taken to be irrelevant under this approach, theories can only be tested by their predictions. It is only possible to test them one variant at a time, and indirectly at that. If the first 1000 are bad predictors, the next 2000 may be excellent. Inevitably, if there is an infinite number of possible theories, some models based on them will, by chance, prove to be excellent predictors in some series of tests, but this does not give us any reason to believe that they will continue to be good predictors in the future, or that they will be good predictors in different situations.

For each variant of theory there are many possible models relating to different situations. Each test of predictions would have to be replicated dozens of times, because of generally accepted problems of testing. A test could easily take four months of a researcher's time, plus several months of other input by a market research company. The test would only be for one model, not even for a single variant of a theory.

¹³ These epistemologies may be linked to methodological approaches like those of Mises (1949), Menger, Robbins (1935), Knight (1940, 1941), Machlup (1963) or Friedman (1953) which argue that the testing or verification of assumptions is unnecessary or undesirable. One school of thought is that theories based on unrealistic assumptions, not simplified assumptions, is actually to be preferred (e.g. Sternthal, 1995). This appears to be based on an unusual interpretation of Friedman (1953).

The view that we must accept any one of an infinite number of possible theories presented until it has been refuted by repeated failure of its tests is absurd: in effect all possible theories would have equal weight.¹⁴ The alternative is not to have any confidence in any variant of a theory until it has been repeatedly tested, still bearing in mind that even close variants in a single theory may produce very different predictions. This has two, apparently insuperable, problems. First, if every one of an infinite number of theories is equally without credibility until its predictions have been tested, how does one decide which it is worth testing? Second, there is the conclusion reached above that testing predictions is somewhere between impractical and impossible.

In contrast, the Theory as Logic and Theory as Truth approaches do not require this corroboration: realistic assumptions plus sound logic give adequate grounds for action: anything logically derived from the truth is true.

Again it is emphasized that variants of the theory may have different sets of fundamental assumptions (which include the fundamental assumptions common to all REM theory approaches to quality) boundary assumptions, *ad hoc* assumptions and logic, so the rejection of one variant as a poor predictor does not imply the rejection of any others: they must be assessed separately.

There has been no attempt at the necessary testing for REM theory.

‘Probably as if’

The ‘as if’ stance can be modified to ‘probably as if’ implying that consumers in aggregate usually, normally, commonly, probably or possibly act as though the REM assumptions on individuals were correct, though it is accepted that they are not correct. Again, the possibility of selecting or rejecting theories on their assumptions has been abandoned.

The theory is not expected to give good predictions in all situations even allowing for the difficulties of indirect testing through models, and even in situations where it sometimes gives good predictions, it sometimes may give bad predictions (it may work only over a certain range of prices, or it may be a matter of pure chance whether it gives good predictions). If anything more is to be said than ‘Models using this theory are accurate predictors when they give accurate predictions, but not otherwise’, a programme of testing is needed. This implies stochastic sampling of situations from a carefully defined parent population. After repeated testing one might come up with the conclusion that variant X appears to give good predictions in Y% of applications to consumer choice in chemists’ shops but to only Z% in ring-road superstores. It is necessary to test each model repeatedly in each situation as a ‘good predictor’ will sometimes give bad predictions. This means many times more tests than for the ‘as if’ theory. The number of tests required to produce this information for even a single variant of the theory requires resources that will never be made available. Certainly there has been no attempt to do it for any branch or group of REM theory.

Accordingly, one is left with an infinite number of possible theories, without any constraint on realism of assumptions or logical correctness, and no way to select which to use. There is initially no corroboration for any one of them, or any reason to select one rather than another for testing. Resources do not permit serious testing of predictions, if indeed it is possible, so no theory ever gets any real corroboration.

It is possible under the ‘as if’ and ‘probably as if’ approaches to say that vast areas of economic theory cannot be used because it has not yet been tested and there is no corroboration for it, but it is doubtful whether any theory escapes this stricture, so no theory is acceptable. The opposite approach, of accepting all possible theories until they have been tested and rejected is equally impractical.

Even if sufficient resources could be mobilized to test one variant of a theory and to accept or reject it, this would give no indication of whether other variants worked or not.

¹⁴ But it exists: see Traill (1995), Hallam (1996).

It might be possible to make a quick check of a sample of situations to see in what proportions of the situations the assumptions of the variant of the theory hold. This is contrary to the spirit of the approach. It is also impossible with some theories, such as REM theory, where it is not possible to say in any instance whether the assumptions do or do not hold.

CONCLUSION

It has been shown that under the most widely used epistemological stances of economics, the REM theories of quality must be rejected as having assumptions which are contrary to observed reality, which rule out all of reality, or which conflict with each other; or as having logical errors or as having no corroboration through the testing of predictions of models based on variants of the theory, or on some combination of these.

It has also been shown that it is possible to reject large groups of theories under some very different epistemologies, though with some it is on the grounds of errors in assumptions and logic, and in others on failure to provide 'corroboration' or 'verification'. The impact can be very different, but all the epistemologies discussed reject REM theories.

This is fatal to all versions of REM theory. It is not possible to make a few changes to existing REM theory and to present an alternative version. The errors identified are at fundamental levels and any change would require reworking the whole theory from first principles, from the first assumptions and logic. There are in any case many alternative theories in general use already. These include the hedonic approach, compensatory models, perceived quality, behavioural, behaviourist and heuristics approaches, which do not depend on REM assumptions. The author's preference is for the composite and complex approaches based on a range of theory and a rigorous observation which is to be found particularly in market economics (e.g. Bowbrick, 1992), the new mainstream economics (e.g. Earl, 1986), and agricultural economics.

REFERENCES

Becker, G.S. "A theory of the allocation of time". Economic Journal 75:493-517, 1965.

Berliant, M., and T. T. Raa, "On the continuum approach of spatial and some local public goods or product differentiation models: some problems" Journal of Economic Theory 55(1) pp95-120, 1991.

Blaug, M., The methodology of economics: or how economists explain, CUP, Cambridge, 1980.

Bowbrick, P., The Economics of Quality, Grades and Brands, Routledge, London 1992.

Brems, H., "Input-output co-efficients as measures of product quality", American Economic Review, 47, 105-18, 1957.

Brems, H., "The interdependence of quality variations, selling effort and price", Quarterly Journal of Economics, 62, 418-40, 1948.

Earl, P., Lifestyle Economics: Consumer Behaviour in a Turbulent World, Brighton, Wheatsheaf, 1986.

Friedman, M., Essays in Positive Economics, Chicago, University of Chicago Press, 1953.

Gorman, W.M., "The demand for related goods", Journal paper J3129, Iowa Agricultural Experiment Station, Nov. 1956 (reprinted as Gorman 1976, 1978)

Heffernan, S.A., "A characteristics definition of financial markets" Journal of Banking and Finance 14 583-609, 1990.

Hendler, R., "Lancaster's new approach to consumer demand and its limitations", American Economic Review. 65 194-200. 1975.

Houthakker, H.S., "Compensated changes in qualities and quantities consumed", Review of Economic Studies, 3 [19] 155-154. 1952.

Hutchison, T.W., The significance and basic postulates of economic theory, 1938

Ironmonger, D.S., New commodities and consumer behaviour, Cambridge, Mass. University Press, 1972.

Johnson, M.D., and C. Fornell, "The nature and methodological implications of the cognitive representation of products", Journal of Consumer Research 14, 214-227, 1987.

Knight, F., "What is "truth" in economics?" Journal of Political Economy. 1940

Knight, F., "The significance and basic postulates of economic theory. A rejoinder" Journal of Political Economy, 49, 750-3.

Ladd, G.W. and M. Zoher, "Model of consumer reaction to product characteristics", Journal of Consumer Research. 4 89-101. 1977.

Lancaster, K.J., "A new approach to consumer theory", Journal of Political Economy. 74 132-157. 1966.

Lancaster, K.J., Variety, equity and efficiency. Columbia studies in Economics no 10. New York and Guildford. Columbia University Press. 1979.

Lancaster, K.J., "Socially optimal product differentiation", American Economic Review. [September] 99-122. 1975.

Lancaster, K.J., Consumer demand: a new approach. Columbia University Press, New York. 1971.

Larue, B., "Is wheat a homogeneous product?" Canadian Journal of Agricultural Economics 39 103-7, 1991.

Lucas, R.E.B., "Hedonic price functions", Economic Inquiry, 13 157-78, 1975.

McDaniels, T.L., M.S.Kamlet and G.W.Fischer, "Risk perception and the value of safety", Risk Analysis 12 (4) 495-503, 1992.

Machlup, F., Essays on economic semantics, Englewood Cliffs, Prentice Hall, 1963.

Mises, L. von, Human action. A treatise on economics, London, William Hodge. 1949.

Muth, "Household production and consumer demand functions", Econometrica, 34:699-708, 1966.

Ortono, G., and F. Scacciati, "New experiments on the endowment effect", Journal of Economic Psychology 13 277-296, 1992.

Popper, K.R., Unended Quest: an intellectual autobiography, Fontana/Collins, 1976.

Popper, K.R., The Logic of Scientific Discovery, Hutchinson, London, 1959, Eighth impression 1975.

Popper, K.R., Objective Knowledge, an evolutionary approach Oxford, Clarendon Press, 1972.

Ratchford, B.T., "Operationalizing economic models of demand for product characteristics", Journal of Consumer Research. 6 [1] 76-84. 1979.

Robbins, L., An essay on the nature and significance of economic science, London, Macmillan, 1935.

Rosen, S., "Hedonic prices and implicit markets: product differentiation in pure competition", Journal of Political Economy. 82 34-5. 1974.

Steenkamp, J-B.E.M., Product Quality, Assen/Maastricht, Van Gorcum 1989.

Sternthal, B., Personal Communication, 1995.

Stigler, G.J., "An analysis of the diet problem" *Journal of Farm Economics*, May 1945.

Thiel, H., "Qualities, prices and budget enquiries", Review of Economic Studies, 19, 129-47, 1952.

Thomas, J.M., "The implicit market for quality: an hedonic analysis", Southern Economic Journal 59 (4) 648-674, 1993.

Trail, B., Personal communication, 1995.

Waugh, F.V., "Quality change influencing vegetable prices", Journal of Farm Economics, 10, 185-96, 1928.

Williams, A.W., " A guide to valuing transport externalities by hedonic means", Transport Reviews 11(4) 311-324, 1991.