6

Uninformative Mathematical Economic Models

> Anyone with the aesthetic sense to recognize the beauty of the proof that the diagonal of a unit square is not the ratio of two integers … will sense the same harmony that resides in Ricardian comparative cost or Walrasian general equilibrium. Economic theory is a mistress of even too tempting grace.
>
> Paul A. Samuelson [1962/66, p. 1680]

> you never get something for nothing and never empirical hypotheses from empty deductive definitions. At best your observation can tell you only that the real world (or some subset of it) is not exploding; your theoretical model or system will always be an idealized representation of the real world with many variables ignored; it may be precisely the ignored variables that keep the real world stable, and it takes a significant act of inductive inference to rule this out and permit the Correspondence Principle to deduce properties of the idealized model.
>
> Paul A. Samuelson [1955, p. 312]

In Chapters 2 and 3, I focused the burden of applying the Popper-Samuelson requirement of testability on determining how many observations it would take to refute a solution statement for any specific model. I did not worry about how one would obtain the needed solution statement (partly because I used primarily linear models that easily provide solution statements). In the mid-1930s, Abraham Wald [1936/51] examined some standard economic models as expressed in some systems of equations. The question he considered was whether the 'solvability' of a system of equations used to represent a set of ideas put conditions upon the ideas themselves. Of course, this was not merely an arithmetic problem, as his concern was with the consistency and completeness of the logic brought out by the use of mathematical theorems [see Boland 1970b]. The mathematics he used tended to be rather complicated and severe for the 1930s. Later developments, primarily in the 1950s, made it possible to specify the conditions for solvability in a much 'cheaper' manner [Arrow and Debreu 1954, McKenzie 1954, Gale 1955, Kuhn 1956, Nikaido 1956; see also Weintraub 1985].

The profession has moved on, confident that this matter has been settled or at least is well under control. Nevertheless, I think an important question may have been overlooked: Does solvability necessarily imply explanation? Or more generally, of what significance is the solvability of a model which represents a given explanatory theory for the given theory itself? In light of some methodological considerations concerning the reversibility of assumptions and conclusions [e.g. De Alessi 1971], it turns out that 'solvability' in a methodological context of explanation leads to undesirable (unintended) consequences. In short, solvability if properly specified leads to uninformative economic models – models which cannot say anything of empirical significance about the 'real world'.

1. A simple Walrasian general equilibrium model

Léon Walras supposedly began his famous general equilibrium analysis by putting forth a general equilibrium theory that can be represented [Dorfman, Samuelson and Solow 1958, Arrow and Hahn 1971] by the following system of equations:

$$[\Sigma] \quad \equiv \quad \left[\begin{array}{l} \mathbf{R} = \mathbf{A} \bullet \mathbf{X} \\ \mathbf{R} = R_0 \\ \mathbf{X} = \mathbf{D}(\mathbf{P},\mathbf{V}) \\ \mathbf{P} = \mathbf{V} \bullet \mathbf{A} \end{array} \right]$$

where **X** is the vector indicating the quantities of **m** outputs, **P** is the vector of their prices, **R** is a vector indicating the quantities of **n** resource inputs, and **V** is the vector of the values of those inputs. Also, **A** is an **n**×**m** matrix of input-output coefficients and **D**( ) is a vector formed of the appropriate **m** demand functions for the outputs.

The question that Wald could have considered is:

> Does being able to solve the system of equations [Σ] for **P**, **V** and **X** imply that the system itself is an (informative) explanation of **P**, **V** and **X**?

This would, in the minds of many, raise the issue of whether the solution is a description or an explanation. Stanley Wong, in his examination of Samuelson's methodology, does just this by advocating the need for 'informative explanations' [Wong 1973, 1978]. To assist clarity, I will distinguish these concepts in the following unambiguous way:

(a) an *informative explanation* is the 'explanation of the known by the unknown' [Popper 1972, p. 191],

(b) a *description* is the explanation of the known by the known,

where 'explanation' is the name given to the logical relation of the explican to the explicandum [Popper 1972, p. 192] and where 'known' may be interpreted in the usual empirical sense (although it need not be restricted to that). One sense in which an explican will be an 'unknown' is when (necessarily) one of them is a strictly universal statement (e.g. 'all men are mortal'). Note that I am not equating 'explanation' with a demonstrated necessarily true explanation.

It may be that solvability will guarantee only a description but not an informative explanation. The usual position regarding the explanatoriness of a solution is that it is only a matter of logic [e.g. Debreu 1959, pp. vii-viii]. This position may be based on a popular view which simply equates explanation and description, a view attributed by Fritz Machlup [1966] to Samuelson [1963, p. 234; 1964, p. 737; see Clower and Due 1972, p. 15]. For some this is merely because 'prediction' is directly identified with 'explanation' [e.g. Marshall 1920, p. 638; Liebhafsky 1963, pp. 16-18; Ferguson 1972, p. 8]. For others this response is based on a view that economics is a 'deductive science'. Predictions or explanations are 'conditional', i.e. 'if … then' statements [Friedman 1953, De Alessi 1965, Bear and Orr 1967, Lipsey and Steiner 1972]. This would mean that my idea of explanation only requires that the things to be explained are at least logically compatible with the reasons given. For example, Samuelson and most others would allow multiple equilibria in the explanation of some equilibrium phenomena [Samuelson 1947/65, pp. 49, 75ff, 240; Arrow and Hahn 1971, p. 15].

Wald was concerned with the question of the mathematical circumstances under which we can solve this system of equations for unique values of **P**, **V** and **X**. However, Walras' Law implies that there are only $2\mathbf{m}+2\mathbf{n}-1$ independent equations, so that either the uniqueness is only up to the point of relative prices, or we can add an independent equation, such as:

$$\sum_{i=1}^{\mathbf{m}} \mathbf{P}_i = 1$$

or merely:

$$\mathbf{P}_1 = 1$$

which turns the relative prices into 'unique' values.

All this equation counting is of no concern here. Wald's paper began by refuting the ever-popular notion that assuring the equality between the number of independent equations and the number of 'unknowns' (i.e. endogenous variables) is necessary and sufficient to assure the existence and uniqueness of the solution of those equations. Wald went on to deal with specific representations of Walrasian theory and proved that Wald's proposed set of conditions on the **D**s and **A**s would provide the needed assurance of solvability [Wald 1936/51, Boland 1970b].

2. Methodological requirements of explanatory models

The view that explanatoriness is only a matter of logical entailments unfortunately leads to the following methodological problem over the reversal of assumptions and conclusions: What if the intended one-way relationship from a particular combination of states of **R**, **D** and **A** to a particular combination of values of **P**, **V** and **X** is logically sufficient for the relationship to also go the other way, as well? That is, what if:

RDA → **PVX**, *and* **PVX** → **RDA**?

Can we consider the occurrence of a particular **PVX** to be an explanation or, within the context of the assumptions of our theory, to be 'causes' of the occurrence of the particular 'given' state of **RDA**? Surely not.

The basic idea of explanation operative in economics today has for a long time been the one based on the distinction between endogenous and exogenous variables [see Koopmans 1950b, p. 393; 1953, pp. 27, 40-4; Marschak 1953, p. 10; Hood and Koopmans 1953, p. 115]. That is, **R**, **D** and **A** influence **P**, **V** and **X** (within the model) but **R**, **D** and **A** are determined outside the model – i.e. **P**, **V** and **X** do not influence **R**, **D** or **A** (within the model). Although the coincidence of a particular **PVX** point with a particular **RDA** point may be interesting, it is not an explanation. As almost everyone would agree, such a circularity means that we have at best a 'pure description'. This is sometimes explicitly stated [e.g. Alchian and Allen 1967] and it is implicit in textbook discussions of the 'post hoc, ergo propter hoc' fallacy. This circularity leads to the conclusion that such reversible (complete) explanations are merely what economists usually call 'tautologies' [cf. Agassi 1971a]. When the set of conceivable **RDA**s has the same dimension as the set of conceivable **PVX**s, what I have called a pure description merely becomes a renaming routine. For example, a bundle of goods may be described by its list of goods themselves (and their quantities) or by its inherent 'characteristics' (and their quantities) following Lancaster [1966]. The matrix relating the two means of description is merely a renaming process. This placid view of complete explanation is, I think, the primary result of viewing mathematical functions as explanations.

The concept of reversibility that I have been discussing is a methodological failure if we wish to maintain the irreversibility character ascribed to the one-way influence between endogenous and exogenous variables, namely the explanatory power that results from making that distinction. It may mean that explanation is not a matter of logic alone but also involves *ad hoc* the proper use of our reversible models – i.e. the inclusion of extra *ad hoc* rules in addition to the logic of the mathematized economic models. Note, however, these needed rules are stronger than the mere 'semantic rules' discussed by linguistically oriented methodologists [see Massey 1965, p. 1159; De Alessi 1965, p. 474; cf. Simon 1953, pp. 65 and 74]. Such 'ad hocery' is quite unacceptable if we want the exogenous vs endogenous distinction to be a significant aspect of our explanations.

The irreversibility aspect of exogenous variables needs to be emphasized. It is just this condition which separates pure descriptions from explanations. This interpretation is counter to Samuelson's attempt to cast doubt on such a separation by casting doubt on one particular argument for that separation [Samuelson 1965, p. 1167]. Let us examine this separation more closely. Consider a model of our theory of explanation in which there is an intimate correspondence between mappings and explanations (see Figure 6.1). In a way, whenever the model is solvable, the Walrasian model [Σ] 'maps' from a point in the set of all conceivable combinations of **R**s, **D**s and **A**s to a point in the set of all conceivable combinations of **P**s, **V**s and **X**s. Wald's question concerns whether a unique **PVX** point exists for any given **RDA** point. With this in mind, I think economic theorists presume the following:

(1) Explanatoriness requires that the 'mapping' be 'well defined' and 'onto'.

In mathematical economics, explanatoriness is a matter of establishing (respectively) the consistency and completeness of the model [see Boland 1970b, Wong 1973]. The mapping must be 'onto' because one has not completely explained why **P**, **V** and **X** are what they are unless one has also explained why they are not what they are not (relative to what is conceivable). Although this was suggested by Donald Gordon [1955a, p. 150], it is too often overlooked, even though it seems to be essential [Nikaido 1960/70, p. 268].

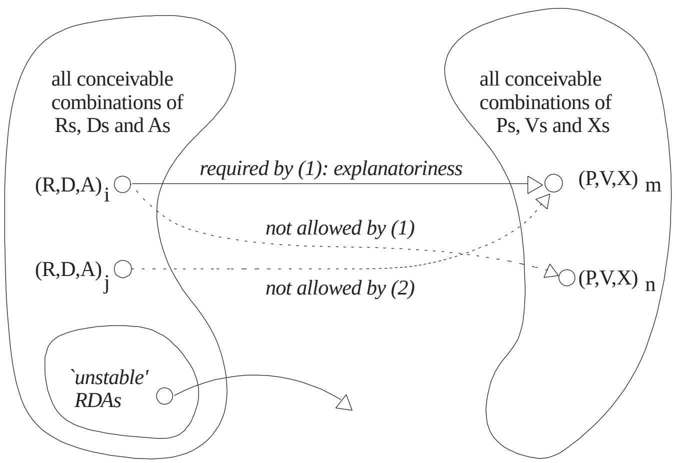

Figure 6.1 *Explaining* P, V *and* X

3. Methodological requirements of informative models

Now there is another, usually unstated, methodological requirement for any explanatory 'mapping' – we want it to be 'informative'. This is a fundamental part of the 'exogenous' character of the **RDA**s and the relationships between **RDA**s and **PVX**s. These relationships are the functions to which Gordon [1955a] refers. The logic of explanation in economics is such that the postulated reasons (for why **P**, **V** and **X** are what they are) consists of the following: (a) the world operates as indicated by the assumed relationships of our model and (b) the exogenous **R**, **D** and **A** happen to be what they are. Had **R**, **D** or **A** been anything else, we would have observed different values for **P**, **V** and **X**. This then is merely a special case of a rational argument for something. The only direct test of this explanation is to determine (independently) whether the real world is as purported – i.e. whether **R**, **D** and **A** are those states which, when plugged into our model, yield the values of the **PVX** point we began with, the occurrence of which we are trying to explain [Gordon 1955a, p. 152].

Note that the common view that one tests an explanation by attempting to predict other **PVX** points on the basis of changing **R**, **D** or **A** presumes a knowledge of **R**, **D** and **A**. But, on the grounds of requirement (1) *alone*, the only means of testing the adequacy of an explanation would be to see if the representation of the explanatory mapping is well defined (and onto). That is, one would look to see if it was true that there had been a change in the **PVX** point when there was no change in the **RDA** point. If this situation were observed then one would know that their explanation of **P**, **V** and **X** is inadequate [Hansen 1970, p. 4]. The 'onto' part requires that something be said about *all* possible **PVX** points.

If requirement (1) were the only basis for a theory of explanation in economics, one could see why most economists might consider such a means of testing to be a practical impossibility in a world of exogenously changing **RDA**s. However, testing can be more than just testing the logical adequacy of the postulated reasons. Once the **RDA** point is known, we have everything needed to directly test the reasons – i.e. for the comparison of the known **RDA** point with the **RDA** point that would be logically compatible with the values of the **PVX** point we wish to explain. That **RDA** point would have to be known in order to know whether it has changed, so long as the knowledge is independent of our knowing **P**, **V** and **X**. This independence is not the case in econometrics, where we use **P**, **V**, **X**, **R** and **D** to determine what the elements of **A** are. Such considerations imply a second requirement:

(2) Informativeness (i.e. testability as falsifiability) requires the mapping to be 'one-one'.

In other words, for any **PVX** point there is only one **RDA** point whichwhen plugged into the model yields that given **PVX** point.

4. The methodological dilemma at issue

If the mapping from **RDA**s represents the explanation of **PVX** and is, in turn, represented by a function, as would be the case with a system of equations, then that function would have to be both ‘onto’ and ‘one-one’. This means that the function is an isomorphism, or a pure description. For our purposes, it is indistinguishable from what economists call a tautology. Moreover, such an explanation would be logically reversible hence contrary to the stated intentions [Simon 1953, p. 52]. Consequently, the methodological claim that the **PVX** point is explained (by a combination of the model and the occurrence of a particular **RDA** point) is not completely represented by the logic of the mathematized model whenever our concept of explanation must incorporate the economic theorist’s desired distinction between exogenous and endogenous and still satisfy the methodological requirements of explanatoriness and testability that is found in economics literature.

Herbert Simon addressed the problem, capturing, in a non-arbitrary way, the explanatoriness of a model by analyzing the causal structure of a model. His solution was to order the equations in a chain of exogeneity whereby the endogenous variables of one equation may be exogenous to the next equation [p. 65]. Exogeneity of the entire model rested with the exogenous variables of the first equation. If we were to add Simon’s needed *ad hoc* specification of one set of variables as exogenous and the remainder as endogenous, we would only beg the question that our models purport to answer.

5. The Super Correspondence Principle

The discussion so far should bring us close to the realization that the problem of explanatoriness, the problem of identification (which is usually solved by equating the number of exogenous and endogenous variables – see Chapter 3), and the problem of stability [Boland 1977b, 1986] which follows from Samuelson’s ‘correspondence principle’ (the testability of equilibrium models requires a specification of disequilibrium dynamics that would provide an explanation for why the equilibrium state was achieved) are not separate issues as most economic

theorists would seem to think. Rather, these three allegedly separate problems are all variants of one methodological problem – namely, the problem of the informativeness (beyond available data) of our economic theories and models. No one of the allegedly separate problems is solved unless all of them are solved. It is what might be called the *Super Correspondence Principle*.

The informativeness of the explanation of **PVX** would be assured when it is realized that the stability analysis of a model must allow for some **RDA**s for which there does not exist ‘equilibrium’ **PVX**s. That is, the set of conceivable (independently of the given theory) combinations of **R**s, **D**s and **A**s must contain at least one subset of combinations which do not ‘map’ to any **PVX** point. On the basis of an independent concept of **R**, **D** or **A** (specifically, one which does not assume that the given theory of **PVX** is true), unless it is logically possible to identify conceivably observable **RDA** combinations which would imply ‘unstable’ or inconsistent circumstances for the determination of the **PVX**s, there is no assurance that anything has been explained.

6. Falsifiability to the rescue

Such a consideration leads to the addition of an extra clause to requirement (2) – namely, the empirical falsifiability requirement that there are points in the **RDA** set which do not map into the **PVX** set. With this addition the requirement now is that the independently conceivable data will not only contain counter-examples (false ‘if-then’ predictions) but will necessarily extend beyond what could be known from the observance of confirming data (true ‘if-then’ predictions).

Two things should be noted. The first is simple: Any *ad hoc* specification of exogeneity to some particular variables will not be equivalent to the addition of the extra clause to requirement (2) as the latter cannot be satisfied *ad hoc* but must be done *within* the theory of **PVX**. The second is more important: Once the extra clause is added to requirement (2), mathematical models consisting only of functional relationships will never be able to satisfy the augmented second requirement. This is because any function which is ‘well defined’, ‘onto’ and ‘one-one’ will imply the existence of the unique inverse, which is itself such a function. That would mean that the inverse is ‘onto’, which is a direct contradiction of the extra clause. Note that by defining **RDA** such that it is consistent with all **PVX**s, the ‘explanation’ is rendered uninformative since the theory would be circular and the circularity implies the reversibility of assumptions and conclusions.

In conclusion, I note that Wald and those mathematical economists who followed him were concerned with the mathematical requirements of solvability of mathematized economic models but may have failed to see the problem in the context of a complete theory of explanation which is able to distinguish explanations from descriptions or other pure exercises in logic. As a result, the common presupposition that explanation is a matter of logic alone leads to the view that 'a superior description … can be given the honorific title of explanationÌ' [Samuelson 1965, p. 1171] even though such a view contradicts the ordinary economic theorist's rationale for distinguishing between exogenous and endogenous variables.

7

On the Impossibility of Testability in Modern Economics

> to test, refute, or 'verify' a meaningful proposition is a tough empirical job and no amount of flossy deduction can obviate this.
>
> Paul A. Samuelson [1955, pp. 310-11]

> The only point that can be affirmed with confidence is that a model for a theory is not the theory itself.
>
> Ernest Nagel [1961, p. 116]

Philosophers of science have drawn fine distinctions between the ideas of testability, falsifiability, verifiability, refutability, and the like, but little of their literature seems relevant for economists. Economists today rarely make much of a distinction between testability and falsifiability (or refutability). Following Samuelson [1947/65], and bolstered by Popper [1959/61], in recent times the economics profession seems to have adopted one particular view of the philosophy of science to the exclusion of all other views.

The philosopher Karl Popper and the economic theorist Paul Samuelson have made falsifiability the keystone of their respective philosophies of science. Popper tells us that a theory is 'scientific only if it is capable of being *tested* by experience' [1959/61, p. 40]. However, Popper argued that, as a matter of logic, testability cannot be verifiability. Specifically, verification, the demonstration of the truth of a statement or theory, is impossible any time we have to prove that a *strictly universal* statement is true – *even if it is true* [p. 70]. For example, in proving that '*all* swans are white' we must guarantee that no future swan will be non-white. This consideration does not preclude falsifiability, as one can prove that such a statement is false with the demonstration of one or more counter-examples [p. 69, see Chapter 2 above]. Samuelson is more direct as he claims that a scientific theory must be empirically falsifiable [1947/65, p. 14; 1953, p. 1]. So both Popper and Samuelson use empirical falsifiability as a demarcation criterion to identify scientific theories.

The foundation of Popper's logical view of testability is the elementary asymmetry recognized in ordinary quantificational logic [Quine 1965]. Specifically, there is an asymmetry between the testability of statements which use the universal quantifiers 'all' or 'every' and the testability of those which use the existential quantifiers 'some' or 'at least one'. This *quantificational asymmetry* will be fundamental in the view to be examined here. The principle concern will be the view that in economics we must be concerned to put forth statements or theories which, as Samuelson [1947/65, p. 4] says, 'could conceivably be refuted, if only under ideal conditions'. As we shall see, Popper's logic sword cuts both ways when we build *models* of our theories.

I think it is important to recognize that, contrary to the popular perception of proponents of 'falsificationism' [e.g. Blaug 1980], testability is more often demanded by 'verificationists' than by 'Popperians' who supposedly demand 'falsifiability' as opposed to 'verifiability'. One would not want to waste time trying to verify a tautology, would one?[1]

When ordinary economists promote the requirement of testability, their purpose is not always obvious. Does testability of a theory mean that it is conceivably false (thus, not a tautology) or that it is conceivably falsifiable (thus, it is possible for the real world to be so constructed that some possible observation would contradict the theory *as a whole*) or does it mean something else?

In this chapter, I examine the basis for testability in modern economics where testability is to be viewed as the possibility of *empirical refutation*, that is, the possibility of observing conceivable facts which when conjoined with the theory in question forms a logical contradiction. This chapter is intended to challenge those economists who would have us utter nothing but testable (empirically refutable) statements *and* who also put a premium on logical arguments. The examination is concerned with two basic questions: (1) When is a statement (e.g. prediction or assumption) considered to be false? And (2) when is a theory as a whole considered to be false?

The discussion will focus exclusively on the now common view which presumes that one tests a theory by building a representative model of it and then tests that model. My argument will be an outline of a proof that it is impossible to test convincingly any economic theory if only models of the theory are tested (this proof will hold even when the theory is not tautological). The basis for the proof will be (a) the common presumption that a test of a model representing a theory constitutes a test of that theory (i.e. nothing more is required), (b) the methodological role of models (i.e. they involve additional assumptions in order to test a theory), (c) the rather standard view of testing (i.e. tests are always relative to testing conventions), and (d) the nature of logic (i.e. it is ambiguous regarding which assumption 'causes' a false prediction).

Economics today is thought to be capable of being 'scientific' mainly because economic theorists now have accepted the need to express their ideas in terms of specific models [see Sassower 1985, Ch. 2]. Mathematical models of economic theories supposedly have enabled us to perform tests of economic theories. The central question of this chapter is: Does refuting a specific model of a given theory *necessarily* refute the theory represented by that model? If not, then the popular methodological requirement that our theories be testable would seem to be rather puzzling at best. It will be argued here that for fundamental logical reasons it is impossible to test an economic theory by constructing specific *models* of that theory which are more empirically restrictive than the theory itself.

1. Tautology vs testability

Despite Samuelson's 1955 warning, little has been said in our literature that examines the logic of testing in economics. A possible exception is Bear and Orr [1967], but even their examination was limited to the discussion surrounding Friedman's instrumentalistic view concerning the need to test assumptions [Friedman 1953, see also Boland 1979a].

---

[1] However, the economic theorist Cliff Lloyd held that falsifiability is necessary but not sufficient for testability [Lloyd 1969, p. 87]. Some Popperians, however, do not always relate testability so strongly with falsifiability, as this distinction seems to involve extra-logical matters such as the purposes for testing. For example, Kurt Klappholz and Joseph Agassi [1959, p. 65] argued that 'Tests can easily be performed without risk of falsification…; but nothing can be learnt from such testing'. This latter distinction will not be the issue here. Nor will the weaker view of testability which only requires that data deducible from a theory be 'comparable' with empirical evidence [cf. Archibald 1966].

As I have briefly noted above, the current positive concern for testability may stem instead from the older methodological prescription that we should strive to verify our theories, assumptions or predictions. The problem that might face any theorist or model builder is that a tautology could be ‘verified’ even though one might think it is not meaningful or informative. That is, a successful verification might not be meaningful. To assure that any verification is meaningful one should make sure one’s theories are not tautologies. On the other hand, if a theory is testable then any verification would be meaningful. This prescription persists despite the above noted methodological arguments which seem to undermine the possibility of a successful verification of a testable (non-tautological) theory by showing that such a verification would go beyond what is empirically possible.

The requirement of testability has survived because most model builders presume it to be the sole means of avoiding tautologies. Actually, to avoid tautologies, all we need do is establish that our theories are conceivably false [Samuelson 1947/65, pp. 4, 22 and 172]. Testability may be a stronger requirement if viewed only from the basis of the logical properties of a theory. It has been argued that any statement which is conceivably falsifiable is, by logical necessity, testable since a counter-example which would falsify the theory could quite well be viewed as the outcome of a successful test [Finger 1971]. I have argued elsewhere that in economics (though perhaps not in physics) testability is quite different from conceivable falsifiability [Boland 1977b] but it is best to leave all such sophistication aside for the present.

2. Test criteria as conventions

Given that it is not always clear what ordinary economists mean by testability, as a first step in my argument against the possibility of direct model-based testability in modern economics, I wish to consider a distinction between testability and falsifiability. Falsification is a logical property of *successful* tests [Popper 1959/61, pp. 32-3] although it is not exclusive to them. A falsification is a proof that a statement in question is false. The primary logical aspect of a test is setting up a situation where the conjunction of what is being tested and the frame of reference against which it is being tested either yields a logical contradiction or it does not [see Boland 1983]. It would be ideal if that frame of reference were indisputable facts or observations, such that the proof of falsity

would be beyond question and hence absolute. But few of us would adopt such a naive view today. It is this non-existence of absolute facts which leads us to distinguish between testing and falsification. Testing itself is a willful act, one for which there are problems that extend beyond the logical question of falsification. Testing (at least today) is done in a manner designed to convince someone else (perhaps one’s colleagues) of the truth or falsity of a particular theory of some type.[2] When testing and falsification are not considered identical, it is not uncommon to find that what distinguishes testing from purely logical falsification is that testing, in any allegedly scientific enterprise such as economics, is a ‘sociological’ problem – a matter of social conventions [e.g. Lloyd 1965, p. 20; Blaug 1968, p. 7]. So, to keep things clear I will say that falsification is considered to be absolute whereas a test is relative to the testing standards agreed on (i.e. the criteria and the frame of reference).[3]

Problems could arise if we accept this distinction between falsification and testing. Whenever a social agreement *alone* must be relied on for testing conventions we must, logically, allow for any arbitrary convention [Lloyd 1969, p. 89] but this might also allow for possible dishonesty or ‘self-interested phantasies’ [Zeuthen 1957, p. 6]. To avoid the possibility of dishonesty we could revert to requiring absolute logical proofs of any falsification of a test statement (even when it is only within the relative context of the accepted conventions) [Lloyd 1965, p. 23]. This, however, leads to an infinite regress as every logical proof must itself use some assumptions. Thus the proof can be questioned, and the proof of the proof can be questioned, and so on. Testing as a sociological problem would seem to mean that we must choose between a dangerous arbitrariness or a hopeless infinite regress.

If the famous historian of science, Thomas Kuhn, is correct, the existence of a community of scientists seems to create another option. Any logical proof of falsification of a test statement is considered relative to the accepted ‘world view’ as manifested in our revealed ‘paradigms’ [Kuhn 1962/70]. The paradigms are the pool of available assumptions

2 Most practical ‘testing’ today is done in the context of ‘safeguards’ for society against dangerous drugs, building techniques, etc. This concept of testing easily applies to economic policy – but this, too, is not the use of testing that concerns us here because that is not what theorists such as Samuelson discuss.

3 An example would be an acceptable definition of a ‘swan’ or a ‘firm’, or an acceptable statistical parametric indicator [cf. Lloyd 1965, p. 23].

used to construct the logical proof of any falsification. They are the basis of our test conventions and are at any present moment – during everyday workings of 'normal science' – considered beyond question. Benjamin Ward [1972], for example, identifies 'liberal philosophy' as the basic paradigm of neoclassical economic theory. The famous 'impossibility proof' of Kenneth Arrow [1951/63] is based on unquestioned assumptions that are fundamental to that 'liberal philosophy'. The acceptance of the paradigms stops the potential infinite regress. This philosophical strategy is sometimes called 'Fideism' [see Bartley 1964b]. At the same time, the acceptance hopefully avoids the two problems of subjective truth – namely dishonesty and self-delusion. The avoidance is possible because the criteria are slowly developed within the community and only with ample support of the entire community of scientists are they still applied [Kuhn 1962/70, Ward 1972].

Which philosopher or historian of science is quoted to support one's preconceptions of methodology in economics is usually a fad but it is safe to assume for the purposes of this discussion that Kuhn's view of science comes closest to representing what is commonly accepted and practised in economics. On this basis we can say that a false test statement (e.g. a prediction) is one which contradicts the 'facts' as defined by our 'world view' (i.e. our paradigms). To many this would seem to be the only way we can define 'false' and avoid the above problems. Within this context (i.e. Fideism) a successful test can be convincing since the basis for the logical proof is, by common agreement, considered beyond question for the immediate purposes (i.e. normal science). Our 'world view' thus becomes our needed frame of reference – the starting or stopping point of *all* rational arguments. The danger for the would-be methodologist here is that 'theories' might be confused with 'world view' [cf. Weintraub 1985]. Equating the 'world view' with our theories would turn all of economics into a tautology [see Agassi 1971a]. That is, if our basic theories are beyond question – because they are treated as paradigms – there would be nothing to test. The fact that we consider alternative theories (of the firm or of the consumer) means that the standard theory is *not* a paradigm – no matter how standard. To identify our 'world view' we must look much deeper into such things as our ideas of individualism [Simon 1963, p. 230; Ward 1972, p. 26; Boland 1982, Ch. 2], of the independence of decision-making [Morgenstern 1972, p. 1171; Ward 1972, p. 26; Boland 1977b] or even the inherent goodness of Pareto optimality [Morgenstern 1972, p. 1169]. Of course, the details of our 'world view' are not the issue here, only the recognition that such a view exists.

3. The role of logic in testing

Both falsification and successful testing depend on logical proofs of some kind. Before we get to the matter of the role of logic, let us examine the nature of ordinary logic. Beyond our conventions concerning the criteria by which we would agree that a given test statement is false and constrained by the quantificational asymmetry noted above, testing relies on some important asymmetrical properties of logic [e.g. Bear and Orr 1967, p. 190]. These properties follow from the construction of our standard logic. We can say, for a given *set* of premises from which a set of conclusions is said to follow logically, that:

(1) if *all* the premises are true (in some sense) then *all* the conclusions will be true (in the same sense), without exception; and its corollary,

(2) if *any* conclusion is false (not true by the same sense) then *at least one* premise *of the given set* is false – but we do not know which premise is false (when there is more than one) nor do we know how many are false, for they could all be false;

(3) if any premise is false, it does not preclude any of the conclusions from being true; and its corollary,

(4) if any one of the conclusions is true, any of the premises from the given set could still be false.

Truth here is a property of a statement. A statement can be true or false. Truth is not something that can be discussed separately from the statement which exhibits the truth status in question. Nevertheless, it is heuristically useful to pretend that it is a quantity that can be passed around. In this (and only in this) heuristic sense we can say that if logic is correctly applied, the truth of the premises (or assumptions) will be 'passed' on to the conclusions (or predictions). However, even when correctly applied, the truth of the conclusions cannot usually be 'passed' back to the premises. Consequently, there is an asymmetry between the directions of apparent transfer. It is also true of this asymmetrical property that the falsity of the conclusions can be 'passed' back at least to the set of premises, but the falsity of the premises is in no way ‘passed’ to the conclusions (except those conclusions identical to the false premises). A more subtle asymmetry can be pointed out: whereas the truth of the *set* of premises is ‘passed’ unambiguously to *every* conclusion, the falsity of *any one* conclusion is ‘passed’ ambiguously only to the *set* of premises; that is, not necessarily to any particular premise. (All this is said with the realization that the truth of our theory is a separate matter from the knowledge of that truth – which is another asymmetry [see Boland 1982, Ch. 11; 1986, p. 18].)

There does not seem to be any significant dispute about these asymmetries. The consequences of these facts about logic define the role of logic in testing. First, it is impossible to verify a *theory* by examining only the conclusions that may be deduced from it (even though we may show many of them to be true). Second, it is impossible to verify any of the conclusions *indirectly* from the *known* truth of all the premises whenever any one of the premises of the theory happens to be a strictly universal statement (as at least one must be, in order for the theory to explain or predict [Popper 1959/61, p. 60]). As I noted already, this is so merely because we can never *know* when a strictly universal statement is true in terms of empirical facts from which that knowledge is alleged to follow logically. And third, showing any premise to be false says nothing about the truth or falsity of the conclusions.[4] The inability to pin down the source of falsity (of a conclusion) in the *set* of premises, i.e. property (2), is probably *the* major obstacle in testing economic theories. A falsification of one conclusion (e.g. a prediction) need not say much about the theory *if* additional assumptions have been added to the theory to make it testable.[5] Since this is exactly what we do when we build a *model* (see Chapter 1), we need to examine the consequences of using the asymmetries of logic somewhat more closely.

4. The role of models in testing theories

As I noted in Chapter 1, there are two main reasons economists might wish to test their behavioural theories. Applied economists might wish to test for the limits of applicability of their general theories. Pure theorists might wish to determine the truth status of their theories. In economics, theories are made empirically testable by constructing models of our theories [e.g. Marschak 1953, p. 25; Koopmans 1957, p. 173; Papandreou 1958, pp. 8-11; Arrow 1968, p. 639]. As I have tried to show in Chapters 2 and 3, as well as Chapter 6, how one chooses to build a specific model can have profound consequences for the testability of the constructed model. In those chapters we were dealing with only non-stochastic models. Here the discussion is more general. Nevertheless, the essential point is that models are constructed to be more specific than the behavioural theories they represent.

One way to test the limits of applicability of a theory is to apply that theory to a practical real-world situation. Practical real-world situations are very specific. For example, demand elasticities may be known in advance, production cost functions may be easy to specify. With such considerations in mind, it is possible to build a model of a theory that is limited to a specific situation, in the sense that there is no claim that it would pass a test when applied to some other situation with different demand elasticities or cost functions.

Those theorists concerned with the truth status of their theories also build specific models when performing empirical tests. To be specific enough for the purposes of an empirical test (that is, with evidence drawn from the real world), it is almost always necessary to modify the general behavioural theory with extra assumptions (as is routinely done in econometric tests) by representing the behavioural situation with specific mathematical functions [Koopmans 1953, p. 29]. Should the production function be linear or quadratic or what? Should we worry about formal solutions that imply negative output levels when none will ever be observed? Recognizing that there may be random errors in the observations used to construct the test, what is the allowable error in a statement of equality implied by an explicit equation? Such questions must be addressed when building a specific model of a theory whether our objectives are those of applied theorists concerned with the applicability of their theories or of pure theorists concerned with the question of whether their theories constitute true explanations of the economy. To answer these questions is to specify the additional assumptions.

The central question for the theorist who builds models to test the truth status of a theory (or to test the theory’s applicability) concerns the well-known ‘problem’ that follows from the ambiguity caused by the addition of the assumptions needed to specify the model. If a model is

4 An exception is the trivial case pointed out by Samuelson [1963] where the purpose of a set of assumptions is to form a pure description.

5 An exception is the rare case where the *only* additional assumptions were paradigmatic and beyond question [see Boland 1981b].

revealed to be false by a test, does this mean that the theory it represents is false or that it is just a poor representation? To deal with this question, I would like to make clear the ingredients of a model. Initially, I wish to recognize two separate ingredients: the theory itself and the added assumptions used to specify the functions that represent the theory. This gives the following schemata for the concept of a model as used here:

(A) A *set of behavioural assumptions* about people and/or institutions. This set might include, for example, the behavioural proposition $\mathbf{C} = \mathrm{f}(\mathbf{Y})$, where $\partial\mathbf{C}/\partial\mathbf{Y}$ is positive. The *conjunction* of all the behavioural assumptions is what traditionally constitutes a 'theory'.

(B) A *set of simplifying assumptions* about the relationships contained in the above set. For example, the demand function stated in the theory might be specified as a linear function, $\mathbf{C} = a + b\mathbf{Y}$, where a is positive and b is between 0 and 1.

Any model is a conjunction of these two sets of assumptions. The schemata lists these sets in a descending order of autonomy. As discussed in Chapter 2, the nature of the assumptions about functional forms is limited by the number of relevant variables recognized in the theory. And, as discussed in Chapter 3, an econometric study presumes that the two sets of assumptions are true and applies them to actual data to deduce the parametric values recognized by the second set.

That a model is a conjunction of more than just an explanatory theory means that from consideration (2) of the asymmetries of logic (see Section 3, p. 135), we must conclude that when a prediction of a model is falsified we still do not know specifically whether the 'cause' was a basic assumption of the theory or an additional assumption introduced in the model's construction [Lloyd 1965, p. 22]. This is an important consideration for all those who say that the purpose of constructing mathematical models of a theory is to *make* that theory testable.

5. The falsification of theories using models

Now we are close to establishing the major point of this chapter. If we think we are going to test an unspecified theory by showing that it is false on the basis of empirical tests, then we must be prepared to be able

to show that all possible models of the theory are false. In other words, to test the basic behavioural assumptions themselves we must consider all possible ways of modelling them (however simple or complex). Unfortunately, this is a very difficult assignment since there will always be an infinite number of ways of modelling any given theory. In the absence of errors in logic, if *every* one of the modelling assumptions when conjoined with the basic assumptions can be shown to lead to *at least one* false prediction then at least one basic assumption must be false. And if that were not the case, that is if all the assumptions are (non-tautologically) true, then it is *possible* to specify the basic assumptions such that no false predictions could or would ever happen. I wish to stress that I am making this argument without having to resort to extraordinary or otherwise severe concepts of truth and falsity. To the contrary, my argument here presumes that the determination of the truth of any prediction is a matter of ordinary testing conventions (e.g. R^2 of appropriate value) which, as noted in Section 2, need only be consistent with the accepted 'world view' of the community of scientists.

Obviously, the requirement that we must show that *all* possible models are false is impossible – for the same reason that it is impossible to verify a strictly universal statement (the quantificational asymmetry). We must therefore conclude that on this basis (i.e. the relative basis of our 'world view') *the empirical falsification (and thus testability) of any economic theory is impossible whenever the test is based only on an evaluation of models representing the theory*. For future reference I will call this the Ambiguity of Direct Model Refutation. It is also obvious that we cannot have the ability to do something which is impossible. So, the Ambiguity of Direct Model Refutation means that any economists who say that falsifiability is necessary but not sufficient for testability may have to admit to the impossibility of testability in modern economics.

6. The 'bad news'

It is always very difficult to write about methodology. Even if one were to prove that positive supporting evidence in no way decides the truth of a theory, readers will still demand positive evidence of one's proof. Methodologists are not immune to such inconsistencies.

This chapter was motivated partly by the popularity among economists of the view that our major methodological concern is, and should be, the testability of our theories. (I have heard of more than one theoretical manuscript being rejected by editors because it could not be shown that it contributed to an increase in the testability of standard demand theory.) Increased testability is, by implication, the primary criterion of scientific progress.

I have deliberately referred to very few examples in this chapter because I think the possibility of direct testability is a matter of logical inconsistency. Besides, one cannot prove impossibility with a long list of examples. The view that one must show an example of an error to show that an error of logic exists is itself an error of logic! I am concerned with the logic of the commonplace *argument* for testability in economics. It is sufficient to show that a logical error exists. Note that by property (2) of logic (p. 135), showing one conclusion (or prediction) to be false indicates that one (or more) of the assumptions is false *but that condition presumes the absence of logical error*.

In summary, the foregoing constitutes a theory of testing in modern economics. It is based on: (1) the asymmetries of logic which are beyond question here, (2) the assumption that economists are followers of something like Kuhn's view of science at least to the extent that the acceptance of test results depends on the use of conventionally accepted test criteria, and (3) the empirical assumption that economists think their theories are not directly testable except by constructing (mathematical) *models of their theories*. It is logically impossible to combine these assumptions with the Popper-Samuelson doctrine that we must utter only testable economic statements. What are we to do? Perhaps all economists interested in scientific rigour need to examine their view that testability is the sole means of avoiding tautologies. There simply are more ways than one to avoid tautologies. In short, on the safe assumption that we are unwilling to abandon ordinary logic, we may have to yield either to the falsity of the strategy of relying on Kuhn's view of science which presumes that testing must be based only on socially accepted testing conventions (see Chapter 5) or to the falsity of Samuelson's methodology which avoids tautologies only by requiring testability – or to the falsity of both.

8

Model Specifications, Stochasticism and Convincing Tests in Economics

> Refutations have often been regarded as establishing the failure of a scientist, or at least of his theory. It should be stressed that this is an inductivist error. Every refutation should be regarded as a great success; not merely a success of the scientist who refuted the theory, but also of the scientist who created the refuted theory and who thus in the first instance suggested, if only indirectly, the refuting experiment.
>
> Karl R. Popper [1965, p. 243]

> Even in the most narrowly technical matters of scientific discussion economists have a shared set of convictions about what makes an argument strong, but a set which they have not examined, which they can communicate to graduate students only tacitly, and which contains many elements embarrassing to the official rhetoric. A good example is the typical procedure in econometrics.
>
> Donald N. McCloskey [1983, p. 494]

There are two basic problems relating to discussions of methodology. Too much is taken for granted and too little has been worked out with the thoroughness we have come to expect in discussions of other aspects of economics. Throughout this book I have attempted to analyze critically the explicit process of building models of our traditional economic theories. Model building is so widely practiced in economics that it becomes very difficult to question the soundness of the process. Even when we accept the traditional theories of the textbooks we still must make decisions regarding the process of applying them to the real world. As such the process of model building is certainly not automatic. There may be infinitely many potential models that could be used to represent any given theory. However, it does not matter whether we view model building as a process of choosing between existing complete models or a process of building from the ground up, so to speak, assumption by assumption.

In Chapters 6 and 7, I have dealt with specific problems that arise in the process of model building, problems which are peculiar to models and to their relationship with the given theory. The question 'What is the purpose for building any model?' must always be asked. No single model will serve all purposes. The models we build to describe or explain the real world will not necessarily be usable for testing the given theory. Similarly, models built for quick application to policy issues may not always be sufficiently realistic for plausible explanations or convincing tests.

Sometimes one might suspect that these considerations may not be well understood or appreciated by most economists. Too often it is assumed that a model which has been shown to be successful for one purpose (e.g. description of data) will automatically be sufficient for another (e.g. testing). Throughout this chapter, I will be examining two specific types of erroneous methodological positions often taken regarding testing and the nature of models. The first position is the one taken by economists who claim to have performed tests of economic theories by testing models whose specifications represent 'interpretations' of those theories. More is required than testing numerous models or interpretations of those theories. The second popular position is the one which views the real world as a necessarily 'stochastic environment'. The problem here is that it is our models which are stochastic rather than the world we wish to explain.

1. Falsifiability lives in modern economics

Critics of Popper's philosophy of science or Samuelson's methodology were probably encouraged in Chapter 7 by my pessimistic look at the possibility of refuting a theory by testing only a model of that theory. Many of Popper's critics will continue to argue that in actual practice ordinary economists do not refute their theories. Refutability, the critics might say (in concert with Chapter 5), is a conventionalist criterion for theory choice and any progress in economics has not been due to economists refuting their theories and replacing them by better ones – that is, by economists practicing what the critics call Popper's

'falsificationism' (e.g. Daniel Hausman, Wade Hands, Bruce Caldwell, etc.). Economic methodologists who may be so eager to dismiss Popper's view of science are unable to see the forest for the trees. Perhaps they spend too much time talking to philosophers rather than to practicing economists.

Modern economics is characterized more by the activities of model builders than by debates among economic philosophers and ideologists. Modern economists see themselves engaged in an ongoing saga of advances in model building. The behavioural theory that ordinary economists attempt to model has not changed in several decades, but various models come and go. Roy Weintraub [1985, pp. 112-17 and 121] argues eloquently in his history of neo-Walrasian economics that there was steady progress from the 1930s to the high plateau in the work of Arrow and Debreu [1954] which continues to be refined even today. What I have been calling the set of behavioural assumptions is what Weintraub and others call the 'core' or 'hard core' of a research programme. He said that there has been significant progress in the 'hardening of the hard core'. The hardening is a result of the axiomatization of general equilibrium theory followed by a long series of adjustments to the list of behavioural assumptions. The adjustments have been primarily ones which reduce the list to just those necessary and sufficient to represent any Walrasian general equilibrium.

Critics of neoclassical economics may wish to say that measuring the extent to which the core has been hardened sometimes requires a powerful microscope. The progress in general equilibrium theory is much more visible when it is seen to involve the rejection (rather than empirical refutation) of various modelling techniques. In the 1930s the use of calculus-based techniques were commonplace. The utility or production functions were specified to be everywhere continuous and differentiable in order to complete a model of general equilibrium. The functions were specified as such to fulfill the requirements of Brouwer's fixed point theorem which was commonly used to prove the existence of a general equilibrium [see Wald 1936/51]. Calculus was eventually rejected as a basis of general equilibrium modelling. In later models the continuous utility and production functions were replaced with 'upper semi-continuous' set-theoretic correspondences. These less demanding specifications were allowed by Kakutani's fixed point theorem [see Debreu 1959]. The testing standards – if we wish to see the process in these terms – were the criteria employed by formalist mathematicians.

Any model which cannot be shown to meet the 'standards of rigor of the contemporary formalist school of mathematics' [Debreu 1959, p. viii] is to be rejected in the same way empirical models are rejected when they do not meet the standards set by the currently accepted testing conventions.

As I have noted in Chapter 1, economists today are more concerned with testing models than testing theories. Models are routinely rejected without ever calling into question the basic set of behavioural assumptions or ever seeing a need for adjustments [see Boland 1981b]. The linear algebra of activity analysis and linear programming used extensively during the 1950s and 1960s has virtually disappeared from the literature. Game theory was briefly popular in the 1960s and seems to have reappeared in the 1980s. None of the changes in modelling techniques has resulted in changes in the basic behavioural assumptions.

Editors of the **Journal of Political Economy** and the **Journal of Money, Credit, and Banking** have often complained that failed attempts to fit data have not been reported. The problem, as these complaints imply, is that the actual refutations of models and modelling techniques are not observable phenomena. I think the widespread practice of simply not reporting failed empirical models has misled the critics of Popper to think that falsificationism is unrealistic. Certainly, the rejection of linear models in favour of quadratic models implies that some economists consider linear models of certain phenomena to be false. Similarly, if we go beyond Chapter 7 by recognizing that any chosen estimation convention is an integral part of one's econometric model [see Boland 1977a] then we can also see that economists who reject ordinary least-squares (OLS) in favour of generalized least-squares (GLS) or two-stage least-squares (2SLS) as means of estimating a model's parameters do so because they have found models based on OLS estimates to be false in some important respect. Of course, there are many unreported models which researchers have deemed false and thus unusable for the purpose of model building. These observations lead me to conclude that if we view the practice of economics to be building and testing models rather than the more lofty pursuit of testing general theories of economics, then falsification lives regardless of the views of some critics of Popper's philosophy of science.

2. Overcoming Ambiguity of Direct Model Refutations

Through the eyes of practicing economists it is easy to see that the critics of Popper's so-called falsificationism are wrong. Ordinary practicing economists, who see their task as one of building and applying models of neoclassical economics, will testify that the occurrence of *model* refutation is a common experience. Nevertheless, spending one's time refuting particular models without ever addressing the question of testing neoclassical theory itself seems less than satisfactory. Surely one can go beyond the problems inherent in the Ambiguity of Direct Model Refutation. In the remainder of this chapter I present an approach to testing which to some extent does overcome the Ambiguity of Direct Model Refutation. For anyone trying to refute a theory it will require a little extra effort before they rush off to employ the latest modelling techniques.

2.1. Critical interpretations are merely models

Our central concern in Chapter 7 was the issue of testing with models. In this context, I argued only that the falsification of a model of a theory does not necessarily imply the falsification of the theory itself and I called this the Ambiguity of Direct Model Refutation. The methodological problem at issue concerned the logical relationship between models and theories and the limitations imposed by the principles of ordinary logic. In this light I noted that all testing involves adding extra assumptions to the theory being tested. That is, it involves building a model of the theory in order to test the theory. Surely, one cannot expect to be able to observe just one false model and thereby prove the theory itself to be false. If one thinks that theories are refutable in this way, then one's job would appear to be much too easy. For example, one could always append a known false extra assumption to a theory and thereby construct a model which is automatically false. Certainly, such testing would be 'unfair' at best. Just as surely we would have no reason to expect that proponents of a theory would accept such a 'refutation' of their theory. So, in what way does building a model of a theory constitute a test *of the theory*?

Many economists seem to think that the act of building a model always constitutes a test of a theory because the act of specification amounts to an interpretation of the theory. For example, in a critique of

Milton Friedman's famous essay on methodology, Tjallings Koopmans claimed that if any one interpretation of a theory is false then the theory must be false [Koopmans 1957, p. 138]. This method of criticism presumes that 'interpretation' is a process equivalent to a logical derivation from a set of given postulates without the addition of any other assumptions. Probably only mathematicians would make such a claim since the pure mathematical models used by economists are always presumed to be logically complete. With complete models, the only possibility of a false interpretation would be due to an error in the logical derivation. Surely, there are other ways to produce a false interpretation. If so, under what circumstances does a possibly false interpretation of someone's view constitute a criticism of that view?

Despite Koopmans's presumption, the ordinary sense of the word 'interpretation' (like 'model building') always involves additional assumptions (e.g. 'I assume by this word you mean …'). Moreover, any assumption could be false. Putting numbers in place of letters in mathematical models is an overt act of 'specifying' the equations. Most interpretations require such specifications of the variables. Such specification involves at least an assumption about their dimension or scale and this type of specification, too, involves possibly false assumptions. In other words, a model is merely a mode of interpretation. More important, a model or an interpretation can lead to a successful direct criticism (or test) only when what has been added is known to be true. This is the root of the problem. Testing a theory by adding assumptions and finding that the resulting model does not 'fit the facts' does not usually allow us to conclude that the theory was at fault since our added assumptions may not be true.

Some readers may say that the real root of this problem is that it is always difficult or impossible to determine when any assumption is true. They may be right and I will consider how model builders cope with this question. First, let us assume that it is possible to make true observations such that there is no ambiguity concerning whether a model 'fits the facts', so that I can show that testing theories with models is not completely impossible. The apparent impossibility of testing theories with models is due entirely to not going far enough. It will be argued subsequently that if one builds a model of the theory and also builds a model of a conceivable counter-example to the *theory* in question, then, using ordinary test conventions, convincing tests can be performed!

2.2. Testing with models of counter-examples

Prior to discussing the difficulties encountered when performing a convincing test in this way, we need to investigate what constitutes a counter-example and when a model of a counter-example constitutes a refutation of the theory. Rather than building a model of a theory to see whether it 'fits' the available data, considering counter-examples amounts to another approach to testing theories by building models. Before we proceed to build a model by adding what I earlier called 'simplifying assumptions' (i.e. extra assumptions concerning the functional relationship between the endogenous and exogenous variables recognized by the 'behavioural assumptions'), we might try to identify one or more propositions that are directly denied by the behavioural assumptions alone (i.e. without benefit of further specifications).

At first this approach seems too easy. Consider a theory which has several endogenous variables but has only one exogenous variable. Following my discussion in Chapter 6, we could say that any two observations of the values of the endogenous and exogenous variables which show that the values of the exogenous variable did not change, but also show that one or more of the endogenous variables did change, constitutes a refutation of that theory. For endogenous variables to change at least one exogenous variable must change. We can thus say that the observation of changes in endogenous variables without changes in the posited exogenous variables constitutes a counter-example for the theory in question. In the performance of this test we can see that two observations (of all variables) may constitute a refutation.

There are other conceivable counter-examples which may require more. When we consider theories which recognize many exogenous variables, things get much more complex and the minimum number of observations can grow large, as I have shown in Chapters 2 and 3. What constitutes a counter-example is also limited by considerations of quantificational logic. To use the philosopher's example, if our theory is merely that all swans are white, then the observation of just one non-white swan would constitute a counter-example. If our theory were, instead, that there is at least one pink swan, then the observation of a counter-example is impossible. The counter-example in this case amounts to a collection of observations sufficiently large to prove that all swans are non-pink.

Strictly speaking, one does not observe a counter-example directly. Instead, one builds a model of a conceivable counter-example *relevant* for the theory such that verifying the model would necessarily refute the theory. This requirement of relevance is apparently not widely appreciated. It is often violated in discussions of Giffen effects. Would the observation of the Giffen effect logically ever be considered a refutation of traditional ordinal demand theory? At first blush everyone might answer 'yes'. But on logical grounds such a question is very misleading since it would presume that we have a complete theory of the downward-sloping demand curve – i.e. of the so-called Law of Demand. It may have been the intended purpose of demand theory to explain why demand curves are always downward sloping [Hicks 1956, p. 59] but ordinal demand theory never succeeded in doing so [Samuelson 1953, pp. 1-2]. Simply stated, the existence of a Giffen effect is not completely denied by ordinal demand theory, hence its observation cannot be considered a refutation [see Samuelson 1948].[1]

In general terms, whether a particular observation constitutes a test (i.e. a refutation or a verification) of a given theory necessarily depends on what that theory logically affirms or denies. Such dependence (or 'relevance') is never a matter of judgement. It is always a matter of logic. What constitutes a test always depends on what is put at stake within the theory being tested. Whenever a theory is claimed to be true and informative, it must be denying specific observations. The more informative the theory, the more observations denied. This connection between informativeness and the number of conceivable counter-examples is the keystone of both Popper's philosophy of science and Samuelson's methodology of searching for 'operationally meaningful propositions'. But more important, it is the significance of what is denied by a theory that determines how much is at stake.

Let us assume away some of our irritants. Let us assume (1) that relevance is not a problem, (2) that we can test theories without having to build models, and (3) that, for the sake of argument, the logical consistency of the set of assumptions constituting the theory or model has not yet been established. Now let us consider a *simultaneous test* of a theory and one of its many counter-examples. On the one hand, in Chapter 2 we recognized that it is impossible to verify a theory by showing that the theory 'fits' one observation of all of its variables – that is, by finding a 'good fit' with the available data – since there is no guarantee that more observations tomorrow will also 'fit'.[2] However, if a counter-example does 'fit' the data (e.g. an observed change in endogenous variables without a change in the exogenous variables) then, so long as we accept the observations as true statements, we would have to admit that any logically complete and consistent theory which denies the counter-example in question has been refuted. That is, in any combination of the theory and this counter-example, both cannot be true.

Consider now the four possible outcomes of a combined simultaneous test of the theory and one of its counter-examples. If neither the theory nor the counter-example fits the available data then we could easily argue that the theory must not be consistent. If both the theory and its counter-example fit the available data then again it is easy to argue that the theory could not be consistent. Of course, these conclusions are based on the acceptance of the observations as true. If the theory is logically consistent then we would expect that any combined simultaneous test of the theory and one of its counter-examples will yield a fit of either the theory or its counter-example – i.e. at least one, but not both. When it is the counter-example that fits, the theory is obviously refuted – either directly because the counter-example is 'verified' or indirectly by saying that even if the theory somehow fits, it would have revealed an inconsistency. When the theory fits but the counter-example does not, then not much has been accomplished. On the one hand, such an event is a minimum condition for logical consistency.

1 Alternatively, it can be argued that Giffen effects are contrary to our traditional theory of prices [see Boland 1977d]. Demand theory itself is traditionally offered as logical support for our equilibrium theory of prices. Elsewhere I have gone further to argue that downward-sloping demand curves are necessary for a stable equilibrium in a world of truly independent decision-makers [see Boland 1977b, 1986]. In this sense ordinal demand theory is intended to be a complete set of reasons for why demand curves are downward sloping. And in particular, those reasons are required to be consistent with independent decision-making. As is well known, the traditional demand theory is only able to tell us when Giffen effects will occur (e.g. the implications of the Slutsky relations – a Giffen effect implies a counter income effect that is stronger than the substitution effect of a change in price). Thus, apart from price theory, Giffen effects are not denied and the simple observation of a Giffen effect alone would not constitute a test of ordinal demand theory, no matter what one means by 'testing'. Such testing in this case is simply not relevant.

2 For example, observing only white swans to date does not assure us that tomorrow we will not see a non-white swan.

On the other hand, it is still only a single fit and (as I have already noted) there is no guarantee that the theory will fit future observations (or that other possible counter-examples will fit the current data).

What is important about this combined approach to testing is that, if we accept the observations as being true, we can overcome the problem of the Ambiguity of Direct Model Refutation. To see this we need to reconsider the arguments of Chapter 7 where it was noted that showing that a specific model of a theory does not yield a ‘good fit’ will not (by itself) prove that the theory being modelled is false until one has proven that there does not exist some other model of the theory which does yield a ‘good fit’. While a bad fitting model *of the theory* does not constitute a refutation (even though we accept the observations as true) a good fitting model *of the counter-example of the theory* may constitute a refutation when the observations are considered true. To see this let us again assume the behavioural theory is logically consistent so that either the theory is true or its counter-example is true but not both. Again, when we are using the same data, there are four possible outcomes of a combined simultaneous test of the model of the theory itself and a model of one of its counter-examples. Whenever models of the theory and its counter-example both fit the data, we know there is something wrong with the modelling. If they both do not fit then not much has been accomplished since a bad fit of either the theory or the counter-example runs afoul of the Ambiguity of Direct Model Refutation.

Test model of:		
Theory	*Counter-example*	TEST RESULT
good-fit	good-fit	model inconsistency
good-fit	bad-fit	corroboration
bad-fit	good-fit	refutation
bad-fit	bad-fit	ambiguous

Table 1

Whenever the model of the counter-example fits and the model of the theory does not then this is a strong case against the theory, although we cannot be sure there is no modelling problem. Avoidance of a refutation would require at least a critical examination of the modelling methodology. When the model of the theory fits but the model of the counter-example does not then we have a situation which Popper [1965, p. 220] calls a ‘corroboration’. Going beyond Popper we can say that a corroboration would occur whenever the combined simultaneous test of a theory and its counter-example runs the risk of yielding a refutation but the behavioural theory manages to survive. A corroboration means that a refutation could have occurred but did not. These four outcomes are summarized in Table 1.

3. Stochasticism and econometric models

Having now argued that convincing refutations are logically possible – at least, in principle – we should see whether my argument is compromised by the consideration of the difficulties involved in the acceptance of observations as true statements. Any refutation of a theory based on a model of a counter-example still requires the acceptance of the truth of the refuting observation or evidence. As one of my early critics noted, ‘The quest for truth and validity is indeed a noble venture. However, the economist exists in a stochastic environment’ [Smith 1969, p. 81].

3.1. Stochastic models vs stochastic worlds

The problem with ‘stochasticism’ is that it takes too much for granted. Modern economists are very fond of claiming (like Professor Smith) that the world is a ‘stochastic environment’. This concept of the world is, I think, very misleading. Before examining the significance of stochasticism for my argument (namely, that testing is possible when we include tests of counter-examples), I offer a brief theory of stochasticism in modern economics. My purpose is to show that stochasticism involves model building since it requires an explicit modelling assumption which is possibly false, and thus stochasticism should not be taken for granted. Following this, I will use my brief theory to discuss how econometricians deal with stochasticism.

As I stated in Chapter 1, there are two ‘worlds’: the ‘real’ world which we observe and the ‘ideal’ world of the theory or mathematical model which we construct. When we say the theory (or model) is ‘true’ we mean that the real and the ideal worlds exactly correspond. Many will argue that there are obvious reasons why, even with true theories, the correspondence will not be exact (e.g. errors of measurement, irrational human behaviour, etc.). For these reasons, modern economists build ‘stochastic models’ which explicitly accommodate the stochastic nature of the correspondence. For example, we can assume that the measurement errors leave the observations in a normal random distribution about the true values of the ideal world. This means that the correspondence itself is the stochastic element of the model.

It should be noted, thus, that it is the model which is stochastic rather than the world or the ‘environment’. Any test of a stochastic model is as much a test of the assumed correspondence as it is of the theory itself. Following our discussion in Chapter 1, *one can choose to see the world as being necessarily stochastic only if one assumes beyond question that one’s model is true (and fixed) and thus that any variability of the correspondence is due entirely to the unexplainable changes in the real world*. Thus, stochasticism can be seen to put the truth of our theories beyond question.

I think there is always a serious danger of intellectual dishonesty in asserting that the environment is stochastic. We assume that the ‘assumptions’ of our theory or model are true because we do not know them to be true. Thus there is no reason for any of them to be put beyond question, as stochasticism seems to presume.

3.2. Econometrics as a study of stochastic models

Of course, stochasticism itself is put beyond question in the study of econometric models. Econometrics was a research programme founded in the early 1930s to address the obvious need to be able to confront stochastic statistical data with exact models of economic theories. The usual statistical analysis that one would have learned in a typical mathematics department was not always appropriate for the intended research programme. In the early 1940s an entirely different approach was proposed. The idea then was to make the statistical analysis part of economic theory itself [Haavelmo 1941/44, see also Koopmans 1941, Mann and Wald 1943, and Haavelmo 1943]. While there is some danger in seeing this as an endorsement of stochasticism, Haavelmo was quite aware of the limitations of such an approach and was careful to stress that the approach necessitated separating our stochastic models from our exact theories. Moreover, he stressed that his approach required a thorough commitment to stochastic modelling with no hope of returning to the world of exact models [see Haavelmo 1941/44, pp. 55-9].

Few econometricians seem willing to go all the way with Haavelmo and thus still wish to see a possibility of stochastic models being helpful in the assessment of exact theories and models [e.g. Klein 1957]. Nevertheless, many lessons seem to have been learned from Haavelmo’s manifesto, the most important of which is to stress the importance of systematically recognizing that if our theories of the interrelationship of economic variables are true then we ought not treat the observation of one variable as independent of the observations of other variables. The question raised is whether and to what extent are the observation errors also interdependent.

The possibility of interdependent observation errors is manifested in the common emphasis on ‘stochastic models in which the error elements are associated with separate structural equations rather than with individual economic variables’ [Leontief 1948, p. 402]. For example, rather than using an exact (i.e. non-stochastic) equation such as:

$$\mathbf{C} = a + b\mathbf{Y}$$

we would explicitly recognize a ‘disturbance term’, $\mathbf{e}$, as follows:

$$\mathbf{C} = a + b\mathbf{Y} + \mathbf{e}$$

The disturbance term accounts for errors in the equation as a whole rather than just for errors resulting from observations of **C** and **Y**. Unfortunately, it also introduces or recognizes other ways of accounting for errors. My colleague, Peter Kennedy, itemizes three ways econometricians account for errors represented by the disturbance term: measurement error, specification error, and what he calls ‘human indeterminacy’ [Kennedy 1979/85, p. 3]. Measurement error is what I call ‘observation error’. A specification error is one possible consequence of decisions made by the model builder. While the occurrence of observation errors is external and independent of the model, specification errors are internal and completely dependent on the nature of the constructed model. By impounding specification errors with observation errors into the disturbance term, econometricians make it extremely difficult to discuss the truth status of one's modelling assumptions. The recognition of 'human indeterminacy' is, of course, allowance for those econometricians who believe in stochasticism. Since I am rejecting stochasticism, 'human indeterminacy' is an unacceptable means of accounting for the magnitude of the disturbance term.

If one restricts econometric model building to practical problems then one would have to say, in Friedman's instrumentalist terms [see Boland 1979a], that the truth status of the model is less important than the usefulness of the results of its application. If one restricts econometrics to instrumentalist methodology, then there may be no need to separate internal inaccuracies introduced by specification errors from the external inaccuracies caused by observation errors. However, if the truth status of our economic theories and models does matter, then econometric modelling which does not treat observation and specification errors separately will not obviously be an appropriate tool for analysis. Stated another way, whenever the truth status of our theoretical and modelling assumptions is at issue, *the only acceptable means of accounting for errors is the recognition of external 'measurement errors'*. Moreover, when the truth status of a model matters, specification errors are always unacceptable and always imply a false *model* (but not necessarily a false theory as I noted in Chapter 7).

Nevertheless, the important point to be retained is that, since economic models are primarily concerned with explicit interrelationships between observable variables, whether the errors of observation are interconnected externally may be an important source of information about the data being used. This is obviously an important consideration for some typical macroeconomic data. For example, consider the standard textbook equation:

$$\mathbf{Y} = \mathbf{C} + \mathbf{I} + \mathbf{G}.$$

To the extent that observations of **C**, **I** and **G** may be the result of simple income accounting, we can say that whatever is not considered **C** or **G** can be accounted for by calling it **I**. In this elementary case, the observations are by construction interrelated. Assuming we are dealing with models for which such 'identities' have been eliminated, any information we have about errors of observation will not usually imply an interdependence and thus makes it all the more important to treat observational errors separately from specification errors, particularly when we wish to assess the possibility of specification errors.

3.3. Alternative views of testing with models

It might be tempting for some readers to confuse what I am calling 'combined simultaneous testing' with so-called 'non-nested hypothesis testing'. This would be a serious error. Models are said to be non-nested when 'they have separate parametric families and one model cannot be obtained from the others as a limiting process' [Pesaran 1974, p. 155]. One reason for avoiding the confusion is that any non-nested models being tested are *models* of competing theories, such that it is *possible* for all to be false [see Kennedy 1979/85, p. 70; MacKinnon 1983, p. 87] and given my arguments in Chapter 7, for nothing to be accomplished. Depending on test procedure, they all can be confirmed, too. In the case of combined simultaneous tests, the essential counter-example is determined solely on the basis of the single behavioural theory in question (i.e. a single list of exogenous and endogenous variables and thus a single parametric family). More important, the counter-example is not just any contrary theory, as in the case of non-nested models, it is a statement whose truth status is denied by the original behavioural theory. Thus, when models of the theory and its counter-example both fail or both fit the data using the same test procedure, at least we have demonstrated a shortcoming in the modelling method or in the logic of the original behavioural theory.

It is equally important to avoid confusing a combined simultaneous test with what is called an 'encompassing test' [see Hendry 1983]. Ideally, an encompassing test would show that one model of a theory is superior to another model of the same theory by showing that the modelling techniques used in the superior model allows for the explanation of (i.e. encompasses) the results obtained by the inferior model. Both models explain the same data but use different modelling assumptions. The reason for avoiding the confusion is simply that a counter-example is not a competing explanation and moreover, the same modelling assumptions are used to model the theory and the counter-example.

Since the same modelling assumptions are used in a combined simultaneous test and the counter-example is not a competing model or theory, the econometric perspective of either encompassing tests or tests of non-nested alternatives is not obviously a relevant basis for considering combined simultaneous testing.[3]

4. Asymmetries in tests based on stochastic models

Now, with stochasticism and the related aspects of econometrics put in their proper places, let me nevertheless accommodate stochastic models in my theory of convincing tests. The central question here is whether the recognition of *stochastic models* undermines my theory of convincing tests or, as I shall argue, actually emphasizes the need for combined simultaneous tests using counter-examples.

The key question that necessitates the recognition of stochastic models is still the acknowledgement that observation statements are seldom exactly true. Recall that my discussion in Sections 1 and 2 *assumed* that observation statements were (exactly) true. Given that assumption, whenever a model of the counter-example was said to fit the available data, we knew that the compound statement, which is formed of the counter-example plus modelling assumptions, was a true statement. Since the truth of the theory would deny the possibility of our ever building a relevant model of a counter-example of that theory which would fit the data, it was concluded that whenever the counter-example did fit, the theory must be false. Now what happens when the determination of a good fit is not exact (due to inaccuracies of the observations used to determine the fit)?

4.1. A simple example

Consider a simple one-equation model which represents the theory that the level of consumption (**C**) is a *linear* function of the level of national income (**Y**):

$$\mathbf{C} = a + b\mathbf{Y}$$

The question at issue will be whether the specification of a two-variable linear model represents the true relationship between **C** and **Y**. However, for the purposes of this elementary discussion, we will say we know that there are no other relevant variables so that the only issue is the linearity of the model. Let us say we have made three observations to determine if

3 Nevertheless, many aspects of non-nested hypothesis testing methodology may still apply to the combined simultaneous testing procedure.

the linear relationship holds. With two observations we could deduce values for *a* and *b* which are the same for both observations – that is, we solve the pair of simultaneous equations (one equation represents one observation):

$$\mathbf{C}_1 = a + b\mathbf{Y}_1$$
$$\mathbf{C}_2 = a + b\mathbf{Y}_2$$

The third observation is used to test the deduced values of *a* and *b*. The question is whether the calculated **C** which equals $(a + b\mathbf{Y}_3)$ also equals the observed $\mathbf{C}_3$? In Chapters 2 and 3 any difference between the calculated **C** and the observed **C** would both constitute a counter-example and be immediately interpreted as a refutation of the model. But that was primarily due to the assumption that observations were always exactly true.

Let us relax this assumption somewhat by saying the observations are not exact but, unlike ordinary econometric modelling, let us also say we have some independent knowledge of the possible observation errors. If we knew the observations could be wrong by no more than 10 percent, then our criterion for interpreting the third observation must accommodate errors of as much as 10 percent. But most important, if we allow for errors in the determination of whether the third observation constitutes a refutation of the linearity of the equation, we will run the risk of incorrectly claiming that the counter-example fits and thus falsely claiming a refutation of the theory. Similarly, we run the risk of incorrectly claiming that the third observation *confirms* the linearity assumption when in reality the relationship is non-linear. What needs to be appreciated when there are errors in observations is that failure to confirm may not constitute a confirmation of a counter-example. With errors in observations, both the theory and its counter-example could fail to be confirmed by the same observations whenever we make allowances for errors. This will depend on how we decide whether we have confirmation.

To illustrate this asymmetry, let us say that we can make two correct observations but all subsequent observations will be subject to errors of as much as (but no more than) 10 percent. For example, if we correctly observe that $\mathbf{C}_1 = 10$, $\mathbf{Y}_1 = 20$, $\mathbf{C}_2 = 12$ and $\mathbf{Y}_2 = 30$, then by assuming linearity we can deduce that $a = 6$ and $b = 0.2$. Now let us say that at the time of our third observation the (unknown) true value of $\mathbf{Y}_3$ is 40 but our third observation is inaccurate, so that we observe $\mathbf{Y}_3 = 44$. At the same time, we also observe that $\mathbf{C}_3 = 12.6$. Both observed variables are off by about 10 percent. If the true relationship is linear then the true value for $\mathbf{C}_3$ is 14 – but if the true relationship is non-linear, then the true value of $\mathbf{C}_3$ could differ from 14. Assuming linearity is true, our calculated **C** will be 14.8 which differs from the *observed* **C** by more than 17 percent, even though neither observation is more than 10 percent wrong. Depending on how we choose to interpret this, we might incorrectly conclude that **C** and **Y** are not linearly related when in reality they are.

For the sake of discussion, let us say we doubt that *both* observations would be off by as much as 10 percent so we will interpret a 17 percent calculated difference as a 'bad fit' with regard to our linearity assumption. However, a bad fit in this case does not mean that we have proven that the true model is non-linear. All that we have concluded is that the linearity assumption is not confirmed. For us to conclude that the linearity assumption is false we have to decide what would constitute a counter-example as well as a good fit for a counter-example.

4.2. Disconfirming vs non-confirming observations

In my example I said that it is known that observations could differ from the true values of **C** and **Y** by as much as 10 percent and thus when making our third observation, the calculated and observed values of **C** could be found to differ by as much as 17 percent without necessarily proving that the true relationship is non-linear. By recognizing that non-confirmations of linearity are not necessarily confirmations of non-linearity, it is always possible when adopting conservative test criteria based on single observations that both the theory (linearity) and the counter-example (non-linearity) will fail to be confirmed. Thus a test based on a single observation is not usually considered a very convincing test. This is so even though a single observation of, say, a 20 percent calculated difference in our simple example would constitute a refutation while a zero error does not constitute a proof that the relationship is linear, since the next observation might not be errorless.

Anything short of the maximum possible error in the calculated difference leaves the results of the test doubtful. Nevertheless, we may wish to interpret the test based on any notions we might have about the acceptability of errors. Specifically, we might think that a claim that linearity is confirmed based on a 17 percent allowable error is too risky. Even a 15 percent error might be considered too risky for a confirmation

of linearity. We might take the position that while a 15 percent error does not constitute a proof that the model is not linear, such an observation casts serious doubt on the model's linearity. Let us call this interpretation (of the observation) a 'disconfirmation' of the linear model. Similarly, an error of 5 percent may be too risky for a conclusion that the counter-example is confirmed and thereby that the assumption of linearity is definitely false. In this case, the observation may be interpreted as a disconfirmation of the counter-example.

It is important here not to confuse the disconfirmation of a theory with the confirmation of its counter-example. Equally important, we ought not confuse 'not disconfirmed' with a 'confirmation'. While a calculated difference greater than 18 percent may constitute a proof of non-linearity when we know the observations cannot be more than 10 percent wrong, using 18 percent as a test criteria seems too severe. So we need to choose a convenient standard to interpret the calculated difference.

On the one hand, if we are looking for a confirmation of the counter-example, we may wish to say that a calculated error of 15 percent is sufficient for us to conclude that the linearity assumption is false but an error of less than 10 percent is not sufficient, and thus the counter-example is not-confirmed. If we are looking for a disconfirmation of the counter-example, we might say an error of less than 5 percent leads to the conclusion that the counter-example is disconfirmed but an error over 10 percent leads us to declare the counter-example to be not-disconfirmed. On the other hand, a similar disparity can be created when we are directly assessing the linearity assumption. If we are looking for a confirmation of the linearity assumption, we may wish to say that a calculated error of less than 2 percent is sufficient for us to conclude that the linearity assumption is confirmed but an error of more than 10 percent is not sufficient, so that the linearity assumption is not-confirmed. If we are looking for a disconfirmation of the linearity assumption, we might say an error of more than 15 percent leads us to conclude that the linearity assumption is disconfirmed but an error between 5 and 10 percent leads us to declare the linearity assumption to be not-disconfirmed.

Here, of course, I am arbitrarily assigning numbers to the allowable or required criteria for the purposes of discussion. Any actual criteria will be decided on the basis of what we know about the nature of the observation errors and the nature of the actual theory being tested. As my simple example illustrates, it is easy to adopt very different criteria of rejection or of acceptance. I am using the words 'confirmed' and 'disconfirmed' rather than 'true' and 'false' to bring out the essential asymmetry. In the true-false case, 'not-true' means false and 'not-false' means true (so long as we do not deny the axiom of the excluded middle [see Boland 1979a, 1982]). Here, it should be clear that 'not-confirmed' does not necessarily mean disconfirmed and 'not-disconfirmed' does not mean confirmed whenever there is a wide range of possible errors.

In terms familiar to those who have read any elementary statistics book, we have to decide which is more important: avoiding the rejection of a true assumption or avoiding the acceptance of a false assumption. Statisticians refer to these as Type I and Type II errors. What criterion is used to define Type I or Type II errors is still a matter of judgement with a heavy element of arbitrariness. Selecting a criterion which makes it easier to avoid one type of error will usually make it easier to incur the other type of error. Furthermore, whether we use a criterion of 5 percent or 10 percent as allowable deviation in calculated values may be determined more by the economics of the situation than by one's philosophy of science. The cost of making one type of error based on a narrow range of 5 percent may be greater than the other type of error based on a range of 10 percent. When dealing with matters of social policy it may be considered safer to have low standards of accepting a false linearity assumption and high standards for rejecting a true linearity assumption. Since there is usually ample room for doubt, linear models are often easier to apply to practical problems. It all depends on what we are looking for or are willing to accept.

My distinction between disconfirmations and non-confirmations (or maybe even between confirmations and disconfirmations) may not be clear to those familiar only with the concept of hypothesis testing found in statistics textbooks. Once one has chosen to avoid, say, Type I error, then any failure to confirm the counter-example is automatically interpreted as a confirmation of the theory. Furthermore, exclusive concern for Type I error leads to the exclusive use of confirmation criteria. Concern for Type II error would have us use disconfirmation criteria instead. If for any reason we are unwilling to choose between Type I and Type II error, then we will need to be able to distinguish between disconfirmations and non-confirmations.

4.3. Confirmation vs disconfirmation test criteria

The possible asymmetry between confirmation and disconfirmation criteria needs to be seen against the background of the problems I have already discussed concerning the process of testing theories using models of those theories. Even when we considered observations to be without errors, we still could not expect to be able to refute a theory by refuting just one singular model of that theory. However, I did show in Section 2.2 that if we simultaneously test a model of a theory and a model of its counter-example, it is possible to say what would constitute a refutation of a theory. Specifically, a refutation would occur when the model of the theory fails to fit the data while the model of a counter-example does fit. Now, what is added to this by entertaining the possibility of observational errors?

If we were to base our combined simultaneous test on a single observation of the theory and a coincident observation of its counter-example, we would be wise to adopt rather conservative criteria of acceptance or rejection – maybe, as in my simple example, something like 2 percent for confirming observations of linearity vs 15 percent for a confirmation of an observation of the counter-example. The difficulty here is that a single observation test is one-dimensional. It is necessary then to distinguish between a 'confirming observation' and a 'confirmation' which may require many confirming observations. Similarly, a 'disconfirming observation' is distinguished from a 'disconfirmation' which may require many disconfirming observations.

Since observation errors are possible and we might not wish to jump to a conclusion on a single observation, let us now repeat the third observation (of **C** and **Y**) 19 more times. This new dimension (the number of observations) raises a new question for decision: how many non-confirming observations will we allow in a confirmation? No more than 1 in 20? Maybe 2 in 20? Given that observation errors are possible, let us consider alternative postures concerning how to interpret the results of 20 observations.[4] Our test criteria and our posture must, of

[4] Note that requiring a minimum of 20 stochastic observations to play the same role of one non-stochastic observation means that a stochastic version of a non-stochastic model (such as one from Chapter 3) which has a P-dimension of, say, 30 would now have an effective P-dimension of 600. This means a model for which it would have taken at least a year to construct a refutation would now require at least 20 years to refute!

course, be decided before making the observations if we wish to avoid unnecessary skepticism. The following represent four different and illustrative postures that employ only confirmation/non-confirmation criteria for the assessment of observations:

(1) We might say that whenever 5 or more of the 20 observations are convincing *confirming observations* of linearity (no more than 2 percent calculated difference, as discussed in Section 4.1) we will conclude that the linear model is *confirmed*, otherwise it is *not confirmed*.

(2) We might say that whenever 5 or more of the 20 observations are convincing *confirming observations* of non-linearity (at least 15 percent calculated difference) we will conclude that a model of a counter-example of the linear model is *confirmed*, otherwise it is *not confirmed*.

(3) We might say that whenever 5 or more of the 20 observations are not convincing *confirming observations* of linearity (more than 2 percent calculated difference) we will conclude that the linear model is *disconfirmed*, otherwise it is *not disconfirmed*.

(4) We might say that whenever 5 or more of the 20 observations are not convincing *confirming observations* of non-linearity (less than 15 percent calculated difference) we will conclude that a counter-example of the linear model is *disconfirmed*, otherwise it is *not disconfirmed*.

Given that our criteria for convincing observations might be considered extreme (2 percent or less in one case and at least 15 in the other), it may be reasonable not to expect a large proportion of observations to be meeting either criterion. Thus, we have an asymmetry between the confirmation of a counter-example and a disconfirmation of the theory itself. Even though we have employed a confirmation/non-confirmation criterion (to assess observations), in order to define the four interpretation postures we still need to decide whether we are more interested in finding disconfirmations or confirmations – although there may not be any non-arbitrary way to do so.

Let me illustrate the possible consequences of all this for combined simultaneous tests of the model of a theory and a model of its counter-example. If we recognize that 'not-confirmed' does not imply disconfirmed, then to illustrate the possible outcome, depending on whether we are looking for confirmations or disconfirmations as defined in statements (1) to (4) above, we need two tables. In Table 2, the presumption is that the socially acceptable testing conventions only identify a confirmation, as in the case of desiring to avoid Type I errors. And in Table 3 it is presumed that only disconfirmations are identified (avoiding Type II errors).

Confirmation-based test model of:		
Theory	*Counter-example*	TEST RESULT
confirmed	confirmed	inconclusive
confirmed	not-confirmed	weak conditional corroboration
not-confirmed	confirmed	conditional refutation
not-confirmed	not-confirmed	inconclusive

Table 2

Disconfirmation-based test model of:		
Theory	*Counter-example*	TEST RESULT
not-disconfirmed	not-disconfirmed	inconclusive
not-disconfirmed	disconfirmed	conditional corroboration
disconfirmed	not-disconfirmed	weak conditional refutation
disconfirmed	disconfirmed	inconclusive

Table 3

In Tables 2 and 3, I have noted that all corroborations or refutations must be considered conditional. The condition is that the interpretation of the result is always dependent on the acceptance of the specified test criteria used. In the case of my simple example above, the criteria involve the possibly extreme limit of a 2 percent acceptable error between the calculated and observed **C**. Other criteria are possible such as limiting the ratio of acceptable to unacceptable errors in the given number of observations made. In both tables the inconclusive results may cause one to question the test criteria in a single equation model. In multiple equation models inconclusive results might also suggest that the model could be either incomplete or inconsistent.

As long as one is willing (a) to not demand unconditional refutations, (b) to adopt standard views of testing and thus commit oneself to which type of error (I or II) to avoid, and (c) to commit oneself to use either a confirmation or a disconfirmation criterion for the evaluation of observations, then I think by making all tests of a theory combined simultaneous tests of the model of the theory and of at least one counter-example to the theory, refutations in economics are in principle possible, albeit conditional.[5]

4.4. The irrelevance of the probability approach

So far, I have not said anything about probabilities. Many readers will find this irritating because they think probabilities are essential for a discussion of conclusions drawn from inaccurate observations, that is, from stochastic models. While the probability approach to economics [e.g. Haavelmo 1941/44] may appear to solve some of these problems, it too often masks from view the logical structure that defines the methodological problem at hand. If we wish to discuss things in probability terms then, instead of saying that errors of observation could be as much as 10 percent, we could *assume* that when we repeatedly make the third observation, the possible errors for this observation will be distributed in a manner we associate with the Gaussian 'normal distribution' (i.e. the bell-shaped curve). If we also assume that the average value of the observation is the true observation, then the formal mathematical properties of such a curve can be used to calculate the probability that the observations will be incorrect in more than, say, 5 percent of the observations. In doing so, we facilitate a calculation of potential damage done by incorrectly accepting a fit. If we have no reason to assume that errors are normally distributed or if we know something about the observation process independent of both the model and the testing process, then some probability presumptions may be a major source of difficulty. I suspect that the primary reason for promoting probabilistic approaches to economics is that they provide a basis for formalizing the arbitrary decisions regarding the choice of confirmation or disconfirmation criteria. Some people simply feel better when necessary arbitrariness is formalized.

Nevertheless, for some practical problems where the assessment of benefits and costs are unproblematic (and there is no independent means of assessing the accuracy of observations), characterizing the occurrence of errors to be governed by a given probability distribution can be very helpful. But if we do not know how the errors are distributed, more questions may be begged than are answered [cf. Swamy, Conway, and von zur Muehlen 1985]. I think for the purposes of understanding the difficulties in forming conclusions when there are errors in observations, it is better not to confuse stochastic models with probabilistic models. As I have attempted to show in this section, the problems and means of avoiding the Ambiguity of Direct Model Refutation *do not require* a probabilistic approach to testing stochastic models.

5. 'Normative' vs 'positive' methodology

While the many critics of Popper's philosophy of science or of Samuelson's methodology may have been encouraged by Chapter 7 and its pessimistic look at the possibility of refuting a theory, it is unlikely that those critics will be very pleased with my conclusion that refutations are possible whenever we agree on the testing conventions and test both the model of the theory and a model of its counter-example. My purpose here is not, however, to defend Popper or Samuelson but merely to represent what I think is a viable interpretation of what is possible, given what ordinary economic model builders seem willing to accept in terms of testing conventions or criteria.

Perhaps, as the critics charge, mainstream economists ought to be attempting to refute neoclassical economic theory. However, on the basis

[5] Some of my students have made elementary applications of this approach to testing in economics. What is most striking from their applications, where they have repeated previously reported tests of mainstream economic theories, is that in almost every case the reported results do not correspond to the decisive categories but to the inconclusive results. For another explanation of this approach to testing and how it can be used, see Bennett [1981]; Jensen, Kamath and Bennett [1987].

of Chapter 7 we can certainly understand why economists would not waste their time attempting to build models of an economic theory merely to refute it. Economists ought not to be scolded for not doing impossible tasks. Many model builders in economics will see themselves, nevertheless, engaged in an ongoing programme of inventing and refuting an endless series of specific models. It is important for methodologists, who claim to be explaining what mainstream economists are actually doing, to attempt to construct a view of testing that corresponds to what is actually practiced. In this chapter I have presented such a view of testing in economics, but I have not stopped there. I have offered a view of testing that overcomes the obstacles to direct tests of economic theories without requiring any substantial changes in the socially accepted testing conventions currently used by practicing model builders in economics.

EPILOGUE

Methodology after Samuelson: Lessons for Methodologists

> **Foundations of Economic Analysis** had successes in generating a wide variety of substantive theories. But what interested its young author most … was the success it could achieve in formulating a general theory of economic theories.
>
> Paul A. Samuelson [1983, p. xxvii]

> a good deal of [the Conference's] time was devoted to methodological discussions; Professor Popper cast a long shadow over our proceedings! This was regrettable since most of the papers had something to teach us while the methodological arguments had not. We had all been through them since undergraduate days … these things are as much a matter of temperament and what one likes doing as they are of philosophy of which most of us are pretty ignorant. I simply record that, in my view, the Conference would have been better than it was if we had spent more time on what people were saying in their papers than on what they ought to have been saying.
>
> Frank H. Hahn [1965a, p. xi]

In this closing chapter I wish to share with methodologists some of the lessons I have learned from a research programme in applied methodology that has spanned more than twenty years. What I have learned from my specific research programme (concerning how apparently innocuous modelling assumptions and techniques can affect the testability of one's model) should be evident in the previous eight chapters. So, in this chapter I wish to discuss what I learned about the methodology of economics in general with special emphasis given to Samuelson's impact on the field of economic methodology.

Often I have taken the opportunity to point out that Paul Samuelson's book, **Foundations of Economic Analysis**, was intentionally a methodology book. Here I wish also to point out that it represented very good methodology, despite its being an operationalist version of what is often alleged to be Karl Popper's falsificationist methodology.

Unlike many of today's methodologists, Professor Samuelson did more than just talk about his methodology. Like the good methodologist that he is, Samuelson proceeded to show how to implement his methodology. In this sense, I think we can all agree that Samuelson has been a most successful methodologist. Here I shall examine the consequences of Samuelson's impressive success as a practicing methodologist for the profession of economic methodologists.

Samuelson repeatedly noted that his book demonstrates how the development of economic theory is intimately intertwined with successful efforts to meet one methodological goal – namely, theories progress by creating 'operationally meaningful propositions'. Explaining the development of economic theory is always an intellectual juggling act. There are two balls to juggle – one's theory of economic phenomena and one's theory of economic theories. Theorists as methodologists are always trying to keep these two balls in the air. And as with all juggling acts, if we concentrate on just one ball, the other ball will usually come crashing down.

Now, I wish to extend this two-ball approach to recognize explicitly that methodology as a 'general theory of economic theories' always involves both rhetoric and the sociology of science. It involves both because they are both necessary and because rhetoric and sociology should not be considered separately. Rather than a mere two-ball juggling act, we will need to consider a four-ball juggling approach to economic methodology. Again, the important point to keep in mind is that whenever we concentrate on one ball at the expense of the others, the other balls will surely be dropped.

1. The interdependence of sociology and rhetoric

There are some methodologists today who argue that we should concentrate exclusively on rhetoric of economics (e.g. Arjo Klamer and followers of Donald McCloskey) while others argue that we should concentrate on the sociology of economics (e.g. Bob Coats). I think both of these arguments are misleading from the perspective of methodology since sociology and rhetoric are interdependent. Moreover, both are implicitly dealt with in any methodological study. Before analyzing the juggling skills of any good methodologist, let me examine what I think is the necessary interdependence of sociology and rhetoric.

The well-known Canadian communications theorist, Marshall McLuhan, is famous for pointing out that *how* we say something is usually more informative than *what* we say. I think this is true for both economic methodology and economic methodologists. To understand McLuhan's observation and its relevance for economic methodology, I wish to consider the following question: How can the method of presentation by itself ever be informative? The answer which I wish to defend here is as follows: The message of one's statement or argument is dictated to a great extent by the medium of its presentation.

1.1. The medium IS the message

The effectiveness of any method of presenting a statement depends profoundly on the nature of the audience to whom the statement is directed. Some methods are better than others. If the audience is a convention of religious fundamentalists, quoting from the approved scripture will be essential. If the audience is a meeting of mathematical economists, one must put statements in the form of 'propositions', 'lemmas' and the other paraphernalia of mathematical formalism. If the audience is a gathering of …

Well, I think you get the idea. What you say will be considered 'informative' only when it is properly stated. What is proper is not a matter for free choice since it is dictated by the tastes of the intended audience.

Rhetoric is the study of what it actually takes to convince a given audience. That is, rhetoric is concerned with the requirements of 'effectiveness'. If every audience were made up of independently thinking individuals – ones like those studied by neoclassical economists – then perhaps rhetoric would be a simple matter of psychology [see Boland 1982]. Unfortunately, only one out of twenty people are independent thinkers [Wilson 1963, p. 318]. This is not a matter of psychology but a recognition that our educational system does not promote independent thinking. When I say that most people are not independent thinkers, I am merely saying that what most people think depends on what other people think. Such an interdependence is directly a matter of sociology since whenever one studies rhetoric, one is implicitly (if not explicitly) studying sociology.

Sociology is, among other things, a study of the basis for interdependent decision-making. That is, it studies how one individual's decisions depend on expectations of what other people will do and expect [see Boland 1979b]. Social institutions and conventions provide a convenient basis for forming such expectations. When approaching an audience of economists, for example, we take for granted that they know the contents of a standard principles textbook and thus expect them to know what a demand curve is. A more obvious example is the requirement that one must understand the language which the audience speaks if one ever hopes to get a point across. But rhetoric is more than a matter of language. Successful rhetoric involves using the conventional 'truths' to build convincing arguments. Institutions and published norms of social behaviour are important sources of the information needed to make everyday decisions. Interdependence of decision-makers at some level must involve the method of how one decision-maker is convinced by the actions or arguments of another decision-maker. Thus the study of sociology always involves the study of rhetoric. Despite what some neoclassical economists want us to believe, information that is only about prices is never enough – but I will leave this digression for a future study.

1.2. Methodology as a juggling act

Recall that Samuelson, the methodologist, said that he wished to be successful 'in formulating a general theory of economic theories' – that is, in explaining what economic theorists do. What I have been saying is that to be successful, a methodologist must be a skillful juggler. The methodologist not only must (1) understand the logic of explanation, but must (2) understand economic theory, (3) understand what the audience for that explanation knows or takes for granted, and (4) know what it would take to convince that audience. This is no easy juggling act because these four requirements must be met in a logically consistent manner. Moreover, any methodologist who tries to deal with these requirements by keeping them separate – perhaps to reap the benefits of a division of labour – is not likely to be successful or appreciated by anyone.

Let us consider the application of a division of labour to methodology. Contrary to what we teach in our principles courses, a division of labour

does not always lead to an optimum. Nevertheless, it is too tempting to think that each of the essential elements of methodology would benefit from the expertise of specialists. For example, it is quite common for methodologists today to consult a philosophy of science textbook to obtain a solid foundation for an understanding of the logic of explanation. Unfortunately, this approach to understanding falsely presumes a solid unanimity among philosophers of science. Even so, some young economic methodologists think that a spectacular demonstration of an understanding of one particular philosophy of science is all that it takes to be a methodologist. Let us leave aside the question of how boring it would be for a juggler to juggle just one object. For us the question is, who is the audience for such a demonstration? This is an important question since, as I said above, the sociology and rhetoric of that audience will dictate the appropriate method and substance for one's demonstration.

eSamuelson may not be immune from this criticism of misplaced specialization. Samuelson seems to suggest that he was merely implementing an 'operationalist' methodology, namely the one which says that an acceptable explanation of a phenomenon requires us to specify how to measure it. If Samuelson really was implementing such a philosophically prescribed methodology, as he seemed to suggest, then Joan Robinson's criticism challenging a unique measure of capital would have destroyed Samuelson's version of economics. But her critique failed to convince everyone simply because Samuelson was going beyond the philosophers of science rather than blindly following them. More important, the audience to which Samuelson directed his **Foundations of Economic Analysis** did not understand the philosophy of operationalism sufficiently to comprehend the substance of the debate and thus were unable to appreciate his methodology or the logic of Mrs Robinson's critique. Unfortunately, his audience still is convinced more by form than substance and continue to think he won but they seem unable to explain why.

2. History and reality

I would like to continue to illustrate the pitfalls of relying on specialists in such other things as rhetoric, sociology or mathematics but I must move along. The time has come for me to attempt to practice what I am preaching in this closing chapter, so, I turn to the sociology of the economics profession with an eye on its rhetoric, paying particular attention to the role of methodology and the methodologist. We need to examine the evolution of the sociology of our profession because it contributed more to Samuelson's success as a methodologist than did the veracity of his particular opinions about methodology.

To begin let me acknowledge what I think is the dominant empirical feature of the sociology of our profession since about 1960. Most mainstream economists are convinced that methodological discussions are a waste of time [e.g. see Hahn 1965a]. Of course, mainstream economists are probably right but let us suspend our judgement and first try to understand this reality. The key word is 'mainstream' and the question is: Why is methodology no longer a part of the profession's mainstream? I think this is an important question and, to illustrate, let me ask a couple more.

The first question is: How many 'top ten' North American universities' economics programmes are represented at the typical History of Economics Society meetings (the only conference in North America where methodology is openly and regularly discussed)? To answer this question I consulted the official list of participants for the 1986 meetings of this august society. I could not find even one person from Harvard, Stanford, MIT, UCLA, Yale, Princeton, Pennsylvania, Berkeley or Chicago. And just to add some perspective to this question, I looked up each member of MIT's faculty to see where they obtained their PhDs and of the thirty-four listed in the **American Economic Association Handbook**, ten were from Harvard, nine from MIT, and one or two from each of Princeton, Chicago, Pennsylvania and Yale. The only school represented in both lists is Columbia which is understandable since that was the location of the 1986 History of Economics Society meetings. For all practical purposes one can conclude that there is nothing in common between the interests of the members of the History of Economics Society who promote the study of economic methodology and the professors in the mainstream of the economics profession.

The second question is: How many methodology papers were published in the top economics journals between 1967, the publication year for the last major contribution [i.e. Bear and Orr 1967] to the methodological discussion initiated by Samuelson's famous critique of Friedman's methodology, and 1979, the year that I, with the help of Mark Perlman, stirred up the same hornet's nest with my infamous **Journal of Economic Literature** article? If by 'top journals' we mean the **American Economic Review** or the **Journal of Political Economy**, then the answer is *two* articles. The only articles that stand out from that period are the 1971 **Journal of Political Economy** article by Louis De Alessi and the 1973 **American Economic Review** article by Stanley Wong. The important point here is that for twelve years methodology was virtually banned from the pages of the leading journals. There was a brief period of about five years when Robert Clower allowed methodologists some limited access to the **American Economic Review**, but unfortunately the door seems to be closed again.

So, the sociological facts of our economics profession are that, with the possible exception of Samuelson, there are no methodologists in the top mainstream economics departments and with the exception of a brief moment in the sun during the early 1980s, there has been virtually no methodology in the leading journals since Samuelson's final word on 'theory and realism' [Samuelson 1965]. I claim that if one wishes to be successful as a professional methodologist today, then one must somehow understand and overcome these sociological facts.

2.1. Understanding the sociology of the economics profession

Since we are all products of the profession we wish to study, we should be able to draw upon our own experiences. My graduate education, or should I say training, in economics was exclusively in mathematical economics. This was inevitable since I came to economics from an undergraduate programme in mechanical engineering. I point this out because it is just this type of educational programming that systematically rules out any consideration of the philosophical questions inherent in the study of methodology. When I announced to my undergraduate teachers that I wished to study methodology they patiently explained to me the error of my ways. So off I went to study mathematical economics since I supposedly had the requisite engineering mentality to do so.

As we all know, what was once just a special area in economics has since grown to dominate all of economics. Today, virtually all graduates of the leading mainstream schools are well-trained economics engineers. Their teachers send them off into the world to seek out 'operationally meaningful' (i.e. refutable) economics propositions. This normative prescription is their only concession to doing something other than strict applied economics. What should be apparent here is that mainstream economics today is the reification and embodiment of nothing other than

Samuelson's **Foundations**. The pre-eminent problem facing someone who wishes to study methodology is not that methodology has been banned from the workplace but rather that there is only one methodology. Methodology is no longer an interesting research topic for the mainstream simply because the choice has been made and there is nothing more to argue.

Why has Samuelson's methodology been so successful in dominating the mainstream profession's methodological choices? Of course, this will have to be answered by some sort of revealed preference analysis. I think the overwhelming aspect of Samuelson's **Foundations** and its embodied methodology is his implicit rejection of any need to appeal to the authority of philosophers. I applaud his anti-authoritarianism. However, cynics might wish to point out that he rejects philosophy only because he wished us to appeal to the authority of mathematicians, but let us leave that touchy issue aside for now. If we keep our eyes wide open, I think it is easy to see that the methodological motto over the door of every mainstream economics department is:

WHEN IT COMES TO METHODOLOGY, TALK IS CHEAP,
SO LET YOUR ACTIONS DO THE TALKING.

In small print just under this motto is the additional normative proscription: PHILOSOPHERS SHOULD MIND THEIR OWN BUSINESS.

2.2. Overcoming the sociology of the profession

Clearly, if I am correct about why Samuelson's methodology is so dominating, would-be methodologists will have to avoid invoking the authority of philosophers if they want the study of methodology to be respected and accepted in the mainstream. By this statement, I am not trying to suggest that the mainstream has made a correct methodological choice – I am only trying to be realistic about what it takes to be successful in the mainstream of economics.

An obvious alternative to my suggestion is for methodologists to group together into a special-interest group or subdiscipline much as the mathematical economists did in the 1930s. Perhaps we can convince one of the leading journals to devote part of their journal space to us or maybe we can even have our own journal. This way, it might be argued, methodologists can get their papers published and thus they can get tenure like the mainstream economists do. While this surely is a more enjoyable way to spend one's time as a methodologist than continually banging one's head on the doors of mainstream departments, I am still not optimistic about its possibility of success. There are two reasons for this. One concerns the question of the inappropriateness of the division of labour which I have already discussed. The other reason concerns the obvious need to demonstrate standards of scholarship that invite respect.

Those methodologists who promote methodology as a separate subdiscipline run the risk of suggesting that one can successfully study methodology as a topic separate from the rest of economics. The worst of this danger is the temptation to invoke the perceived standards of the philosophy profession. There is no reason in the world to think that a philosopher is a better judge of what is appropriate to convince an audience of economists than the economists themselves. As I have been arguing, the audience matters most. Besides, some of the concerns of philosophers are silly by anyone's standards. But most important, if we surrender to the temptation to use ready-made arguments from philosophy books, we turn away from the lesson to be learned from Samuelson's success. One of the major reasons why Samuelson's methodology was so successful is that he openly rejected the usual type of methodology [see Metzler 1948]. Actually, what he rejected was professional philosophy and its authoritarian and prescriptive tendencies.

3. Lessons for would-be methodologists

Many of us think that the methodological issues embodied in neoclassical economic theory are not dead issues and are still worthy of further discussion and criticism. The key question is: How are professional methodologists ever going to survive and prosper when so many mainstream economists think methodological discussion is a waste of time? I have been arguing that to answer this question we need to study the history of Samuelson's success as a methodologist rather than promote philosophers such as Popper. Moreover, the lessons we learn from Samuelson's success we must never violate. Let me now summarize the lessons I think I have learned.

Lesson 1: Hahn is typical

The opinions of Frank Hahn, quoted at the beginning of this chapter, are typical of mainstream economics. While it might be possible at the department of economics of Podunk University to find a receptive audience which will delight in hearing a paper offering a spirited rendition of the maxims of your favourite philosopher of science, hardly any mainstream economist will read anything more than its title. No matter how well your paper is written, no major journal will waste its time or funds having it reviewed.

Lesson 2: Cookbook methodology is unappetizing

Mainstream economists react very negatively to papers which offer cookbook recipes for ‘proper’ and ‘improper’ scientific methods. Such papers turn the average economist off because they involve preaching to economists that they must view economic methodology in accordance with the author’s favourite philosopher of science. In my student days Karl Popper was the object of worship. Today the fad is Imre Lakatos. Maybe tomorrow it will be Ian Hacking. It will not matter who the current hero or heroine is, mainstream economists will not be interested.

Lesson 3: Methodology does not always matter

What these negative lessons tell us is that we cannot presume that there is an automatic audience for philosophy of economics or for any bag of methodological judgements. What can we conclude that might have some positive flavour? From the history of Samuelson’s success I think we can conclude that actions speak louder than words. Rather than extolling the virtues of rhetorical methods or literary criticism, demonstrate how the mainstream economist can learn from such methodology. Rather than extolling the virtues of sociological analysis, demonstrate how the mainstream economist can benefit from such analysis. Rather than extolling the virtues of the philosophy of science of Hacking or Lakatos, demonstrate for the mainstream economists why such a discussion will ever matter.

Samuelson’s success was made possible by his ability to demonstrate how his view of methodology matters in the development of economic theory. What impressed his audience was not his dazzling display of philosophical cleverness, but rather the demonstration of his thorough understanding of economic theory. His success amounts to an outstanding juggling act. He demonstrated how his view of methodology is inseparable from his understanding of economic theory. And, as I have claimed, his success depended on his clear understanding of both the sociology of the economics profession and its rhetoric. I think the audience was very impressed with his skillful juggling act. Moreover, their preferences have been clearly revealed for all would-be methodologists to see.

Bibliography

Abramovitz, M. (ed.) [1959] **Allocation of Economic Resources** (Stanford: Stanford University Press)

Agassi, J. [1959] ‘Corroboration versus induction’ **British Journal for the Philosophy of Science, 36**, 311–17

Agassi, J. [1961] ‘The role of corroboration in Popper’s methodology’ **Australian Journal of Philosophy, 34**, 82–91

Agassi, J. [1963] **Towards an Historiography of Science, History and Theory, Beiheft 2** (The Hague: Mouton)

Agassi, J. [1966a] ‘Sensationalism’ **Mind, 75**, 1–24

Agassi, J. [1966b] ‘The confusion between science and technology in the standard philosophies of science’ **Technology and Culture, 7**, 348–66

Agassi, J. [1967] ‘Planning for success: a reply to Professor Wisdom’ **Technology and Culture, 8**, 78–81

Agassi, J. [1971a] ‘Tautology and testability in economics’ **Philosophy of Social Science, 1**, 49–63

Agassi, J. [1971b] ‘Tristram Shandy, Pierre Menard, and all that’ **Inquiry, 14**, 152–81

Alchian, A. and W. Allen [1967] **University Economics** (Belmont: Wadsworth)

Allen, W. [1977] ‘Economics, economists, and economic policy: modern American experiences’ **History of Political Economy, 9**, 48–88

Archibald, G.C. [1960] ‘Testing marginal productivity theory’ **Review of Economic Studies, 27**, 210–13

Archibald, G.C. [1961] ‘Chamberlin versus Chicago’ **Review of Economic Studies, 29**, 1–28

Archibald, G.C. [1966] ‘Refutation or comparison’ **British Journal for the Philosophy of Science, 17**, 279–96

Arrow, K. [1951/63] **Social Choice and Individual Values** (New York: Wiley)

Arrow, K. [1959] ‘Towards a theory of price adjustment’ in Abramovitz [1959], 41–51

Arrow, K. [1968] ‘Mathematical models in the social sciences’ in Brodbeck [1968], 635–67

Arrow, K. and G. Debreu [1954] ‘Existence of an equilibrium for a competitive economy’ **Econometrica, 22**, 265–90

Arrow, K. and F. Hahn [1971] **General Competitive Analysis** (San Francisco: Holden-Day)

Bartley, W.W. [1964a] **The Retreat to Commitment** (London: Chatto & Windus)

Bartley, W.W. [1964b] ‘Rationality vs the Theory of Rationality’ in Bunge [1964], 3–31

Bartley, W.W. [1968] ‘Theories of demarcation between science and metaphysics’ in Lakatos and Musgrave [1968], 40–64

Bator, F.M. [1957] 'The simple analytics of welfare maximization' **American Economic Review, 47**, 22–59
Bear, D. and D. Orr [1967] 'Logic and expediency in economic theorizing' **Journal of Political Economy, 75**, 188–96
Bell, R. [1912] **A Treatise on the Coordinate Geometry of Three Dimensions** (London: Macmillan)
Bennett, R. [1981] **An Empirical Test of some Post-Keynesian Income Distribution Theories**, PhD thesis, Simon Fraser University, Burnaby, B.C.
Blanché, R. [1965] **Axiomatics** (New York: The Free Press of Glencoe)
Blaug, M. [1968] **Economic Theory in Retrospect** 2nd edn (Homewood: Irwin)
Blaug, M. [1978] **Economic Theory in Retrospect** 3rd edn (Cambridge: Cambridge University Press)
Blaug, M. [1980] **The Methodology of Economics** (Cambridge: Cambridge University Press)
Bohm, P. [1967] 'On the theory of "second best"' **Review of Economic Studies, 34**, 81
Boland, L. [1966] **On the Methodology of Economic Model Building**, PhD thesis, University of Illinois, Urbana, Ill.
Boland, L. [1968] 'The identification problem and the validity of economic models' **South African Journal of Economics, 36**, 236–40
Boland, L. [1969] 'Economic understanding and understanding economics' **South African Journal of Economics, 37**, 144–60
Boland, L. [1970a] 'Conventionalism and economic theory' **Philosophy of Science, 37**, 239–48
Boland, L. [1970b] 'Axiomatic analysis and economic understanding' **Australian Economic Papers, 9**, 62–75
Boland, L. [1971] 'Methodology as an exercise in economic analysis' **Philosophy of Science, 38**, 105–17
Boland, L. [1974] 'Lexicographic orderings, multiple criteria, and "ad hocery"' **Australian Economic Papers, 13**, 152–7
Boland, L. [1977a] 'Testability in economic science' **South African Journal of Economics, 45**, 93–105
Boland, L. [1977b] 'Time, testability, and equilibrium stability' **Atlantic Economic Journal, 5**, 39–47
Boland, L. [1977c] 'Model specifications and stochasticism in economic methodology' **South African Journal of Economics, 45**, 182–9
Boland, L. [1977d] 'Giffen goods, market prices and testability' **Australian Economic Papers, 16**, 72–85
Boland, L. [1979a] 'A critique of Friedman's critics' **Journal of Economic Literature, 17**, 503–22
Boland, L. [1979b] 'Knowledge and the role of institutions in economic theory' **Journal of Economic Issues, 8**, 957–72
Boland, L. [1980] 'Friedman's methodology vs. conventional empiricism: a reply to Rotwein' **Journal of Economic Literature, 18**, 1555–7
Boland, L. [1981a] 'Satisficing in methodology: a reply to Rendigs Fels' **Journal of Economic Literature, 19**, 84–6
Boland, L. [1981b] 'On the futility of criticizing the neoclassical maximization hypothesis' **American Economic Review, 71**, 1031–6
Boland, L. [1982] **Foundations of Economic Method** (London: Geo. Allen & Unwin)
Boland, L. [1983] 'Reply to Caldwell' **American Economic Review, 73**, 828–30
Boland, L. [1984] 'Methodology: reply' **American Economic Review, 74**, 795–7
Boland, L. [1986] **Methodology for a New Microeconomics** (Boston: Allen & Unwin)
Brems, H. [1959] **Output, Employment, Capital and Growth: A Quantitative Analysis** (New York: Harper)
Briefs, H. [1960] **Three Views of Method in Economics** (Washington: Georgetown University Press)
Brodbeck, M. (ed.) [1968] **Readings in the Philosophy of the Social Sciences** (London: Collier-Macmillan)
Bronfenbrenner, M. [1966] 'A "middlebrow" introduction to economic methodology' in Krupp [1966], 4–24
Brown, K. (ed.) [1965] **Hobbes' Studies** (Oxford: Basil Blackwell)
Bunge, M. (ed.) [1964] **The Critical Approach in Science and Philosophy** (London: Collier-Macmillan)
Caldwell, B. [1982] **Beyond Positivism** (London: Geo. Allen & Unwin)
Chamberlin, E. [1933] **The Theory of Monopolistic Competition** (Cambridge: Harvard University Press)
Charlesworth, J. (ed.) [1963] **Mathematics and the Social Sciences** (Philadelphia: American Academy of Political and Social Sciences)
Clower, R. [1959] 'Some theory of an ignorant monopolist' **Economic Journal, 69**, 705–16
Clower, R. and J. Due [1972] **Microeconomics** (Homewood: Irwin)
De Alessi, L. [1965] 'Economic theory as a language' **Quarterly Journal of Economics, 19**, 472–7
De Alessi, L. [1971] 'Reversals of assumptions and implications' **Journal of Political Economy, 79**, 867–77
Debreu, G. [1959] **Theory of Value: An Axiomatic Analysis of Economic Equilibrium** (New York: Wiley)
Dorfman, R., P. Samuelson and R. Solow [1958] **Linear Programming and Economic Analysis** (New York: McGraw-Hill)
Duhem, P. [1906/62] **The Aim and Structure of Physical Theory** (New York: Atheneum)
Eddington, A. [1958] **Philosophy of Physical Science** (Ann Arbor: University of Michigan Press)
Einstein, A. [1950] **Out of My Later Years** (New York: The Wisdom Library)
Einstein, A. and L. Infeld [1938/61] **The Evolution of Physics: The Growth of Ideas from Early Concepts to Relativity and Quanta** (New York: Simon & Schuster)
Ellis, H. (ed.) [1948] **A Survey of Contemporary Economics** (Homewood: Irwin)

Enthoven, A. [1963] 'Economic analysis of the Department of Defense' **American Economic Review, Papers and Proceedings, 53**, 413–23
Ferguson, C. [1972] **Microeconomic Theory**, 3rd edn (Homewood: Irwin)
Finger, J. [1971] 'Is equilibrium an operational concept?' **Economic Journal, 81**, 609–12
Fisher, F. [1966] **The Identification Problem in Econometrics** (New York: McGraw-Hill)
Friedman, M. [1953] 'The methodology of positive economics' in **Essays in Positive Economics** (Chicago: University of Chicago Press), 3–43
Gale, D. [1955] 'The law of supply and demand' **Mathematica Scandinavica, 3**, 155–69
Georgescu-Roegen, N. [1954] 'Choice and revealed preference' **Southern Economic Journal, 21**, 119–30
Georgescu-Roegen, N. [1966] **Analytical Economics: Issues and Problems** (Cambridge: Harvard University Press)
Georgescu-Roegen, N. [1971] **The Entropy Law and the Economic Process** (Cambridge: Harvard University Press)
Goldberger, A. [1964] **Econometric Theory** (New York: Wiley)
Gordon, D. [1955a] 'Operational propositions in economic theory' **Journal of Political Economy, 63**, 150–62
Gordon, D. [1955b] 'Professor Samuelson on operationalism in economic theory' **Quarterly Journal of Economics, 63**, 305–10
Grubel, H. and L. Boland [1986] 'On the efficient use of mathematics in economics: some theory, facts and results of an opinion survey' **Kyklos, 39**, 419–42
Haavelmo, T. [1941/44] 'Probability approach to econometrics' **Econometrica, 12**, Supplement
Haavelmo, T. [1943] 'The statistical implication of a system of simultaneous equations' **Econometrica, 11**, 1–12
Hahn, F. [1965a] 'Introduction' in Hahn and Brechling [1965], xi–xv
Hahn, F. [1965b] 'On some problems of proving the existence of an equilibrium in a monetary economy' in Hahn and Brechling [1965], 126–35
Hahn, F. [1973] **On the Notion of Equilibrium in Economics** (Cambridge: Cambridge University Press)
Hahn, F. and F.P.R. Brechling (eds) [1965] **The Theory of Interest Rates: Proceedings of a Conference Held by the International Economics Association** (London: Macmillan)
Hansen, B. [1970] **A Survey of General Equilibrium Systems** (New York: McGraw-Hill)
Hendry, D. [1983] 'Comment' (on MacKinnon [1983]) **Econometric Reviews, 2**, 111–14
Hicks, J. [1939/46] **Value and Capital** (Oxford: Clarendon Press)
Hicks, J. [1956] **A Revision of Demand Theory** (Oxford: Clarendon Press)
Hicks, J. [1979] **Causality in Economics** (Oxford: Basil Blackwell)
Hood, W. and T. Koopmans (eds) [1953] **Studies in Econometric Method** (New York: Wiley)
Hurewicz, W. and H. Wallman [1948] **Dimension Theory** (Princeton: Princeton University Press)
Hutchison, T. [1938] **The Significance and Basic Postulates of Economic Theory** (London: Macmillan)
Hutchison, T. [1956] 'Professor Machlup on verification in economics' **Southern Economic Journal, 22**, 476–83
Hutchison, T. [1960] 'Methodological prescriptions in economics: a reply' **Economica, 27 (NS)**, 158–61
Jensen, K., S. Kamath and R. Bennett [1987] 'Money in the production function: an alternative test procedure' **Eastern Economic Journal, 13**, 259–69
Johnston, J. [1963] **Econometric Methods** (New York: McGraw-Hill)
Kaldor, N. [1972] 'The irrelevance of equilibrium economics' **Economic Journal, 82**, 1237–55
Kennedy, P. [1979/85] **A Guide to Econometrics**, 2nd edn (Oxford: Basil Blackwell)
Keynes, J.M. [1936] **General Theory of Employment, Interest and Money** (New York: Harcourt, Brace & World)
Klappholz, K. and J. Agassi [1959] 'Methodological prescriptions in economics' **Economica, 26 (NS)**, 60–74
Klein, L. [1957] 'The scope and limitations of econometrics' **Applied Statistics, 6**, 1–17
Koopmans, T. [1941] 'The logic of econometric business cycle research' **Journal of Political Economy, 49**, 157–81
Koopmans, T. (ed.) [1950a] **Statistical Inference in Dynamic Economic Models** (New York: Wiley)
Koopmans, T. [1950b] 'When is an equation system complete for statistical purposes?' in Koopmans [1950a], 393–409
Koopmans, T. [1953] 'Identification problems in economic model construction' in Hood and Koopmans [1953], 27–48
Koopmans, T. [1957] **Three Essays on the State of Economic Science** (New York: McGraw-Hill)
Koopmans T. and W. Hood [1953] 'The estimation of simultaneous linear economic relationships' in Hood and Koopmans [1953], 112–99
Krupp, S. [1966] **The Structure of Economic Science: Essays on Methodology** (Englewood Cliffs: Prentice-Hall)
Kuenne, R. [1963] **The Theory of General Economic Equilibrium** (Princeton: Princeton University Press)
Kuenne, R. (ed.) [1967] **Monopolistic Competition Theory: Studies in Impact** (New York: Wiley)
Kuhn, H. [1956] 'On a theorem of Wald' in Kuhn and Tucker [1965], 265–74
Kuhn, H. and A.W. Tucker (eds) [1956] **Linear Inequalities and Related Systems** (Princeton: Princeton University Press)
Kuhn, T. [1962/70] **The Structure of Scientific Revolutions** (Chicago: University of Chicago Press)
Lakatos, I. and A. Musgrave (eds) [1968] **Problems in the Philosophy of Science** (Amsterdam: North Holland)

Lancaster, K. [1966] ‘A new approach to consumer theory’ **Journal of Political Economy, 74**, 132–57
Leontief, W. [1948] ‘Econometrics’ in Ellis [1948], 388–411
Lerner, A. [1944] **The Economics of Control: Principles of Welfare Economics** (London: Macmillan)
Liebhafsky, H. [1963] **The Nature of Price Theory** (Homewood: Dorsey Press)
Lipsey, R. [1963] **An Introduction to Positive Economics** (London: Weidenfeld & Nicolson)
Lipsey, R. and K. Lancaster [1956-7] ‘The general theory of second best’ **Review of Economic Studies, 24**, 11–32
Lipsey, R. and P. Steiner [1972] **Economics** (New York: Harper & Row)
Lloyd, C. [1965] ‘On the falsifiability of traditional demand theory’ **Metroeconomica, 17**, 17–23
Lloyd, C. [1967] **Microeconomic Analysis** (Homewood: Irwin)
Lloyd, C. [1969] ‘Ceteris paribus, etc.’ **Metroeconomica, 21**, 86–9
Loasby, B. [1981] ‘Hypothesis and paradigm in the theory of the firm’ **Economic Journal, 81**, 863–85
Machlup, F. [1955] ‘The problem of verification in economics’ **Southern Economic Journal, 22**, 1–21
Machlup, F. [1966] ‘Operationalism and pure theory in economics’ in Krupp [1966], 53–67
MacKinnon, J. [1983] ‘Model specification tests against non-nested alternatives’ **Econometric Reviews, 2**, 85–110
Mann, H. and A. Wald [1943] ‘On the statistical treatment of linear stochastic difference equations’ **Econometrica, 11**, 173–220
de Marchi, N. [1985/88] ‘Popper and the LSE economists’ in de Marchi [1988], 139–66
de Marchi, N. (ed.) [1988] **The Popperian Legacy in Economics** (Cambridge: Cambridge University Press)
Marschak, J. [1953] ‘Economic measurements for policy and prediction’ in Hood and Koopmans [1953], 1–26
Marshall, A. [1890/1920] **Principles of Economics**, 8th edn (London: Macmillan)
Massey, G. [1965] ‘Professor Samuelson on theory and realism: comment’ **American Economic Review, 55**, 1155–64
McCloskey, D. [1983] ‘The rhetoric of economics’ **Journal of Economic Literature, 21**, 481–517
McKenzie, L. [1954] ‘On equilibrium in Graham’s model of world trade and other competitive systems’ **Econometrica, 22**, 147–61
McManus, M. [1959] ‘Comments on the general theory of the second best’ **Review of Economic Studies, 24**, 209–24
Meade, J. [1955] **Trade and Welfare** (Oxford: Oxford University Press)
Metzler, L. [1948] review of Samuelson [1947/65], **American Economic Review, 38**, 905–10
Mishan, E. [1960] ‘Survey of welfare economics’ **Economic Journal, 70**, 197–265
Mishan, E. [1964] **Welfare Economics** (New York: Random House)
Morgenstern, O. [1963] ‘Limits to the uses of mathematics in economics’ in Charlesworth [1963], 12–29
Morgenstern, O. [1972] ‘Thirteen critical points in contemporary economic theory: an interpretation’ **Journal of Economic Literature, 10**, 1163–89
Nagel, E. [1961] **Structure of Science: Problems in the Logic of Scientific Explanation** (New York: Harcourt Brace)
Nikaido, H. [1956] ‘On the classical multilateral exchange problem’ **Metroeconomica, 8**, 135–45
Nikaido, H. [1960/70] **Introduction to Sets and Mappings in Modern Economics** (Amsterdam: North Holland)
Papandreou, A. [1958] **Economics as a Science** (New York: Lippincott Company)
Papandreou, A. [1963] ‘Theory construction and empirical meaning in economics’ **American Economic Review, Papers and Proceedings, 53**, 205–10
Pesaran, M. [1974] ‘On the general problem of model selection’ **Review of Economic Studies, 41**, 153–71
Pigou, A. [1962] **The Economics of Welfare** (London: Macmillan)
Poincaré, H. [1905/52] **Science and Hypothesis** (New York: Dover)
Popper, K. [1945/62] **The Open Society and Its Enemies** (London: Routledge & Kegan Paul)
Popper, K. [1959/61] **The Logic of Scientific Discovery** (New York: Science Editions, Inc.)
Popper, K. [1965] **Conjectures and Refutations: The Growth of Scientific Knowledge** (New York: Basic Books)
Popper, K. [1972] **Objective Knowledge** (London: Oxford University Press)
Quine, W. [1965] **Elementary Logic** rev. edn (New York: Harper & Row)
Quine, W. [1972] **Methods of Logic** (New York: Holt, Rinehart & Winston)
Richardson, G. [1959] ‘Equilibrium, expectations and information’ **Economic Journal, 69**, 225–37
Robinson, J. [1933] **Economics of Imperfect Competition** (London: Macmillan)
Robinson, J. [1962] **Economic Philosophy** (London: Watts)
Rotwein, E. [1959] ‘On "The methodology of positive economics"’ **Quarterly Journal of Economics, 73**, 554–75
Rotwein, E. [1966] ‘Mathematical economics: the empirical view and an appeal for pluralism’ in Krupp [1966], 102–13
Salmon, G. [1928] **A Treatise on the Analytic Geometry of Three Dimensions** (London: Longmans, Green & Company)
Samuelson, P. [1938] ‘The empirical implications of utility analysis’ **Econometrica, 6**, 344–56
Samuelson, P. [1947/65] **Foundations of Economic Analysis** (New York: Atheneum)
Samuelson, P. [1947-8] ‘Some implications of "linearity"’ **Review of Economic Studies, 15**, 88–90
Samuelson, P. [1948] ‘Consumption theory in terms of revealed preference’ **Economica, 15 (NS)**, 243–53

Samuelson, P. [1950a] 'The problem of integrability in utility theory' **Economica, 17 (NS)**, 355–85

Samuelson, P. [1950b] 'Evaluation of real national income' **Oxford Economic Papers, 2 (NS)**, 1–29

Samuelson, P. [1952] 'Economic theory and mathematics – an appraisal' **American Economic Review, 42**, 56–66

Samuelson, P. [1953] 'Consumption theorems in terms of overcompensation rather than indifference comparisons' **Economica, 20 (NS)**, 1–9

Samuelson, P. [1955] 'Professor Samuelson on operationalism in economic theory: comment' **Quarterly Journal of Economics, 63**, 310–14

Samuelson, P. [1962/66] 'Problems of the American economy: an economist's view' Stamp Memorial Lecture (London: The Athlone Press) reprinted in Stiglitz [1966], 1656–81

Samuelson, P. [1963] 'Problems of methodology: discussion' **American Economic Review, Papers and Proceedings, 53**, 231–6

Samuelson, P. [1963/66] 'Modern economic realities and individualism' in Stiglitz [1966], 1407–18

Samuelson, P. [1964] 'Theory and realism: a reply' **American Economic Review, 54**, 736–9

Samuelson, P. [1965] 'Professor Samuelson on theory and realism: reply' **American Economic Review, 55**, 1164–72

Samuelson, P. [1967] 'Monopolistic competition revolution' in Kuenne [1967], 105–38

Samuelson, P. [1983] **Foundations of Economic Analysis**, 3rd edn (Cambridge: Harvard University Press)

Samuelson, P. and A. Scott [1971] **Economics**, 3rd Canadian edn, (Toronto: McGraw-Hill)

Sassower, R. [1985] **Philosophy of Economics: A Critique of Demarcation** (New York: University Press of America)

Shapiro, H. [1973] 'Is verification possible? The evaluation of large econometric models' **American Journal of Agricultural Economics, 55**, 250–8

Simon, H. [1953] 'Causal ordering and identifiability' in Hood and Koopmans [1953], 49–74

Simon, H. [1963] 'Problems of methodology: discussion' **American Economic Review, Papers and Proceedings, 53**, 229–31

Smith, V.K. [1969] 'The identification problem and the validity of economic models: a comment' **South African Journal of Economics, 37**, 81

Stigler, G. [1949] **Five Lectures on Economic Problems** (London: Macmillan)

Stigler, G. [1963] 'Archibald vs. Chicago' **Review of Economic Studies, 30**, 63–4

Stiglitz, J. (ed.) [1966] **Collected Papers of Paul A. Samuelson** (Cambridge: MIT Press)

Swamy, P.A.V.B., R. Conway and P. von zur Muehlen [1985] 'The foundations of econometrics – are there any?' **Econometric Reviews, 4**, 1–61

Tarascio, V. and B. Caldwell [1979] 'Theory choice in economics: philosophy and practice' **Journal of Economic Issues, 13**, 983–1006

Tinbergen, J. [1956/67] **Economic Policy: Principles and Design** (Amsterdam: North Holland)

Wald, A. [1936/51] 'On some systems of equations of mathematical economics' **Econometrica, 19**, 368–403

Ward, B. [1972] **What's Wrong with Economics?** (New York: Basic Books)

Wartofsky, M. [1968] **Conceptual Foundations of Scientific Thought** (New York: Collier-Macmillan)

Watkins, J. [1957] 'Between analytic and empirical' **Philosophy, 32**, 112–31

Watkins, J. [1965] 'Philosophy and politics in Hobbes' in Brown [1965], 237–62

Wedeking, G. [1969] 'Duhem, Quine and Grunbaum on falsification' **Philosophy of Science, 36**, 375–80

Weintraub, E.R. [1985] **General Equilibrium Analysis** (Cambridge: Cambridge University Press)

Wilson, C. [1963] **The Outsider** (London: Pan Books)

Wong, S. [1973] 'The "F-twist" and the methodology of Paul Samuelson' **American Economic Review, 63**, 312–25

Wong, S. [1978] **The Foundations of Paul Samuelson's Revealed Preference Theory** (London: Routledge & Kegan Paul)

Woods, F. [1961] **Higher Geometry: An Introduction to Advanced Methods in Analytic Geometry** (New York: Dover)

Zeuthen, F. [1957] **Economic Theory and Method** (Cambridge: Harvard University Press)

Names Index

Subject Index