Jei JOURNAL OF ECONOMIC ISSUES
Vol. XIII No. 4 December 1979

Theory Choice in Economics: Philosophy and Practice

Vincent J. Tarascio
and
Bruce Caldwell

It is generally accepted that the construction and use of theories is an essential and ineliminable aspect of the pursuit of scientific knowledge.[1] It is also widely recognized that for any given phenomenon under study, there exist alternative theoretical frameworks by which it may be investigated. The question arises: How does one choose among such alternative frameworks in cases when two or more theories are offered in explanation of the same phenomenon?[2] Is choice in such matters an entirely subjective affair, or are there objective criteria by which alternative theoretical constructions may be compared, evaluated, and ultimately ranked? In a phrase, does a set of canons for the rational appraisal of theories exist in science?

The usual answer is that such canons do indeed exist, that there are well-known procedures for weighing the comparative strengths and weaknesses of competing theories. The criteria employed for theory choice are often divided into two categories: empirical and nonempirical or logical. Empirical criteria which have been popular with economists have been of two sorts: the examination of a theory's assumptions for testability and

---

The authors are, respectively, Professor of Economics, University of North Carolina, Chapel Hill, and Assistant Professor of Economics, University of North Carolina, Greensboro.

"realism," and the comparison of a theory's predictions with reality. To meet the empirical criterion of acceptability, then, a theory's assumptions or predictions should be testable and highly confirmed in a large number of independent test situations. Nonempirical criteria, which are used to evaluate aspects of theoretical structure and form, include logical consistency, simplicity, elegance, generality, theoretical support, and others.

The empirical and nonempirical criteria of theory choice may be used for evaluation in the following way. Theories are ranked according to their relative degrees of empirical confirmation and by how well they satisfy the various nonempirical standards of form and structure. Disconfirmed, illogical, and cumbersome theories are rejected or reworked; highly confirmed, mutually consistent, fruitful, and elegant structures are retained. By this process, the frontiers of scientific knowledge are carefully, but inexorably, expanded.

This account is a pleasing one, for it seems to assure that theory appraisal takes place on rational grounds. A number of philosophers of science have claimed that theory choice is rarely so rational, but such claims will not concern us here.[3] The question which this article addresses is whether or not theory choice in economics is effected according to the canons of theory appraisal outlined above. To answer this question, the empirical and nonempirical criteria must be carefully examined to see what role they play in economic science. In the first section of this article, one of the empirical criteria of acceptability, predictive adequacy, is investigated; since that task is quite involved, an analysis of the testing of theories by their assumptions must be omitted. Although many economists have written about prediction in economics, the treatment here breaks new ground by stressing the implications of such arguments for the question of theory choice and by unifying many seemingly disparate arguments regarding prediction in economics within a single hypothesis testing framework. Because economists have almost totally ignored the nonempirical criteria, new ground is also broken in the second part of the article, in which a number of such criteria are defined and their usefulness for theory choice evaluated. In the remaining sections the nature of the theory selection process is examined, with the emphasis on practice rather than on methodological prescription.

The Criterion of Predictive Adequacy

Economists have always taken discussions of the various empirical criteria of theory evaluation very seriously. Many positions have been advocated: Assumptions, some would claim, must be testable, or conceivably testable, or highly confirmed, or realistic; and critics would deny the claims. In comparison with the assumptions debate, the literature dealing with the role of prediction in economics is even larger and more diverse. Why have empirical tests of hypotheses received so much attention from economists and economic methodologists? Our discipline's fascination with the testability question is doubtless due to the powerful influence which positivism has exerted, not only on the methodology of economics, but also on the methodologies of most of the natural and social sciences in the last half century. Modern positivism is a collection of models, prescriptions, and procedural rules, all of which are meant to delimit the scope and methods of legitimate scientific activity.[4] Positivism acknowledges the central importance of deductive reasoning and theory building in science, but attaches even greater significance to the prescription that all scientific hypotheses must be tested against data, the latter consisting of the brute atomic facts of the phenomenal world. While the general notion of testing knowledge statements has been a major theme in Western epistemology for centuries, many specific recommendations which indicate just how such testing should be carried out have their origins in twentieth-century positivist thought.

As noted earlier, some economists have suggested that theories be empirically evaluated by comparing the testability of their assumptions. Others insist that the "realism" of assumptions is the key. Still others would have attention focused on the predictive ability of theories. Each one of these areas has been subjected to thorough critical scrutiny by economists; it would be presumptuous to try to deal with all of them in a single article. All discussions of theory testing which focus on assumptions are therefore omitted from what follows.

Where does one begin in choosing among competing theories on the basis of their predictions? First, the predictions themselves must be roughly comparable in terms of quantitative, qualitative, and temporal ranges of acceptability. Second, theories are then ranked according to their relative strengths of confirmation. This measuring of the strength of arguments is the primary task of inductive logic; just as the task of deductive logic is to discover whether deductive arguments are valid or invalid, inductive logic attempts to rank systematically the relative confirmation of inductive arguments. Philosopher Carl Hempel [1966] mentions the following criteria of confirmation, all of which count in favor of a hypothesis: quantity of favorable test outcomes, precision of procedures of observation and measurement, variety of supporting evidence, and confirmation by new test implications. By applying these criteria, the relative confirmation of competing theories may be established.

However, as Hempel himself admits, the most highly confirmed theory need not be the "best" one, if one means by that the one which is closest to the "true" theory.[5] Simply put, even a perfectly confirmed theory need not be true. Of course, the problem could be semantically circumvented if one *defined* prediction as the ultimate goal of science; then the theory which predicted best would be the best theory. Such an approach can easily lead one into trouble, however. For example, a "theory" of the business cycle based on a totally specious correlation would have to be chosen over an econometric model's implicit theoretical structure if the former were better at prediction. Clearly, scientists have always wanted something more than predictive adequacy from their theories. Extreme predictivism, although present in positivist methodological rhetoric, has never been seriously practiced in any science.

There are many reasons why theory choice based solely on the relative strengths of confirmation of competing hypotheses does not guarantee that the best theory will be chosen. For confirmation, one need only turn to the philosophy of science and examine the conclusions drawn from discussions of David Hume's riddle of induction, Hempelian and Goodmanian paradoxes of confirmation, or the possibilities for constructing an inductive logic that is both justifiable and workable.[6] While such discussions are fascinating to the professional philosopher, their practical relevance for the working scientist, particularly the economist, is not always immediately apparent. An alternative approach is to investigate the problems which arise from the conditional form of all scientific hypothesis testing.

Following Hempel and Paul Oppenheim, every conditional hypothesis is composed of two parts, an explanandum and an explanans: "By the explanandum we understand the sentence describing the phenomenon to be explained (not that phenomenon itself); by the explanans, the class of those sentences which are adduced to account for the phenomenon" [1948, pp. 136–37]. The explanans contains two subclasses: sentences comprising a list of initial conditions which must obtain, and those representing either general or statistical laws, which are themselves highly confirmed. The explanandum statement is the prediction of the hypothesis. Depending on whether the laws invoked are of universal or statistical form, the explanandum statement will follow either deductively or with "a high logical, or inductive, probability," given, of course, that the initial conditions also obtain [Hempel 1963, p. 110].

Positivists have cited the fact that all scientific hypotheses are of conditional form to defend the unity of science thesis: If all sciences use conditional hypotheses, then they follow the same method, and differences among them are only matters of degree. Few would deny that all legitimate sciences make use of some kind of conditional hypotheses, that is, of the type if P, then Q. There is not unanimous agreement, however, that all sciences are capable of meeting the conditions specified in the explanans-explanandum scenario. Indeed, the differences between those sciences which can employ that framework in describing their hypothesizing and those for which it obfuscates more than it illuminates may be a more fruitful basis for distinguishing among the sciences than is the usual natural-social dichotomization.[7] Where would economics fall? To answer that question, various aspects of hypothesis testing in economics must be investigated.

Initial Conditions

Logically, it is impossible to specify all of the initial conditions which would have to hold in any given test situation. As a practical matter, however, one can have confidence that a test result gives an accurate indication of the truth or falsity of an hypothesis if initial conditions are finite in number, empirically specifiable, technologically realizeable, and met. This may be possible in some test situations in economics; few would claim it is the general case. A host of exogenous variables exist which can affect the outcome of an economic experiment; wars, fads, the weather, and political actions are only a few of the many noneconomic variables of whose effects economists may not even be aware, much less able to anticipate. Economists employ causal, closed models in which exogenous variables are conveniently impounded in *ceteris paribus*. However, such models may be too simple to capture the complexity of economic reality; the use of *ceteris paribus* may lend a "degree of determinedness" to a model which does not exist in the subject matter [Wilber and Harrison 1978, pp. 67–68]. If the economic system is, as one critic proposes, "essentially open," then it cannot be captured by a closed model [Schoeffler 1955].[8]

The existence of a large number of noneconomic variables which may impinge on the predictions generated by economic hypotheses need not necessarily be a problem if, in each test case, they are specified beforehand and can be checked afterward to see whether they have, in fact, been met. In a discipline such as economics, this prescription makes hypothesis testing a more troublesome activity than it might appear at first. Unless this procedure is followed, however, one's confidence that either a confirming or disconfirming instance has anything to do with the truth or falsity of an hypothesis is diminished.

General Laws and Change

The evolution of economic institutions and of the cultures in which economic relationships are embedded has been a favorite theme of those who criticize the abstract generalizing of standard economists, based as it is on various laws of economics. Hence, members of the German Historical School, institutionalists, certain Marxists, and their sundry contemporaries have all insisted that an evolutionary or historical approach should replace those methods which have dominated economics since the time of David Ricardo. Austrian economists, who share with the orthodox an admiration of deductive reasoning, but who reject the modern emphasis on prediction and the use of econometric techniques, insist that changes in economic reality reduce econometric results to little more than "the economic history of the recent past" [von Mises 1962, p. 74]. Concomitant with such critiques are alternative methodological frameworks for analyzing economics; it is perhaps for this reason that these groups have had a relatively small impact on the mainstream of economic thought, particularly in this century.

Significantly, less radical economists have made similar charges based on the changing nature of economic reality. Employing a Popperian model of scientific explanation and prediction which is quite similar to the Hempelian model advanced above, T. W. Hutchison writes: "Since very few or no fully adequate scientific laws, in the physico-chemical or natural scientific sense, have been established in economics, on which economists can base predictions, what are used, *and have to be used,* for predictive purposes are *trends, tendencies, or patterns,* expressed in empirical or historical generalizations of less than universal validity, restricted by local and temporal limits" [1977, pp. 19–20]. Hutchison cites as an example that, even if the elasticity of demand for herrings over a period of years lay between 1.2 and 1.45, "it surely cannot be claimed to be a universal law, that in all markets, in all countries, at all times, the elasticity of demand for herrings is, and has always been, between 1.2 and 1.45" (p. 21). In the case of more general laws, such as the law of demand, the absence of checkable initial conditions, especially regarding tastes, prices of other goods, and price expectations, effectively reduces one's ability to test such laws (p. 15). Hutchison concludes that the primary contribution of economists "must inevitably come from *trend-spotting,* not by deduction from laws" (p. 22).

Similar claims have been made by other economists. Oskar Morgenstern [1972] claims that "theory-absorption," the fact that the kind of economic theory that is known to economic agents has an effect on their actions, is a methodological problem "worthy of careful attention" (p. 707). Gerald Garb [1964] points out that certain economic phenomena are best characterized as nonrecurring or infrequently occurring events, which imposes severe constraints on the applicability of causal models for their interpretation. Wassily Leontief [1971] argues that economic change constantly affects the form and parameters of the structural equations assumed in economic models. Alfred Chalk [1970] emphasizes the interdependence and constant growth of wants as contributors to instability. Robert Heilbroner [1970] notes that behavioral data, although fairly stable in the long run, are "highly unpredictable" in the short run, and that while production possibilities are fairly constant in the short run, technical change makes long-run prediction nearly impossible.

Many of these critics have used these arguments to underscore the inadequacy of economic predictions for policy decisions. The methodological implications for the question of theory choice are equally dismal. The absence of easily checkable initial conditions noted in the previous section imply that neither confirmation nor disconfirmation should carry much weight in choosing among theories, unless test procedures are so rigorous (in terms of specifying and checking initial conditions) that they ensure that test results can be trusted. The absence of universal laws means that even highly confirmed theories which have worked well in the past need not be applicable in the future, that "laws" may work well in some cases, but fail in others.[9]

Other Problems

There are other problems confronting economists who wish to evaluate theories empirically. One is the question of incomplete or "dirty" data. This has been pinpointed by Wilber and Harrison: "Both the methods of collection and construction of economic data are unreliable. Typically, economic data are statistically constructed and are not conceptually the same as the corresponding variables in the theory. Therefore, econometricians and statisticians engage in data massaging. If a test disconfirms a hypothesis, the investigator can always blame the data—they have been massaged, either too much or not enough" [1978, p. 69]. Such a view ignores the tremendous leaps forward in the collection and interpretation of data which have occurred within economics in this century. However, a case can be made that advances in model construction and in the development of statistical *techniques* have far outdistanced the more mundane work of gathering and interpreting economic statistics. Wassily Leontief, for one, believes this is the case, and does not view the trend favorably. "Continued preoccupation with imaginary, hypothetical, rather than with observable reality has gradually led to a distortion of the informal valuation scale used in our academic community to assess and to rank the scientific performance of its members. Empirical analysis, according to this scale, gets a lower rating than formal mathematical reasoning" [1971, p. 3].

A second problem, stressed by M. A. Katousian [1974], is that certain economic propositions are presented in such a way as to be irrefutable. What does it mean to say that utility will *eventually* diminish or that cost functions will *eventually* rise if the points at which those eventualities are supposed to occur are never specified? How can one ever determine whether or not a market is in equilibrium? If factors are compensated in the short run, in what sense does the concept of long-run product exhaustion have any empirical meaning?[10]

Finally, the poor predictive record of economists, documented by various authors [Schoeffler 1955; Hutchison 1977; Jewkes 1978], certainly supports the view that predictive adequacy may not be the hallmark of economic science. The results of such studies have relevance for those concerned with policy failures in economics; they do not, however, speak to the question of theory choice. What is needed are studies of how theories have been and are chosen in economics; on what grounds, for example, have economists decided that a subjective theory of value is preferable to a land or labor theory of value? Such studies have appeared [Latsis, ed., 1976]; more are certainly necessary.

Summary

As was mentioned above, some who question the workability of predictive adequacy as the empirical criterion of theory choice have advocated alternative methodological approaches for economics. It seems to us that many such attempts are insightful as far as criticism goes, but fail in proposing adequate alternatives.[11] In any case, no new programs are presented here. The aim of this section is modest: to suggest that theory choice based on a predictive empirical criterion of adequacy can be problematical. For theory choice on empirical grounds to be workable in any

discipline, general laws must be present, initial conditions should be relatively few in number and easily checkable, and data must be trustworthy and complete. There are numerous instances in economics in which such requirements are not satisfied. As a result, it is too often possible for one's favored theories to be confirmed, while disconfirming instances are brushed aside with the claim that *ceteris* were not *paribus*. Similarly, it is too easy to punch holes in the confirming instances of an opponent's theory: Were the data really clean? Does the prediction really justify the assertion of the theory? And so on. When instances of confirmation and disconfirmation hold so little weight, the danger arises that the results of empirical "tests" are used only for rationalization and justification, not theory choice. On such occasions, an objective empirical criterion of theory appraisal is not being employed.

It must be stressed that there are many instances in which the interpretation of a test result in economics is not problematical. This is probably not the case, however, when one must choose between two or more legitimately competing theories, each of which has mustered some evidential support. As a result, few debates in such fields as industrial organization or macroeconomics have been *resolved* on empirical grounds, although empirical testing is prolific. This does not mean that empirical studies are useless; they provide increments of confirming and disconfirming instances. Yet their value for resolving such issues should not be overestimated.

Nonempirical Criteria

Given the difficulties inherent in garnering and interpreting evidential support for theories, it makes sense to ask whether other, nonempirical criteria for judging the acceptability of theories may be invoked. Unfortunately, no systematic treatment of this topic exists in the economics literature.[12] We must turn, therefore, to philosophers of science for a listing of nonempirical criteria.[13] Among many that could be considered, perhaps seven are most important. The first is logical consistency, which requires that no axioms or relationships postulated within a theoretical structure may contradict other relations or axioms in the structure, and that no mutually incompatible theorems may be deducible from the postulated axioms and relations. Logical consistency is probably the oldest and most generally accepted of the nonempirical criteria of acceptability. The second is elegance, perhaps the most subjective standard. It focuses on the beauty and aesthetic appeal of a theoretical structure. In Henry Margenau's elegant prose, "this regulative maxim separates what is ugly and cumbersome from sweeping ideas that carry élan and give pleasure on comprehension" [1966, p. 33]. The third is extensibility: A theory is to be preferred if it allows extension through deductions into other areas of investigation. The fourth, generality, maintains that a theory which incorporates an existing and well-established body of knowledge into a single unified framework is to be judged superior. The fifth nonempirical criterion involves theoretical support, or multiple connectedness. If a new hypothesis fits in well with an established theoretical structure, it gains in acceptability. Hempel [1966, p. 39] cites the generalization of Johann Balmer's formula as a positive instance of the application of this principle; its critical application is exemplified by the story of a Dr. Caldwell of Iowa:

> The credibility of a hypothesis will be adversely affected if it conflicts with hypotheses or theories that are accepted at the time as well-confirmed. In the *New York Medical Record* for 1877, a Dr. Caldwell of Iowa, reporting on an exhumation he claims to have witnessed, asserts that the hair and the beard of a man who had been buried clean-shaven, had burst the coffin and grown through the cracks. Although presented by a presumptive eyewitness, this statement will be rejected without much hesitation because it conflicts with well-established findings about the extent to which human hair continues to grow after death [Hempel 1966, pp. 39–40].

The sixth criterion concerns fertility, fruitfulness, and heuristic value. Theories which suggest new areas or methods of investigation, or new approaches to old problems, are to be judged favorably. The seventh criterion, simplicity, is another ancient standard. It merely states that the simpler and more economical of two theories is to be preferred.

Can theory choice be based on nonempirical grounds? An affirmative answer requires that the criteria outlined above be both justifiable and capable of straightforward application. Problems exist on both counts.

Most of the criteria are justifiable only on an intuitive basis. We generally *like* our theories to cohere well, or exhibit properties of simplicity or elegance, and that predilection is offered as justification. But clearly this will not do. A closer examination of our nonempirical criteria preferences indicates that many are based on metaphysical assumptions. The principle of simplicity, for example, has been justified on the grounds that nature is orderly, which clearly presupposes a metaphysics.[14] Other methodological justifications (such as Karl Popper's, which states that simpler theories are more "falsifiable") depend on the results of individual tests and are thus themselves subject to practical "falsification."[15] Similar criticisms could be advanced about other nonempirical criteria of theory

choice, simply because criteria which impose constraints on the form and structure of theories implicitly presume a certain form and structure of the phenomenal world. The criteria are then justified because they guarantee the use of theories which are somehow optimal for the study of phenomenal reality. Any such justification assumes that one *knows* how reality is structured, and in making that assumption, one has entered the realm of metaphysics.

A likely response is, so what? After all, what matters to the working scientist is not justification, but workability. Even if the choice of nonempirical criteria is arbitrary (that is, no ultimate foundation for that choice exists), could not their applicability be justification enough? If all economists agreed, for example, that economic theories should be logically consistent, and elegant, and so forth, would not such agreement be sufficient justification for retaining those criteria? Such an approach requires only that the criteria be easily applied in judging theories (that is, that one *know* whether or not a given theory does meet a given criterion) and, further, that most economists agree about the value of the various criteria. This approach circumvents the problem of justification and attacks the issue of workability in a "truth-by-consensus" manner. But it seems that even this defensive strategem encounters difficulties in economic science.

It is not always easy to determine whether a given hypothesis or theory meets criteria of acceptability. Some are so loosely defined that subjective interpretation is inevitable. Whether or not a theory is elegant is clearly a matter of opinion; heuristic value, too, depends greatly on what the observer feels are valuable areas of investigation.[16]

Some criteria cannot be employed for short-run theory choice because they are often only distinguishable in retrospect. This seems to be the case for extensibility and generality; for example, a generation passed before economists realized that the tools of marginal utility analysis were general ones that could fruitfully be extended into such areas as production and distribution theory. The claim of greater generality is further hindered by the fact that future research may invalidate it. J. M. Keynes's *General Theory,* an obvious advance (in his eyes) over his "Classical" predecessors, was dubbed a "special case" in the 1950s by Don Patinkin [1966, chapters 8–13] and other founders of the neoclassical synthesis. Then, in the 1960s, revisionist Keynesians [Clower 1965; Leijonhufvud 1968] reinterpreted Keynes's work as more general again, since it isolated as explanatory principles the facts that expectations may not be realized and that information is costly. Theories which assume costless information and perfect expectation are limiting, special cases.

Even such ancient criteria as logical consistency and simplicity are not sacrosanct. A case may be made that it is logically inconsistent to assume that one knows the interest rate when one determines the value of capital, and then state with utter equanimity that the interest rate is determined by the marginal productivity of capital, which, of course, can only be determined if one knows the value of capital.[17] And even simplicity, which has held the veneration of scientists since the time of William of Ockham, may encounter problems. As Hempel [1966, pp. 40–42] points out, simplicity is relative to a certain (often mathematical) background, and choice of the background is arbitrary. For example, let us posit three hypotheses:

$$\begin{array}{ll} \text{H.1} & v = u^4 - 6u^3 + 11u^2 - 5u + 2; \\ \text{H.2} & v = u^5 - 4u^4 - u^3 + 16u^2 - 11u + 2; \\ \text{H.3} & v = u + 2, \end{array}$$

where $u = 0, 1, 2, 3$, and $v = 2, 3, 4, 5$, respectively. One would usually think of H.3 as being the simplest, but only if we define simplicity in terms of the order of a polynomial. If our background is in polar coordinates, H.3 would be more complex (since it describes a spiral) than, say, $v \cos (u - a) = p$, the polar equation for a straight line. It thus seems that identifying which theories meet which criteria can be annoyingly difficult.

An even more important barrier to the application of these criteria is the fact that no theory exhibits all of the criteria listed above. Some hypotheses are fruitful and suggestive but are insufficiently formalized; thus they do not meet the criteria of, say, logical consistency or elegance. Others may advance our understanding of a particular problem but may do little to satisfy generality or extensibility. That no theories meet all of the criteria makes theory choice on nonempirical grounds problematical, for competing theories may be incommensurable in terms of those criteria. This opens the door to a selective application of the nonempirical criteria of acceptability. Proponents of well-established theories, for example, might stress logical consistency, elegance, and multiple connectedness; proponents of alternatives might stress the fruitfulness or greater realism of their theories. Both camps, one assumes, would claim greater generality on the grounds that their theories cover areas which are not included in the domain of investigation of alternative formulations.

The implications of this discussion of applying nonempirical criteria to theory choice are as disheartening as the earlier discussion of predictive adequacy; it seems that these criteria can only be used for the *justification* or *rationalization* of theories, rather than for the *choice* from among competing theories. The only exception would occur when two theories share a number of the same attributes, and one is shown to be superior in terms

of some of these attributes. For other cases, evaluation usually entails debates over which criteria should be employed, and agreement over which attributes a theory might possess usually occurs only in retrospect. Arguments have even been offered that certain criteria should not be employed; Milton Friedman's attack against "realism of assumptions" is perhaps the most well-known example in economics. Indeed, even certain entrenched criteria have been subject to similar broadsides. Philosopher Paul Feyerabend [1970], to cite only one example, has argued that logical consistency and theoretical connectedness are arbitrarily strict and, therefore, inappropriate criteria for judging new, alternative theories; two defining characteristics of such theories are that they challenge existing approaches and that they are first expresed in rough form.

How Does Theory Choice Take Place?

Once it is realized that an infinite number of theories can be proposed to explain a given phenomenon or a given set of data, it makes sense to enumerate objective empirical and nonempirical criteria of acceptability by which theory choice might be effected. Because the subject matter of economics does not always allow a finite number of checkable initial conditions, well-established general laws, and data which are both complete and unambiguously interpretable, theory choice on predictive grounds may yield inconclusive results. Similar problems are encountered when we attempt theory choice on nonempirical bases. Failure to reach anything but an arbitrary justification of such criteria prevents us from ranking them in terms of relative importance; it is often difficult to tell whether a theory objectively meets certain criteria; and because no theory meets all the criteria listed, those stressed as being the most relevant and significant by any one individual may not be independent of that person's preferred theory.[18]

Theory choice (as opposed to theory justification or rationalization) on objective grounds appears to be elusive. *Yet, crucially, theory choice does occur in economics and other sciences.* A methodological question of the most far-reaching consequences is: *On what basis does theory choice occur?*

Philosophers of science within the recent "growth of knowledge" tradition have grappled with this and other problems for at least two decades. Some of the more prominent are Thomas Kuhn, Imre Lakatos, Paul K. Feyerabend, and Stephen Toulmin; their approaches are diverse, but they all agree that rational theory choice in the short run is problematical, except in the simplist of situations.[19] The growth of knowledge tradition also challenges in many fundamental areas the tenets of positivist philosophy of science, which has dominated the discipline since the days of the Vienna Circle. Whereas positivists concern themselves with the elaboration of universal models and procedural rules which they believe aptly characterize legitimate scientific practice, the newer analyses emphasize the growth of knowledge over time, the dynamics of change within individual disciplines, and the actual practices of scientists. Universality is qualified by specificity; immutable verities are challenged by the recognition of changing standards of investigation and patterns of thought; logical analysis is supplemented by and checked against the study of history; the "context of discovery" is treated on an equal footing with the "context of justification."[20]

This literature has had some influence on economics; the analyses of Kuhn and Lakatos, in particular, have been embraced by historians of economic thought as tools for investigating the evolution of economic ideas. As might be expected, there is little agreement as to which approach best describes the intellectual history of economics.[21] Questions have also been raised about the legitimacy of applying models of historical change formulated with the natural sciences in mind to a discipline such as economics.[22] Such historical-descriptive debates, although interesting, often obscure the implications of the growth of knowledge analyses for the question of theory choice, the issue with which we are concerned here. Although some work in this area has been done [Latsis 1976], much more is needed before the importance of the growth of knowledge approach for economic methodology can be evaluated.

The remaining sections of this article attempt another description of the selection process by which contributions to economic knowledge are determined. While incorporating elements from both the standard and growth of knowledge traditions regarding theory choice, it differs from both in emphasizing the distinction between the objective existence of a contribution and its subjective acceptance as a part of a body of knowledge, and by stressing the actual processes which are involved in the selection process. (Hence the dichotomy between philosophy and practice.) Let us examine the nature of this selection process.

Some Initial Conceptions

We begin with four initial conceptions. First, a body of professional knowledge exists in the form of a literature.[23] It may be defined as consisting of all writings on economics from the beginning to the present, or

it may be limited to a "current" corpus. These definitions are also extensible to a body of "specialized" literature.

Second, because of human limitations, no individual researcher is able to "absorb" completely the contents of the existing literature, past or present. Instead, he picks and chooses according to his own impression as to what is important and useful for his purpose. For the same reason, it is often not possible for an author's meaning to be conveyed exactly to his readers.[24] Because they do not and cannot have access to the writer's total understanding of the subject, much is left to interpretation.[25] This is a general condition: the transfer of written information or knowledge is not exact, but selective; not quantitative, but qualitative.[26]

Third, each science is characterized by a prevailing set of scientific and professional norms which determine its character. The former may involve principles to be followed, scope of study, and choice of analytical methods. Among the latter in economics, for example, are the norm of allocative efficiency and others, such as forms of presentation of results, and so forth. Together, these two kinds of norms constitute a Kuhnian type of "disciplinary matrix" [Kuhn 1970]. We shall refer to these as disciplinary norms. The practices of individual scientists more or less conform to the generally accepted conventions of their science.[27]

Fourth, there is a formal system for submission and evaluation of research. This system is peer-group oriented.[28]

Given these initial conceptions, what is the nature of the process of theory selection? Let us discuss the most familiar and then proceed to the less familiar. We begin with the individual researcher and examine the various stages of selection during the process.

The Evaluative Stage of the Selection Process

Let us suppose that the body of knowledge represented by an existing literature possesses an *objective* existence, apart from those who created it. A researcher cannot deal with the literature in its entirety, so he picks and chooses according to his impression of what is important for his purpose. He may add his own immediate (subjective) experience to this distillation. In other words, his research is cast within the framework of his *subjective* state of knowledge.[29]

The program of *all* science has been to "objectify" subjective knowledge. This involves conceptualizing and systematizing subjective experience (knowledge) in the form of theory so that it can be conveyed to others. Once a theory is developed, it stands on its own, quite apart from its inventor. Nevertheless, theory is not the exclusive domain of science, although scientific theories tend to be more rigorous than nonscientific theories. But rigorous theory, as a means of "objectifying" subjective experience, is not sufficient for positivists. As we have seen, positivism has developed certain objective criteria (tests, rules, procedures, and so forth), with an emphasis on theory *choice* as an additional means of "objectifying" subjective experience. We recognize the existence of "scientific" norms pertaining to theoretical and empirical research. Examples in economics are conventions deriving from mathematical, statistical, and econometric *theories*.[30]

A researcher is free to cast his research, theoretical or empirical, within the framework of his subjective state of knowledge, but he is also constrained to observing the disciplinary norms obtaining in his science. Minimum standards of professional competence and training exist in every science and are a necessary condition for favorable evaluation of research. From the point of view of the individual, these norms are "objective" in that they exist independently of his own judgment and are taken as a given part of scientific activity. Nevertheless, they are norms, and nothing more.[31]

To the extent that the evaluation of research is confined to such norms, it may be called "objective." For example, a referee applies a set of generally accepted standards in evaluating research. Nevertheless, acceptable professional workmanship does not in itself assure an acceptable contribution to the literature. This point has been overlooked by positivists. There is, in addition, the requirement of *significance,* which is interpretive and hence subjective in nature. A research paper may be "technically correct" and at the same time lack significance.

Significance involves two related aspects. The referee compares the researcher's "subjective" state of knowledge with his own. If the latter has included all "relevant" material from the existing body of literature, his work is judged "well motivated" from the subjective viewpoint of the referee, that is, the researcher has demonstrated a "grasp of the literature." The referee then judges the research results. Such terms as new, novel, interesting, or their opposites, usually are applied. It is crucial to understand that the referee's judgment on such matters often involves notions which resemble methodological criteria, such as simplicity, generality, and fruitfulness. But equally as crucial, as Kuhn has emphasized, referees rarely have any explicit knowledge of methodology; rather, they learn their methodology by doing science. Thus, such judgments of results reflect the *subjective* state of knowledge of the *referee.* Hence, the assessment of significance is subjective.

Once a contribution is deemed acceptable, someone must decide whether or not it will be published. Economists, more than any other scientists, should realize how economic constraints limit available space. Although this economic constraint differs from the purely scientific and professional considerations discussed above, it is an important part of the process.[32] At this level, selection involves such matters as general interest, topical aspects, ranking according to the qualitative judgments of referees, and so forth, which are subjective in nature. If a contribution survives this hurdle, it becomes part of the existing literature.

The formal system for submission and evaluation of research involves a selection process in which subjective factors play an important, if not dominant, role. The objective factors (those external to the individual, that is, disciplinary norms) are hardly more than a necessary prerequisite for an acceptance of theories. Indeed, the more technically oriented a discipline, the more widespread is similar technical training among its members, and the less important are these prerequisites in theory choice, because they are more or less taken for granted.

Objective Existence and Subjective Knowledge

Once a contribution becomes part of the objective body of knowledge (the existing scientific literature), it is presumed to be an objective contribution.[33] To the extent that it, or a part of it, becomes part of the subjective state of knowledge of other researchers, it can be judged to be such a contribution.[34] Or it may be ignored, in which case its contribution to the subjective state of knowledge of others is nil.

To the extent that contributions are initially or eventually ignored, they become a part of the garbage heap of knowledge; they have an objective existence, but are of no use or interest to anyone. As long as they exist they are potential candidates for "rediscovery," but with the passage of time, this possibility becomes remote. Since economic knowledge is produced continually, earlier contributions become remote, not only in time, but also in mind. One result is much unnecessary "originality" in economic science deriving from an ignorance of the past, as historians of economics have pointed out. The nature of the selection process is such that economics has been a less cumulatively progressive science than, say, physics.

Most important, however, is that the criteria for "rejecting" such theories are not those which have emerged in positivist philosophy of science, that is, disconfirmation, falsification, and so forth, either at the formal stage of evaluation or the postpublication stage of objective existence. Both parts of the selection process involve primarily subjective factors of an interpretive nature. Indeed, the objective factors are relatively unimportant in the first stage and irrelevant in the second. In the latter instance the objective fact of existence is unimportant. What matters in both stages are the significance of theory from the subjective viewpoint of individual referees and the relevance of theory from the subjective viewpoint of individual researchers.

We can now understand why, in deference to positivists, controversies and disagreements in economics regarding the (subjective) significance and relevance of theories are seldom resolved definitively, according to positivist criteria, but instead expire from exhaustion or boredom. This is not to say that such controversies are a waste of time because they are often impossible to resolve, objectively. They provide a critical environment which acts as a check on what might otherwise become fanciful flights of reason because of the subjective nature of scientific knowledge. Criticism is an important aspect of the growth of knowledge. Controversy in the social sciences is not a perverse situation which will disappear with the "development" of the social sciences, but a normal condition, given the predominately subjective character of the selection process.

Social and Positivist Sciences

Much of what constitutes positivist philosophy of science grew out of and pertains to the physical sciences. We have also argued that theory choice as conceived by positivist philosophy applies more to the physical than the social sciences.[35] This is not to argue that certain methods of the physical sciences are not appropriate for the social sciences (mathematics, statistical applications, and so forth), but such methods are useful as methods of presenting, not "verifying," theories in the social sciences.[36] For much of what constitutes research and knowledge in the latter fields stands or falls on the basis of subjective significance and relevance, respectively. Although these concepts are not as precise and definitive as positivist concepts of empirical validity, they are at present, and have been in the past, the practical criteria in practice. The social sciences, thus, are more interpretive disciplines than is often admitted by the advocates of positivist science. And if the past and present are any indication of the future, it appears that they will continue to be interpretive regardless of the degree of technical sophistication. Hence, the social sciences occupy a middle ground between the humanities and the natural sciences. To the extent that quantitative techniques are used in research and presentation, they correspond to the natural sciences; to the extent that interpretation is used in the selection of theories, they have more in common with the humanities. Failure to distinguish between these two aspects of the social sciences in general, and economics in particular, has resulted in methodological prescriptions which seldom conform to the practice of theory choice.

Notes

1. *Theory* is a notoriously difficult concept to define. It is used here in a very loose sense, incorporating anything from a singular statement to a full-blown set of interdependent hypotheses, the latter concept being comparable to a Lakatosian "research programme." The criticisms raised here about the workability of various criteria of theory choice are most applicable when theory is taken to mean more than simply a singular statement.
   Carl Hempel [1962] shows why theories are essential and ineliminable in science. Israel Scheffler [1970] labels those opposed to this view "eliminative fictionalists."
2. Such a question can arise only if one believes that alternative theoretical frameworks can be compared; our approach therefore denies any extreme form of theory dependence. For a balanced discussion of this philosophical issue, see Peter Achinstein [1968, chapter 4].
3. See Thomas Kuhn [1970], Imre Lakatos [1970], and Paul Feyerabend [1970]. More will be said about these philosophers later.
4. For a more complete description of positivism, especially in its relation to twentieth-century methodological thought in economics, see Bruce Caldwell [1979].
5. Hempel [1966, pp. 45–46]. Following Karl Popper's reading of Alfred Tarski, a true theory is defined as one which "corresponds with the facts." See Popper [1962, chapters 3 and 10].
6. See the discussions on confirmation and inductive logic in Baruch Brody [1970], Brian Skyrmes [1975], and Suppe [1977, Introduction and Afterword].
7. A somewhat similar proposal can be found in Book I, chapter 3, of G. L. S. Schackle [1972].
8. This point was made by Hutchison as early as 1938; it has been remarked upon by a number of economic methodologists since. See, for example, Hutchison [1938], Fritz Machlup [1955; 1966], and Emile Grunberg [1978].
9. Thomas Malthus's "law" of population, if stated in a testable form, did not adequately describe the Western European demographic experience of the nineteenth and twentieth centuries. Malthusian doctrines were revived in the 1960s because of their apparent relevance for certain less developed countries. Are the parson's observations laws or trends? Can certain economic relationships hold in one period, and not in another?
10. It is a little surprising that Marxists, whose analyses have long been subject to dismissal by orthodox critics on the grounds that their leader's long-run predictions were irrefutable, had not made this counterargument

long ago. It took, instead, an economist steeped in Popperian philosophy of science to make this point.

11. See the comments by Vincent Tarascio [1977]. A counter-example is the proposal by Charles Wilber and Robert Harrison [1978], which outlines the limitations and potential contributions of an institutionalist approach to scientific explanation. The Austrian renaissance may provide another as it becomes more constructive.
12. A number of economists have *mentioned* these criteria; Milton Friedman [1953, p. 10], for example, notes that simplicity, fruitfulness, logical completeness, and consistency are additional criteria which should be invoked when choosing among theories which are equivalent in predictive adequacy. D. A. Collard [1964] goes farther and suggests a "lexicographical ordering" of the nonempirical criteria, but stops there.
13. See Hempel [1966] and Henry Margenau [1966].
14. Hempel [1966, pp. 40–45 and references cited there] offers a more complete treatment of the simplicity criterion.
15. Ibid., pp. 44–45. "Falsification" is used rather than disconfirmation to retain the flavor of Popper's original discussion [1934, English translation 1959].
16. This problem might be avoidable if more strict definitions of the various criteria were possible. The difficulties of such a task may be one reason why it has yet to be seriously attempted in economics.
17. This brief reference to the Cambridge debates, as well as the other examples cited from the history of thought, are meant to suggest areas for further methodological research rather than the conclusions of settled debates in economics. The interpretations suggested in the text could well be revised as a result of such research. The general point, that the determination of whether or not a theory meets various nonempirical criteria of acceptability is difficult to achieve, would only be strengthened by such revisions.
18. This statement could be easily disconfirmed; all we need find is a "neoclassical" theorist who does not believe that logical consistency, elegance, and theoretical support are virtues in theory evaluation, or a New Left "theorist" who insists on the importance of such criteria over, say, "realism."
19. Such situations occur during Kuhnian "normal science" or when hypotheses in the "protective belt" of a Lakatosian research program are tested. See Kuhn [1970] and Lakatos [1970].
20. See Suppe [1977] for an excellent discussion of the differences between the positivist and growth of knowledge approaches.
21. For example, A. W. Coats [1969] says there has been one dominant paradigm in economics; Michel de Vroey [1975] feels the transition from classical to neoclassical economics was a Kuhnian revolution; Jorg Baumberger [1977] claims there have been no Kuhnian revolutions in economics; and Dudley Dillard [1978] counts five revolutions in England between the time of Adam Smith and J. M. Keynes.
22. See, for example, Leonard Kunin and Stirton Weaver [1971] and T. W. Hutchison [1977].

23. One may argue that knowledge is both written and unwritten, so that defining a state of knowledge as the existing literature is too narrow. Nevertheless, most professional knowledge eventually takes a published form.
24. Writing by necessity involves condensation, and the judgment of the writer as to what is important in conveying his ideas often leaves much implicit. For example, Keynes's *General Theory* is there for everyone to read; yet, many disputes regarding the meaning of certain passages indicate that Keynes was not able to convey the meaning of those passages clearly.
25. A related issue regarding the problem of total understanding is the time dependent nature of interpretation. For example, it is obvious that some current interpretations of Keynes's *General Theory* are projections of the present knowledge into the past, and it is doubtful that Keynes possessed a perception of the future states of knowledge such as that implied in, for example, "disequilibrium" theory.
26. This is essentially a problem of communication. The limitation of human capability is such that individuals must always simplify, condense, and summarize information being received in oral or written form.
27. Indeed, the term *discipline* implies conformance to certain norms.
28. This system varies somewhat in practice. We have in mind that associated with the publication of scholarly books and research articles involving reviewers. This fourth conception is not essential for our analysis, but provides some concreteness. We shall see that even in the absence of a formal system of submission and reviewing, the nature of the selection process remains unchanged.
29. The problem of the inception of ideas is little understood. We know little about how individuals rearrange what they know or think they know to form a new synthesis or "combination." All we know is that the process is subjective.
30. This does not mean that such theories were or are accepted by *economists* on the basis of positivist criteria, although such criteria *may* in varying degrees have been a factor in their development. We say "may" because it is doubtful that, for example, the theory of calculus can be verified empirically. Also, it does not follow that the *application* of these theories in economic research satisfies those same criteria, as we have argued earlier.
31. The norms discussed in this and preceding paragraphs include those which would be found in a Kuhnian "disciplinary matrix," that is, symbolic generalizations, models, shared values, and exemplars. See Kuhn [1970, pp. 181–87]. Such norms are different from methodological criteria discussed earlier, and the reader should keep such differences in mind in this section.
32. For example, Leon Walras found it necessary to publish most of his works at his own expense. Although some degree of self-financed, individually subsidized, or special interest publication goes on today, most scientific literature is subject to external (to the individual) economic constraints, as every journal editor knows.

33. This second step in the selection process is logically independent of the first step (a formal system of submission and evaluation) discussed in the preceding section. Indeed, our analysis does not depend on the first step, but it was introduced because it conforms more or less to practice.
34. The term *contribution* is used here in a factual sense. This is why it is called "objective." Evidence of such would be citations in the literature.
35. This characterization of the physical sciences as following positivist methodological precepts may be overstated; the works of Kuhn and Lakatos, which challenge positivism, draw their examples from the natural sciences. Nevertheless, this does not affect our argument that the case for positivism is even weaker in the social sciences.
36. Again, we must refer the reader to our earlier discussion for elaborations on this position.

References

Achinstein, Peter. 1968. *Concepts of Science: A Philosophical Analysis*. Baltimore: Johns Hopkins Press.

Baumberger, Jorg. 1977. "No Kuhnian Revolutions in Economics." *Journal of Economic Issues* 11 (March): 1–20.

Brody, Baruch, ed. 1970. *Readings in the Philosophy of Science*. Englewood Cliffs, N.J.: Prentice-Hall.

Caldwell, Bruce. 1980. "Positivist Philosophy of Science and the Methodology of Economics." *Journal of Economic Issues* 14 (March).

Chalk, Alfred. 1970. "Concepts of Change and the Role of Predictability in Economics." *History of Political Economy* 2 (Spring): 97–117.

Clower, R. W. 1965. "The Keynesian Counterrevolution: A Theoretical Appraisal." In *The Theory of Interest Rates,* edited by F. H. Hahn and F. P. R. Brechling. London: Macmillan.

Collard, D. A. 1964. "Swans, Falling Bodies, and Five-Legged Dogs." *Quarterly Journal of Economics* 78 (November): 645–46.

DeVroey, Michel. 1975. "The Transition from Classical to Neoclassical Economics: A Scientific Revolution." *Journal of Economic Issues* 9 (September): 415–39.

Dillard, Dudley. 1978. "Revolutions in Economic Theory." *Southern Economic Journal* 44 (April): 705–24.

Feyerabend, Paul K. 1970. "How to Be a Good Empiricist—A Plea for Tolerance in Matters Epistemological." In *Readings,* edited by Brody. Pp. 325–35.

————. 1975. *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: New Left Review.

Friedman, Milton. 1953. *Essays in Positive Economics*. Chicago: University of Chicago Press.

Garb, Gerald. 1964. "The Problem of Causality in Economics." *Kyklos* 17: 594–609.

Grunberg, Emile. 1978. " 'Complexity' and 'Open Systems' in Economic Discourse." *Journal of Economic Issues* 12 (September): 541–60.

Heilbroner, Robert. 1970. "On the Limits of Economic Prediction." *Diogenes* 70 (April): 27–40.

Hempel, Carl G. 1958. "The Theoretician's Dilemma." In *Minnesota Studies in the Philosophy of Science,* edited by Herbert Feigl, Grover Maxwell, and Michael Scriven. Minneapolis: University of Minnesota Press. Volume 2, pp. 37–98.

————. 1963. "Explanation and Prediction by Covering Laws." In *Philosophy of Science: The Delaware Seminar,* edited by Bernard Baumrin. New York: John Wiley and Sons. Volume 1, pp. 107–33.

————. 1966. *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice-Hall.

Hempel, Carl, and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15: 135–75.

Hutchison, Terence W. 1938. *The Significance and Basic Postulates of Economic Theory*. London: Macmillan.

————. 1977. *Knowledge and Ignorance in Economics*. Chicago: University of Chicago Press.

Jewkes, John. 1978. *A Return to Free Market Economics: Critical Essays on Government Intervention*. New York: Holmes and Meier.

Katousian, M. A. 1974. "Scientific Methods and Positive Economics." *Scottish Journal of Political Economy* 21 (November): 279–86.

Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*. 2d ed. Chicago: University of Chicago Press.

Kunin, Leonard, and F. Stirton Weaver. 1971. "On the Structure of Scientific Revolutions in Economics." *History of Political Economy* 3 (Fall): 391–97.

Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes." In *Criticism and the Growth of Knowledge,* edited by Imre Lakatos and Alan Musgrave. London: Cambridge University Press.

Latsis, Spiro, ed. 1976. *Method and Appraisal in Economics*. London: Cambridge University Press.

Leijonhufvud, Axel. 1968. *On Keynesian Economics and the Economics of Keynes*. New York: Oxford University Press.

Leontief, Wassily. 1971. "Theoretical Assumptions and Nonobserved Facts." *American Economic Review* 61 (March): 1–7.

Machlup, Fritz. 1955. "The Problem of Verification in Economics." *Southern Economic Journal* 22 (July): 1–21.

————. 1966. "Operationalism and Pure Theory in Economics." In *The Structure of Economic Science: Essays on Methodology,* edited by Sherman R. Krupp. Englewood Cliffs, N.J.: Prentice-Hall.

Margenau, Henry. 1966. "What Is a Theory?" In *The Structure of Economic Science,* edited by Sherman R. Krupp. Englewood Cliffs, N.J.: Prentice-Hall.

Morgenstern, Oskar. 1972. "Descriptive, Predictive and Normative Theory." *Kyklos* 25: 699–714.

Patinkin, Don. 1965. *Money, Interest and Prices*. 2d ed. New York: Harper and Row.

Popper, Karl. 1959. *The Logic of Scientific Discovery*. New York: Harper and Row.

————. 1962. *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Basic Books.

Schackle, G. L. S. 1972. *Epistemics and Economics: A Critique of Economic Doctrines*. Cambridge: the University Press.
Scheffler, Israel. 1970. "The Fictionalist View of Scientific Theories." In *Readings*, edited by Brody. Pp. 211–22.
Schoeffler, Sidney. 1955. *The Failure of Economics: A Diagnostic Study*. Cambridge, Mass.: Harvard University Press.
Skyrms, Brian. 1975. *Choice and Chance: An Introduction to Inductive Logic*. 2d ed. Encino, California: Dickenson.
Suppe, Frederick, ed. 1977. *The Structure of Scientific Theories*. 2d ed. Urbana: University of Illinois Press.
Tarascio, Vincent. 1977. "Theories of Behavior and Public Policy." *Spoudai* 27: 279–90.
von Mises, Ludwig. 1962. *The Ultimate Foundation of Economic Science*. Kansas City: Sheed Andrews and McMeel.
Wilber, Charles, and Robert Harrison. 1978. "The Methodological Basis of Institutional Economics: Pattern Model, Storytelling, and Holism." *Journal of Economic Issues* 12 (March): 61–89.