# 5 Models and theories in economics

Chapters 1-4 presented fundamental neoclassical theory - "equilibrium theory" - and explored how it is incorporated into partial and general equilibrium theories. The discussion showed how recourse to theory unifies and systematizes empirical generalizations, and it provided some of the flavor of theoretical work in microeconomics, general equilibrium theory, and welfare economics. We caught some glimpses of the challenging task of reformulating relevant parts of equilibrium theory, common simplifications, and specifications of the epistemological, institutional, and other circumstances so as to deduce enlightening theorems. We witnessed the significance and centrality of the theoretical enterprise of neoclassical economics.

But we did not comprehend that enterprise in a philosophically satisfying way, for nothing was said to connect the description of theoretical practice in microeconomics, general equilibrium theory, and welfare economics to general philosophical theses concerning the nature, role, and importance of theories in science. Indeed the discussion in the previous chapters has probably also been jarring to economists with its old-fashioned talk of laws and theories. For economists prefer to speak of "models" rather than "theories" and of "behavioral postulates" rather than "laws." Why? What are models and how are they related to theories? Should one not regard their central clauses or generalizations as laws?

My view of theories and models and of global theory structure in economics is controversial. I hope to show that it helps one to understand how economists theorize. This view is philosophically cogent and, I shall argue, enables one to describe theoretical practice in economics more compactly and intelligibly than do alternative philosophical accounts.

## 5.1 Logical positivism and the nature of scientific theory

Recent philosophical work concerning scientific theories either grows out of or reacts against the view of scientific theories developed by the





logical positivists (A.1.1, A.6). According to the logical positivists, scientific theories are basically sets of *sentences*, which are closed under logical deduction. These sentences should be expressed in a formal language, such as the first-order predicate calculus. Sentences are syntactic objects, and their identity is independent of their interpretation. "$(x)(Fx \vee \sim Gx)$" is a sentence. Its logical notation is a precise way of saying "Everything is $F$ or not $G$." If "$F$" is interpreted as the predicate "mortal" and "$G$" is interpreted as the predicated "human," then the interpreted sentence is true. If "$F$" is interpreted as "blue" and "$G$" as "red," then the interpreted sentence is false. Deducibility is a relationship between sentences, which is independent of their interpretation, and, by focusing on the sentences of which a theory is composed, one can investigate the deductive consequences of these sentences without semantic distraction.

Obviously scientific theories are not uninterpreted. There is always some "semantics," in particular, a standard interpretation, which consists of a specification of a domain for the variables (such as "$x$" in "$(x)(Fx \vee \sim Gx)$"), an assignment of "extensions" (sets of entities of which the predicates are true) to the predicates, of entities to constants, and of functions to function symbols. An interpretation constitutes a *model* of the theory if and only if it makes all the sentences come out true. A model is an ensemble of entities with various properties which a theory, when appropriately interpreted, is true of. Apart from its standard interpretation (under which the theory may or may not be true of some portion of the real world), there may be other interpretations and other models, which may be useful in the development and assessment of the theory. This logician's notion of a "model" is *not* what economists mean when they talk of models.

Since the sentences in physical theories often cannot be interpreted as making claims about observable things and properties, the semantics of such theories is philosophically problematic. The "correspondence rules" discussed by the logical positivists (A.7) are supposed to provide an interpretation of the sentences in a scientific theory that do not make claims about observables.

Alfred Mackay has provided a particularly nice illustration of this notion of theory and model in his book on Arrow's theorem (1980). Recall (pp. 64-5 above) that Arrow proved that there is no way to derive a social preference ordering from individual preference orderings, which satisfies the following four conditions:

1 It must provide some social preference ranking for any profile of individual preferences (collective rationality).

2 If everybody prefers social option $x$ to $y$, then $x$ must be socially preferred to $y$ (weak Pareto principle).
3 There is no individual whose preferences are decisive over all options regardless of the preferences of others (non-dictatorship).
4 The social ranking of $x$ and $y$ must depend on nothing except the individual rankings of $x$ and $y$ (independence of irrelevant alternatives).

Arrow's proof, like all proofs, results from the syntax of the axioms, not their interpretation, and there may be alternative interpretations. MacKay (1980) proposed that one consider the problem of deriving an overall ranking of athletic performance in a multi-event athletic competition such as a decathlon on the basis of the ordering of accomplishments in individual events. Arrow's four conditions on social choice translate into the following four conditions on a multiathlon scoring system:

1 It must provide some overall ranking for any profile of finishes in individual athletic events (collective rationality).
2 If athlete $x$ beats $y$ in every event, then $x$ must rank higher than $y$ in the overall ranking (weak Pareto principle).
3 There is no individual event, the outcome of which is decisive regardless of how competitors perform in other events (non-dictatorship).
4 The overall ranking of athletes $x$ and $y$ must depend on nothing except how they rank in the individual events (independence of irrelevant alternatives).

When Arrow proved his theorem, he also proved that there is no system of multiathlon scoring that satisfies all these conditions.

Since Arrow's conditions cannot be simultaneously satisfied, their conjunction has no model, but there are, as we have seen, at least two significant and interesting interpretations. By separating syntax and semantics, one economizes on logical effort, and one can see precisely the formal identity of the distinct problems of scoring athletic events and making social choices. By seeing theories as syntactic objects and by formalizing them, one might, in the positivist's view, put logic to work, gain just such an economy of logical effort, and recognize the formal connections between distinct problems. How much improved might science be! An application to economics of the logical positivist's view of scientific theories can be found in Papandreou (1958, 1963).

## 5.2 Semantic and predicate views of theories

The positivists' syntactic view of scientific theories faces difficulties that has led philosophers such as Bas van Fraassen (1980) and Frederick



Suppe (1974, 1988) to propose in its stead a semantic view of theories. First, if one identifies a theory with a particular syntactic object (or with a class of syntactic objects with certain similarities of shapes or symbols), then any reformulation of a theory or even translation of a theory into a different language counts as a different theory. While there are ways around the objection, they undercut its appeal. Second, it is difficult to express scientific theories in formal languages and awkward, challenging, and time-consuming to do proofs in existing formal languages. Scientists do not waste their time this way. Third, one can argue, as Bas van Fraassen (1980) most effectively has, that the positivist emphasis on language is misplaced. One's focus should be on the content of scientific theories, on the models of which they are true, and on the relations among such models, not on the sentences used to express the theories. Indeed van Fraassen argues that some significant relations cannot be expressed within a syntactic view of theories (1980, p. 44). On grounds such as these, van Fraassen and Suppe urge us to regard scientific theories as the set of models of which the sentences in any particular formulation are true rather than as anything linguistic or sentential at all.

One might question whether the semantic view of theories differs more than linguistically from the syntactic view that it attempts to replace. What the logical positivists called the set of models of which a theory is true, the semantic theorists call "theories," and what the logical positivists called "theories," the semantic theorists call sets of sentences that are true of theories. But the relabeling is significant, because it redirects philosophical interest from sentences to things. There is also some question about whether one can accurately interpret "theories" in van Fraassen's and Suppe's sense as merely sets of models of which theories in the positivist's sense are true.[1] Apart from some awkwardness, which the semantic view shares with its predecessor and from some puzzles mentioned in the last footnote, I see nothing "wrong" with the semantic view of theories. But it does not fit the practice of economics as well as the alternative I prefer.

That third alternative is to regard scientific theories not as purely syntactic or purely semantic, but simply as a set of lawlike and interpreted *statements* (or as an equivalence class of such sets to allow one to count reformulations and restatements of theories as the same theories).

[1] If theories in the positivist sense are logically consistent, but not true in their standard interpretation, it is typically easy to find some alternative arbitrary interpretation that will specify a model. But the sets of models that constitute theories in Suppe's and van Fraassen's sense seem not to include such arbitrary concoctions. Instead they stress the ensembles of *possible* entities that may provide models in the positivist's sense. One reason why I am reluctant to accept this view of scientific theories is that it forces one immediately to confront confusing metaphysical questions concerning possible "worlds." I am indebted to Mark Bauder for help with these points.

Although my view might seem little different than that espoused by the logical positivists, it owes at least as much to a fourth view of scientific theories developed and defended by Patrick Suppes (1957, chapter 12), Joseph Sneed (1971), and Wolfgang Stegmueller (1976, 1979). Ronald Giere provides a useful simplified exposition of this fourth view in his *Understanding Scientific Reasoning* (1979 [1982]). Before saying more about my notions of theories and models, let us consider this fourth view.

In Suppes' view, scientific theories should be regarded as *predicates.* So they are linguistic entities, unlike the "theories" of the semantic theorists, but they are not *sentences.* The empirical claims of science consist of assertions that employ these predicates. Suppes argues that scientific theories are set-theoretic predicates, because he hopes to provide set-theoretical formal restatements of scientific theories. Since I am not concerned to formalize scientific theories, I shall not follow Suppes here. Other writers on economic methodology have provided formal reconstructions of economics modeled after the work of Suppes and, particularly, Sneed (see Händler 1980, Stegmueller, Balzer, and Spohn 1982, Hands 1985c, and Balzer and Hamminga 1989).

In Giere's simplified presentation (1979, chapter 5), scientific theories are definitions of predicates rather than predicates themselves. Newton's laws of motion and his law of gravitation define, for example, what Giere calls "a classical particle system." The predicate, "is a classical particle system," is true of something if and only if Newton's laws of motion and gravitation are true of it. The predicates which constitute scientific theories are not purely syntactic, for the terms in Newton's laws - body, force, distance, etc. - all have interpretations. But the interpretations of these terms do not determine (though they do constrain) the extension of the new predicate, that is, of the theory, in this sense of "theory." Reformulations of a theory in this sense that do not change its extension should not count as theory changes.

On Giere's view of scientific theories, the basic statements of what I have called "the basic equilibrium model" define a new predicate "is an economic equilibrium system" or a new kind of system, "an economic equilibrium system." An actual economy is an economic equilibrium system if and only if the laws of consumer choice theory and of the theory of the firm are true of it, and an equilibrium obtains. The simple consumption system of section 2.4 and the simple competitive production system of section 3.3 are explicitly formulated as such definitions.

On this view of scientific theories, there is no point in asking whether the claims of a theory are true or whether a theory provides reliable predictions. Predicates cannot be true or false or provide any predictions.



Definitions are trivially true, but also do not enable one to make any predictions.

Since science does more than provide definitions, the proposing of theories, in this view, is only one part of science. The other crucial part is proposing *theoretical hypotheses,* which assert that the new term is true of some actual system. Newton not only defined a classical particle system; he also offered the theoretical hypothesis that the solar system is a classical particle system. Economists do more than merely define an economic equilibrium system. In using microeconomic theory to explain or to predict, they also assert or imply that some actual economic objects, at least to some degree of approximation, constitute economic equilibrium systems.

This account of scientific theories idealizes, for in reality theorizing and modeling are not sharply separated, and there is often little point in attempting to pry them apart.[2] This account of scientific theories may also appear awkward, but much of the awkwardness can be avoided by an important terminological change. What Suppes, Sneed, Stegmueller, and Giere (in 1979) call a "theory," I shall call a "model." I shall then use the term "theory" for a set of connected lawlike assertions. Although terminological changes court confusion, this one is worth the risks, for it brings the language of this abstract discussion of scientific theories into close accord with the usage of economists and avoids the paradoxical denial that scientific theories make claims about the world.[3]

Economists use the term "model" in many ways (Machlup 1960, p. 569). Although some economic models are also models in other senses of the term, I know of none in theoretical economics which cannot be characterized as a predicate or as a definition of a predicate.[4] Taking models as definitions permits one to develop a cogent interpretation of economic models. Note that this sense of "model" is distinct from the logical positivist's notion. In their notion, a model is an interpretation of the sentences of a theory such that they all come out to be true. Models of the sort I am talking about, in contrast, are definitions and are constituted by sets of assumptions. They have nothing to do with the semantic interpretation of theories as sets of sentences. If the predicates

[2] Indeed the claim that they could be sharply separated would run afoul of Quine's critique of the analytic-synthetic distinction (appendix, p. 302).

[3] Suppes notes that scientists frequently use the term "model" as I shall to mean what Suppes means by "theory" (1957, p. 254). In the second edition of his *Understanding Scientific Reasoning* [1982], Giere changed his terminology in just the way I am recommending.

[4] Econometricians use the term "model" differently to contrast partially unspecified claims about some phenomena to fully specified "structures" (Marschak 1969). I shall not be concerned with the econometricians' notion of models.

models define have an extension - if actual systems satisfy these definitions - then with the proper interpretation of some "theory" in the logical positivist's sense - that is, of some set of sentences - the actual systems will be "models" in the positivist's sense of those sentences. But models in the sense of "model" specified here have no simple relation to models in the positivist's sense.[5]

Science consists not only of model making, but also of offering theoretical hypotheses that maintain that a model applies to the world. In defining a simple consumption system and offering the theoretical hypothesis that the quadruple consisting of Alice, coffee, the everything-else composite commodity, and Alice's income is a simple consumption system (pp. 34f), one is asserting that all the assumptions of the model are true of the relevant aspects of reality - that is, one is asserting that coffee is infinitely divisible, that Alice possesses a concave, increasing, and differentiable utility function, and so on.

From a theoretical hypothesis one infers what I call "closures" of the assumptions of the model (Hausman 1981a, pp. 47-8). The model Giere calls a "classical particle system" contains, for example, the assumption that any two bodies attract one another with a force inversely proportional to the square of the distance between them. Although the terms in the assumption are not uninterpreted, the assumption does not say what domain or system of entities it applies to. From the theoretical hypothesis that the solar system is a classical particle system, one can infer a closure of the assumption - that any two bodies in the solar system attract one another with a force inversely proportional to the square of the distance between them. In a closure of the assumption the domain is specified and in some cases the interpretation of the specific predicates in the assumption is sharpened. From a theoretical hypothesis one "recovers" the assumptions of the model as assertions about the world. A theoretical hypothesis entails closures of all the assumptions of the model. Closures of assumptions are genuine statements that are true or false.

For example, one might take the claims in chapter 1, that agent's preferences are complete, continuous, and transitive and that agents choose the option they most prefer, as providing a model of rationality. In doing so, one is just defining rationality. One is not saying that people's preferences are in fact complete, continuous, or transitive. One is not saying whether people are really utility maximizers. All one is doing is defining one notion of rationality. Having done only that, one has said nothing about the world, but, if the model is fecund, one has provided the means for making assertions about the world. One might, for example,

[5] I am indebted to Mark Bauder here.



discover that in certain domains people are not rational, or one might maintain that people are largely rational in certain sorts of decision-making activities. The latter claim is, of course, equivalent to saying that with respect to those decision-making activities people's preferences are complete, continuous, and transitive and they choose the option they most prefer. But formulating the model not only provides a useful abbreviation, it makes possible conceptual, logical, and mathematical explorations of the consequences of rationality so defined. Every model is, in a sense, a detour, but some models are very useful detours that greatly increase our conceptual resources.

The differences between models and theories can be displayed in the following table:

| Models | Theories (Descriptions, explanations, predictions) |
|---|---|
| definitions of predicates, concepts, or systems | sets of lawlike assertions |
| trivially true or neither true nor false | true or false |
| point is conceptual exploration | point is to make claims about the world |
| assess mathematically or conceptually, untestable | assess empirically testable |
| consist of assumptions | consist of assertions |

A model plus a *general* theoretical hypothesis asserting that the assumptions of the model are true of some portion of the world results in a theory. Some theoretical hypotheses, on the other hand, state that a particular real-world system, such as the solar system or the quadruple, <Alice, coffee, everything-else, Alice's income> belongs to the extension of the predicate defined by the model. When a theoretical hypothesis is such a singular statement, one might call the resulting set of closures of the assumptions of the model *an applied or restricted theory.* To say that New York stock brokers are rational is to offer an applied or restricted theory; one is asserting that the predicate defined in the model of rationality applies to a particular hunk of the world. Some restricted theories have much narrower scope than others, and indeed it may sometimes be misleading to speak of "theories."

Philosophers are sometimes attracted to the predicate view of theories (which I am calling "models"), because they are instrumentalists (A.2). They see the goal of theorizing not as discovering theoretical truths, but as discovering tools that enable one to predict and to control phenomena. From an instrumentalist perspective, one virtue of the predicate view of theories is that it permits one to avoid judging whether Newton's law of gravitation, for example, is a universal law. Instead one can merely judge, case by case, whether it is true of particular ensembles of bodies.

Although instrumentalists may thus make use of a predicate view of theories, they are mistaken if they think the view constitutes an argument for instrumentalism. Theoretical hypotheses need not be restricted to singular claims about individual systems. The closures of assumptions they imply may be general laws. Adopting a view of models as predicates or as definitions of predicates does not itself commit one to any thesis concerning the aims of science or whether general theoretical claims may be true.

Furthermore, instrumentalists are on dangerous ground if they tie their instrumentalism to a strategy of restricting the scope of generalizations. The methodological injunction to seek generalizations with broad scope is an important part of scientific practice. It explains, for example, why unsuccessful tests of a generalization are taken as casting doubt on the generalization rather than as merely revealing the limits to its scope. Without seeking broad scope and regarding successful generalizations as achieving it, how could scientists or engineers ever rely on laws in domains in which they have not been specifically tested?

In summary, models are definitions of kinds of systems, and they make no assertions. It is a category mistake to ask whether they are true or to attempt to test them. Their point lies in conceptual exploration and in providing the conceptual means for making claims that can be tested and can be said to be true or false. Theories are sets of systematically related lawlike statements. They do make true or false assertions about the world, and they can sometimes be tested. When one offers a general theoretical hypothesis asserting that something is the kind of system defined by a model, then one is enunciating a theory. The kind of assertion that results when one offers a theoretical hypothesis depends on the kind of theoretical hypothesis. A model may be used to state a general theory, to explain or to predict, or merely to state a fact about an individual.

## 5.3 Theories and models in economics

One might wonder what purpose this detour through the predicate view of theories has served. Since the activities of theorizing and of exploring



models are constantly intertwined in fact, why bother with what I am calling "models," instead of considering theories directly?

Developing theoretical knowledge is not just discovering correlations among properties that are already understood. An absolutely crucial part of the scientific enterprise – a part that was underemphasized by the logical positivists – is the construction of new concepts, of new ways of classifying phenomena. Even extremely simple models, such as the model of a simple consumption system, provide such concepts.

Concepts or terms are important to empirical scientists only insofar as they may enable them to say informative things about the phenomena under study. But scientists may nevertheless wish partly to *separate* questions concerning their conceptual apparatus from questions concerning the extent to which that apparatus applies to the world. That is, they may sometimes wish to investigate the properties of models without worrying about whether those models depict or apply to any aspect of reality.

In defining a model of a simple consumption system and in proving that the individual's consumption will lie at the point of tangency between some indifference curve and the budget constraint, one is not making any claims about the world. Nor need theorists regard themselves as revealing mysterious hypothetical truths concerning hypothetical situations. They are merely constructing concepts and employing mathematics and logic to explore further properties which are implied by the definitions they have offered. Such model building and theorem proving does not presuppose that one believes that the particular model is of any use in understanding the world. An economist might, for example, be intrigued with a mathematical question or attempt to discredit certain assumptions by revealing their consequences.

Insofar as one is only working with a model, one can dismiss any questions about the realism of the assumptions one makes. But remember that the reason is that one is saying *nothing* about the world. The irrelevance of questions about the realism of the assumptions to the mathematical investigation of properties of models has nothing to do with any questions concerning the assessment of scientific theories. Empirical assessment is out of order simply because there is nothing to assess: no empirical claims have been made.[6] Insofar as one is only working with a model, one's efforts are purely conceptual or mathematical. One is only developing a complicated concept or definition.

[6] In this discussion I am thus not in any way joining in Friedman's (1953c) or Machlup's (1955) defenses of "unrealistic assumptions" discussed below in chapter 9.

Max Weber's "ideal types" can, I think, be construed as models in the sense presented here. In a famous passage, Weber introduces the notion of an ideal type as follows:

> This conceptual pattern brings together certain relationships and events of historical life into a complex, which is conceived as an internally consistent system. Substantively, this construct in itself is like a utopia which has been arrived at by the analytical accentuation of certain elements of reality. Its relationship to the empirical data consists solely in the fact that where market-conditioned relationships of the type referred to by the abstract construct are discovered or suspected to exist in reality to some extent, we can make the *characteristic* features of this relationship pragmatically *clear* and *understandable* by reference to an *ideal-type.* This procedure can be indispensable for heuristic as well as expository purposes. The ideal typical concept will help to develop our skill in imputation in *research*: it *is* no "hypothesis" but it offers guidance to the construction of hypotheses. It is not a *description* of reality but it aims to give unambiguous means of expression to such a description....In its conceptual purity, this mental construct (*Gedankenbild*) cannot be found empirically anywhere in reality. It is a *utopia.* Historical research faces the task of determining in each individual case, the extent to which this ideal-construction approximates to or diverges from reality, to what extent for example, the economic structure of a certain city is to be classified as a "city-economy." (1904, p. 90)

Although Weber's ideal types fit my general characterization of models, they have special features, too. "Laws" play a lesser role than in models such as Giere's "classical particle system." What is important to Weber is not only more detail, but the specification of a sort of *system* or *entity,* not an abstract nomological structure. Most economists are less concerned with historical detail than was Weber and most are willing to use the term "model" sometimes much as Giere does. For example, most economists would find nothing strange or awkward in regarding continuity, completeness, transitivity, and utility maximization as a model of rationality. But economic models are generally definitions of (hypothetical) economies or markets, not of purely nomological concepts.

This observation may be of some general importance, for it suggests that the unit of theoretical analysis in economics is frequently not laws or theories but their *application* to particular ensembles of agents, markets, and institutions. Models are not themselves empirical applications, but they have the same structure. Economists are often concerned with developing applications of theory, not theory itself (section 6.4); and they are concerned with particular, albeit often stylized, circumstances. In these regards they are more like chemists than physicists (A.14).

Indeed, models in economics serve many purposes and are of many kinds. Models such as the simple consumption system of section 2.4 are crutches or pedagogical devices rather than conceptual innovations. Such models (which I have elsewhere called "special case" models (1981a, pp. 48–51)) simplify features of more general models and make them vivid. They are particularly useful for illustrating or evaluating more general models. "Model" is a particularly apt term for such constructions, because they resemble descriptions of the physical models that engineers build. Just as one can illustrate, develop, teach, and test claims about the properties of airplanes by means of scale models, so one can illustrate, develop, teach, and test features of theories and general models by means of special case models. But the use of special case models to assess theories and general models is controversial (Hempel 1965, p. 165; Popper 1968, pp. 442–56). For, unlike wind-tunnel tests on airplane models, for example, special case models do not provide us with occasions for the acquisition of new perceptual beliefs. They only help us to bring to bear the beliefs we already have.



The fact that theoretical economics is devoted to the exploration of models does not distinguish economics from other sciences. In theoretical work, *all* scientists attempt to exclude the complications of reality. As Galileo showed, the only good way to theorize is to develop models (1632, 1638). But, largely because of the possibility of creating simplified experimental circumstances, closures of assumptions in models in the natural sciences may often be regarded as truths of different degrees of universality. Model building in the natural sciences thus appears to be less distinct from empirical investigations.

In economics the problems of application are thornier. Even though models in economics need not be as abstract as those which characterize mainstream theorizing, they will never apply to economic reality cleanly. Insofar as one has any hopes for economic theory at all, there will always be some need to divorce conceptual development and empirical application. "Unrealistic" model-making is unavoidable for theoretically inclined economists.

The distinction between models and theories helps one to understand the attitude of economists toward what I called "equilibrium theory." Most are uncomfortable thinking of consumer choice theory, the theory of the firm, and the claim that equilibrium is attained as lawlike assertions that are either true or false. They prefer to think of these "laws" as "behavior postulates," as the most fundamental assumptions of the discipline, not as assertions. Given the obvious difficulties in regarding these claims as laws, one can sympathize with this attitude, and many economists are not committed to the truth of all these "behavioral postulates."

But questions of assessment must ultimately be faced. If economists did not believe that there was a great deal of truth to these "laws," if they only worked with "the basic equilibrium model" without any commitment to "equilibrium *theory*," then their practice would be mysterious. Unless one is to conclude that economists are uninterested in explaining or predicting economic phenomena, one must attribute to them either a commitment (much qualified and hedged) to the truth of the claims of equilibrium theory, or the conviction that the conclusions would still follow if the false assumptions were replaced with true ones (A.2).

As these last paragraphs and indeed the first four chapters suggest, it is unhelpful to regard neoclassical economics as a collection of separate and unconnected models or theories. Without understanding what unites and directs specific theoretical endeavors, one understands very little about economic theorizing. There are more global questions about theorizing in economics to which we now need to turn.

# 6 The structure and strategy of economics

Over the last two decades philosophers interested in scientific theory have been concerned not only to improve the logical empiricist's view of scientific theories, but to supplement it with accounts of the broader structures which shape individual theories and are in turn shaped by particular theoretical achievements. The best known of these accounts have been presented by Thomas Kuhn and Imre Lakatos.[1] Before offering my own abstract characterization of the structure and strategy of economic theorizing, let me consider whether their accounts can help with this task.

## 6.1 Disciplinary matrices

Although few philosophers of science are satisfied with his particular formulations,[2] Thomas Kuhn (1970) deserves credit for first devoting sustained attention to such "metatheoretical structures," which he initially called "paradigms," then "disciplinary matrices" (1970, Postscript, 1974). Disciplinary matrices are the constellation of beliefs, presumptions, heuristics, and values that tie together the theoretical efforts of practitioners of some discipline. When Kuhn speaks of a "discipline" or a "community," he has in mind specific theoretical enterprises which involve perhaps a few dozen scientists. But I shall not be stretching his remarks in an unusual way if I take them as applying to microeconomics as a whole.

In Kuhn's view, disciplinary matrices consist of four main components: (1) "symbolic generalizations," (2) metaphysical and heuristic commitments, (3) values, and (4) "exemplars." Symbolic generalizations

[1] Other philosophers have offered significant theories of global theory structure, but their work has had little influence in economics. Laudan's (1977) account should be of more interest to economists because of his emphasis on conceptual problems. See also Shapere 1974, 1984, 1985. Morgenbesser's distinctions between schemata and theories and between theories *of* a subject matter and theories *for* a discipline (1956, chapter 1) anticipate much of this later discussion.

[2] For criticism, see Scheffler 1967, Shapere 1964, and Suppe 1977. For an account of the ambiguities of the term, "paradigm," see Masterman 1970.