The Methodology of Econometrics

Kevin D. Hoover
First Draft, 3 January 2005
Revised, 15 February 2005

Department of Economics
University of California
One Shields Avenue
Davis, California 96616-8578
U.S.A.

Tel. (530) 752-2129
Fax (530) 752-9382
E-mail kdhoover@ucdavis.edu

Prepared for the Palgrave *Handbooks of Econometrics, volume 1: Theoretical Econometrics*.

Abstract of the

Methodology of Econometrics

by

Kevin D. Hoover

The methodology of econometrics is not the study of particular econometric techniques, but a meta-study of how econometrics contributes to economic science. As such it is part of the philosophy of science. The essay begins by reviewing the salient points of the main approaches to the philosophy of science – particularly, logical positivism, Popper's falsificationism, Lakatos methodology of scientific research programs, and the semantic approach – and orients econometrics within them. The principal methodological issues for econometrics are the application of probability theory to economics and the mapping between economic theory and probability models. Both are raised in Haavelmo's (1944) seminal essay. Using that essay as a touchstone, the various recent approaches to econometrics are surveyed – those of the Cowles Commission, the vector autoregression program, the LSE approach, calibration, and a set of common, but heterogeneous approaches encapsulated as the "textbook econometrics." Finally, the essay considers the light shed by econometric methodology on the main epistemological and ontological questions raised in the philosophy of science.

Keywords: econometrics, methodology, philosophy of science, Popper, Lakatos, semantic approach, Haavelmo, vector autoregression, LSE approach, calibration, Cowles Commission

JEL Codes: B41, C10

The Methodology of Econometrics

I. What is Methodology?

Methods vs. Methodology

The etymology of *methodology* implies that it is the study of methods. Taken narrowly it would neatly describe the typical econometrics textbook or the subject matter of most econometrics journals. And of course, it is not infrequently used this way, as when an article on an applied topic has an “econometric-methodology” section. Yet, “methodology” has come to have a broader meaning in the philosophy of science – both in the physical and life sciences and in economics and the social sciences. Methodology is not the study of particular methods but a meta-study of the ways in which particular methods contribute to the overall scientific enterprise.

Mark Blaug (1992, p. xii) defines the *methodology of economics* as

> . . . a study of the relationship between theoretical concepts and warranted conclusions about the real world; in particular, methodology is that branch of economics where we examine the ways in which economists justify their theories and the reasons they offer for preferring one theory over another; methodology is both a descriptive discipline – "this *is* what most economists do" – and a prescriptive one – "this is what economists *should* do to advance economics" . . .

Recent economic methodology has taken a "naturalistic turn"; it stresses the descriptive rather than the prescriptive function. (The subtitle to Blaug's book is the thoroughly descriptive "Or How Economists Explain," despite its frequently prescriptive content.)

The naturalistic turn is partly a reaction to a long-standing charge from economists that those who cannot do economics turn to methodology (e.g., Fisher 1933; Hahn 1992a, b). The examples of methodological work by great practitioners – John Stuart Mill, Lionel Robbins, Milton Friedman, and Paul Samuelson, among others – exposes the charge as a canard (see Hoover 1995b).

Nevertheless, until very recently much methodology has been concerned with questions such as what is the line of demarcation between science and non-science, that seem distant to the quotidian concerns of economists. The place of methodology is to address the larger conceptual and philosophical questions that arise in the everyday practice of economics. The naturalistic turn is partly an attempt to realign methodological thinking with that practice.

The naturalistic turn arises out of the humility of methodologists, who are unwilling to prescribe to economists. Ironically, the greater attention to the fine details of a field that followed in the wake of the naturalistic turn provides a sounder basis for prescription than previously. The philosopher Alexander Rosenberg asserts what might be called the "continuity thesis": the practice and methodology of economics or other sciences are not different fields but the same field addressed from different levels of generality:

> If theory adjudicates the rules of science, then so does philosophy [methodology]. In the absence of demarcation, philosophy is just very general, very abstract science and has the same kind of prescriptive force for the practice of science as any scientific theory. Because of its generality and abstractness it will have less detailed bearing on day-to-day science than, say, prescriptions about the calibration of pH meters, but it must have the same kind of bearing. [Rosenberg 1992, p. 11]

This history of econometrics in the 20th century (see Epstein 1987; Morgan 1990; Hendry and Morgan 1995) is full of methodological discussions. The key documents of the systematization of the field in the 1940s and early 1950s, Haavelmo (1944) and the Cowles Commission volumes (Koopmans 1950; and Hood and Koopmans 1953) are particularly rich. From the 1950s to about 1980 econometricians tended to focus on technical developments (methods) rather than larger conceptual issues (methodology). Interest in econometric methodology has been reborn in the subsequent period in a variety of articles (e.g., Hendry 1980; Sims 1980; Leamer 1983; McAleer, Pagan, and Volcker 1985; Pagan 1987; Aldrich 1989; Hoover 1990, 1994a,b, 1995a, b; Kydland and Prescott 1991; and Spanos 1995,) and books (e.g., Spanos 1986, 1999; Darnell and Evans 1990; Lawson 1997; Cartwright 1999; Magnus and Morgan 1999; Granger 1999; Keuzenkamp 2000; Stigum 2003; and Zellner, Keuzenkamp, and McAleer 2001).

What is Econometrics?

While econometrics may go back as far as the work of Davenant and King in the 17th century, it did not come into self-consciousness as a separate field until the founding of the Econometric Society in 1933. The society defined econometrics as “economic theory in its relation to statistics and mathematics” and its object as the “unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems” (cited by Frisch 1933, p. 1). The central problem for the field over the next seven decades has been just how to combine economic theory, mathematics, and statistics.

The term *econometrics* has come to refer mainly to the statistical apex of the economic theory-mathematics-statistics triangle, but it is a statistics that is centrally conditioned by economic theory. Econometric methodology could mean merely the methodology of the statistics that happens to be used in economics; yet it typically means something beyond that.

Perhaps the number one concern of a methodology of statistics is the proper interpretation of probability and how it applies to data. Here, the debate between classical and Bayesian interpretations of probability is central. This is, of course, an important issue for econometric methodology as well, yet we will largely ignore it here. First, it predates econometrics as a separate field, and so belongs to a largely methodological or philosophical discussion. Second, it is addressed in the articles on Bayesianism elsewhere in this volume. And third, it is not the issue that distinguishes econometrics from other uses of statistics in non-economic contexts.

So how does econometrics differ from statistics applied to economic subjects? There are at least two answers, each controversial in its own way. The philosopher Nancy Cartwright gives a clear statement of the first. Econometrics, she believes, provides a uniquely revealing application of statistics because, unlike, say, sociology, “economics is a discipline with a theory” (Cartwright 1989, p. 14). Cartwright (1999, ch.

7) argues that probabilities are not there for the taking, but are characteristics of quite particular set-ups (e.g., of roulette tables or particular configurations of unstable atoms). Only in such designed set-ups do objects display well behaved probabilities. The role of economic theory (and of quantum-mechanical theory) is to provide the conditions that articulate such a well-defined set-up: a *nomological* (or law-generating) *machine*. Cartwright’s views resonate with those of many econometricians, who believe that economic theory must provide the identification needed to render statistics economically interpretable. They also suffer from the problem of all *a priori* approaches: we must have the *right* theory to define the nomological machine or to identify the model, but if the inferential direction runs only from theory to data, how could we ever use empirical evidence to determine which theory is right? (See Hoover 2001a, ch. 4; 2001b, lecture 2; 2002 for critical accounts of Cartwright’s views).

The economist James Heckman provides a second answer for what distinguishes econometrics from statistics: econometrics focuses on establishing causation, while statistics is content with correlation. Heckman writes:

> *Most econometric theory* adapts methods originally developed in statistics. The major exception to this rule is the econometric analysis of the identification problem and the companion analyses of structural equations, causality, and economic policy evaluation. [Heckman 2000, p. 45, emphasis added.]
>
> . . .
>
> *The major contributions* of twentieth century econometrics to knowledge were the definition of causal parameters . . . the analysis of what is required to recover causal parameters from data . . . and clarification of the role of causal parameters in policy evaluation . . . [Heckman 2000, p. 45, abstract, emphasis added.]

Although drawing on the same insights as Cartwright about identification conditions, Heckman’s own applied work makes it clear that the inferential path runs from data to theory. Identification secured through “natural experiments” is used to establish which causal links ought to be reflected in the theory.

The idea that econometrics is a science of causes is attractive (see Hoover 1990, 2001a), but it is almost certainly unhistorical (Hoover 2004). Most econometricians through most of the post-World-War-II period have subscribed to a vision closer to Cartwright’s than to Heckman’s. Still, in the last twenty years, there has been a causal revival among both micro- and macro-econometricians.

Econometric Themes

The different approaches to distinguishing econometrics from statistics suggest two thematic questions that will help to organize our discussion. First, do social sciences generally, and economics particularly, provide suitable material for the application of probability?

A long tradition in economics – clearly expressed in Mill (see Hausman 1992, ch. 6), Marshall (see Stigler 1999, ch. 1), and Robbins (1937) – holds that social relationships are too complex, too multifarious, and too infected with capricious human choice to generate enduring, stable relationships that could be modeled by tractable probability distributions. These views were an important element of Keynes’ (1939) criticism of Tinbergen’s (1939) early econometric business-cycle model.

Although these views did not put a stop to the statistical analysis of economic data in the pre-World-War-II period, they were reinforced by developments within statistics itself, such as Yule’s (1926) analysis of nonsense regressions. Although we now view Yule as having laid the groundwork for later developments in cointegration analysis, the discovery that high correlations in time series were frequently artifacts appeared to undercut the goal of applying probability theory to economic data.

The theory of probability, especially in the work of R.A. Fisher (e.g., 1930, 1935), had come to be seen as intimately tied to the design of experiments. Stigler (1999, ch. 10), for example, attributes the early adoption (relative to economics) of probabilistic statistics in psychology, especially by the polymath philosopher C.S. Peirce in the 1880s, to the development of a well-defined program of psychological experimentation. Although experimental economics is *now* a flourishing field, economics – of its very nature – provides a more restricted scope for experimentation than many other human, life, and physical sciences. The dominant place of *a priori* theory in economics is partly a reflection of the paucity of experimental opportunities.

The second thematic question then concerns theory: how do econometric procedures applied to empirical data relate to economic theory? We will postpone further consideration of the first thematic question until section III. This question, however, ties directly into broader issues in the methodology and philosophy of science. We turn to these next.

II. The Place of Econometrics in a Scientific Methodology

What Can Econometrics Do?

One of the central questions in the philosophy of science concerns the relationship of empirical evidence to theoretical understanding. Econometrics naturally stands on the evidential end of this relationship. There are at least four roles for econometrics.

First, the most obvious is that econometrics is used to *test* an implication of a theory.

Second, econometrics may be used to *measure* unknown values of theoretically defined parameters or unobservable variables. In the extreme case, we might think of econometrics as giving flesh to a phenomenal law – that is, directly measuring a basic relationship posited by economic theory.

These two roles place theory ahead of evidence. In the first case theory proposes, evidence disposes. In the second, theory is essential to define or identify the object of measurement.

Third, econometrics may be used to *predict* the value of a variable. Prediction may be based directly on a prior economic theory or it may be an atheoretical statistical exercise. Prediction presumes a background uniformity that warrants projection of a relationship out of sample. This is, in itself, a weak assumption. Still, a theoretical account is frequently offered to buttress the distinction between accidental and genuine regularities. Without it, econometricians might be seen as little removed from stock-market chartists.

Fourth, econometrics may be used to *characterize* a relationship or phenomenon. It packages the data in a way that reveals relationships that, in turn, become the fodder for theory.

The Main Approaches to Scientific Methodology

However econometrics may used as a source of evidence, the question of the relationship of evidence to theory is one that applies broadly to science, not simply to economics.

(Excellent general discussions of the philosophy of science are found in Newton-Smith 1981, Caldwell 1982, Blaug 1992, and Hausman 1992.)

The approaches to scientific methodology that appear most relevant to econometrics are variants of, or successors to, *logical positivism*, a philosophical school that grew out of the Vienna Circle in the 1920s, a group of philosophers, and physical and social scientists, including Otto Neurath, Herbert Feigl, Karl Menger, Kurt Gödel, and Rudolph Carnap, among others. Positivists viewed scientific knowledge as having two sources: deductive inference from indisputable axioms and inductive inference from empirical data. Following David Hume (1777), they dismissed beliefs not founded on these two bases as unscientific metaphysics.

Hacking identifies six instincts that characterize logical positivist methodologies:

> (1) An emphasis upon *verification* (or some variant such as *falsification*): Significant propositions are those whose truth or falsehood can be settled in some way. (2) *Pro-observation*: What we can see, feel, touch, and the like, provides the best content or foundation for all the rest of our non-mathematical knowledge. (3) *Anti-cause*: There is no causality in nature, over and above the constancy with which events of one kind are followed by events of another kind. (4) *Downplaying explanation*s: Explanations may help to organize phenomena, but do not provide any deeper answers to *Why* questions except to say that the phenomena regularly occur in such and such a way. (5) *Anti-theoretical entities*: Positivists tend to be non-realists, not only because they restrict reality to the observable but also because they are against causes and are dubious about explanations. . . (6) Positivists sum up items (1) to (5) by being *against metaphysics.* [Hacking 1983, pp. 41-42]

Variants on logical positivism came to dominate Anglo-American philosophy of science by the 1960s. Although it has been the object of substantial criticism and revision, especially since Thomas Kuhn’s *The Structure of Scientific Revolutions* (1962), the sensibility of logical positivism provides the implicit philosophical background to most empirical economics. We now consider some of the main variants and descendants of logical positivism.

A. The Received View

Sometimes known as the *covering-law model of explanation* or as the *hypothetical-deductive method*, the *received view* understands scientific theories as networks of scientific laws, which are themselves understood as true, universal generalizations (Hoover 2001b, lecture 2). Explanations then take a *deductive-nomological form*: the relevant (or "covering") set of laws, taken with a set of initial conditions, allows us to deduce an empirically relevant conclusion. If the conclusion is yet to be observed, the deduction provides a prediction; if it is already known, an explanation. Explanation and prediction are symmetrical: explanation is just prediction directed to the past.

The received view agrees with Hacking's six logical-positivist instincts nearly perfectly. It nonetheless is not suitable in its pure form for many sciences including economics. Actual economic behavior is influenced by more factors than theory can typically account for. We often say, then, that the theory holds *ceteris paribus*. Too often, however, *ceteris paribus* clauses act as unfulfilled promissory notes: not knowing just what to hold constant, theoretical conclusions are conditional on factors that are not only unspecified, but unimagined.

To avoid the pitfalls of *ceteris paribus*, advocates of the received view propose a weaker, but more applicable, desideratum: the *inductive-statistical explanation*, in which laws hold only probabilistically and the inferences are not to what happens, but to the probability that it happens. The role for econometrics in such explanations is obvious.

B. Popper’s Falsificationism

Advocates of the received view typically used evidence positively: a theory predicts an outcome; if the outcome occurs, the theory gains support; the more support, the more warranted our belief in the theory. Such verificationism, however, falls foul of one of the oldest problems in epistemology – Hume’s (1739, book II, part III) riddle of induction: on what grounds can we justify a universal generalization from a collection of particular cases (e.g., how can we legitimately infer that the demand curve slopes down just because, when $price_j$ rose, demand for $good_j$ fell, for a large number of *j*’s.

Numerous solutions to Hume’s riddle rely on auxiliary universal generalizations, such as the uniformity of nature, that themselves stand in need of justification. Economists are inclined to view the downward-sloping demand curve (accounting separately for income effects) as a theorem, based on axioms that are known more or less intutitively. But unless we are willing to accept the subjectivist approach of the extreme Austrians (e.g., Mises 1966) that places the main principles of economics on a par with the axioms of mathematics – that is, unless we are willing to regard economics as a non-empirical science – then Hume’s riddle is not easily dismissed.

Hume’s objection to induction has an analogue in deductive logic. The following invalid syllogism illustrates the *fallacy of affirming the consequent*: *A* implies *B*; *B* is true; therefore *A* is true. The philosopher Karl Popper (1959) sees all induction as an application of this fallacy. Hume’s riddle, he believes, cannot be solved; it can only be dissolved by adopting another strategy. In place of the invalid syllogism, Popper proposes a valid one – *modus tollens*: *A* implies *B*; *B* is false; therefore *A* is false.

Evidence, no matter how many positive cases are collected, cannot establish the truth of a theory; it can establish only its falsehood. We should not adopt theories because they enjoy a high level of affirmative support, but instead reject those that are falsified. Science proceeds not through generalization from data but through a series of theoretical conjectures, tests, and empirical rejections.

Popper’s strategy is clearest when the deductions from a theory are deterministic. Yet, as we already saw in the case of the received view, economics almost certainly requires statistical rather than deterministic conclusions. But in that case, what counts as a falsifying instance? Very quickly we must appeal to the conventions of statistical testing, such as the critical value. We can now see that the question of which convention to adopt (should we use the ubiquitous 5 percent test? or 1 percent? or 10 percent? or *X* percent?) is no longer a merely technical question but one that lies near the heart of a key methodological issue. Popper does not resolve it.

Even if we could agree practically on what to count as falsification, a number of commentators have observed that economists rarely play the test-and-reject game consistently. Blaug (1992, p. 241) likens the research strategy of the typical economist to “playing tennis with the net down.” Summers (1991) and Keuzenkamp and Magnus (1995) claim that no scientifically significant proposition has every been decided on the basis of a statistical test. Where Blaug (1992, p. 244) calls for a redoubled commitment to serious, as opposed to “innocuous falsificationism,” Summers, and Keuzenkamp and Magnus reject the falsificationist strategy altogether.

Another key problem with falsificationism is that it rules theories out without giving guidance on which theories to use. Suppose we entertain two theories with some importantly different implications and neither have been rejected. Popper's view is that all theories are conjectural and never proven, but which should we use in the meantime? Popper argues that the theory with the richest content, in the sense that it rules out the most possible states of the world and is, therefore, easiest to reject, ought to be taken up for testing. Intuitively, the theory that survives the most rigorous tests is the best corroborated and, as a methodological rule, ought to be used even while we continue to try to reject it. Popper's (1959, appendix *ix) attempts to provide a formal measure of the degree of corroboration suffered from apparently insoluble technical problems that cast the whole strategy into doubt (Newton-Smith 1981, pp. 59-65; Keuzenkamp 2000, p. 60-62).

The problem is worse than even this perhaps technical failing. Theories are generally incomplete, and testing cannot proceed without auxiliary hypotheses. The failure of a test cannot be directed to the core theory (or some particular proposition within it) but acts on the whole complex of theory and auxiliary propositions. As Duhem (1906) and Quine (1951) have famously pointed out, there is considerable latitude in how the whole complex is adjusted in the face of contrary evidence. If all the adjustments are made to auxiliary hypotheses, then the core theory may never be threatened, even by an unambiguous falsification. Popper gives us little useful guidance on how to proceed constructively.

C. Lakatos's Methodology of Scientific Research Programs

A measure of the degree of corroboration is intended to guide the choice among conjectured theories. But how do we come to these conjectures? Popper's methodology offers no guidance and relegates theory construction to the realms of psychology, aesthetics, and unstructured imagination. In general, falsificationism does little to connect empirical evidence to the positive development of theoretical understanding.

In part to address these issues, Popper’s student Imré Lakatos (1970) proposed the *methodology of scientific research programs*. Lakatos viewed science not as a set of competing theories but as a set of competing programs, each defined by a set of essential propositions, known as the *hard core*. Every theory within a program consists of this hard core, a set of *negative* and *positive heuristics* for theory development, and a set of propositions that extend the hard core into a complete theory known as the *protective belt*.

In Lakatos’s view, Popper’s methodological rule to reject a theory that is falsified is useless because every theory is falsified on some dimension. Instead, he proposes to judge the success of a research program both by what it explains (novel facts predicted, anomalies resolved) and by what it fails to explain (anomalies discovered or reinstated). Programs develop by adjusting the protective belt according to the methodological guidance provided by the positive and negative heuristics, always leaving the hard core intact.

One program is superior to another when it explains the anomalies of the other program and successfully predicts more novel facts. Importantly, the standard is not static. A program is *progressive* and ought to receive support (for example, in the form of funding and professional esteem) when it shows ever increasing ability to predict novel facts. A program that fails to keep up or one that becomes mired in constructing accounts of its own anomalies without sufficient payoff in novel predictions ought to lose ground relative to more progressive programs.

More than with falsificationism, the methodology of scientific research programs employs all four of the possible uses of econometrics cited earlier: testing; measurement or instantiation of a phenomenal law; prediction; characterization or discovery of empirical relationships or phenomena. Yet Lakatos's methodology was only briefly popular in economics, starting in the mid-1970s (Latsis 1976). It has proved too difficult to fit the activities of economists into a set of competing programs with clearly differentiated hard cores: Is all neoclassical economics a single program? Or do we count general equilibrium theory, game theory, and macroeconomic theory as separate programs? If so, how do we account for cross-fertilization among programs that appears to suggest that no proposition is really immune from revision? (See Hoover 1991a.) Equally, an uncontroversial definition of the key notion of the "novel fact" has proved problematic (Hands 1993a). In the end, Lakatos's methodology has provided a useful vision of the scientific process while failing to offer a secure basis for prescription on how to advance the field (see particularly the variety of assessments in Blaug and De Marchi 1991).

D. The Semantic Approach to Scientific Theory

Although Lakatos's methodology is prescriptive, it is also naturalistic in the sense that it attempts to provide a positive account for the constructive activities of scientists. An increasingly popular methodological account is provided by the *semantic approach to scientific theories*. The received view was a *syntactic* account in that the central action concerned the deductive relationship between theories as formal systems and their implications. Once one agreed to certain theoretical axioms, the deductive work could be done algebraically, as it were, without reference to the meanings of the key terms. Advocates of the semantic approach reject such a content-neutral approach to science.

The key notion in the semantic approach is the *model*, in the sense of the content of a formalization. For example, a formal system might involve the minimization of an objective function subject to a constraint. A model for that system might involve identifying its terms with factor inputs, output, and various prices. A formal system might have more than one model. In the classic accounts of the semantic approach, a theory is a formal system plus the set of models with which it is consistent (Suppes 1962; Suppe 1989).

Recently, Davis (2000, 2005) and Chao (2002, 2005) have explored the role of econometrics in a semantic approach to economics. Perhaps the most thorough account is given by Stigum (2003). Stigum emphasizes that modeling takes place on two levels, which he regards as separate “worlds.”

The world of theory deals with theoretical models with crisp concepts. Theoretical deductions generate only theoretical conclusions.

The world of data is the realm of econometrics and statistics. Data are what they are, but the interesting properties of data are the relationships among them, which are not immediately evident but must themselves be modeled. Ideally, one would want to know the true probability distribution that governs the data. In practice, a particular probability model may be a more or less successful as an approximation; but its success must be judged on grounds other than correspondence to the truth, since that yardstick is not directly accessible.

The world of theory and the world of data are not immediately commensurable. Stigum and other advocates of the semantic approach argue that they must be connected by bridge principles that stipulate the mapping between theoretical variables, such as national output, unemployment, or money supply, and their statistical counterparts, such as GDP as reported in the national income and product accounts, the official unemployment rate, or M1. Bridge principles will not always be as straightforward as mere assignment of a data category to a theoretical concept. A theoretical concept such as the natural rate of unemployment, for instance, can be tied to data only through commitment to a model-dependent measuring strategy.

The object of the semantic approach is neither exclusively the falsification nor the verification of competing theories – although each may play a role; rather, it is the construction of richer models, more complete at both the theoretical and data levels and more consistent between the levels. This is not a mechanical enterprise, since the failures of the program may be remedied by adjustments to either the theoretical or data models or to the bridge principles that connect them.

Stigum's account belongs to the tradition that embeds models into detailed logical formalisms and relies on logical theory to check consistency. A "softer" model-theoretical account is provided in the work of Giere (2000), Cartwright (1999), and the contributors to Morgan and Morrison's (1999) (Boumans 1999, in particular, addresses economics directly). One of the primary advantages of such a softer account is that it is more naturalistic, drawing more immediately on, and more immediately applicable to, the practices of workaday scientists, economists, and econometricians.

III. Do Probability Models Apply to Economic Data?

A model-theoretic methodology makes it easier to cast the question of the applicability of the theory of probability to economic data into high relief. The ordinary practices of econometrics assume with little reflection that probability models apply quite naturally to econometric data. It is hard to recapture the once widespread resistance to accepting this as natural. The issues at stake in the earlier debates are, however, by no means settled and account for some of the continuing methodological differences.

The Probability Approach to Econometrics

Trygve Haavelmo’s “Probability Approach to Econometrics” (1944) is the *locus crucis* of econometrics and can be seen as marking the beginning of its modern period. Before Haavelmo, a long tradition questioned the applicability of probability models to economic data. Histograms of raw economic data rarely displayed a bell curve, and many economists doubted that economic behavior was sufficiently uniform to be beat into the mould of any standard probability model. Economics was not, of course, unique in this respect. But in other fields controlled experimentation was used to generate data that could be suitably described using probability models. Fisher’s work on experimental design, especially for agricultural applications, dominated statistical thinking in the first half of the 20^{th} century (e.g., Fisher 1935). Fisher’s view was that, without a controlled experiment, a probability model was simply inappropriate.

Econometricians nonetheless plowed ahead conducting ever more statistical investigations. Still, the absence of an acceptable methodological foundation for their activities undermined confidence in their results. Haavelmo’s insight was that properly accounting for the naturally occurring variations in economically important factors could act as a surrogate for the missing explicit experimental controls. If a regression equation were properly specified, then the residual errors would necessarily conform to a well-defined probability distribution. Such conformity, then, provides a test of proper specification.

Economic theory ideally explains which factors are the appropriate ones in an econometric specification. If the articulation of the theory and the data are compatible, then the observed data can be seen as independent realizations of the actual, complicated underlying process that generates the data – what econometricians refer to as the data-generating process and what Cartwright (1999, ch. 3) calls a nomological machine. When statistical account is taken of the appropriate observed factors, repeated, controlled experiments are not necessary, even in a time-series context.

Regression, rather than, say, correlation or analysis of variance, is the natural statistical technique in Haavelmo's program, because its coefficient estimates assign values to the importance of each factor that corresponds to the relationships implied by economic theory. The correspondences are, of course, not necessarily straightforward. Haavelmo saw the data-generating process (to use the modern term) as a complex system, so that measurement of the influence of individual factors (e.g., a price elasticity or the marginal propensity to consume) demanded attention to the system characteristics of the data-generating process. This *identification problem* had already been addressed by a variety of economists, including Haavelmo (1943) himself (Morgan 1990, ch. 6; Hendry and Morgan 1995, part III). Now, however, he had exposed its conceptual roots.

Haavelmo clearly believed that a complex economic system truly lay behind observed economic data. He nonetheless insisted that econometric work does not treat reality directly but, instead, involves our own constructions (models) that may vary in scope. The scope of our model depends in part on our particular interests and in part on the need to identify autonomous relationships – that is, ones that remain stable in the face of natural or intentional interventions (Aldrich 1989). Haavelmo’s implicit orientation to viewing theory and statistics in terms of models places his work in substantial sympathy with the semantic approach to scientific theories.

Responses to Haavelmo

Before Haavelmo, linear regression models had been developed in the context of several distinct traditions. In Fisher’s approach to experimental design, regressions were used to sort out the effects of independently varying factors in controlled and randomized experiments. The emphasis was on testing whether or not each of these factors had a hypothesized effect. Data mining was taboo because it violated the assumptions of randomness and independence. An experiment might be redesigned and rerun, but the data should be analyzed only in accordance with its original design. Information gathered from experimentation could test particular hypotheses, and the results of the tests could be used to inform the construction of theories. A clear route existed from data to theory.

In contrast, the theory-of-errors approach, traceable to Gauss and Legendre, took regression to be a tool of approximation of known theoretical relationships applied to empirical data, such as observations of planetary motions (Stigler 1986, chs. 1 and 4;

1999, ch. 17). In astronomical work, the underlying theory was presumed to be true; the role of the regression was not to test but to measure in the face of unavoidable observational errors. Data mining is acceptable as a means of obtaining better approximations.

Both the experimental-design and the theory-of-errors approaches have informed econometricians reactions to Haavelmo. On the one hand, “textbook” econometrics takes theory as a surrogate for experimentation, which bows considerably to Fisher. On the other hand, if theory is to stand in the place of experimental controls, it must be the right theory. Where Fisher was free to redesign and rerun experiments that failed on some statistical criteria, theory cannot be easily redesigned, as the evidence presupposed its truth in the first place (what Leamer 1978, p. 4, refers to as the “axiom of correct specification”).

Spanos (1995) argues that, as a result of this commitment to prior theory, “textbook” econometrics has adopted the strategy of reacting to failures of errors to reflect statistical assumptions by modifying those assumptions rather than by modifying theory. The use of Cochrane-Orcutt transformations of error terms in the face of serially correlated error terms provides one of many examples of the “robust-estimation” approach. This approach treats all violations of the statistical assumptions that one would expect to see from a well-designed experiment as failures to model the error terms correctly rather than, as Fisher would have it, a failure to institute adequate controls.

Paradoxically, such a strategy tends to enshrine a one-way relationship from theory to data more appropriate to the theory-of-errors approach than to the experimental-design approach. Hoover (1994b) refers to this strategy as “econometrics as measurement.”

Another approach, which Spanos (1995) refers to as “probabilistic reduction” and Hoover (1994b) as “econometrics as observation,” reacts differently to Haavelmo (also see Spanos, chapter ??, in the current volume). (Spanos traces the approach to the biometric tradition of Galton and Pearson.) A complete, true theory necessarily induces desirable statistical properties in data: independent, serially uncorrelated, white noise. The object of econometrics, therefore, should be to find compact representations of the data that deliver these properties without loss of information. These representations are the statistical regularities that theory must explain. Again, as with the experimental-design approach, a clear path is opened from data to theory. Where the robust-estimation approach seeks to mitigate the effect of any deviations from desirable statistical properties, the probabilistic-reduction approach seeks to characterize the systematic elements of the statistical model in such a way that the desirable properties arise naturally. Such an approach is not only compatible with data mining; it requires it.

IV. The Main Econometric Methodologies

Haavelmo’s probability approach emphasizes the relationship between economic theory and statistical modeling of data. Different econometric methodologies can be classified according to different roles that they assign to theory and to the degree of independence from theory that they assign to characterizations of data.

The Cowles Commission

The work of the Cowles Commission in the late 1940s and early 1950s was an outgrowth of Haavelmo's seminal monograph (Koopmans 1950; Hood and Koopmans 1953). The Cowles Commission was particularly concerned with the mapping between theory and data – that is, with the *identification problem*. It worked out the theory of identification to a high degree of completeness. The general problem is illustrated by the example used in most econometrics textbooks: given that we have only realized data on prices and quantities and both supply and demand relate prices to quantities, how can we separately identify the supply and demand curve? The Cowles Commission solution relies on economic theory to propose restrictions on the form of estimated regressions that permit us to break down the observational equivalence between supply and demand curves. Most commonly these restrictions take the form of exogenous variables that appear in one equation but not in another.

As with Haavelmo, the Cowles Commission methodology was subject to alternative interpretations. In the "measurement-without-theory" debate with Vining, Koopmans, for instance, strongly maintained that theory must be prior to data (Hendry and Morgan 1995, ch. 43). Data could not be interpreted without theoretical presuppositions. Such an approach implied that the object of econometrics was purely one of measurement and not of exploration and discovery. Koopmans' position places the empiricist in a vicious circle: how do we obtain empirically justified theory if empirical observation can only take place on the supposition of a true background theory?

Not all members of the Cowles Commission group adopted Koopmans' hardline. Simon (1953), for instance, in a paper showing the equivalence of the identified system with a causally ordered system, argued that experiments (natural or otherwise) could be used to distinguish between otherwise observationally equivalent systems (Hoover 1990, 1991b, 2001a).

The profession was more influenced by Koopmans than by Simon on this point. From the 1950s through the 1970s, the estimation of models consisting of theoretically identified systems of equations was the gold standard of applied econometrics. Much of the work in theoretical econometrics focused on developing appropriate systems estimators.

Vector Autoregressions

Some critics, notably Liu (1960), noticed early on that the number of restrictions needed to identify large-scale macroeconomic models far exceeded the number that economic theory could be confidently relied upon to provide. The point was driven home in Christopher Sims’ “Macroeconomics and Reality” (1980), in which Sims referred to the restrictions typically employed by macromodelers as “incredible.”

Sims’ proposal was to remove the pretence of applying theoretical structure to the data and, instead, to use unrestricted systems of reduced form equations (or *vector autoregressions* or *VAR*s) to model the responses of variables to shocks. Each equation in a VAR system regresses one variable on its own lags and the lags of all the other variables. Such a procedure still requires a form of identification. The reduced-form errors are generally intercorrelated; distinct shocks for each equation require that they be orthogonalized. Sims proposed to “identify” the shocks using a Choleski decomposition to normalize the system. Such a transformation renders the covariance matrix of the error terms diagonal and establishes a hierarchy (a triangular or Wold-causal order) among the variables such that contemporaneous shocks to higher ordered variables feed through to lower ordered variables but not vice versa. Such orderings are arbitrary, in the sense that there are as many triangular orders as there are permutations of the variables (i.e., if there are n variables, there are $n!$ possible orders). What is more, orthogonalized shocks can be achieved in orderings that are overidentified (or non-triangular) – that is that have more restrictions than the $n(n-1)/2$ needed to establish a just-identified Wold-causal order.

Initially, Sims regarded the choice of causal order as unproblematic. But under criticism from Cooley and LeRoy (1985) and Leamer (1985), among others, Sims (1986) came to accept that different causal orders had different implications for impulse-response functions and for innovation accounting and were, therefore, analytically significant. Once the need to commit to a particular causal order was accepted, the VAR transformed to reflect a particular contemporaneous causal order became known as a *structural* VAR (*SVAR*).

The same issue arises, then, for Sims' SVAR methodology as for the Cowles Commission's structural-modeling methodology: what restrictions are to be imposed and what makes them credible? Standard economic theory rarely implies Wold-causal, or other recursive, orderings among variables – simultaneity is the norm. VAR practitioners have typically appealed to very informal, casual arguments to justify particular orderings.

The LSE Approach

The London School of Economics (LSE) methodology originated in the work of Denis Sargan and is now strongly associated with David Hendry and various colleagues and students widely dispersed among academic institutions, mainly in Britain and Europe (see Mizon 1995 for a systematic discussion). The LSE methodology is closely related to a wider range of work on integrated and cointegrated systems originating in the work of Engle and Granger at the University of California, San Diego and of Johansen and Juselius at the University of Copenhagen (Juselius 1999). Where Koopmans’ strongly apriorist version of the Cowles Commission’s methodology and Sims’ SVAR methodology belong more to the theory-of-errors approach, the LSE program is a species of the probabilistic-reduction genus. As with the VAR methodology, the LSE methodology stresses dynamic specification with special attention to the lag structures. There are, however, some key differences: it pays particular attention to stationarity and cointegration; and it is not content with profligate parameterizations but seeks parsimonious specifications that nevertheless deliver errors with good statistical properties (white-noise innovations).

The leading principle of the LSE approach is consistently to apply the theory of encompassing (Mizon 1984, Hendry 1988). Roughly, one specification encompasses another if it carries all of the information of the other specification in a more parsimonious form. An easy way to think about encompassing is to consider two competing specifications for the same dependent variable. Both can be nested in a joint model formed from the nonredundant set of regressors found in the two specifications. If one of the specifications is a valid restriction of this joint model and the other not, then the one *encompasses* the other. The LSE approach proceeds in a series of horse-races. Any specification is maintained only tentatively. Any proposed alternative specification is judged on its ability to encompass the reigning specification.

While encompassing is the key idea of the LSE methodology, most attention has been paid to Hendry's *general-to-specific modeling strategy*. The general-to-specific strategy derives in large measure from Hendry's vision of the economy as governed by a true data-generating process – the complex probability distribution that actually governs the realizations of economic variables. From this staring point, Hendry develops a theory of data reduction (Hendry 1995, ch. 9). The origins of the theory of reduction are found in the analysis of exogeneity in the seminal article of Engle, Hendry, and Richard (1983). The central question is always how can the data be characterized in a way that is partial or simpler than the true data-generating process without loss of information relative to the questions of interest. The theory defines the conditions under which the key steps in the reduction would be legitimate.

Practically, the general-to-specific approach involves starting with as broad a general specification as possible and then searching over the space of possible restrictions to find the most parsimonious specification. At each step in a sequential reduction (usually along multiple paths), the statistical properties of the errors are tested, the validity of the reduction is tested statistically both against the immediate predecessor and the general specification, and encompassing is tested against all otherwise satisfactory alternative specifications.

Despite strong advocacy among econometricians of the LSE school, the general-to-specific strategy is not an essential element of the methodology. The data-generating process assumed at the start of any search is "local" and not the true one. Its specification is based on common sense, the availability of data (both the particular variables and the number of observations, which dictates the degrees of freedom), and exploratory data analysis. Since there is no direct access to the true data-generating process, there is no way to demonstrate that the local data-generating process is itself a legitimate reduction of the true one.

It is always possible that an alternative specification exists that is not, in fact, nested in the local data-generating process. In this case, the local process is supplemented with the alternative specification to form a new local data-generating process, and the search is run again. But this is a specific-to-general strategy.

Although not essential, the general-to-specific strategy retains a strong heuristic justification. It ensures that the space of alternative specifications is fully explored, minimizing the danger that relevant competing specifications are ignored, and ensures that no information is lost relative to the general specification.

Critics of the LSE approach (Faust and Whiteman 1995, 1997) argue, among other things, that the general-to-specific approach is vitiated because it is a form of data-mining in which the large number of sequential tests render the reported test statistics uninterpretable. This objection is supported by studies of data-mining algorithms that show large size distortions (e.g., Lovell 1983). The LSE response is to note that there are two sources of error that need to be taken into account: (i) size distortions (the cost of search); and (ii) statistics based on a misspecified regression are not likely to reflect those of a correct specification (the cost of misspecification). Critics of data-mining tend to stress the first and ignore the second. Yet Monte Carlo studies of the efficacy of general-to-specific search shows that the costs of search are small: size tends to be close to the nominal size of the exclusion tests used in the search; power achieves a high fraction of the power given knowledge of the correct specification; and the ability to recover the true specification is high (Hoover and Perez 1999, 2004; Hendry and Krolzig 1999 and Krolzig and Hendry 2001).

The different conclusions reached about the effectiveness of data-mining appear to arise because of differences between the general-to-specific algorithm and the relatively simpler algorithms tested by Lovell and others (e.g., maximium R^2, step-wise regression, maxi-min t). These simpler search algorithms do not enforce any requirement that particular specifications pass reasonable tests of statistical adequacy. They are thus more susceptible to specification error.

A theorem due to White (1990, pp. 379-380) states that, for a fixed set of specifications and a battery of specification tests, as the sample size grows toward infinity and increasingly smaller test sizes are employed, the test battery will – with a probability approaching unity – select the correct specification from the set. White’s theorem implies that type I and type II error both fall asymptotically to zero. White’s theorem says that, given enough data, only the true specification will survive a stringent enough set of tests. The true specification survives precisely because the true specification is necessarily, in the long run, the fittest specification. (See Hoover and Perez 2000 for further discussion.) The theorem suggests that a stringent battery of tests should help to reduce the costs of misspecification; in practice, the Monte Carlo studies indicate that these dominate the costs of search.

Hendry refers to the LSE methodology and the general-to-specific strategy as a “progressive research strategy” with an explicit reference to Lakatos (Hendry 2000, pp. 116, 117, 363-364, and 440; cf. Mizon 1995). The methodology is Lakatosian in spirit in that it applies the encompassing principle repeatedly, so that the appropriate unit of appraisal is not, say, an individual regression model but a family of models incorporating the relevant information as it becomes available. Hendry stresses that the true test of a specification that has been constructed to meet statistical design criteria is its success on new data – much like Lakatos's requirement of successful prediction of novel facts. Though Lakatosian in spirit, the LSE approach has no particular commitment to the details of Lakatos's methodology, with its emphasis on hard core propositions, protective belts, and heuristics for development of suitable theories.

Like the VAR approach in its initial formulation, the LSE approach stands on the side of probabilistic reduction rather than the theory of errors. Theory plays a part in helping to define the range of variables that should be included in the local data-generating process and in choosing interpretable transformations of those variables (that is, as a bridge principle), but Koopmans' notion that a complete, *a priori* theoretical articulation must precede statistical investigation is rejected. Although data may be packaged in more or less illuminating ways, it is the job of theory in the LSE view to conform to, and explain, the facts of the data, not of data to conform to the presuppositions of theory.

Calibration

The calibration methodology is the polar opposite of the LSE methodology: it maintains a commitment to prior core economic theory above all. Calibration is largely associated with Finn Kydland and Edward Prescott's (1991) program of quantifying dynamic general-equilibrium macroeconomic models, though it is closely related to the methodology of computable general-equilibrium models common in trade, development, and taxation literatures (Mansur and Whalley 1984).

A calibrated model starts with a theoretical model – for Kydland and Prescott generally a representative-agent rational-expectations model of the business cycle or growth – and completes it by assigning numerical values to the key parameters. These values are not estimated through systems-equations methods according to the Cowles Commission program (as, for example, in Hansen and Sargent 1980). Instead, they are drawn from considerations of national-accounting, the "great ratios," unrelated statistical estimations, common sense, experience, and other informal sources. Once parameterized, the calibrated model is validated through simulation. Do the simulated data display patterns of covariances that adequately mimic the patterns found in the actual data? Once validated, calibrated models are used to explain historical economic performance and for policy analysis. (See Hartley, Hoover, and Salyer 1997; 1998, chapters 1 and 2, for a detail critical description of the calibration methodology.)

Many econometricians question whether calibration, with its rejection of statistical estimation, can be counted as an econometric methodology. Kydland and Prescott (1991) vigorously defend its standing as *bona fide* econometrics, arguing that it fits clearly into the original vision of the Econometric Society of econometrics as "economic theory in its relation to statistics and mathematics." Some econometricians have attempted to recast calibration into a form more in keeping with mainstream understandings of the term (e.g., Gregory and Smith 1991, 1993).

A calibration methodology is superior to econometric alternatives in Kydland and Prescott's view. First, they regard basic economic theory as established, secure knowledge that need not be tested. Second, they believe that the goal for quantitative economics should be the construction of artificial economies or models that sufficiently well mimic key dimensions of the actual economy that can be used as a test bed for counterfactual experiments. Third, they acknowledge that such models are only approximations to complex reality that will successfully mimic the actual economy only along a limited set of chosen dimensions. This limited focus, however, rules out standard likelihood-based statistical estimates of model parameters, since those methods penalize the models for failing to fit on dimensions irrelevant to the intended use of the models.

Although the notion that models built for one purpose may be inadequate for others is a sound one (Hoover 1995a), the calibration methodology raises some significant doubts. Kydland and Prescott reject the application of standard methods of statistical estimation to their models because they see them as an application of the Koopmans’ variant of the Cowles Commission methodology, which seeks to estimate directly completely articulated theoretical models. The purpose-built models that they advance are necessarily incompletely articulated and are, therefore, necessarily easy to reject according to Cowles Commission standards.[1] But this raises a key question never clearly answered by calibrationists: what is the standard for assessing the adequacy of the values assigned to the parameters of a calibrated model?

The second doubt is similar. Models are assessed by a comparison of descriptive statistics between simulated and actual data. The standard implicit in most calibration

[1] “Purpose-built” is the literal meaning of *ad hoc*, but this latter term has become such a term of abuse among economists that it is doubtful that Kydland and Prescott would willingly describe their theory-based models as *ad hoc* (see Hands 1993b).

exercises is “looks good to Ed.” The question of a non-subjective standard for judging, or even of guidance for assessing, the quality of the output of calibrated models has rarely – if ever – been addressed in the literature. A corollary is that there is no established method for adjudicating between the claims of competing calibrated models.

It is surprising that advocates of calibration methodologies have not noticed that their objections to the Cowles Commission methodology do not apply to the LSE methodology or to any probabilistic-reduction approach. An adequate statistical characterization of data may help to supply robust parameterizations for calibrated models. And Hoover (1994a) has proposed the application of the encompassing principle in a manner that would help to adjudicate among competing calibrated model while preserving Kydland and Prescott’s insight that a theoretical model may be a valuable tool of historical assessment and policy guidance even though incompletely articulated.

Textbook Econometrics

By the middle of the 1950s, the Cowles Commission program had seized the methodological high ground. The focus of econometric research shifted away from high level conceptual questions to ground-level concern with the development of particular estimators and tests. Still, Cowles Commission attitudes continued to dominate econometric thought, even in the face of the failure of systems-estimation of macroeconomic models or of microeconomic demand systems to live up to the promise of the Cowles Commission methodology.

Econometrics textbooks and many of the applications of econometrics to applied problems reverted to single-equation regression models. Much of the applied work was atheoretical or – at best – weakly theoretically justified. The robust-estimation approaches that were part of the post-Cowles Commission developments dominated textbook econometrics (Spanos 1995).

Other applied econometricians took the notion that a single equation was always embedded in a larger system more seriously, addressing it through the application of instrumental-variables estimators to obtain consistent estimates. In most such applications, Koopmans’ notion that theory must be prior to data dominates: the choice of instruments was guided by *a priori* considerations; the goal continued to be estimation in the theory-of-errors tradition.

Recently – especially in the applied labor-economics literature – a data-first approach has captured substantial support. The goal is to use “natural experiments” – e.g., changes in institutional arrangements – as instruments to help identify causal effects (Angrist and Krueger 2001). Here, contrary to Koopmans, the information obtained from statistical processing of the data restricts the class of acceptable explanatory theories.[2]

V. Some Key Issues in Econometric Methodology

The main philosophical approaches to science discussed in section II are variants (or, at least, descendants) of logical positivism. Most econometricians are positivists in the very broad sense of finding the source of scientific knowledge in either logical deductions from secure premises or in empirical observation. Yet few are logical positivists in the

---

[2] Time, space, and ignorance compel me to omit detailed discussion of another approach to econometrics: *simplicity*. See Keuzenkamp (2000, ch. 5) and Zellner, Keuzenkamp, and McAleer (2001) for detailed discussion.

sense of subscribing to all of Hacking's six characteristics cited in section II. Considering how econometrics and its different methodologies relate to Hacking's list provides a way of tying together the somewhat separate discussions of the main philosophical approaches to science with the main econometric methodologies.

1. *Emphasis on verification or falsification.* Essentially logical positivism is characterized as mainly concerned with the testing of theories. But econometrics has many goals that involve little or no testing of theories. Even an econometrician such as Hendry (1980, pp. 27-28), who says that the three golden rules of econometrics are "test, test, test," is not concerned so much with the direct test of theories as with achieving statistical adequacy. Econometrics in the Cowles Commission program is mainly concerned with the measurement of theoretically articulated parameters. And other approaches often seek to establish the best statistical models for data reduction or forecasting. The criticism of Summers (1991) that no econometric test ever decided an economic question or the challenge of Keuzenkamp and Magnus (1995) to produce an example of a case in which a statistical test was decisive for an economic question miss the point: neither verification nor falsification is typically the sharp result of single tests; rather empirical research is gradually redirected from lines less consonant to lines more consonant with the accumulated weight of evidence. Econometrics thus operates in a Lakatosian spirit, albeit with no commitment to the fine details of Lakatos's methodology.
2. *Pro-observation*. Virtually all econometricians could be described as pro-observation, but the important methodological question is what exactly counts as an observation. Raw data, such as collected by national statistical agencies, are not observations in the

relevant sense; for what economic theory predicts is typically relationships among these raw data, although forecasting along the lines of the covering-law model is also possible.

The Cowles Commission program (at least in Koopmans' variant) sees true observations only when adequately identified by *a priori* theory. Calibrationists, who are otherwise skeptical of the Cowles Commission methodology, share this understanding, which is what justifies the title of Prescott's paper, "Theory Ahead of Business Cycle Measurement" (1986).

The data-mining procedures of the LSE school can be seen as an effort to get good observations, to focus the telescope as it were (Hoover 1994b). Leamer's book, *Specification Searches* (1978) proceeds in a similar spirit. Leamer's (1983) *extreme-bounds analysis* tries to find observations that are robust in the sense that parameters of interest maintain the direction of their influence however their magnitude may change. Despite similar motivations, there remains a substantive argument over the relative efficacy of extreme-bounds analysis relative to the LSE methodology (McAleer, Pagan and Volcker 1985; Pagan 1987; Hoover and Perez 2004).

3. *Anti-cause*. Causality is principally about the structure of influence of one variable over another (Hoover 2001a). Up to the heyday of the Cowles Commission in the early 1950s, conceptual discussions of causes were commonplace and causal talk was ubiquitous (Hoover 2004). In the subsequent period, causal language nearly dropped out of econometric discussion. It revived somewhat with the development of Granger-causality tests. But "causality," as Granger uses it, is not closely related to structural notions (Granger 1980). The absence of causal language does not imply an

abandonment of causality. Cause is a diagnostic notion: when we flip a switch we would not normally say "I caused the light to come on"; but if the light failed to come on we would say "I wondered what caused that?" At the theoretical extreme, the calibrationist methodology is fundamentally about a commitment to models that reflect particular structures of economic influence. Though the word "cause" rarely turns up in calibrationist discourse, it takes a deeply causal view of the world and is in no sense anti-cause.

Zellner (1979) has advocated using the philosopher Feigl's (1953, p. 408) notion that cause as a shorthand for "predictability according to law" (also see Keuzenkamp 2000, p. 240). Unfortunately, such a proposal begs the question; for it assumes that there is a clear concept of a law – applicable to economics – that can stand as surrogate for causal structure and that prediction is the sole goal of econometrics, neither of which is the case (Hoover 2001a, chapter 4; 2001b, lecture 2).

4. *Downplaying explanation*s. The received view starts with universal laws and deduces predictions from them. The symmetry thesis holds that there is no fundamental difference between explanation and prediction. Yet if prediction is not the principal objective of applied econometric investigation, then some of the same problems arise here as they did in the discussion of causality: Are there economic laws? And, if so, where do they come from? The view that economic theory is given *a priori* to the econometrician, far from downplaying explanation, places emphasis on it. Elster (1994) argues that there are few, if any, universal empirical regularities, but that we can come to understand economic mechanisms that explain rather than predict. Such a view is compatible with the LSE and other probabilistic-reduction approaches. Their

object is not to seek universal regularities, but regularities that arise in particular institutional and temporal contexts, which may in turn become the object of theoretical explanation. The same general strategy underlies the natural-experiment approach in applied microeconomics, as well as the strategy of behavioral finance, which seeks to construct an account of financial behavior from experimental and other empirical observations rather than from *a priori* assumptions of economic rationality.

5. *Anti-theoretical entities*. To the true anti-realist or instrumentalist, theory is a convenient summary of the relationships among empirical data. Some econometricians entertain such positions. Keuzenkamp (2000, p. 216), for instance, maintains that the object of econometrics is not discovery but intervention leading to predictable outcomes. He deeply opposes the LSE methodology with the slogan, “the Data Generation Process does not exist” (Keuzenkamp 2000, p. 214). Others, such as Lawson (1997), are extreme realists who believe that that complexity of the economy renders even local stable regularities a virtual impossibility. Most econometricians appear to fall somewhere in between these extremes. Once again, those who insist on the primacy of *a priori* theory can hardly think of that theory as an instrumental data summary without a real referent; while those who seek probabilistic reductions generally do so on the assumption that the data-generating process does, in fact, exist. The two camps differ over how knowledge is to be made secure, but not over the ontological status of economic entities.

VI. Wither Econometric Methodology?

Hacking’s sixth point merely states that the first five taken together amount to a rejection of metaphysics on the part of logical positivists. Of the five, econometricians can be reliably expected to agree with one: the attitude of being pro-observation. It would be unreasonable, then, to deduce that econometricians are typically against metaphysics. Nonetheless, metaphysical questions rarely arise explicitly in econometric discourse. In fact, econometric methodology is a largely underdeveloped field in which practicing econometricians address methodological issues at best implicitly. That the field is ripe for substantial development is clear from two points that have arisen in this survey.

First, the main approaches to the philosophy of science, with the exception of the semantic approach broadly conceived, do not square well with the failure of the main econometric methodologies to conform to a broadly logical positivist program. One key research program, then, is to develop a methodology of economics which clarifies the role of econometrics in a larger philosophy of science.

Second, while the there are connections and continuities among the different econometric methodologies, there are also deep divisions that do not turn on merely technical matters. These divisions cry out for conceptual resolution – a task better suited to explicit methodological thought than to the tacit methodology found in the pages of *Econometrica.*

References

Aldrich, John. (1989) “Autonomy,” *Oxford Economic Papers*.41(1), 15-34.

Angrist, J. and A. Krueger (2001): “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments,” *Journal of Economic Perspectives* 15(4), pp. 69-85.

Blaug, Mark. (1992). *The Methodology of Economics: Or How Economists Explain*, 2nd edition. Cambridge: Cambridge University Press.

Blaug Mark and Neil de Marchi, editors. (1991) *Appraising Economic Theories: Studies in the Application of the Methodology of Research Programs*, Aldershot: Edward Elgar, 1991.

Boumans, Marcel. (1999) ”Built-in Justification”, in Morgan and Morrison (1999), pp. 66-96.

Caldwell, Bruce J. (1982) *Beyond Positivism: Economic Methodology in the Twentieth Century*. London: Allen & Unwin.

Cartwright, Nancy. (1989) *Nature's Capacities and Their Measurement*. Oxford: Clarendon Press.

Cartwright, Nancy. (1999) *The Dappled World*. Cambridge: Cambridge University Press.

Chao, Hsiang-Ke. (2002) *Representation and Structure: The Methodology of Econometric Models of Consumption*. Ph.D dissertation, University of Amsterdam.

Chao, Hsiang-Ke. (2005) “*A Misconception of Semantic Conception of Econometrics?*,” *Journal of Economic Methodology*, forthcoming.

Cooley, Thomas F., Stephen F. LeRoy (1985) “Atheoretical Macroeconometrics: A Critique,” *Journal of Monetary Economics* 16(3), 283-308.

Darnell, Adrian C. and J. Lynne Evans. (1990) *The Limits of Econometrics.* Aldershot: Edward Elgar.

Davis, George C. (2000) “A Semantic Conception of Haavelmo’s Structure of Econometrics”, *Economics and Philosophy* 16(2), pp. 205-28.

Davis, George (2005) “Clarifying the ‘Puzzle’ Between Textbook and LSE Approaches to Econometrics: A Comment on Cook’s Kuhnian Perspective on Econometric Modelling”, *Journal of Economic Methodology*, forthcoming.

Duhem, Pierre. (1906) *The Aim and Structure of Physical Theory*, Louis de Broglie, translator. Princeton: Princeton University Press, 1954.

Elster, Jon. (1994) “A Plea for Mechanisms,” unpublished manuscript, Oslo; cited in Stigum (2003).

Engle, Robert F., David F. Hendry and Jean-François Richard. (1983) “Exogeneity,” *Econometrica* 51(2), 277-304.

Epstein, Roy J. (1987) *A History of Econometrics*. Amsterdam: North-Holland.

Faust, Jon and Charles H. Whiteman. (1995) “Commentary [on Grayham E. Mizon ‘Progressive Modeling of Macroeconomic Times Series: The LSE Methodology’],” in Hoover (1995c), pp. 171-180.

Faust, Jon and Charles H. Whiteman. (1997) “General-to Specific Procedures for Fitting a Data-Admissible, Theory-Inspired, Congruent, Parsimonious, Encompassing, Weakly-Exogenous, Identified, Structural Model to the DGP: A Translation and Critique,” *Carnegie-Rochester Conference Series on Economic Policy,* Vol. 47, December.

Feigl, Herbert. (1953) “Notes on Causality,” in Herbert Feigl and Mary Brodbeck (eds.) *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts, pp. 408-418.

Fisher, I. (1933) “Statistics in the Service of Economics,” *Journal of the American Statistical Association* 28(181), 1-13.

Fisher, Ronald A. (1930) *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

Fisher, Ronald A. (1935) *The Design of Experiments.* Edinburgh: Oliver and Boyd.

Frisch, Ragnar. (1933) “Editor's Note.” *Econometrica* 1(1), pp. 1-4.

Giere, Ronald N. (2000) “Theories”, in W. H. Newton-Smith (ed.), *A Companion to the Philosophy of Science*, Oxford: Blackwell.

Granger, C.W.J. (1980) “Testing for Causality: A Personal Viewpoint,” *Journal of Economic Dynamics and Control* 2(4), 329-352.

Granger, C.W.J. (1999) *Empirical Modeling in Economics: Specification and Evaluation.* Cambridge: Cambridge University Press.

Gregory, Allan W. and Gregor W. Smith. (1991) “Calibration as Testing: Inference in Simulated Macroeconomic Models,” *Journal of Business and Economic Statistics* 9(3), 297-303.

Gregory, Allan W. and Gregor W. Smith (1993). “Statistical Aspects of Calibration in Macroeconomics,” in G.S. Maddala, C.R. Rao and H.D. Vinod (eds.) *Handbook of Statistics*, vol. 11. Amsterdam: North Holland, 1993, 703-719.

Haavelmo, Trgyve. (1943) “The Statistical Implications of a System of Simultaneous Equations,” *Econometrica* 11(1), pp. 1-12.

Haavelmo, Trgyve. (1944) “The Probability Approach in Econometrics,” *Econometrica* 12 (supplement), July.

Hacking, Ian. (1983) *Representing and Intervening*. Cambridge: Cambridge University Press.

Hahn, Frank. (1992a). “Reflections,” *Royal Economic Society Newsletter*, No. 77, April.

Hahn, F. (1992b). “Answer to Backhouse: Yes .” *Royal Economic Society Newsletter*, No. 78, July.

Hands, D. Wade. (1993a) “The Problem of Excess Content: Economics, Novelty, and a Long Popperian Tale,” in *Testing, Rationality, and Progress: Essays on the Popperian Tradition in Economic Methodology*. Lanham, MD: Rowman and Littlefield, pp. 125-41

Hands, D. Wade. (1993b) “Ad Hocness in Economics and the Popperian Tradition,” in *Testing, Rationality, and Progress: Essays on the Popperian Tradition in Economic Methodology*. Lanham, MD: Rowman and Littlefield, 83-99

Hansen, Lars Peter and Thomas J. Sargent. (1980) “Formulating and Estimating Dynamic Linear Rational Expectations Models,” in Robert E. Lucas, Jr. and Thomas J. Sargent (eds.) *Rational Expectations and Econometric Practice*. London: George Allen & Unwin, 1981.

Hartley, James E., Kevin D. Hoover and Kevin D. Salyer. (1997) “The Limits of Business Cycle Research: Assessing the Real Business Cycle Model,” *Oxford Review of Economic Policy* 13(3), 34-54.

Hartley, James E., Kevin D. Hoover and Kevin D. Salyer, editors. (1998) *Real Business Cycles: A Reader*. London: Routledge.

Hausman, D.M. (1992). *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.

Heckman, James J. (2000) “Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective,” *Quarterly Journal of Economics* 115(1), 45-97.

Hendry, David F. (1980) “Econometrics: Alchemy or Science?,” in Hendry (2000), pp. 1-28.

Hendry, David F. (1988) “Encompassing,” *National Institute Economic Review*, (August), 88-92.

Hendry, David F. (1995) *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, David F. (2000) *Econometrics: Alchemy or Science*, 2nd edition. Oxford: Blackwell.

Hendry, David F. and Hans-Martin Krolzig. (1999) “Improving on ‘Data Mining Reconsidered’ by K.D. Hoover and S.J. Perez,” *Econometrics Journal*, 2(2), 202-218.

Hendry, David F. and Mary S. Morgan, editors. (1995) *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press.

Hood, W.C. and T.C. Koopmans. (1953) (eds.) *Studies in Econometric Method*, Cowles Commission Monograph 14. New York: Wiley.

Hoover, Kevin D. (1990) “The Logic of Causal Inference: Econometrics and the Conditional Analysis of Causation,” *Economics and Philosophy* 6(2), 207-234.

Hoover, Kevin D. (1991a) “Scientific Research Program or Tribe? A Joint Appraisal of Lakatos and the New Classical Macroeconomics,” Blaug and De Marchi (1991).

Hoover, Kevin D. (1991b) “The Causal Direction Between Money and Prices: An Alternative Approach,” *Journal of Monetary Economics* 27(3), 381-423.

Hoover, Kevin D. (1994a) “Six Queries About Idealization in an Empirical Context,” *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 38, 43-53.

Hoover, Kevin D. (1994b) “Econometrics as Observation: The Lucas Critique and the Nature of Econometric Inference,” *Journal of Economic Methodology* 1(1), 65-80.

Hoover, Kevin D. (1995a) “Facts and Artifacts: Calibration and the Empirical Assessment of Real-Business-Cycle Models,” *Oxford Economic Papers* 47(1), 24-44.

Hoover, Kevin D. (1995b) “Why Does Methodology Matter for Economics?” *Economic Journal* 105(430), 715-734.

Hoover, Kevin D. (ed.) (1995c) *Macroeconometrics: Developments, Tensions, and Prospects.* Dordrecht: Kluwer.

Hoover, Kevin D. (2001a) *Causality in Macroeconomics.* Cambridge: Cambridge University Press.

Hoover, Kevin D. (2001b) *The Methodology of Empirical Macroeconomics.* Cambridge: Cambridge University Press.

Hoover, Kevin D. (2002) “Econometrics and Reality,” in Uskali Mäki (ed.) *Fact and Fiction in Economics: Models, Realism, and Social Construction.* Cambridge: Cambridge University Press.

Hoover, Kevin D. (2004) “Lost Causes,” *Journal of the History of Economic Thought*,” 26(2), 149-64.

Hoover, Kevin D. and Stephen J. Perez. (1999) “Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search,” *Econometrics Journal* 2(2), 167-191.

Hoover, Kevin D. and Stephen J. Perez. (2000) “Three Attitudes Towards Data Mining,” *Journal of Economic Methodology* 7(2), 195-210

Hoover, Kevin D. and Stephen J. Perez. (2004) “Truth and Robustness in Cross Country Growth Regressions,” *Oxford Bulletin of Economics and Statistics* 66(5), 765-798.

Hume, David. (1739) *A Treatise of Human Nature.* Page numbers refer to the edition edited by L.A. Selby-Bigge. Oxford: Clarendon Press, 1888.

Hume, David. (1777) *An Enquiry Concerning Human Understanding*. Page numbers refer to L.A. Selby-Bigge, editor. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 2nd edition. Oxford: Clarendon Press, 1902.

Juselius, Katarina. (1999) “Models and Relations in Economics and Econometrics,” *Journal of Economic Methodology* 6(2), 259-290.

Keuzenkamp, Hugo A. (2000) *Probability, Econometrics, and Truth: The Methodology of Econometrics.* Cambridge: Cambridge University Press.

Keuzenkamp, Hugo A. and Jan R. Magnus. (1995) “On Tests and Significance in Economics,” *Journal of Econometrics* 67(1), 5-24.

Keynes, John Maynard. (1939) “Professor Tinbergen’s Method,” *Economic Journal* 49(195), 558-568.

Koopmans, T.C. (1950) *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10. New York: Wiley.

Krolzig, Hans-Martin and David F. Hendry. (2001) “Computer Automation of General-to-Specific Model Selection Procedures,” *Journal of Economic Dynamics and Control.* 25(6-7), 831-66.

Kuhn, Thomas (1962) *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press, 1970.

Kydland, Finn E. and Edward C. Prescott. (1991) “The Econometrics of the General Equilibrium Approach to Business Cycles,” *Scandinavian Journal of Economics*, 93(2), 161--78

Lakatos, Imré. (1970) “Falsification and the Methodology of Scientific Research Programmes,” in Lakatos and Alan Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 1970.

Latsis, S. J. (1976) *Method and Appraisal in Economics*. Cambridge: Cambridge University Press.

Lawson, Tony. (1997) *Economics and Reality*. London: Routledge.

Leamer, Edward E. (1978) *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. John Wiley: Boston.

Leamer, Edward E. (1983) “Let’s Take the Con Out of Econometrics,” *American Economic Review* 73(1), 31-43.

Leamer, Edward E. (1985) “Vector Autoregressions for Causal Inference?” *Carnegie Rochester Conference Series on Public Policy* 22, 255-303.

Liu, T. (1960) “Underidentification, Structural Estimation, and Forecasting,” *Econometrica* 28(4), pages 855–65.

Lovell, Michael C. (1983) “Data Mining,” *The Review of Economics and Statistics*, 65(1), 1-12.

Magnus, Jan R. and Mary S. Morgan (eds.) (1999) *Methodology and Tacit Knowledge*. New York: Wiley.

Mansur, Ahsan and John Whalley. (1984) “Numerical Specification of Applied General-Equilibrium Models: Estimation, Calibration, and Data,” in Herbert E. Scarf and John B. Shoven (eds.) *Applied General Equilibrium Analysis*. Cambridge: Cambridge University Press.

McAleer, Michael, Adrian R. Pagan, and Paul A. Volker. (1985) “What Will Take the Con Out of Econometrics,” *American Economic Review*, 75(3), 293-307.

Mises, Ludwig von. (1966) *Human Action: A Treatise on Economics*, 3rd. edition. Chicago: Henry Regnery.

Mizon, Grayham E. (1984) “The Encompassing Approach in Econometrics,” in D.F. Hendry and K.F. Wallis (eds.) *Econometrics and Quantitative Economics.* Oxford: Basil Blackwell, pp. 135-172.

Mizon, Grayham E. (1995) “Progressive Modelling of Economic Time Series: The LSE Methodology,” in Hoover (1995c), pp. 107-170.

Morgan, Mary S. (1990) *The History of Econometric Ideas*. Cambridge: Cambridge University Press.

Morgan, Mary S. and Margaret Morrison (eds.) (1999) *Models as Mediators*. Cambridge: Cambridge University Press.

Newton-Smith, William. (1981) *The Rationality of Science*. Boston: Routledge & Kegan Paul.

Pagan, Adrian (1987) “Three Econometric Methodologies: A Critical Appraisal,” *Journal of Economic Surveys* 1(1), 3-24.

Popper, Karl. (1959). *The Logic of Scientific Discovery*. London: Hutchison.

Prescott, Edward C. (1986) "Theory Ahead of Business Cycle Measurement," Federal *Reserve Bank of Minneapolis Quarterly Review* 10(4), 9-22.

Quine, Willard V.O. (1951) "Two Dogmas of Empiricism," in *From a Logical Point of View*, 2nd ed. Cambridge: Harvard University Press, 1961.

Robbins, L. (1937) *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.

Rosenberg, A. (1992) *Economics: Mathematical Politics or Science of Diminishing Returns?* Chicago: Chicago University Press.

Simon, Herbert A. (1953) "Causal Ordering and Identifiability," in Herbert A. Simon, *Models of Man*. New York: Wiley 1957, ch. 1.

Sims, Christopher A. (1980). "Macroeconomics and Reality," *Econometrica* 48(1), 1-48.

Sims, Christopher A. (1986) "Are Forecasting Models Usable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review* 10(1), 2-15.

Spanos, Aris. (1986) *Statistical Foundations of Econometric Modelling*. Cambridge: Cambridge University Press.

Spanos, Aris. (1995) "On Theory Testing in Econometrics: Modeling with Nonexperimental Data," *Journal of Econometrics* 67(1), 189-226.

Spanos, Aris. (1999) *Probability Theory and Statistical Inference : Econometric Modeling with Observational Data*. Cambridge: Cambridge University Press.

Stigler, Stephen M. (1986) *The History of Statistics: Measurement of Uncertainty Before 1900*. Cambridge, MA: Belknap Press.

Stigler, Stephen M. (1999) *Statistics on the Table*. Cambridge, MA: Harvard University Press.

Stigum, Bernt P. (2003) *Econometrics and the Philosophy of Economics: Theory-Data Confrontations in Economics.* Princeton: Princeton University Press.

Summers, Lawrence H. (1991) "The Scientific Illusion in Empirical Macroeconomics," *Scandanavian Journal of Economics* 93(2), 129-48

Suppe, Frederick (1989) *The Semantic Conception of Theories and Scientific Realism*, Urbana: University of Illinois Press.

Suppes, Patrick (1962) "Models of Data", in Nagel, Ernest, Suppes, Patrick and Tarski, Alfred (eds.) *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford: Stanford University Press, pp. 252-61.

Tinbergen, Jan. (1939) *Statistical Testing of Business-Cycle Theories,* Vol. II: *Business Cycles in the United States of America, 1919-1932*. Geneva: League of Nations.

White, Halbert. (1990) "A Consistent Model Selection Procedure Based on *m*-testing," in C.W.J Granger (ed.) (1990) *Modelling Economic Series: Readings in Econometric Methodology*. Oxford: Clarendon Press, pp. 369-383.

Yule, George. U. (1926): "Why Do We Sometimes Get Nonsense Correlations Between Time Series? A Study of Sampling and the Nature of Time Series" (with discussion), *Journal of the Royal Statistical Society*, 89(1), 1-64.

Zellner, Arnold A. (1979) "Causality and Econometrics," in Karl Brunner and Allan H. Meltzer (eds.) *Three Aspects of Policy Making: Knowledge, Data and Institutions*, Carnegie-Rochester Conference Series on Public Policy, vol. 10. Amsterdam, North-Holland, pp. 9-54.

Zellner, Arnold, Hugo A. Keuzenkamp, and Michael McAleer (eds.) (2001) *Simplicity, Inference, and Modeling: Keeping It Sophisticatedly Simple.* Cambridge: Cambridge University Press.