CONFIRMING MAINSTREAM ECONOMIC THEORY†

Daniel M. HAUSMAN*

* Department of Philosophy, University of Wisconsin, 600 N. Park Street, Madison WI 53706, USA. E-mail: D.HAUSMAN@lse.ac.uk

BIBLID [0495-4548 (1998) 13: 32; p. 261-278]

ABSTRACT: This essay is concerned with the special difficulties that arise in testing and appraising mainstream economic theory. I argue that, like other theories designed to apply to complex open systems, it is very hard to confirm mainsteam economics. Parts can be tested and appraised, but the theory is only very weakly supported by evidence[1].

Keywords: Economic theory, mainstream, confirmation, testing, appraising, evidence.

CONTENTS

1. Equilibrium theory and its casual appraisal

The fundamental theory of mainstream economics, which I call "equilibrium theory" consists of four components. First there is a theory of rationality. People's preferences are rational if they are complete and transitive, and people's choices are rational if there is no feasible alternative they prefer to the one they choose. An agent *A*'s preferences are complete if for all alternatives *x* and *y*, either *A* prefers *x* to *y*, or *A* prefers *y* to *x*, or *A* is indifferent between *x* and *y*. *A*'s preferences are transitive if for all alternatives *x*, *y* and *z*, if *A* prefers *x* to *y* and *A* prefers *y* to *z*, then *A* prefers *x* to *z* (and similarly for indifference). If *A*'s preferences are complete and transitive and satisfy a further technical condition, then they can be represented by a utility function that assigns a larger number to *x* than *y* whenever *A* prefers *x* to *y* and assigns *x* and *y* the same number whenever *A* is indifferent between them. The condition on the rationality of choice, that *A* never chooses *x* if there is some feasible alternative *y* that *A* prefers to *x*, can be restated as the claim that *A* is a utility maximizer. This account of rational preference and choice says nothing about actual behavior. It becomes a part of "positive" economic theory when economists add the assertion that people are (at least to some extent) rational in this sense. The first component of mainstream economics thus asserts that people's preferences are in fact complete and transitive and that people's choices in fact follow their preferences.

The second component of equilibrium theory consists of two generalizations concerning the content of preferences: that people prefer larger commodity bundles to smaller ones and that people's preferences for commodity bundles show diminishing marginal rates of substitution -that is (roughly), people will be willing to pay less money for another unit of some commodity if they already have a great deal of that commodity than if they have very little of it. Notice that the generalization that people prefer larger commodity bundles to smaller bundles, coupled with the definition of a commodity, implies that individuals are self-interested. Their choices are determined by the size of their own consumption bundles.

The third component of equilibrium theory consists of two generalizations concerning firms. The first is the so-called law of diminishing returns: in the neighborhood of actual output, the first partial derivative of output with respect to every input is positive, and the second partial derivative is negative. With more labor on a fixed piece of land with fixed seed and fertilizer output can be increased, but each additional unit of labor has a smaller positive effect. The second generalization states that firms (or those who run them) seek to maximize net returns. It is also frequently assumed that there are constant or non-increasing returns to scale, but this assumption is arguably a simplification rather than a fundamental principle.

The fourth component of the theory is that equilibrium will be reached - that all markets will clear with no excess demand or, except at zero price, excess supply. This proposition is not stated as an axiom of the theory. When it shows itself explicitly, it is usually as the consequent of some theorem. But it is not a fact that just happens to be provable in most mainstream models. On the contrary the models are constructed so that the existence of equilibrium will follow.

How can equilibrium theory be tested? How can be it confirmed?[2]. Since its claims appear to concern matters of observation, one might suppose that one could simply look and see whether they are true or false. One can "look and see" if people's preferences are complete: Give people a choice between two alternatives, and they will usually pick one. People may have a hard time making up their mind, but they do not often say that they cannot choose. But they sometimes do, and it seems that preferences are not always complete. Choices are usually transitive, but sometimes a person *A* chooses *x* when given a choice between *x* and *y*, *y* when given a choice between *y* and *z*, and *z* when given a choice between *x* and *z*. In such cases either *A* does not choose what he or she prefers, or *A*'s preferences are not transitive. One can "look and see" whether people always prefer larger commodity bundles to smaller ones. Obviously people very often do. They hunt for bargains and respond to sales. Yet sometimes people pay higher prices at local stores and give money away to charities. In these cases, it seems that they are choosing a smaller bundle of commodities and money over a larger one. One can "look and see" whether firms maximize profits. Obviously firms often seize new opportunities to increase profits, but do they always equate marginal revenue and marginal cost? Surveys suggest that many do not calculate marginal revenue and marginal cost and that others have additional objectives, such as loyalty to their employees. Finally it seems hard to reconcile a world with high unemployment with the existence of equilibrium.

These remarks on the casual direct testing of equilibrium theory are, of course, terribly naive. Economists need not take people's reports -whether they concern their preference rankings or the objectives of their firms- at face value. Those paying higher prices for apparently the same goods may be in effect be purchasing better service or greater convenience. Charitable contributions may be investments in later sales, employment or insurance rather than deliberate choice of less consumption. Given pervasive uncertainties and decision-making costs, rules of thumb that appear to be at variance with profit maximizing may in fact implement it. Apparent unemployment may be consumption of leisure or extended job search. Casual observation is far from decisive.

But economists do learn something from casual observation, including observation of themselves. In particular, they learn that there is a good deal of truth to the fundamental propositions of equilibrium theory. The claims that people's preferences are transitive, that people prefer more commodities to fewer, that there are diminishing marginal rates of substitution, that there are diminishing returns in production, and that firms seek prefer larger net returns to smaller are not wildly at variance with the facts. They appear to capture significant causal factors at work in economic life. At the same time, economists also learn that these principles are not exceptionless universal laws. Sometimes there are intransitivities. Sometimes people care about other things than the size of their consumption bundle. Sometimes firms care about other things than profits.

2. *Confirming inexact "laws"*

If one held a crude view of science, one could take this last point as ending discussion. No matter how much "truth there is to" the principles of equilibrium theory, they are not true universal generalizations. Hence the purported "laws" of economics are not laws at all, and equilibrium theory belongs in the dust bin or back on the drawing board. Such a view is too crude. The claim that people are not satiated -that they prefer larger commodity bundles to smaller- states an important truth, even though the universal generalization "Everyone always prefers larger commodity bundles to smaller" is false. The proper task for the methodologist is to find out what sort of truth is being stated here, rather than to give up in disgust because economics does not fit a simplistic model of science.

There are a variety of ways of making precise the substance of the claim that people on the whole or typically prefer larger commodity bundles to smaller. This is not the occasion to discuss the complexities of interpreting such claims[3]. Instead I shall simply state my view and then explore its implications for the confirmation of equilibrium theory. I believe that one should interpret claims such as "People prefer more commodities to fewer" as carrying (implicit) *ceteris paribus* qualifications. What one is really saying is "Everyone always prefers more commodities to fewer *ceteris paribus* -that is, in the absence of interferences or disturbing causes. Economists can list some of those disturbing causes, but not all of them. The *ceteris paribus* qualification is therefore vague. But it is not so vague as to make the qualified claim empty, uninteresting, or untestable. There is no guarantee that economists will not find someone who does not prefer more commodities to fewer; yet for whom no disturbing cause can be found. In J.S. Mill's view, which has been reiterated forcefully by Cartwright (1989, ch. 4, 5), the pursuit of more commodities is one cause of behavior and it results in a *tendency* that is visible when other competiting causal factors are absent or cancel one another out.

I believe that my proposal is the best way to make precise the claim that the pursuit of more commodities is a causal tendency: one should take the claim that people prefer more commodities to fewer to be qualified with a *ceteris paribus* clause. Whether or not this interpretation is accepted, one has to be wary of the view, to which Mill himself was attracted, that everyday experience conclusively establishes that the pursuit of more commodities is a genuine causal tendency. This goes to the opposite extreme compared to the simplistic philosopher, who would toss out equilibrium theory because its principles are not true universal generalizations. Everyday experience does not show how *important* this alleged tendency is or to what extent it explains generalizations such as the law of demand. Nor does everyday experience demonstrate the correctness of the *ceteris paribus* law. Confirming equilibrium theory requires a great deal more than the observation that people seem pretty keen on accumulating commodities.

The problems involved in confirming equilibrium theory are in this way beginning to take shape. First, and foremost, the principles of equilibrium theory are *inexact.* They state tendencies, and I construe this to mean that they carry vague *ceteris paribus* qualifications. Since the principles of the theory are inexact (or if, one prefers, statements of tendencies only), the implications of the theory will not be exact either. Consequently, one cannot refute the theory by finding data that are inconsistent with what one deduces from unqualified statements of the principles along with initial conditions. Second, whatever course of testing one has in mind, it is important to recognize that the theory begins with a good deal of initial plausibility. The theory is not a set of wild conjectures. It consists of statements such as "people's preferences are transitive" and "people prefer more commodities to fewer," which economists know to correspond to a great deal of experience.

The philosophical theory of theory assessment is an area of contemporary controversy. Bayesian views appear to be in ascendency, but they are subject to serious criticisms. There is no philosophical consensus concerning how theories should be appraised, which a philosopher of economics can rely upon. I shall instead examine how different views on assessment apply to economics, criticize those that are implausible or unhelpful, and see what one can learn from what remains.

Like most philosophers, economists, and indeed "the man or woman in the street," I am an empiricist about theory assessment: the evidence that ultimately leads one to accept or to reject claims about the world is observational evidence. Economists believe that individuals generally prefer more commodities to fewer because this claim is borne out by experience, and if this claim were not "supported" by experience, economists would be inclined to give it up. What is it for a claim to be "supported" by experience -or in this sense confirmed? Two different questions should be distinguished:

1. The problem of evidence: How does observational evidence provide *any* confirmation or disconfirmation (no matter how weak) of scientific hypotheses?

2. The problem of acceptance or choice: When are hypotheses strongly confirmed or disconfirmed on the basis of the results of observation and experiment?

My discussion will focus on answers to first question.

3. The hypothetico-deductive method

The dominant view of how one tests scientific theories used to be the so-called "hypothetico-deductive (HD) method." (Notice that the name may mislead: this is an *inductive* method.) Reduced to its bare bones, this method consists of the following four steps:

1. *Formulate* some hypothesis or theory *H*.

2. *Deduce* some "prediction" or observable claim, *P*, from *H* conjoined with a variety of other statements. These other statements will include descriptions of initial conditions, other theories, and *ceteris paribus* ("other things being equal") clauses.

3. *Test P*. (One tests *H* only indirectly by means of the HD method.) Testing may involve complicated experimentation or simple observation.

4. *Judge* whether *H* is confirmed or disconfirmed depending on the nature of *P* and whether *P* turned out to be true or false. *"Confirmed" does not mean "proven" or "true," nor does "disconfirmed" mean "disproven" or "false," for false hypotheses may have true implications, and the falsity of)* P *may be due to some premise from which)* P *is derived other than* H.

These steps can be amplified or modified in a variety of ways. For example, in the case of statistical theories, it may not be possible to *deduce P* or, if it is, the testing of *P* may be problematic. If one seeks *good* evidence, as opposed to just *some* evidence, one will look for predictions that background knowledge or alternative theories render unlikely (see Giere 1983). But the four steps listed above capture the essential features of hypothetico-deductivism. A "prediction" is any testable implication. It need not be about the future and its truth may already be known. The HD method is an account of how evidence supports or disconfirms some hypothesis -however strong or weak that support may be. It is not an account of what it is for a hypothesis to be *well-supported*.

Karl Popper defends a radical variant of hypothetico-deductivism, which denies that theories can ever be confirmed. They remain forever conjectural. All one can ascertain is whether or not they pass tests, and even that judgment is risky, since success or failure may be due to the additional statements needed to derive the prediction, not to the hypothesis itself. Progress in science depends on a willingness to reject theories that fail tests and to subject those that pass to ever more rigorous tests. Popper's view is untenable. It is inconsistent with the actual conduct of science and absurd as an account of how science should proceed. Rather than elucidating what confirmation is, it denies that confirmation exists; and it consequently demands that scientists never make use of information concerning how well confirmed theories are. Although Imre Lakatos rejected many of Popper's theses, he remained an adamant critic of what he called "justificationism" - the view that evidence could to some extent confirm theories. His account of assessment in science is consequently as untenable as Popper's[4].

The above sketch of the HD method enables one to formulate more precisely the special problems of theory assessment in economics. Suppose one wants to test a principle of equilibrium theory such as "people prefer more commodities to fewer" or an implication of equilibrium theory such as the law of demand. The law of demand states that a change in the price of some commodity or service causes (*ceteris paribus*) a change in quantity demanded in the opposite direction. When the price of gasoline goes up, people will demand less of it. From (a) the law of demand, (b) a statement describing a price change, (c) a *ceteris paribus* assumption, and (d) various assumptions about the reliability of the statistical data one is using, one can deduce a prediction about demand data. And one can then observe whether the prediction is true. Although there are practical problems in carrying out the first three steps of the HD method, there seems to be no fundamental philosophical difficulty in implementing them.

The point of the HD method lies, however, in step 4, in deciding whether the evidence supports the hypothesis, and, ideally, to what extent.

It is at this step that thorny problems in assessing economic theories show themselves. Suppose one finds that price and demand both decrease. Such apparently disconfirming data are readily available. Ought one to regard the law of demand as disconfirmed? Hardly. For demand also depends on other factors. That is why the law of demand states only that a change in price will, *ceteris paribus*, cause a change in quantity demanded. Given the multitude of "disturbing causes" in economics and the difficulty of performing controlled experiments to weed these out, it seems that little can be learned from experience.

The general philosophical difficulties facing the HD method mainly involve the notion of "evidential relevance." The hypothesis to be tested must play an essential role in the second, deductive step, or else, trivially, any hypothesis can be confirmed by conjoining it to some confirmable theory. But it is not easy to spell out adequately the notion of an "essential role." One might say that h is confirmed by e only if h and some theory T imply e, but T does not imply e by itself (Schlesinger 1976; Horwich 1978). Unfortunately, this condition can be satisfied trivially. Just take any hypothesis h and any true observation report, e, and let T be the true "theory": "if h then e." Then h is essential to the deduction of e from T, but not confirmed by e. In their search for criteria of "cognitive significance" the logical positivists devoted considerable effort to such problems, only to conclude eventually that they had no formal solution.

One appealing way to improve upon the HD method is summarized by the slogan that confirmation involves "inference to the best explanation." On this view, the second step of the HD method is the source of the problem. Rather than merely *deducing* some proposition P, one looks for a proposition P that the hypothesis *explains* better than any alternative does. The truth of P then confirms the hypothesis. This account of confirmation relies heavily on the theory of explanation, which is a troubled area of philosophy. If explanation is conceived of as deductive-nomological (Hempel 1965), then confirmation as inference to the best explanation collapses into the general HD method. No other account of explanation is generally accepted.

Although the problems of evidential relevance are serious, they seem to be "merely philosophical." Real economists do not cook up arbitrary theories such as "If efficiency-wage theory is true, then some apples are red" in order to defend efficiency-wage theory with irrelevant evidence. But the problem cannot be dismissed, for a philosophical account of the relationship between theory and evidence ought to explain *how* scientists avoid these pitfalls of the HD method. In addition, the HD method provides no good way of explicating how evidence may be regarded as relevant only to particular *parts* of scientific theories. When predictions are faulty, *where* does the blame lie, and, when predictions are successful, *which* particular hypotheses should take the credit? The HD method tempts one to an unhelpful and untenable holism that assigns praise and blame only to whole amalgams of theories and auxiliary assumptions.

Both the specific difficulties about employing the HD method in economics and the general philosophical difficulties of evidential relevance are linked to the so-called Quine-Duhem Problem. Pierre Duhem, particularly in *The Aim and Structure of Physical Theory* (1906), pointed out that one never tests significant scientific propositions by themselves. As the HD schemea illustrates, testing an hypothesis involves deriving a prediction from a conjunction of many propositions, of which the hypothesis is only one. Even if one could capture formally the requirement that the hypothesis be essential to the deduction, there would still be the problem that a predictive failure could be due to the falsity of one of these other propositions. Consequently, one can always "save" any given hypothesis by casting the blame on some other claim. Moreover, if one takes the further step, which Quine endorses, of rejecting the distinction between analytic and synthetic statements and the notion of necessary truth, then the predictive failure could be due to a "mistaken definition" or perhaps even to the use of the "wrong logic."

If the Quine-Duhem problem is posed as a purely logical difficulty, then it may not be in practice very serious. For example, careful weighing of metals and their oxides refutes the view that metals are a compound of their oxides and phlogiston -but not if phlogiston has negative mass. In principle, one can continue to hold on to the theory of phlogiston and maintain that phlogiston has negative mass. In practice, the evidence against phlogiston is compelling. But if, as in economics, one is unable to place much confidence in the other premises needed to derive a prediction P from an hypothesis H, then there is a serious practical problem. To get a definite prediction concerning aggregate market behavior from a principle or implication of equilibrium theory usually requires a long list of implausible simplifications concerning knowledge, the shape and stability of utility functions, the divisibility of commodities, the speed of adjustment to shocks, and so forth. These subsidiaries assumptions are individually and collectively so dubious that a successful prediction provides negligible confirmation for the principle or implication one is testing, and an unsuccessful prediction gives one little reason to regard the principle or implication as disconfirmed. It becomes almost impossible to learn from experience.

Many economists mistakenly believe that Milton Friedman provides a way out of these difficulties. Friedman regards the goals of economics and of science in general as practical. Economists should seek theories that will provide valid and substantive predictions concerning the phenomena they are interested in. Since economists are interested in market behavior, they should pay no attention to the results of surveys or to other tests of the "assumptions" of their theories. The only question they should be concerned with is whether their theories provide valid predictions concerning market phenomena.

This advice ignores the difficulties and in fact aggravates them. If economists look only at market data, they will find that some theories provide roughly accurate predictions (which occasionally go astray) in some circumstances. Such data will not tell them how to diagnose predictive failures and it will not tell them whether to rely on a particular hypothesis in some new application. Friedman is urging economists to restrict their attention to data that bear very weakly on their theories and to ignore any other data. Economists who follow his advice will never acquire good reason to rely on their theories in new applications or to improve them. Economists need either other less ambiguous data or some better way of bringing market data to bear on their theories.

4. Bayesian philosophy of science

Actual testing and appraisal in science makes heavy use of substantive scientific commitments. Scientists generally know what phenomena a hypothesis ought to account for and what data are relevant. Arbitrary "theories" such as "if *h* then *e*," where *h* is any hypothesis and *e* is any true observation report are never formulated and would not be taken seriously if they were. By paying attention to heuristic rules that are central to "paradigms" or "research programs," one can perhaps compensate for the formal weaknesses of the HD method. Furthermore, Lakatos insists that all tests involve comparisons of competing theories. This view goes too far, but it is an instructive exaggeration. Many problems of confirmational relevance are simplified when one is seeking evidence that will discriminate between competing hypotheses.

Although there is a good deal of truth to these last observations, they do not lend themselves easily to systematic development. Too much depends on the details of particular "paradigms" or "research programs." Even if discipline-specific knowledge is of great importance, it is hard to believe that there is *nothing* more that can be said in general concerning the relation of theory to evidence.

One way to say more, which current commands a great deal of philosophical attention and which should be attractive to economists, is espoused by so-called Bayesian philosophers of science (Dorling 1972, Eells 1982, Hesse 1974, Horwich 1982, Howson and Urbach 1989, Rosenkranz 1977, 1983). Crucial to Bayesian philosophy of science is the view that individuals assign subjective probabilities -degrees of confidence- to propositions, including propositions stating hypotheses and propositions stating evidence, and that they update these probability judgments in response to new observations or experimental results in accordance with Bayes' Theorem. Let h be some hypothesis and e be an evidence proposition. Then Bayes' Theorem can be stated as $\Pr(h/e) = \Pr(h).\Pr(e/h)/\Pr(e)$, where "$\Pr(h/e)$" denotes the conditional probability of h given e. $\Pr(h)$ is the "prior probability" of h, the probability an agent assigns to h prior to possessing evidence e. $\Pr(h/e)$ is called the posterior probability. Many Bayesians argue that it is the probability an agent should assign to h after possessing evidence e. $\Pr(e/h)$, the conditional probability of the evidence given the hypothesis is called the "likelihood" of the hypothesis given the evidence. When the hypothesis implies e, the likelihood is one. These probabilities are degrees of belief or confidence. There are disagreements about how these should relate to knowledge of objective frequencies. The dominant view, "subjectivist" or "personalist" Bayesianism, is very permissive about subjective probabilities. For a less subjectivist contemporary variant, see Rosenkranz (1977, 1983). Notice that Bayes' theorem itself is a trivial consequence of the definition of conditional probability -$\Pr(h/e)=\Pr(e\&h)/\Pr(e)$. What is distinctively Bayesian is the interpretation of the theorem as a rule for updating degrees of belief.

The virtues of this formal fiddling are first that it gives substantive discipline-specific knowledge a definite role in confirmation via the prior probabilities. In a cooked-up theory such as "If h then e," where h is any arbitrary hypothesis and e any true statement of evidence, h will have a low prior probability, and, since e is already known to be true, the denominator will be unity and the posterior probability will be the same as the prior probability.

Second, unlike the general hypothetico-deductive scheme, which is mute on how *much* some bit of evidence confirms a hypothesis, the Bayesian account apparently permits a simple metric of degree of confirmation: $\Pr(h/e) - \Pr(h)$[5]. It also neatly explains why "hard tests" of hypotheses - tests in which $\Pr(e)$ is low- strongly confirm them, while mere addition of instances of the same kind contribute little confirmation. Various other intuitive confirmation principles follow (see Eells 1982, ch. 2). For example, logical truths or contradictions cannot be confirmed. If $\Pr(h) < 1$, then if e entails h, e confirms h. If $\Pr(h)$ and $\Pr(e)$ are not zero or one, then if h entails e, e confirms h. Moreover, the Bayesian account motivates these conditions on confirmation in a natural way. Scientists want to avoid errors, and an increased probability suggests that error is less likely.

In the simplest Bayesian vision, the prior probabilities and the likelihoods are known. Plausible hypotheses have non-negligible priors, and good tests have high likelihoods and low prior probabilities of the evidence. When a test is carried out, the probability of h is up-dated as $\Pr(h/e)$ or $\Pr(h/\sim e)$, depending on whether e or $\sim e$ is observed. Since the likelihood, $\Pr(e/h)$ is known, so is the evidential relevance of e to h, and there is no Quine-Duhem problem.

All this seems too good to be true, and it is. If the priors and the likelihoods were known, and one had reason not to change them in response to test results, then it would be easy to know exactly what was the significance of an observation or experimental result. But this is just to say that if the problems of confirmation and theory assessment were solved, then they would be solved. For the Quine-Duhem problem and the problem of evidential relevance arise precisely because scientists do not know exactly how evidence bears on the individual conjuncts from which some prediction is deduced -that is- because scientists do not know the likelihoods of specific hypotheses. Although Bayesian views of theory appraisal are sometimes presented in this oversimplified form, I think it is less deceptive to make explicit what is hidden in the assumptions that the priors and likelihoods are known.

Let us then sketch a more sophisticated Bayesian approach:

1. *Formulate* a hypothesis h that has a substantial prior probability.
2. Calculate $\Pr(h/PT)$ and $\Pr(h/T.\text{not-}P)$ such that:

    (a) $\Pr(T)$ is close to 1
    (b) $\Pr(P/h\&T)$ is high, and

(c) $Pr(P/{\sim}h\&T)$ is low.

3. *Test P.*

4. *Update* the probability one assigns to h as approximately equal to $Pr(h/PT)$ or $Pr(h/T.\text{not-}P)$ depending on whether P is true or not.

This sketch is subject to many additions and refinements. In formulating it, I have emphasized the parallelism between Bayesian and hypothetico-deductivist views. The sketch says nothing about whether hypotheses should ever be "accepted" or "believed." All it talks about is their probability. The schema is idiosyncratic, for Bayesians often suppress the mention of some background theory T and suppose that the likelihood of h ($Pr(P/h)$) is known. But, as I already argued, this habit begs important questions. Notice that nothing is said about what is admissible as a part of "T" here, nor whether the probability of T or the likelihood of $T\&h$ might be revised as a result of this testing. But simple pictures still have their uses.

Despite the many virtues of the Bayesian approach, it faces serious problems, too (see especially Glymour 1980, chapter 3). Simplest to describe is the problem of old evidence. If the truth of e is known, then $Pr(e) = 1$ and e cannot confirm h. But even if new evidence is better than old evidence (which is not obvious), old evidence is not worthless (see Eells 1985, Garber 1983, Howson and Urbach 1989, Kaplan, 1996 and Niiniluoto 1983). Second, the Bayesian account gives precise directions about how to update probability judgments only if judgments of likelihoods ($Pr(e/h)$) are known and may not be revised (as stressed by Miller 1987, pp. 314f). But, as the Quine-Duhem problem suggests, it may be as reasonable to revise judgments of likelihood as to change one's degree of belief in the hypothesis.

This is a serious practical problem in economics. Limitations in the ability to test might make the principles and general implications of equilibrium theory *de facto* unfalsifiable, even if economists were explicitly employing a Bayesian account of confirmation. Let H be a principle of equilibrium theory such as non-satiation and A be the conjunction of all the other statements needed to derive a prediction concerning market behavior, e, from H. The prior probability of H, $Pr(H)$ is much larger than the prior probability of A, $Pr(A)$. That people prefer more commodities to fewer is plausible and supported by everyday experience. Economists in contrast have little confidence in the conjunction A of all the auxiliary assumptions needed to derive an implication concerning market outcomes from *H*. Each of the simplifications and *ceteris paribus* qualifications is improbable, and the probability of the conjunction will consequently be much smaller than that of the separate conjuncts.

How should these probability judgments be influenced by testing? To keep things simple, suppose that *H* and *A* are probabilistically independent of each other -that is, that $Pr(H.A) = Pr(H).Pr(A)$. Personalist Bayesians typically suppose that the relevant probabilities are known, and I shall temporarily go part way with them, although the assumption is fantastic. (I will not, however, assume that the likelihoods of *H* and *A* -Pr(e/H) and Pr(e/A)- are known.) From Bayes' theorem, the independence of *H* and *A*, the fact that $Pr(e/H.A)$ is one, and some simple algebra, one can derive the following two equations:

$$Pr(H/e)/Pr(H) = [Pr(A)/Pr(e)] + Pr(\sim A).[Pr(e/H.\sim A)/Pr(e)] \quad (1)$$

$$Pr(H/\sim e)/Pr(H) = Pr(\sim A).[Pr(\sim e/H.\sim A)/Pr(\sim e)] \quad (2)$$

Since $Pr(A)$ is close to zero and $Pr(\sim A)$ is close to one, the two ratios on the left-hand sides, which one may take as indices of the extent to which *H* is confirmed or disconfirmed by the observation respectively of *e* or not *e*, depend on $Pr(e/H.\sim A)/Pr(e)$ and $Pr(\sim e/H.\sim A)/Pr(\sim e)$. If one believes that, given *H*, *e* is much more probable and $\sim e$ is much less probable than given not-*H*, even if *A* is not true (that is, if one believes that $Pr(e/H)$ is higher than $Pr(e)$ and $Pr(\sim e/H)$ is lower than $Pr(\sim e)$), then the first ratio will be greater than one, and *e* will confirm *H*, while the second ratio will be less than one and $\sim e$ will disconfirm *H*. Typically economists have little idea what $Pr(e/H.\sim A)$ and $Pr(\sim e/H.\sim A)$ are and no reason to believe the former to be larger than $Pr(e)$ or the latter to be smaller than $Pr(\sim e)$. And if $Pr(e/H.\sim A)$ does not differ from $Pr(e)$ and $Pr(\sim e/H.\sim A)$ does not differ from $Pr(\sim e)$, then *H* is neither confirmed by *e* nor disconfirmed by $\sim e$. Given how weakly evidence bears on *H*, the credible "laws" with which economists begin will be *de facto* non-falsifiable.

5. Mill's "deductive method"

Is it impossible to confirm or disconfirm equilibrium theory? If the only data one had were market data, then I have little hope. More powerful statistical techniques and better data sets will help, but I see no way that observations of open systems, such as markets, in which an enormous number and variety of causal factors are relevant, will bear forcefully on theories.

But there are other data! Friedman's influential advice to ignore all other data is a fundamental mistake. It prevents economists from looking at those data that might actually tell them something about the validity of their theories. Economists can conduct surveys, they can observe what goes on within particular firms and markets, and they can carry out experiments. Nothing of these expedients is a panacea. None of them makes testing the principles or implications of equilibrium theory simple and uncontroversial, but each has the potential of generating data that bear forcefully on the truth or falsity of particular parts, or implications, of equilibrium theory.

In a very different context, more than 150 years ago John Stuart Mill (1836; 1843, Bk. VI) recognized the points I have been making. He argues that economists cannot employ "the method *a posteriori*," the method of "specific experience" to assertions concerning markets, because market phenomena reflect the influence of too many causal factors. Instead economists need to employ "the method *a priori*," an *indirect* method of experience. They need to test the fundamental propositions of economics separately in simpler contexts. Their confidence in economic theory derives from confirmation of its basic principles, not from the success or failure of tests of its consequences. In the same way, our confidence in the theory of tides derives from the confirmation of Newton's laws of motion and gravitation, not from the success or failure of tests of consequences of the theory of tides. If the consequences of economic theory or of the theory of tides are standardly way off the mark, then one knows that the causal factors left out of economic or tidal theories are so important that the theories are not useful. Since the principles are confirmed in other contexts, their success or failure with respect to uncontrolled market or tidal behavior neither confirms nor disconfirms them.

Some commentators have read Mill as defending the untenable dogmatic view that evidence from economics can never disconfirm the fundamental laws of economics. He did, I believe, exaggerate how securely the fundamental principles of economics have been established by introspection and casual observation. But I believe that the reason why he held that market data tell economists only when interferences are or are not significant, was that he recognized how weakly those data bear on the theories, not that he thought there was any methodological mistaken in trying to bring them to bear[6]. Market data fail to confirm or disconfirm the principles of equilibrium theory, because the implausibility of the auxiliary assumptions linking the principles of equilibrium theory to that data make the data irrelevant. There is no methodological principle instructing economists to ignore relevant data.

6. Conclusions

Casual observation, surveys, and experiment provide data that confirm or disconfirm the principles of equilibrium theory. Causal observation confirms the claim that people generally prefer more commodities to fewer. Experiment strongly disconfirms the claim that people have stable transitive preference orderings (Hausman 1992, ch. 13). But a further problem concerning the confirmation of equilibrium theory remains. A claim such as "Other things being equal, an unsupported body near the surface of the earth falls with a constant acceleration" does not by itself say anything about what happens when other things are not equal. In mechanics, there are well-established principles of composition that tell us for example that the acceleration of a steel ball in a vat of molasses near a magnet will be determined by the vector sum of the gravitational, frictional, and magnetic forces. If equilibrium theory told one only what happens when the causal factors it identifies operated separately and with no disturbances or interferences, then it would tell one almost nothing. Like Newtonian mechanics, the theory is committed to principles of composition. But unlike mechanics, these principles are not explicitly formulated. (They are instead implicit in the way that models are constructed and used.) These implicit principles of composition are also difficult to test, because they can only be tested in environments with multiple causal factors, which are difficult to control. To believe that the principles of composition hold in the uncontrolled enviroment of markets involves a leap of faith.

The truth is, I believe, that it is extremely difficult to test a theory that is designed to apply to a complex open system. (How well can physicists predict and explain the path of a falling leaf?) One can test axioms concerning individual causal factors under controlled circumstances, and one can test principles governing simple compositions of causal factors. But ultimately one has to see whether the theory "works," and because the system one is concerned with is open and complicated, the fit between theory and phenomena is likely to be rough. That rough fit will tell one little about the worth of the theory, about how to explain away apparent disconfirmations, and about how to improve the theory.

Notes

† I would like to thank the Charlottenburg Trust for support of this research.

1 This essay represents a further development of theses defended in (Hausman 1992), and parts are borrowed from chapter 12 and section 10 of the appendix.

2 There are obviously many other philosophical questions to be asked about this theory. See (Hausman 1992) for a book-length examination.

3 For further discussion, see (Hausman 1992, ch. 8); (Cartwright 1989); and (Maki 1996).

4 For development of these arguments against Popper and Lakatos, see (Hausman 1992, ch. 10 and 11).

5 To say that e confirms h if and only if $Pr(h/e) > Pr(h)$ leads to paradoxical conclusions. As Salmon points out, e and e' may both confirm h, yet the conjunction of e and e' disconfirm h (1975, p. 104) and e may confirm h and confirm k yet disconfirm (h or k) (1975, p. 117)! Salmon's view (pp. 121-122) is that the qualitative notion of evidence confirming or disconfirming a theory ought to be superseded by a quantitative notion of degree of confirmation. Kaplan, on the other hand, argues compellingly that the quantitative measure of degree of confirmation is fatally flawed by a variant of the problem of old evidence (1996, pp. 75-85).

6 For more detailed argument, see (Hausman 1992, ch. 12).

BIBLIOGRAPHY

Cartwright, Nancy: 1989, *Nature's Capacities and their Measurement*, Oxford, Oxford University Press.
Dorling, J.: 1972, 'Bayesianism and the Rationality of Scientific Inference', *British Journal for the Philosophy of Science* 23, 181-90.
Duhem, Pierre: 1906, *The Aim and Structure of Scientific Theories*, tr. P. Wiener, Princeton, Princeton University Press, 1954.
Eells, Ellery: 1982, *Rational Decision and Causality*, Cambridge, Cambridge University Press.
--: 1985, 'Problems of Old Evidence', *Pacific Philosophical Quarterly* 66, 283-302.
Friedman, Milton: 1953, 'The Methodology of Positive Economics', in *Essays in Positive Economics*, Chicago, University of Chicago Press, pp. 3-43.
Garber, D.: 1983, 'Old Evidence and Logical Omniscience in Bayesian Confirmation Theory', in Earman (1983), pp. 99-131.
Giere, Ronald: 1983, 'Testing Theoretical Hypotheses', in Earman (1983), pp. 269-98.
Glymour, C.: 1980, *Theory and Evidence*, Princeton, Princeton University Press.
Hausman, Daniel: 1992, *The Inexact and Separate Science of Economics*, Cambridge, Cambridge University Press.
Hempel, C.: 1965, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York, Free Press.

Hesse, M.: 1974, *The Structure of Scientific Inference*, Cambridge, Cambridge University Press.
Horwich, Paul: 1978, 'An Appraisal of Glymour's Confirmation Theory', *Journal of Philosophy* 75, 98-113.
--: 1982, *Probability and Evidence*, Cambridge, Cambridge University Press.
Howson, C. and P. Urbach: 1989, *Scientific Reasoning: The Bayesian Approach*, LaSalle, IL, Open Court.
Kaplan, Mark: 1996, *Decision Theory as Philosophy*, Cambridge, Cambridge University Press.
Mäki, Uskali: 1996, 'Two Portraits of Economics', *Journal of Economic Methodology* 3 1-38.
Mill, John Stuart: 1836, 'On the Definition of Political Economy and the Method of Investigation Proper to It', repr. in *Collected Works of John Stuart Mill*, vol. 4, Toronto, University of Toronto Press, 1967.
--: 1843, *A System of Logic*, London, Longmans, Green & Co., 1949.
Miller, Richard: 1987, *Fact and Method: Explanation, Confirmation and Reality in the Natural and the Social Sciences*, Princeton, Princeton University Press.
Niiniluoto, Ilkka: 1983, 'Novel Facts and Bayesianism', *British Journal for the Philosophy of Science* 34, 375-9.
Rosenkranz, Roger: 1977, *Inference, Method and Decision*, Dordrecht, Reidel.
--: 1983, 'Why Glymour *Is* a Bayesian', in Earman (1983), pp. 69-98.
Salmon, Wesley: 1975, 'Confirmation and Relevance', in G. Maxwell and R. Anderson (eds). (1975), pp. 5-36. Repr. in P. Achinstein (ed.): *The Concept of Evidence*, Oxford, Oxford University Press, pp. 95-123.
Schlesinger, G.: 1976, *Confirmation and Confirmability*, Oxford, Clarendon Press.

Daniel M. Hausman is Professor of Philosophy at the University of Wisconsin-Madison and Lachmann Research Fellow at the London School of Economics. He has been the co-editor of the journal *Economics and Philosophy*. His books on philosophy and methodology of economics include *Capital, Profits, and Prices: An Essay on the Philosophy of Economics* (1981), *The Inexact and Separ[a]te Science of Economics* (1992), and *Essays on Philosophy and Economic Methodology* (1992). He is editor of the anthology *The Philosophy of Economics* (1984, 1994). He has also coauthored with Michael McPherson *Economic Analysis and Moral Philosophy* (1996).