Lecture 16
Use of empirical evidence in economics

Most introductory economics textbooks open with a discussion of the distinction between positive economics and normative economics. The distinction suggests that there is an objective, scientific, value-free economics based on the scientific method that establishes “what is.” This is positive economics. In contrast, normative economics is about “what ought to be.” Normative economics is hopelessly unscientific since it is based on value judgments. At least this is what you get in introductory texts.

Mainstream economics has taken over this distinction between normative and positive statements from the philosophical perspective known as logical positivism. (When logical positivism came to America, it became known as logical empiricism). Under logical empiricism value judgments are inherently subjective, relative to local customs and moral beliefs, based on emotion not reason. The only “meaningful statements” in the philosophy of logical empiricism are statements that can be verified by empirical observation. Thus normative statements are not meaningful statements according to logical empiricism. This means not only is it impossible to discriminate rationally between one moral position and another but statements about ethics are not even intelligible.

This is an extraordinary argument. Consider the statement: “Murder is wrong.” Under the logical empiricist argument it is not possible to make a rational determination on the validity of the statement and it is not even possible to understand the statement; it might as well have been: “XAZYUQIOXZ is wrong”.

Logical empiricism (or logical positivism) has been discredited for a generation among professional philosophers. Some outside critics (John Dewey, for example) focused on the weakness of the positivist position on ethics. However, the demise of positivist philosophy was from internal criticism and it focused on the inadequacy of the logical empiricism for the understanding of science.

Under positivism, scientific knowledge is construed as a set of verified positive statements. But it became clear that the positivist distinction between fact and theory was problematical. Without the positivist dichotomy between fact and theory, scientific theories are meaningless. Thus the death blow for logical positivism was that, if the positivist criterion for meaningfulness is applied to scientific knowledge itself, scientific theories are meaningless.

The philosophical positions taken by academics in social sciences and the humanities often lag decades behind current philosophical discourse. Most economists probably still accept Friedman’s methodology of positive economics as the definitive statement on methodology and philosophy of economics. And the normative/positive distinction may seem to be the least problematic aspect of positivist philosophy to mainstream economists.

It is argued elsewhere that so-called normative statements are statements about ethics, that is, desirable modes of behavior of human beings toward one another. The fact that ethical beliefs vary considerably from one culture or locale does not lead to the conclusion that ethical issues are inherently relativistic and completely beyond rational examination. In fact ethical statements are conditioned on empirical knowledge and
beliefs about causal connections in the empirical world. Rather than a statement about what is, it is more accurate to say that science is a set of (well-tested but fallible) beliefs about what kinds of things there are in the world and how these things work. (Even better, science can be seen as a process not a collection of fixed beliefs. In particular, science is a kind of very effective (but still fallible) process of inquiry.

**Standard textbook version of the scientific method.** Presumably positive economics is what the rest of the typical introductory text will present. The textbook presentation of “scientific method” has something like the following steps: (1) a hypothesis is proposed (2) empirical evidence is used to test the hypothesis with the results either being consistent with the hypothesis or contrary to it; (3a) if the empirical evidence is contrary to the hypothesis, the hypothesis is rejected; (3b) if the empirical evidence supports the hypothesis, the hypothesis is tentatively accepted as part of scientific knowledge; (4a) the rejection of the original hypothesis suggests some modification is needed, perhaps the original hypothesis needs to revised or replaced with something substantially different; (4b) if the empirical evidence supports the hypothesis, there may be further testing of the same hypothesis (since empirical evidence is never absolutely conclusive in supporting a hypothesis); also, the tentatively accepted hypothesis may now serve as assumption used in testing other hypotheses.

The typical text will probably show a simple example with empirical data plotted on a graph and a line calculated to fit to the data that summarizes or abstracts the relation between the two variables plotted on the graph. The empirical (that is, observed) data do not lie on a straight line but are somewhat “scattered” on the graph. (The plotted line is called a regression line since it is usually derived using the method of least squares regression.) The particular data points show observed values of the variables x and y, and \((x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n)\), etc. shown plotted on the graph are some of empirical data available for the variables x and y, which are presumed to have some kind of causal relation. For example, y is assumed to be a function of the variables x. That is, \(y = f(x)\).

<table>
<thead>
<tr>
<th>obs</th>
<th>(x_i)</th>
<th>(y_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(x_1)</td>
<td>(y_1)</td>
</tr>
<tr>
<td>2</td>
<td>(x_2)</td>
<td>(y_2)</td>
</tr>
<tr>
<td>3</td>
<td>(x_3)</td>
<td>(y_3)</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>n</td>
<td>(x_n)</td>
<td>(y_n)</td>
</tr>
</tbody>
</table>
It is important to understand the fundamental assumption underlying this entire process: that there is a basic relation between \( y \) and \( x \) that extends far beyond the particular values of data points observed in this case (the \( (x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n) \), etc. The relation between \( y \) and \( x \) is hypothesized to hold, if not universally, at least over a broad range of situations which would allow various combinations of \( (x_i, y_i) \) to be observed. Typically the relation between \( y \) and \( x \) would be characterized as a linear relation, which could specified as an equation of the form \( y = a + bx \), where \( a \) and \( b \) are unknown parameters.

A very important point to understand is this: the reason that observed values of \( x \) and \( y \) do not lie exactly on the regression line is that the value depends not only upon \( x \) but also upon a _random disturbance term_ \( u \). On the average \( u = 0 \) but it can be a positive or negative number in a particular instance. (The disturbance term \( u \) may be called the error term or the stochastic term.) The basic idea of regression analysis is that there is a true relation \( y = f(x) + u \), where \( u \) is a seemingly chance occurrence caused by measurement error, effects of omitted variables, etc. In other words, without the presence of whatever it is that causes \( u \) to take an non-zero value in a particular observation, all the observed pairs of \( (x_i, y_i) \) would lie exactly on the regression line.

The simplest kind of hypothesis is that, for example, \( (\Delta y/\Delta x) < 0 \). From the general characteristics of linear functions, this translates into the parameter \( b \) having a negative value. This hypothesis is based on the assumption (speculation, hope, consistency with previously accepted theory, etc.) that there is a “true relation,” or mathematically specifiable Law of Nature such that \( y = \alpha + \beta x + \varepsilon \), where \( \varepsilon \) is the true random disturbance term. Thus the true parameter \( \beta \) is given by the fundamental nature things. If the hypothesis is correct, the true parameter \( \beta \) would have a negative value. Fitting the regression line to the particular data points available, will yield a numerical value estimated for the parameter \( b \).

Step (1) is hypothesizing that \( (\Delta y/\Delta x) < 0 \) or that \( \beta < 0 \). Step (2) is deriving an estimated value for the parameter \( b \) from the regression procedure allows an inference to be made about the true value of \( \beta \). If \( b > 0 \), the hypothesis is rejected and the procedure goes to (3a). If \( b < 0 \), the hypothesis is supported and the procedure goes to (3b).

**Critical commentary of the standard version of positive economics.** This (or some dumbed-down version) is pretty much how economics textbooks present the topic of positive economics and the scientific method. Typically little is heard of the scientific method or empirical evidence for the rest of the text as the standard models (and features of those models such as elasticity of demand) are presented in words, diagrams, algebra, etc.). The impression left is that diagrams, algebra, verbal arguments, etc. presented in the textbook are the product of the scientific method. That is, these diagrams, algebra, verbal arguments, etc. are indeed positive economics. While the diagrams, algebra, verbal arguments, etc. may be somewhat abstract or simplified, they are presented as being a statement of “what is.”

The impression that textbook economics is some kind of scientifically objective positive economics is misleading on a number of levels. First of all, the very distinction between purely objective positive economics and normative economics is based on an outdated and discredited philosophical perspective called logical positivism (or logical
empiricism). It is impossible for science to be value free; it is not impossible for individual scientists with greatly differing values to agree about the way things in the world seem to work.

This happens routinely. But it does not happen because of some easily described "scientific method" which -- if followed diligently, robotically? -- leads automatically to scientific knowledge. The idea that there is a single, objectively specifiable scientific method is also highly discredited by those who study the history and philosophy of science. The scientific research that takes place in the actual sciences practiced in the world in which we live is much messier. What scientists share as a community is a commitment to finding out how the world works. This is done by subjecting their theories to the test of empirical evidence in as thorough a way as possible.

Scientific knowledge is always fallible. That is, any part of the body of beliefs characterized as scientific knowledge could be mistaken. There are probably parts that are wrong. It is extremely unlikely that large chunks of well-established scientific knowledge are completely wrong or way off the mark. It is not impossible. One hallmark of genuine science is the recognition and acceptance of fallibilism. The methods of sciences evolve as we have more knowledge and have new and better technology to help in gathering, developing and interpreting empirical evidence.

Very little is said about the source of hypotheses that are allegedly subjected to empirical testing in the application of the scientific method. Most economic theory is very abstract and qualitative, providing only very general specifications of what to include and what characteristics the included items would be expected to have. There is a ubiquitous assumption that equilibrium analysis (based on some form of optimization) is appropriate for all economic models but this does not provide many clues about which variables are important or how those variables would be expected to behave. In other words, economic theory leaves a lot of leeway in the specification of econometric models to be used with empirical evidence to provide statistical estimation of the parameters of the econometric models. This is a mouthful of long words so a look at model specification is instructive.

Model specification. What is being done: using what guidance is available from economic theory, specify a model (in the form of a mathematical equation) and apply the technique of econometric estimation (or some other statistical procedure) with some relevant empirical data. Digital computers allow the rapid calculation of the parameters in the specified model. Statistical theory can then be applied to test hypotheses (which are stated in terms of estimated parameter values).

The elements in model specification typically include: (1) deciding which variables are going to be included in the mathematical model; (2) what the functional form of the mathematical equation is going to be -- e.g., it could be linear, multiplicative, logarithmic, exponential, polynomial, etc.; (3) deciding exactly what part of available data should be used; (4) deciding what count as empirical counterparts for those variables included in the model; and more.

For example on item (4), a model includes "price of X" in the model specification. But often price fluctuates or is not known except as an average price or closing price. If some average price (that is an empirically observable price is used as the empirical counterpart to price of X in the model, over what time period, what spatial area, what variety of similar products, etc. is the average price averaged? There is often an
interaction between the choice of data set and availability of empirical counterparts to model variables. Price data may be available on a daily basis but, if income is also a variable in the model, personal income data may only be available on a quarterly basis. Price data and income data have to be averaged and aggregated in such a way as to have the same time period for each observation.

It should be clear that there is a lot of discretion in the specification of the model used by any one economist in statistical estimation procedures that presumably provide tests that support the collection of positive statements that make up “positive economics.” The up-or-down decision on any hypothesis is not such a simple matter. What is published as peer-reviewed research is also subject to great discretion. The prejudices of editors, the vested interests of academic institutions, intellectual fads, and -- not least important -- the political and ideological implications of economic theories have a strong role in the determination of what econometric testing gets a public hearing.

**The identification problem.** The “identification problem” creates additional difficulties in attempting to use empirical data and statistical estimation (principally through econometric techniques) to test hypotheses. The “identification problem” arises when the process of statistical estimation involves a model that is a system of relations and not just a single relation -- and thus a single equation model considered in isolation from everything else. Almost anything in economic theory involves some kind of system with simultaneous relations. For example, in micro economic theory, quantity demanded and quantity supplied and price are determined simultaneously from the interaction of two relations.

In simplest terms, the “identification problem” is the problem of trying to determine the value of the parameters of the structural equations in multi-equation models. (Structural equations specify the relations that come from economic theory, such as supply and demand relations.) The “identification problem” is the fact that the same set of statistical estimated parameter values can fit a number of different structural equations, even equations from rival theories.

In the technical language of econometric theory, the parameter values in a system of simultaneous structural equations are functions of the reduced form equations. Strictly speaking, the “identification problem” is the problem of uniquely deriving structural form parameters from the reduced form estimates. This may not always be possible. Additionally, the statistical theory used to inferences about parameters may not be applicable to structural parameters. Even beyond this, there may be several different structural models (each consisting of a system of simultaneous equations) that are identified in the reduced form parameters but these different structural models are not consistent with one another.

**Leamer.** The institutions of mainstream economics (which include a high level of philosophical illiteracy) conveniently supports an actual practice where there is very little danger of any accepted part neoclassical economic theory is going to be refuted and little real role for real empirical evidence to refute any currently accepted part of the “core of neoclassical economic theory.” Edward E. Leamer tries to illuminate the mismatch between the official scientific rhetoric in mainstream economics with the actual practice. Leamer’s book is entitled *Specification Searches: Ad Hoc Inferences with*
Nonexperimental Data. His famous article is entitled "Let's Take the 'Con' Out of Econometrics" (1983 American Economic Review).

In simplest terms, he is blowing the whistle on the way that mainstream economists actually use statistical estimation (which is overwhelmingly in the form of econometric estimation) in sharp contrast to what they claim to be doing. In classical statistical theory, estimation techniques are to be applied to test whether a hypothesized model may be the presumed "true model." According classical theory, there is an up-or-down determination of whether the estimated value of the parameter should be rejected as an estimate of the value of the parameter in the "true model."

Rather than this, Leamer suggests that economists are really doing is the repeated re-specification and subsequent statistical re-estimation of the original (or current) model until arriving at a suitable model. This process can be reiterated many times. This process very often has nothing to do with testing some hypothesis (as is claimed in the official version of scientific method). Rather Leamer describes the actual purposes of different kinds of "specification searches." These specification searches include the following: (1) hypothesis-testing searches (choose a "true model"); (2) interpretative searches interpret multi-dimensional empirical evidence; (3) simplication searches (find a more "fruitful model"); (4) proxy search (find a quantitative facsimile); (5) data-selection search (select a data set); (6) post-data construction search (improve an existing model).

The basic consideration is that the same data set cannot serve both as a basis for specification of the model and as empirical evidence to test theoretical hypotheses. In specification searches or types (2) - (6), there is no testing of some hypothesis relevant to the validity of economic theory. Type (1) involves issues related to the theory itself but is not a legitimate empirical test of hypothesis using evidence.

Economics as a science? The often cited problem that economics faces as a science, is that it cannot be an experimental science. For most of the history of economics, rational introspection and apriori pronouncements substituted for systematic scientific inquiry based on methodical use of empirical evidence. In fact up until fairly recently (about fifty or sixty years ago), mainstream economists were discouraged from attempting to use empirical evidence in any kind of systematic way. The core of neoclassical economic theory has changed little since the late nineteenth century but intellectual fads have changed. Arguments based on rational introspection and apriori assertions are out of vogue.

The use of econometrics began about sixties years ago. The intent was to use statistical estimation theory to overcome in the inability of economics to be a science based on controlled laboratory experiments. Consider a biological theory of the form:

\[ y = f(x, z, t, w), \]

and the hypothesis of interest is \( \Delta y / \Delta x < 0 \).

The biologists can control \( z, t, \) and \( w \) in the laboratory so that each observation receives identical amounts of each. The same effect might be possible using non-experimental data where the variables \( z, t, \) and \( w \) take various values and econometric techniques are used to get a kind of artificial (or simulated) control. In practice and in the context of economic analysis, there may be many problems in trying to do this. (The discussion above is intended to show some of the problems that arise.)
But besides this there are two things that scientists do that mainstream economists do not do that. First, in natural sciences, scientists are serious about taking into account accepted scientific knowledge from outside their own discipline that is relevant to theories within their own scientific discipline. For example, biologicalists do not propose a genetic theory that is contrary to knowledge in organic chemistry and geologists take the laws of physics into account when they theorize. For economics, cognitive psychology is analogous to physics in geology. As the article by Herbert A. Simon points out, economists largely ignore cognitive psychology.

Second, scientists painstakingly and systematically gather and refine empirical evidence. They develop new kinds of data and new methods of getting data. And scientists are directly involved in processes involving data collection and improvement. In contrast, mainstream economists are rarely directly involved in processes involving empirical evidence. Friedman’s famous article on methodology is motivated in large part to attack and ridicule the direct collection of micro evidence by economists (e.g., on the internal decision making processes in business firms in the actual economy). Friedman’s arguments about the “realism” (or not) of economic theory is intended to protect a certain category of economic theory while ruling certain kinds of empirical evidence -- e.g., internal survey data from business firms -- out of court.