

Methodological Mistakes and Econometric Consequences

Dr. Asad Zaman[®]

International Islamic University of Islamabad

ABSTRACT

Econometric Methodology is based on logical positivist principles. Since logical positivism has collapsed, it is necessary to re-think these foundations. We show that positivist methodology has led econometricians to a meaningless search for patterns in the data. An alternative methodology which relates observed patterns to real causal structures is proposed.

Key Words: Econometric Methodology, logical positivism, realism, causality, VAR models, Forecasting, surprise, goodness of fit.

JEL Classifications: C18, B40

1. INTRODUCTION

The rise and fall of logical positivism is the most spectacular philosophical story of the twentieth century. Rising to prominence in the second quarter of the twentieth century, it swept away all contenders, and became widely accepted throughout the academia. Logical positivism provided a particular understanding of the nature of knowledge, as well as that of science and of scientific methodology. The foundations of the social sciences were re-formulated in the light of this new understanding of what science is. Later on, it became clear that the central tenets of the positivist philosophy were wrong. Logical positivism had a “spectacular crash,” and there was some dispute about who had “killed” logical positivism¹. As a logical consequence, it became necessary to re-examine the foundations of the social science, and to find new bases on which to re-construct them. This has occurred to differing degrees in different disciplines. One of the most recalcitrant has been economics. As discussed in Zaman (2011), the foundations of economics continue to be based on erroneous logical positivist ideas, and hence require radical revisions.

[®] Asad Zaman, International Institute of Islamic Economic, International Islamic University of Islamabad, (email: asadzaman@alum.mit.edu), Tel. +9251 9257939, Fax: + 9251 9258019.

¹Von Fraasen (1980), a staunch empiricist, writes “Today, however, no one can adhere to any of these philosophical positions to any large extent. Logical positivism, ... even if one is quite charitable ... had a rather spectacular crash.”

It is not our intention in this paper to describe the whole story, which would be lengthy as well as complex and contentious. We will confine our attention to tracing the harmful impact of a limited number of positivist ideas on how econometrics has developed. We will argue that these ideas have led to a wrong methodology being accepted and used in econometrics. This wrong methodology has prevented progress and advance in knowledge. As evidence for this lack of progress, in a talk on the 100th anniversary of the first published regression by Yule, Freedman (1997:113) writes: “For nearly a century, investigators in the social sciences have used regression models to deduce cause-and-effect relationships from patterns of association. . . . In my view, this enterprise has not been successful.”

This paper is structured as follows. Section two below provides a summary of the key positivist ideas, and how they lead to a distorted understanding of science. The third section provides several case studies showing how positivist methodologies lead us to consider the wrong questions. The fourth section shows the contrast and opposition between positivist and realist methodologies. The fifth section sketches some possible alternative methods which could be used to avoid these mistakes and make progress. The sixth section summarized the conclusions.

We will also offer a very tentative and preliminary sketch of alternative approaches that could be more successful.

2. THE PHILOSOPHICAL BACKGROUND

It is an embarrassingly bold claim that for more than fifty years, econometric methodology has been on the wrong track. This is however, just a straightforward consequence of fact that logical positivism has provided the foundations for econometric methodology. Even though positivism has collapsed, this has not led to the rethinking required to build new foundations. The goal of this section is to provide enough details of the history and philosophy of science to make this claim plausible.

2.1. The Rise and Fall of Logical Positivism

Kung (1997) has documented the battle between the Catholic Church and emerging science in the middle ages in Europe. This battle has been enormously important in shaping attitudes² and philosophies on both sides. Religion and science have different core functions – one deals with ideals for human behavior and social organization, while the other deals with properties of matter and energy. Because there is virtually no overlap, peaceful co-existence between the two is natural, and this is indeed what occurred in Islamic history. In contrast, in Europe the project was launched to “deify” science by the claim that all valid knowledge was the sole province of science.

Before the emergence of “philosophy of science,” the question of how to differentiate between scientific knowledge and other kinds of human knowledge was studied as the *demarkation problem*. A guiding intuition was the idea that science deals with the material and observable,

²An application for affiliation by a nearby Jesuit college was rejected at a Cal-Tech faculty meeting in the late 1970’s when a physicist reminded them to “Remember Galileo!”

while religion, metaphysics, philosophy and the like deal with the unobservable and the abstract. Attempts to differentiate between the two domains of knowledge on these grounds failed because science also deals with unobservables like gravity and electrons.

In early twentieth century, Bertrand Russell, a leading atheist philosopher, came up with a solution to this dilemma. In an extremely influential paper, Russell (1905) argued that a sentence containing the word “unicorns” could be interpreted in a way that avoided taking a stance on the issue of whether or not unicorns exist. The “Theory of Descriptions” introduced by Russell argued that the word “unicorn” could be parsed as a description of an object. A description like “a winged horse” could be meaningful whether or not the object described existed.

Logical Positivists seized on this idea to conjecture that all references to unobservables like gravity and electrons were actually just shorthand for regularities in collections of observations. So when we say that there is gravitational attraction between the Earth and the Sun, this is just a compact way of saying that the “Earth moves around the Sun in an elliptical orbit” and the equations of the orbit can be obtained by positing a gravitational force. To formalize this idea, they argued that the meaning of a sentence is equivalent to the set of its observable implications.

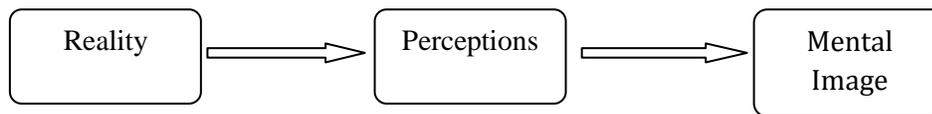
If we accept this idea, then we have a solution to the demarcation problem. Exactly as intuition suggests, science deals purely with observables – apparent references to the unobservables can be reduced to descriptions of patterns in the observables. In contrast, religion deals with unobservables in an essential way. It is not possible to parse religious concepts out of existence by replacing them with equivalents in the form of observables. This means that scientific knowledge is valid, while knowledge imparted by religion is meaningless. This conclusion was deeply satisfying to the pro-science faction, which was the vast majority of western intellectuals, and became widely accepted.

Logical positivism was a set of conjectures about the nature of scientific theory, and more generally about the nature of human knowledge. Some of the finest minds of the twentieth century struggled to prove these conjectures. However these efforts led to the conclusion that “nearly all of it was false,” in the words of A. J. Ayer, one of the leading early exponents of this philosophy. Suppes (1977) is a comprehensive epitaph which discusses the many difficulties with logical positivism in detail. To put the matter in a nutshell, science also involves unobservables in an essential way, and these cannot be reduced to descriptions as per Russell’s prescription.

Unfortunately, the story is far from over. Logical positivism is a complex, sophisticated and subtle set of mistaken ideas about the nature of human knowledge. Its complexity is evidenced by the failure of some of the most brilliant and sophisticated mathematicians and physicists to see that it is all wrong. These wrong ideas were absorbed into the foundational frameworks of social sciences. Because they were taught at universities for decades (and indeed, continue to be taught), they also seeped into the intellectual framework of the general public. Even though the collapse of logical positivism is widely known, few academics know exactly what it was, and why it collapsed. Especially among economists, many continue to espouse ideas derived from positivist precepts without being aware of any difficulties. The goal of this essay is to show that current econometric methodology is fundamentally flawed because it is based on logical positivism. Some directions for progress are also sketched.

2.2. Kantian background to logical positivism

In the next section, we will show that positivists misunderstood the nature of science and scientific methodology. These misunderstanding continue to be widely believed. We must first dig deeper into the foundations of positivist philosophy. These are extensions of certain ideas of Kant which were highly influential. Before Kant, the nature of human knowledge could be described as follows:



Reality impacts upon our limited capabilities of perception via sense data. Our sense data are assembled into an image of reality within our mind. Some of the processing required for the creation of our mental image is done on a pre-conscious level. The child learns to negotiate the real world around him, and to resolve images that he sees into real world objects. Similarly, scientific theory is used to deduce the nature of reality from clues furnished by sense data. According to the sense data generated by our perceptions, it appears that the sun, moon and stars are affixed on a blue sky which lies above the clouds. Some analysis of this sense data leads to the idea that the sky is a celestial sphere which rotates around the Earth. Even deeper processing leads to the conclusion that the first two images are deceptive, and it is actually the Earth which is going around the sun in an orbit. Also, even though sense data show us that the stars look similar to each other, and very different from the Sun and the Earth, scientific and logical analysis reveals that some of the stars we see are just like the Sun in reality, while others are planets like the Earth.

According to this pre-Kantian understanding, knowledge was obtained by using logic and clues furnished by our sense data to acquire an understanding of the nature of reality which generated this sense data. Kant proposed to create a “Copernican Revolution” in philosophy by arguing that it is our perceptions which create reality, rather than the other way around. To understand this paradoxical and apparently nonsensical statement, it is necessary to break it down into more elementary pieces.

1. According to traditional understanding, human knowledge is valid if it provides a correct description of reality, and not valid if it fails to describe reality correctly.
2. Our only access to reality is via our perceptions (sense data). We have no other channels by which we can learn about reality.
3. This means that it is **IMPOSSIBLE** to match a mental model of reality to reality itself and check whether or not it is valid. The true nature of reality (thing-in-itself) is unknowable. This means that the traditional idea about what constitutes human knowledge is wrong. We can never know whether or not our mental model is an accurate match to reality.
4. What we think of as reality is actually our mental model of reality. It is now obvious that this mental model is a product of our perceptions plus the mental processing we do in going from the sense data to the image of reality created in our minds.
5. The “Copernican Revolution” of Kant consists of replacing the impossible-to-solve traditional problem of how to match mental images of reality to reality itself with the

following Kantian alternative: we should study the mental processes by which sense data are converted to images of reality.

This Copernican Revolution of Kant paved the way for the logical positivist understanding of science. Once we believe that electrons and gravity do not exist in reality, but only in the minds of the physicists, it becomes plausible to believe that these terms are just convenient shorthand for describing regularities in collections of observations.

2.3. Nominalist versus Realist Views of Science

Logical positivism led to the following two mistaken ideas about the nature of scientific activity which will be the focus of our attention:

- N1: A scientific law is a regularity or a pattern in the collection of observations. This has been formalized as the idea of a “covering law” by Nagel.
- N2: The scientific method is to look for patterns and regularities in collections of data. There are large numbers of ways to look for patterns. Any pattern we find, if it continues to be observed in subsequent observations, is a candidate for a scientific law.

Both of these widely believed ideas about the nature of science are fundamentally wrong. It would be convenient to term these as a “nominalist” understanding of science. A realist understanding of science is based on the following two alternatives:

- R1: A scientific law reveals causal effects in operation between objects which exist in the real world.
- R2: Scientific methodology consists of using clues from sense data to learn about objects and causal effects which operate in the real world.

Note that Kant and the logical positivists would deny the possibility of R1 & R2 since objects and causal laws which operate in the real world are “unobservable.” Manicas (1987) has discussed the history of how it came to pass that the realist understanding of science was replaced by the nominalist one, even though the latter is wrong.

We now come to a strange and curious twist in the story. The nominalist misunderstanding of the nature and methodology of physical science had no effect on the physical sciences. Physicists were pleased to learn that they had access to the only valid and reliable sources of knowledge. The fact that physicists continued to believe in electrons, gravity, electro-magnetic forces and the like was not of concern. No harm was done by confusing models of reality with reality itself. The philosophy of science described what scientists did; it did not prescribe how scientists ought to act.

In contrast, the social sciences were deeply affected by this positivist misunderstanding of physical science. The tremendous prestige of the physical sciences attracted the admiration and envy of the humanities. Furthermore, it was widely believed that science was the only valid source of knowledge. The incorrect positivist description of science and scientific methodology was taken as a prescription for how to convert studies of human beings into science. This led to the following principles being widely adopted as fundamental principles of social science in the second quarter of the twentieth century:

- S1. Unobservable entities cannot be invoked as explanatory factors. Social science, like the physical sciences, must be built around quantifiable and measurable concepts.
- S2. Morals, values, ideals are unobservable and hence unscientific. Activism, or efforts to create a better society are not part of scientific activity. Social scientists must act as neutral, detached observers seeking to describe laws of motion for human and social activities.

The consequences of adopting these principles have been extremely harmful for the development of social sciences. It is ironic that these principles are based on a misunderstanding of science promoted by logical positivism. Scientists freely utilize unobservables like the strings of string theory, black holes, different types of subatomic particles with strange and unusual properties, whenever necessary to explain observable phenomena. Social scientists strictly avoid the use of unobservables, in an effort to be scientific. Similarly, aesthetic values like elegance and simplicity are often invoked in selection of scientific theories. However social scientists try to maintain an appearance of neutrality and “scientific” objectivity in areas where even the choice of research topic is dictated by value judgments regarding its relevance and importance to human beings.

Before proceeding to discuss the damage caused by the adoption of these erroneous methodological principles in econometrics, we will illustrate the same theme in an assorted collection of areas, just to make it more plausible.

2.4. Errors Due to Misunderstanding of Scientific Methodology

One of the most unlikely candidates for successful application of positivist theories was the field of psychology. How can we study human beings while disregarding the unobservable inner psyche of man? Nonetheless, Skinner argued that the goals of psychology should be prediction and control of observable behavior patterns of human beings. We must move beyond “freedom and dignity” and model humans as robots which can be programmed by appropriate patterns of stimulus and response. Due to the intellectual prestige of logical positivism, behavioral psychology became the dominant paradigm for decades in the mid twentieth century. Paralleling the collapse of positivism, numerous problems with key behaviorist ideas have led to substantial reconsideration. At the most fundamental level, it seems clear that observed behavior plus external environment is not sufficient to explain human action. An influential critique of behaviorism along these lines was developed by Chomsky. He showed that the facility with which children acquire language shows that they must have internal mental structures designed for language. Building on these insights, cognitive psychologists found it necessary to posit

unobservable internal mental structures and emotional states to account for human capabilities at problem solving.

Probability Theory is another field which has been substantially damaged by positivist ideas. The natural interpretation of the idea that the probability of heads on a fair coin flip is one half is the following. There are two possible equally likely states of the world, one with outcome being heads and the other with outcome tails. However this statement refers to unobservable events. Once one of the outcomes occurs, the other is excluded, and there is no possibility of observing an alternative state of the world which did not actually occur. Therefore, even though the statement is perfectly meaningful intuitively, it is meaningless according to the positivist criteria. The current standard definition of probability based on limiting frequency is perhaps the closest observable counterpart to the fundamentally unobservable idea of probability. This is what was offered by the founders of this definition, von Mises and Reichenbach, both of whom were deeply influenced by positivist ideas. The alternative subjectivist definition also uses observable choices over lotteries to define probability in terms of quantifiable beliefs of the observer. Founders of this approach shied away from the unobservable alternate reality due to positivist predilections. Zaman (2008) argues that both of these approaches are unsatisfactory, and the naïve definition of probability based on intuitive ideas works much better, even though it is meaningless according to positivist criteria.

As a third example of errors generated by positivist ideas, consider Samuelson's theory of revealed preference. Economic theory of consumer behavior is based on the idea of utility or preference, the internal satisfaction that a consumer feels from purchasing a good. Using this unobservable means that economics is not a science according to positivist criteria. Samuelson sought to create a new basis for economic theory using the observable choices in place of the unobservable preferences. It was eventually proven that all propositions of consumer theory could be derived from axioms of revealed preference. The Nobel Prize awarded to Samuelson cites his contribution in providing a scientific basis for economics. However, as Wong (2006) has pointed out, there is a fundamental flaw in this approach. The axioms of revealed preference make sense ONLY IF there is an underlying unobservable preference. If the consumer has achieved the Buddhist ideal of complete indifference to worldly affairs, then his choices would be completely arbitrary, not subject to axiomatization. It is only the existence of internal preferences that guarantees consistency and transitivity of choices. Thus the apparent reduction to observable choices in revealed preference theory does not succeed in removing the unobservable utilities from the theory.

3. ECONOMETRIC METHODOLOGY

Hoover (2004) writes that: "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." Current econometric methodology is based on the positivist understanding of science. This methodology is wrong, and needs to be replaced by a realist methodology in order to make progress. This is a miniature version of an argument made by Manicas (1987, 2006) at

greater length and covering a much broader domain. Corresponding to the principles N1 and N2, positivist econometrics can be characterized by:

EN1: Every regression model which fits the data well is a potential economic law. It may be confirmed or rejected according to its abilities to forecast well.

EN2: The methodology of econometrics is to search for good fits to the data.

This can also be called a “nominalist” view – only the data matter, and the reality which generates the data does not. These methodological principles have led to a proliferation of worthless regression models, as we will demonstrate. In order to generate knowledge which accumulates, we need to switch to realist principles:

ER1: Only regression models which capture/represent causal mechanisms are valid.

ER2: We should search for surprising fits, as they provide clues to causal mechanisms operating in the real world. Then we should search for real world causal chains which can provide an explanation for the surprising fit.

We will now explain the problems created by the nominalist methodology, and also how the realist methodology offers solutions to these problems. First, we list four cases where the shift from realist to nominalist methodology manifests itself in econometrics.

3.1. An Attitude Problem: Technological Growth

Barzel (1972) estimated the rate of technological growth by fitting a regression to annual data on winning speeds at the Indy 500 races. Several subsequent papers (cited in Zaman (1996)) do many different types of fancy analyses on this same data set, with the object of obtaining information about rate of technological progress from these winning speeds. The question we want to ponder is: can fancy regression techniques extract more information about technological progress from the winning speeds data set? To put this in the methodological framework of the current paper, consider “reality” as the technology which is certainly one factor which determines the winning speed:

Winning Speed = F(Driver Capabilities, Racing Environment, Car Technology)

A realist methodology would look beyond the observed data to real factors generating the data. It seems abundantly clear that data on car engines would provide the best possible information about technology. Even though data on engine capabilities is easily available, none of the authors utilize this data. This reflects the nominalist attitude looking for patterns in the data rather than the underlying real factors which generate the data. In this particular case, the real factors (i.e. engine data) are easily available, observable, and quantifiable. Nonetheless, methodological blinders lead several analysts to focus on finding patterns in the winning speed data, while none go beyond the data to analyze technological progress as determined by engine capabilities. In

most cases, the real factors are unobservable and unmeasurable, which creates even greater obstacles to utilization of a realist methodology.

3.2. From Cowles Commission to VAR Models

The key contribution of the Cowles Commission was structural simultaneous equations. The economy was characterized by a set of structural relationships all of which held at the same time. Formulation and estimation of such equations required taking a stand on which variables were exogenous, which were endogenous as well as imposition of zero restrictions. All three of these represent causal information about real world structure, which should be provided by theory in advance of modeling. Exogenous variables are causes, endogenous variables are caused, and zero restrictions show absence of causal relationships. Thus the Cowles Commission approach was realist, though it had serious weaknesses and limitations, as discussed in Zaman (2010). Principal among these was the fact that causal structures were specified a priori, and there was no possibility of modifying these in response to evidence provided by the data.

The oil crisis of the 1970's and large forecast failures of econometric models led to substantial decline in the use the simultaneous equation models. Part of the decline can be attributed to the Sims (1980), who argued that the a priori causal information embedded in structural models was unreliable. As an alternative he offered VAR models, which are an ideal representation of the nominalist understanding of econometrics. All variables are treated alike in VAR models, without any reference to the real world structures which generates them.

A recent Master's Thesis received for review by the author illustrates perfectly the problems with VAR models. The thesis estimates a VAR model which includes the variables $W(t)$, $A(t)$, and $R(t)$ which represent wheat output, acreage devoted to wheat production, and the rainfall in year t . Because VAR methodology attempts to be agnostic about causal relations, the VAR model regresses both $W(t)$ and $A(t)$ on lagged values of the same variables $W(t-1)$ and $A(t-1)$, as well as other variables not relevant to our present discussion. Excluding the current acreage devoted to wheat $A(t)$ from the independent variables determining $W(t)$ will strongly bias any estimates. Similarly, any relation between contemporary acreage $A(t)$ and $W(t-1)$, or between $R(t)$ and $A(t-1)$, would seem to be purely a case of spurious correlation. Refusing to think about causal relations leads to attempts to measure correlations which are entirely accidental, and could not contribute to our understanding of real relationships between the variables. In this particular example, the mistake is obvious, because the causal relations are known to us a priori. Similar problems commonly go undetected in the VAR literature, because of the complexity of economic relationships.

The objection being raised here can be articulated more sharply in terms of "autonomy," as discussed by Aldrich (1989). A VAR model measures a collection of partial correlations between a set of variables and their lags. These correlations have differing levels of "autonomy" – independence from possible disturbances and shocks to the system. Some variables X and Y are independent – the measured correlation between the two is entirely spurious. This happens especially frequently when the variables involve time trends, as is now well known. In other cases, the correlations can have varying degrees of autonomy. If consumers decide on current consumption levels on the basis current income without regards to other factors, then the

consumption relation will be highly autonomous – this is called a structural relation in contemporary terminology. The reverse relation, or the regression of Y on C, is a reduced form relation, the parameters of which will be influenced by changes in Investment levels, for example. It is far less autonomous. Different regression relationships have varying levels of autonomy, depending on how resistant they are to interventions and changes in the system. Progress in understanding the economic system occurs when we learn to identify autonomous relationships and differentiate them from accidental ones. This always involves going beyond the observed patterns of correlations to consider the underlying economic structures which generated this pattern.

The Harvard barometer was an attempt to make economic forecasts purely on basis of observed correlations, without any reference to underlying structures generating these correlations. This methodology failed and was abandoned. The VAR models represent the return of these barometric correlations on a grander scale. The observed correlations captured in VAR models can be potentially valuable clues to structure, if they are used in this way. Current econometrics methodology makes the mistake of treating these as the end product of statistical analysis, rather than as a first step towards discovery of structure.

3.3. The Forecast Competitions M1, M2, and M3

A regression model is valid only to the extent that it captures some causal mechanism operating in the real world. Such relationships have “autonomy” – because they capture real world mechanisms, they can be expected to persist beyond the period on which they are observed, and also in face of changing environments. Losing sight of this fact leads one to ask meaningless questions. Perfect examples of this are the forecast competitions which were conducted by the editors of the *International Journal of Forecasting*. There are a large number of techniques which can be used to extrapolate any given time series:

- Exponential Smoothing: Single, Double, Holt-Winters, Damped
- Time Series: ARMA, ARMAX, ARARMA
- Neural Networks: Numerous variants
- State Space Methods: Kalman Filter and variants

Authors were asked to design forecasting algorithms, and these were tested on real world data series. The hope was to learn which of these methods has the best performance in terms of forecasting.

Asking this question makes sense only within the context of a nominalist understanding of science. As long as all methods of pattern finding are regarded as being on potentially equal footing, one can ask whether some methods are more effective than others. From the realist point of view, the question does not make sense. The validity of any of the above methods will depend on how well the resulting model matches the real causal mechanism which generates the data. Since this cannot be assessed within the framework of the competition, we would expect completely random results in this forecast competition, just like the outcomes of a lottery. Some methods should work extremely well for some of series but not for others. Good performance on

the first half of a series would be no guarantee of good performance on the second half. Similarly, good performance in this competition would not say anything about good performance on a different set of time series. In fact this is exactly what was observed. After three competitions, the whole idea was abandoned, as nothing was learned from these competitions, except perhaps that simple methods work just as well as the most sophisticated ones. In fact, as we will discuss, simple methods can be expected to perform somewhat better.

3.4. Lost Causes

In an article of this name, Hoover (2006) documents how utilization of causal terminology declined dramatically in both economics and econometrics in the post-war period. Then, in the nineties there has been a dramatic upsurge. We reproduce one of the graphs depicting his findings in Figure 1 below.³

Hoover (2004) poses the question of why causal language was dropped for several decades and also what accounts for its recent comeback? He goes through several potential explanations, including the influence of logical positivism, as partial explanation. He suggests that the root cause is the rise and dominance of mathematical and theory based approach, to the exclusion of real world data analysis.

Zaman (2010) provides a detailed survey of the extant approaches to causality in the econometrics literature, and argues that all of them are seriously defective. The paper also discusses some promising new emerging lines of research. For present purposes, I would like to argue that

1. Causality is of central importance in economics and econometrics.
2. Current approaches to causality in econometrics, as taught in textbooks, are seriously deficient. This deficiency is due to poor methodology.

³ Figure 1 below is reproduced from Hoover (2004) with kind permission of Cambridge University Press.

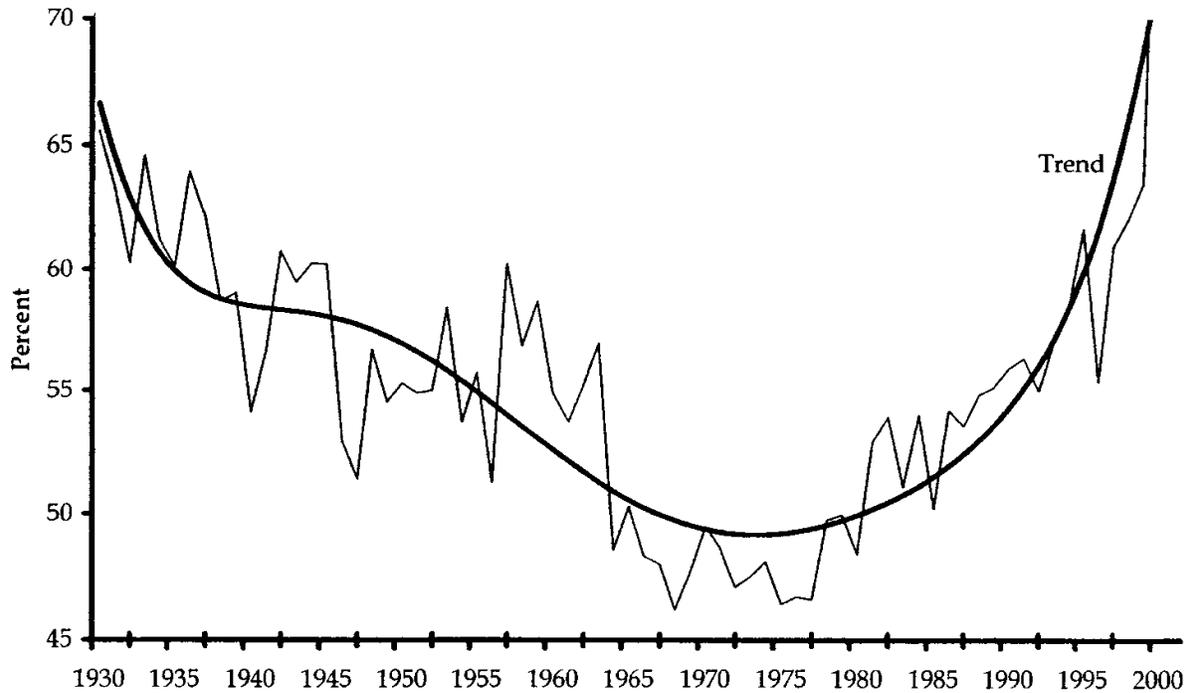


Figure 1. Lost Causes in Economics

Notes: Figure shows articles using the causal family as a fraction of all articles in the JSTOR archive of economics journals, 1930–2001. *Causal Family* = “cause,” “causes,” “causal,” “causally,” “causality,” or “causation.” Searches conducted June 2003. Heavy, smooth line is a sixth-degree polynomial trend.

Instead of re-hashing the detailed explanation and support for the above two propositions provided in Zaman (2010), I would like to provide anecdotal evidence to support and amplify the suggested explanation of Hoover for “lost causes” in econometrics. This also bears directly on the theme of the present paper.

Econometrics was originally thought of as a tool to utilize data to learn about the real world. This is what Hoover (2004) calls the “data first” approach. However, the (statistical & mathematical) “theory first” approach gained the upper hand for several decades in the postwar period. During our graduate training in econometrics, we learned all the mathematical and statistical apparatus required to analyze data. The focus was on methodology of statistical inference. Our graduate courses did not deal with real data, except in a perfunctory way to illustrate the use of statistical techniques. Even though my own Ph.D. thesis was in the area of theoretical econometrics, I did not run a real regression on a real data set until twelve years after graduation. This ties in with our theme that current econometric methodology is positivist, and is concerned purely with finding patterns in the data, and not with analyzing the reality behind the data.

Another good illustration is provided by the VAR methodology of Sims. Calculating the impulse responses to shocks requires a Cholesky decomposition which requires ordering the variables in a causal sequence. Since the whole point of the VAR methodology is to describe the correlation

structure without putting in any “real world” information, Sims originally resisted the suggestion that calculating impulse response requires the input of causal information. However, this point has been reluctantly conceded, and VAR models with the required causal sequencing are called Structural VAR or SVAR models. Examining papers using this methodology, we find that the causal information is put into the paper in a very ad-hoc and perfunctory fashion. That is, the most important and central real world concern is incorporated in a superficial way, while the focus of the analysis is on mathematical dynamics of the resulting system of equations. This reversal of priorities is a consequence of the domination of the positivist methodology. Causality cannot be captured in mathematical terms or in terms of nominal patterns, since it deals essentially with the real world.

The recent upsurge of use of causal language reflects a rising concern with real world data analysis. As Hoover contends, any approach which seriously engages with the data, will be forced to deal with causal concerns. For various reasons, including striking failures of the pattern finding approaches, there has been a rise in data oriented analyses and consequently a deeper engagement with causality in the recent past.

4. FINDING PATTERNS VERSUS LOOKING FOR CLUES

In the previous section, we discussed how the emergence of nominalist methodology has shaped research in econometrics. In this section, we discuss one of the two key differences between the nominalist and the realist approach. When we are looking for patterns, then the criterion is a good fit to the data; thus regression models are evaluated by their goodness of fit. In contrast, when we are looking for clues to understand reality, then we look for surprising fits. Each surprise generated by the data is a clue to a possible real structure which generates this surprise. This section is devoted to clarifying and explaining the difference.

4.1. The opposition between surprise and good fit.

One of the basic problems of econometrics is to discover the nature of a *stochastic* relationship between two variables. For instance, consider a set of data points (y_t, x_t) which reflect a stochastic relation $y_t = f(x_t, \varepsilon_t)$ that we wish to estimate. The function f is unknown, and the errors ε_t reflect several types of uncertainties, principal among which are:

1. Unobservable variables which affect the relationship
2. Relevant variables on which data is not available.
3. Relevant Variables not known to the econometrician to affect the relationship.
4. Compensations for errors in estimating the function f .

Standard operating procedure in this case would be to run a linear regression $y = a + b x + \varepsilon$. As a starting point, this procedure makes sense from the point of view of a realist methodology. Given two apparently independently generated time series, why should there be any relation between them? If there is a strong linear relationship, this is a surprising phenomenon, and worth investigating. Of course a test to assess the p-value of the strength of the relationship under the

null hypothesis that the two series are independent can be important to assess whether or not this is a surprising fit. How to construct a good test is not of relevance to the present discussion.

It is the second step which is missing from current econometrics textbooks, which are all based on positivist methodologies. Once we learn that there is surprising correlation between y and x , we must investigate the reasons for this correlation. But such reasons necessarily involve going beyond the names “ y ” and “ x ” and involve learning what reality is named by these variables. For example, if the two variables are alternative sets of measurements of the same underlying quantity, then we have a clear explanation for the linear relation. Alternatively, there could be a causal relationship such that $x \rightarrow y$, or that $y \rightarrow x$, or that some third variable $z \rightarrow x, y$. In this last case, movements in z would lead to the appearance of a relationship between y and x . There are other possibilities as well; for instance, x might happen to be a good proxy for a cause of y in the time period under consideration. Without being able to differentiate between these myriad possible reasons for significant coefficients in a linear regression, just the observation of a good fit $Y = a + b X$ does not add much to our knowledge of the real world. In an article entitled “Statistical Models and Shoe Leather,” Freedman (1991) pointed out the necessity of expending shoe leather in investigating causal relationships. Real world knowledge acquired by “getting our hands dirty” is required to dis-entangle complex possibilities, and desktop calculations and complex mathematical modeling cannot substitute for this.

Instead of taking it as a first step, as a clue to explore, conventional econometric methodology terminates at the discovery of a good fit. The discovery of a new relationship is announced to the world. Even worse is what happens when the linear regression does not fit well. In this case, we look for more complex models in an effort to fit the data. For the sake of simplicity, assume that we keep increasing the order of the polynomial until we find a good fit between y and x . It is obvious that this will always work, since a suitably high order polynomial will provide a perfect fit to all the data. What is the likelihood that a high order polynomial satisfying $y_t \cong P(x_t)$ represents the true relationship between the two variables? The chances are virtually zero that the polynomial has any meaningful interpretation. This is because following the procedure outlined will lead us to find a perfect fit between any two variables, while we know in advance that not all pairs of variables are related to each other. This is also called “over-fitting”, as a high degree polynomial minimizes or eliminates the errors, incorrectly making them a part of the relationship between y and x .

Since there is no surprise in our ability to find a high degree polynomial fitting the data, the polynomial itself just represents an observed pattern in existing data. Because such a pattern can always be found, we cannot ask the question of why such a pattern exists, and whether it reflects some real world relationship. Whereas finding a good linear fit is surprising, and calls for an explanation, finding a good quadratic fit is less surprising. As we increase the scope of the search and the degrees of the polynomial, the chances increase that the fit of the model is purely an observed pattern which does not reflect any underlying real relationship. The goodness of fit moves in a direction opposite to the surprise generated by the fit. The more models we try, the better the fit and the less the surprise.

Recognition of the problem exists. Many econometrics textbooks advise that only one model should be fitted to the data. This would certainly minimize the possibility of generating a

spurious fit, and the Hendry methodology (to be discussed later) is an honest effort to implement this idea. However, the implication here is that either the causal structure is left unexplored, or that it is specified in advance and not subject to investigation by data. This is a serious gap and deficiency. Equally serious, the prescription of fitting only one model is almost never followed in practice. Some consequences of this bad practice, almost guaranteed to lead to wrong models, are described below.

4.2. Tools to fit anything

We start with a finite data set, and seek to find a model which fits it. The required model must satisfy a large number of restrictions. In the ideal case, a model is given by the theory, and when we apply it, lo and behold, we find a miraculously good fit. This would be a wonderful confirmation of the theory, since it would be rather surprising to find a perfect fit on a first attempt. In practice, this almost never happens. We run through many different models until we find one which confirms our theory. Leamer (1978, 1983) has described the process of fitting a regression model as a “Specification search,” and argued that while this is useful for experimental data, it is either useless or misleading for observational data. This is because the large collection of tools at our disposal virtually guarantees that we can find a suitable model which conforms to whatever criteria we desire. The range of models we can try is infinite dimensional and limited only by our creativity, while the data set is fixed and finite. Test of residuals so strongly recommended by Hendry (1993, p. 24) have been appropriately called “indices of conformity” because they are not really tests. We can and do re-design the model to ensure that the residuals satisfy all tests.

What are the consequences of overfitting, the standard operating procedure in econometrics? As we have argued earlier, overfitting will almost certainly miss any true relationships which exist, because it will build the errors into the function in the process of minimizing them. We provide evidence that we have the “tools to fit anything” – the infinite dimensional variety of theoretical models capable of conforming to any hypothesis about reality can fit any finite dimensional data set.

Since Nelson and Plosser (1982) launched the literature, many authors have attempted to test whether macroeconomic time series are difference stationary or trend stationary. A lot of statistical and economic consequences hinge on the answer. Here is a list of the conclusions of authors⁴ who have studied the US annual GNP series:

Difference stationary: Nelson and Plosser (1982), Murray and Nelson (2002), Kilian and Ohanian (2002)

Trend Stationary: Perron (1989), Zivot and Andrews (1992), Diebold and Senhadji (1996), Papell and Prodan (2003)

Don't know: Rudebusch (1993)

⁴ This material has been taken from Atiq-ur-Rahman (2011), which should be consulted for detailed references to the articles cited in this paragraph.

As is evident, consensus has not emerged, and there has been no accumulation of knowledge with passage of time. In setting up a unit root test, we have a choice of framework within which to test, and choice of test statistics. Atiqurrahman (2011) has shown that these make a crucial difference to the outcome of the test. He has shown that for any time series, we can get whatever result (trend stationarity or difference stationarity) we desire by choosing these two factors in a suitable way.

As a second example, consider papers published which study the export-led growth (ELG) hypothesis for Indonesia. An alternative is the Growth-led Export (GLE). We can also have bi-directional causality (BD), as well as no causality (NC). There exist studies confirming all four hypotheses⁵:

ELG: Jung and Marshall (1985), Ram (1987), Hutchison and Singh (1992), Piazzolo (1996), Xu (1996), Islam (1998), Amir (2004), Liwan and Lau (2007)

GLE: Ahmad and Harnhirun (1992), Hutchison and Singh (1992), Pomponio (1996), Ahmad et al. (1997), Pramadhani et. (2007), Bahmani-Claire (2009).

BD: Bahmani-Oskooee et al. (1991), Dodaro (1993), Ekanayake (1999)

NC: Hutchison and Singh (1992), Ahmad and Harnhirun (1995), Arnade and Vasavada (1995), Riezman et al. (1996), Lihan and Yogi (2003), Nushiwat (2008)

As illustrated above, on economic issues of interest, we can find published results confirming or rejecting almost any hypothesis of interest. For example, whether or not purchasing power parity holds, whether or not debts are sustainable, whether or not markets are efficient, etc. etc. etc. One of the central pillars of macroeconomic theory is the consumption function. There is a huge literature, both theoretical and empirical on estimation of the aggregate consumption function. Thomas (1993) reviews the literature and writes that:

“Perhaps the most worrying aspect of empirical work on aggregate consumption is the regularity with which apparently established equations break down when faced with new data. This has happened repeatedly in the UK since the 1970’s. ... the reader can be forgiven for wondering whether econometricians will ever reach the stage where even in the short run their consumption equations survive the confrontation with new data.”

Magnus (1999) challenged readers to find an empirical study that “significantly changed the way econometricians think about some economic proposition.” We provide a more precise articulation of the challenge to conventional methodology currently under discussion. Our graduate students take courses, pass comprehensive exams, and write theses to qualify for a Ph.D. To ensure that they are adequately grounded in econometrics, suppose we add the following two requirements⁶:

⁵ This material has been taken from Ph.D. Thesis of Hendri (2012), which should be consulted for detailed references to the articles cited in this paragraph.

⁶ This could be called the magnified Magnus challenge.

Test 1: Take any economic theory, and support it by econometric evidence. Or, a simpler version: for any two arbitrarily chosen variables X and Y, produce a regression showing that X is a significant determinant of Y.

Test 2: For any current empirical paper from the literature, reach conclusions opposite to those reached in the paper using standard econometric techniques.

How can we accomplish this? There is huge range of techniques, all of which can be demonstrated as acceptable practice using papers published in top ranked journals. We list some of the major ones:

1. For each theoretical variable can be represented by a wide variety of observable time series. In many cases, a suitable series can be constructed, to suit requirements of the researcher.
2. Additional control variables, dynamic structure, length of lags chosen, provide a large number of models to test for conformity to the desired hypothesis.
3. Large numbers of tests, many known to have low power are available. Formulate an appropriate null hypothesis and fail to reject it using a test of low power.
4. Unit Roots, Nonlinearity, Functional Forms, as well as ad-hoc assumptions create a huge range of possible models to try, one of which will surely work to confirm the desired null hypothesis.

Virtually any outcome can be achieved by varying these parameters. Any professional econometrician worth his salt would easily be able to pass these tests without breaking sweat. Graduate students might have more trouble, but only really incompetent ones would be seriously delayed in graduation by this additional requirement.

Unlike most tests, where passing counts for success, failing these tests is a fundamental requirement for a good methodology. Based on the assumption that currently acceptable conventional methodological practice in econometrics can pass these tests with flying colors, we assert that:

PROPOSITION: Any methodology which can pass tests 1 and 2 is completely useless as a methodology for production of knowledge.

It is immediately obvious that any methodology which can prove *and* disprove all possible economic theories is useless. We now pause to briefly discuss the Hendry methodology (also called the LSE methodology), which has some special characteristics.

4.3. Failure of Hendry's Methodology

Conscientiously followed, Hendry's methodology has a chance of failing the two tests proposed in the previous section. As such, it is a candidate for a viable methodology, even though our tests proposed are very weak – failing both of them is just a first and preliminary step to the possibility of that the methodology might be workable. A comprehensive and detailed examination of

strengths and weaknesses of the Hendry methodology (termed LSE methodology by them) is available in Faust and Whiteman (1997). To avoid length, we assume this as background knowledge, and provide an abbreviated discussion. We discuss the special features of the Hendry methodology, and also why it remains a positivist methodology, and hence is not capable of advancing knowledge.

First, when we start with a general model, and simplify in the systematic way proposed by Hendry, it is possible to control the type I and II error probabilities of the tests. Thus, if we end up with a model satisfying all the properties required by the Hendry methodology, it does possess an element of surprise. As such, it is a worthy candidate for examination as to whether it represents some feature of the real world which makes it valid beyond the data set currently under examination. Unfortunately, Hendry's methodology terminates with the finding of a "congruent" model – what realism requires is explorations of features of reality which account for the congruence; why does the model fit well? Without looking into this question, we cannot make progress in learning about the world.

While the Hendry methodology followed rigorously has the potential of picking up interesting patterns in the data (worthy of further examination), pragmatic concerns vitiate this advantage. If the Hendry methodology leads us to a final model which does not fit very well – does not display parameter constancy for example, the methodology requires us to stop and declare failure. Unfortunately, editors won't accept such papers, and authors must either publish or perish. Virtually no one would stop, and all authors will then introduce new variables (termed "designer" variables by Faust and Whiteman) or twists and turns in the model to try to make it fit. This immediately takes one outside the Hendry methodology, but is routinely done even by the best practitioners. Furthermore, the consequences of this practice are not accounted for, as degrees of freedom lost by testing different models are "reclaimed" in the final model. Because these elements of ad-hockery inevitably find their way into practice, Hendry methodology does not end up doing significantly better than conventional alternatives. Models produced by it routinely break down in forecasting, and no striking discoveries of real constant relationships are on the record, as documented by Faust and Whiteman (1997).

5. REALIST METHODOLOGY: INFERRING STRUCTURE FROM CLUES

A huge number of patterns can be found in any given data set. Some of these patterns are accidental – they pertain purely to the data set at hand and will not extend to any larger data set. Other patterns are real – they result from some underlying real causes which generated the data. These real patterns have varying degrees of autonomy. The first task of the data analyst is to differentiate between the real patterns and the accidental patterns, and to isolate the patterns which are autonomous. This part involves some statistical and mathematical analysis, to assess degree of surprise, as well as statistical significance of departures from the null of no pattern. The second task is to look for the real causes which lead to this autonomy. This second step is inevitably non-mathematical, because it involves going beyond the names of the data set to explorations of the real world. This cannot be taught, except by apprenticeship and examination of case studies of how real causes have been discovered. Freedman (2008) and Freedman et. al. (2007) offers several examples of success and failure in statistical methods used for discovery of causal effects. These case studies clearly establish the importance of qualitative and informal

knowledge about real world. Conventional econometric methodology is a failure because it is merely an attempt to find patterns in the data, without any tools to assess whether or not the given pattern reflects some real forces which shape the data. We provide an elementary example to clarify this distinction.

5.1 Surprising Versus Strong Correlations

Let $\{X_t, Y_t\}$ be a sequence of observations on two variables. Suppose these are *apparently stationary*: by this we mean that the sequence of differences from the means $\{X_t - m\}$ does not fail any tests for being identically distributed⁷. Then if X and Y are correlated, this means that positive values of the two difference series occur together, as do negative values. This is surprising; why should two apparently unrelated series move together? Before proceeding further, it is important to assess whether the non zero correlation is *significantly* different from zero. This is to rule out the possibility that the pattern is due purely to chance variations. Apart from purely statistical aspects, there are procedural aspects which must be taken into account very carefully, whereas current best practice in econometrics treats them in cavalier fashion. A spurious relationship in Zaman (2010) was obtained by searching over all pairs of correlations between ΔC and ΔY for consumption and income series in the WDI database. The maximum correlation over a collection of pairs has a vastly different distribution from a single correlation. The critical values for significance would be substantially higher in this case. Yet it is common practice to ignore the search, and “reclaim” degrees of freedom lost by searching over models when presenting final results. According to the nominalist methodology, the process by which the pattern is found is not relevant, so it seems reasonable to ignore the search process which led to the final regression model. There is an uneasy understanding that the search process does impact on the statistical properties of the model discovered, but it is largely ignored in practice. For a realist methodology, the surprise in finding the model is of the essence. We cannot ignore the search process which led to the model, since it has a huge impact on the surprise. Just like econometricians have learned to adjust standard errors for heteroskedasticity, we need to present search-adjusted p-values of significance. Quite a bit is known about this, but is ignored due to methodological blinders.

Why do we restrict correlation to apparently stationary series? Again this has to do with surprise. Consider the possibility that X and Y are both series with strong increasing trends. Then the mean will occur near the middle, all of the initial set of differences from means from both series will be negative, and all of the later sets of differences will be positive. Then positives and negatives will occur together, and there will be strong correlation between the two series. However, there is no surprise in this movement together – it is a clear consequence of the trend. Thus, even an excellent fit is meaningless, and does not furnish any clues to any underlying reality. This demonstrates the importance of surprise, as opposed to goodness of fit.

⁷ Note that this is an operational definition. Given a series, we can verify whether or not it has the property required – though there is an element of ambiguity because we have not specified all the tests required. As opposed to this, the assumptions that X_t is a stationary series can never be verified, because in all finite data sets, there exist arbitrarily close stationary approximations to the data distribution.

5.2. Using Clues to Uncover Causal Structure:

Having found a surprising and significant pattern in the data, a realist must ask what it means. This step of going beyond the observable to the reality behind the observables was banned by the positivists. To undo the damage caused by positivist thinking, we must re-learn this skill which is currently not part of our textbooks and training in econometrics. In a realist methodology, finding of a pattern is a starting point for research and not the terminal point, as in the nominalist methodology.

As discussed earlier, the concept of probability is inherently unobservable, because it makes reference to possibilities which might occur but do not. This is why we teach absurd and hard to digest definitions to puzzled undergraduates, who have a clear intuitive understanding of the idea that even though the coin came up heads, it could with equal likelihood, have come up tails; see Zaman (2008) for an extended discussion. Once we force them to discard their naïve intuitions and accept the idea that probability refers to limiting frequencies in an impossible to conduct sequence of trials, we force them to abandon intuitions and rely solely on complex and abstruse mathematics for understanding what is basically a simple concept. In exactly the same way, causality is inherently unobservable. Nonetheless, we all have strong intuitions about what it is, and routinely deduce causal effects in real life from a variety of clues. By framing the problem in a way that makes it impossible to apply these intuitions, we impede the progress of knowledge. Oaksford and Chater (2007, Introduction) have described the revolution in artificial intelligence that resulted from abandoning positivist models of human knowledge and reasoning, and adopting more flexible realist models developed by cognitive science. In a similar way, learning to reason about the reality behind the data may resolve the frustration about econometrics expressed by Magnus (1999) and Freedman (1997) cited earlier.

Our goal in this section is to illustrate, via a simple example, how intimately real world knowledge is involved in the deduction of causality from data clues. This means that any process which uses nameless data (X,Y) without knowing what these variables mean, cannot lead to final conclusions about causality. Data can provide clues, even strong clues, but ultimate verification requires deeper investigation than is available from data alone.

Consider the following (artificial) data about admission into Real University (RU). We assume rough equality of backgrounds in the applicant pools of males and females for the reasoning which follows.

Applicants →	Males	Females
Total	1000	1000
Admitted	300	700
Percent	30%	70%

One would immediately guess that RU discriminates against males. The discrepancy in the admission ratio is so large that almost any method for assessing significance of difference would conclude that the null hypothesis of equality of admission ratios would be rejected. Note that we do not consider the possibility that Admission is a cause of Gender – students who are admitted get sex-change operations. Correlations are always symmetric, while causality is always unidirectional, so the pattern in the data can never be a depiction of reality. Next suppose that the university has two major schools, sciences and humanities. Subdivided into these categories, admissions data is as given below:

	Male Applicants		Female Applicants	
School	Social Science	Engineering	Social Science	Engineering
Applicants	100	900	900	100
Admitted	100	200	690	10
Percent	100%	22%	77%	10%

Now it seems clear that the discrimination is against females. In both Social Science and Engineering, male candidates are preferred admits by a large margin. Males apply overwhelmingly to Engineering, which is hard to get into, while females apply overwhelmingly to Social Science, which is easy to get into. This accounts for the difference in overall admit ratios.

The surface pattern of overall admits leads to a clear causal conclusion. Deeper investigation of structure reverses it completely. This shows that we can only get clues to causality from data, and only reach tentative conclusions – there may be hidden variables, which would change the picture. Knowledge about the real world is messy and it is not possible to encapsulate it in elegant mathematical formulae. Suppose we expend some shoe leather and go to the admissions department of the university in question. We find out that all applications are processed uniformly, without regard to which department the candidates have applied to. Candidates are ranked by SAT scores, and admission is given to the top 2000. Alternatively, we learn that candidates choose departments *after* they are admitted. Clearly we would renounce our causal conclusion and think about alternative explanations for the discrepancy in the admission ratios. Also, sitting in an admissions committee meeting at RU would give us important qualitative information about gender bias in admissions in a way that could not be easily captured in any data.

6. CONCLUSIONS

When the proverbial apple fell on Newton’s head, he thought through to the invisible force of gravity which led the apple to fall. Positivist methodology would have led him consider the fit of his apple to the patterns made by apples falling in orchards – these are observable, while gravity is not. Positivist methodology dominated social sciences for the major part of the twentieth century. Positivists eschewed the real underlying structures, and said that scientific laws describe patterns in the observables. The particular, concrete, historical and unique events are therefore

not parts of scientific knowledge. In fact, these play an essential role in the body of human knowledge. As the errors and weaknesses of the positivist methodology became apparent, there has been an effort to break away from this philosophy. Differing amounts of progress has been made in different disciplines, but a lot remains to be done.

In this essay, I have focused on the effects of positivist methodology on econometrics. I have shown that current approaches based on finding patterns in data cannot lead to increases in our knowledge. Data and observations are only partial guides. Striking patterns in the data are clues to real structures. Exploring these clues requires acquiring and organizing information about the real world. There are two keys to developing a realist methodology for econometrics. The first is to measure surprise in how well a data fits a model. The second is to try and look for real causal mechanisms and structures which explain away this surprise. For both of these activities, there exists a substantial amount of existing material. Measurements of surprise via entropy and other techniques have been discussed for decades. There is a very active debate currently going on about how to deduce causal relations from observational data. Even though these activities are central to a realist methodology, currently these research areas are peripheral and marginalized because of the dominant positivist methodology. Progress requires changing priorities.

REFERENCES

- Atiq-ur-Rahman, (2011). "Impact of Model Specification Decisions on Unit Root Tests," *International Econometric Review*, 3(2), 22-33.
- Ayer, A. J. (2008). interview with Brian Magee, www.youtube.com/watch?v=4cnRJGs08hE
- Barzel, Y. (1972). "The rate of technical progress: The "Indianapolis 500" *Journal of Economic Theory*, 4(1), 72-81.
- Freedman, D.A. (1997). From Association to Causation via Regression. In *Causality in Crisis?* ed. V. McKim and S. Turner. University of Notre Dame Press, South Bend, (with discussion) 113-82. Reprinted in *Advances in Applied Mathematics*, 18, 59-110.
- Freedman, D. A (1991). "Statistical Models and Shoe Leather". *Sociological Methodology*, 21, 291-313.
- Freedman, D.A. (2008). On types of scientific enquiry: Nine success stories in medical research. In *The Oxford Handbook of Political Methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady and David Collier. Oxford University Press, Oxford.
- Freedman, D.A.; Pisani, R. and R. Purves (2007). *Statistics*. Norton, New York

Zaman- Methodological Mistakes and Econometric Consequences

- Hendri, S.B. (2012), "Export Led Growth Hypothesis: An Analysis of Indonesia's Trade Policy Pre and Post 1997 Financial Crises" Ph.D. Thesis submitted to IIIE, International Islamic University of Islamabad.
- Hendry, D.F. (1993). The Roles of Economic Theory and Econometrics in Time-Series Economics. In *European Econometric Society*, Invited address, Stockholm.
- Hoover, K. D. (2004). "Lost Causes" *Journal of the History of Economic Thought*, 26(2), 149-164.
- Hoover, K.D. (2005). "The Methodology of Econometrics", article in Palgrave *Handbooks of Econometrics, volume 1: Theoretical Econometrics*.
- Leamer, E. (1978). *Specification Searches: Ad Hoc Inference with Non Experimental Data*. John Wiley and Sons Inc., New York.
- Leamer, E. (1983). "Lets take the con out of Econometrics", *The American Economic Review* 73(1), 31-43.
- Magnus, J.R. (1999). The success of econometrics. *De Economist*, 147(1), 50-58.
- Manicas, P. (1987). *History and Philosophy of the Social Sciences*, Oxford: Basil Blackwell
- Manicas, P. (2006). *A Realist Philosophy of Social Science: Explanation and Understanding*, Cambridge: Cambridge University Press
- Oaksford, M. and Chater, N. (2007). *Bayesian Rationality*. Oxford: Oxford University Press
- Russell, Bertrand. (1905). "On Denoting", *Mind*, 14, 479-493.
- Sims, C.A. (1980). "Macroeconomics and Reality," *Econometrica* 48, 1-48.
- Frederick, S. (1977). *The Structure of Scientific Theories* 2nd edn. Chicago: University of Illinois Press
- Thomas, R. L. (1993). *Introductory Econometrics: Theory and Applications*, 2nd Edition Longman.
- Van Fraassen, Bastiaan, C. (1980). *The Scientific Image*, New York: Oxford University Press
- Wong, S. (2006). *The Foundations of Samuelson's Theory of Revealed Preference*, Routledge,
- Zaman, A. (1996). *Statistical Foundations of Econometric Techniques*, Academic Press.

Zaman, A. (2008). Post Positivist Probability Available at SSRN:
<http://ssrn.com/abstract=1319048> or <http://dx.doi.org/10.2139/ssrn.1319048>

Zaman, A. (2010). "Causal Relations via Econometrics", *International Econometrics Review*,
2(1):36-56

Zaman, A. (2011). The Normative Foundations of Scarcity, to appear in *Real World Economics Review*, 60