ECONOMETRIC METHODOLOGY AND THE SCIENTIFIC STATUS OF ECONOMICS

MICHAEL SHERLOCK

Senior Sophister

To many economists, econometrics is a method of exploring many of the heated debates in a clinical, scientific way. However in this essay, Michael Sherlock argues that, despite the myriad rules and rigidity in the models, econometrics can be seen as a deeply flawed attempt by economists to legitimise their subject in the eyes of broader scientific disciplines. He discusses the weaknesses inherent in the field and explains how they clash with the standard ideas of what constitutes a science.

Introduction

This essay seeks to engage in the debate on the scientific status of economics by considering whether econometric methodology constitutes a scientific process sufficiently similar to that of other sciences in order that the epithet of ‘science’ can be conferred on the discipline of economics. The essay will first attempt to clarify what is meant by the term ‘science’ and situate it within the context of economic history. It will then proceed to discuss the crux of the intellectual debate, considering in turn the various problems that have been identified within econometric methodology. The conclusion will reflect upon the issues raised and take a position as to whether economics can be classified as a science.

Economics as a science: a false hypothesis?

‘Science is a public process. It uses systems of concepts called theories to help interpret and unify observation statements called data, in turn data are used to check or test the theories’ (Hendry, 1980: 388). Although various definitions of the term ‘science’ exist, what is interesting about this one, written by an economist, is that it highlights the importance of *empirical testability* within economic models. Traditionally, classical economics (economics without econometrics) largely consisted of deductive theories, wholly devoid of any real data, which relied on quite complicated mathematics for their existence (consider general competitive equilibrium, a fundamental axiom of microeconomics whose existence was eventually proved using fixed point theorems). Accordingly, classical economics could not be considered a true science in any meaningful sense of the word. This led to an identity crisis in the economics profession which has resulted in the birth of econometrics, a branch of economics which provides a series of methods necessary for the analysis of data.

Pearson (1938) championed the ‘unity of science’ principle which conceived that the essence of any science consists of a scientific method. Ritchie (1923) concurred, arguing that the only constant in science was this scientific method and that while scientific theories are in a constant state of flux, the process used to generate these theories has remained static. This stimulated debate among economists as to whether econometrics provided economics with this much needed ‘scientific process’, thereby providing the discipline with the intellectual legitimacy which it sought. The essence of econometric methodology is the development of a framework which seeks an adequate ‘conjunction of economic theory and actual measurement, using the theory and technique of statistical inference as a bridge pier’ (Haavelmo, cited in Pesaran & Smith, 1992: 9). Ever since the Cowles commission, regression analysis has become the empirical workhorse of econometrics, apparently providing the methodology of the scientific process at last.

‘It must be possible for an empirical scientific system to be refuted by experience’ (Popper, 1959: 41). This statement encapsulates the principle of *falsifiability* - the fact that in order for a theory to be considered scientific it must be capable of being disproved. Much of the controversy surrounding econometric methodology is whether it is capable of testing theories. Ostensibly, it seemed to do so. Nash (2007: 56) highlighted the fact that, from the outset, econometric methodology appeared to graft a scientific method onto mainstream economics ‘as now, apparently, hypotheses can be tested empirically and also falsified, thereby satisfying the scientific method’. In the early days, many commentators were less sanguine and even displayed open scepticism about the ability of econometrics to achieve this objective. Spanos (1986: 660) best articulated this position when he said ‘No economic theory was ever abandoned because it was rejected by some econometric test nor was a clear-cut decision between competing theories made in lieu of such a test’. To disambiguate the position we need to examine in depth the econometric process itself.

A ‘failure to accept’ the econometric methodology

Koutsoyiannis (1973) has identified the following steps as the core of econometric methodology: formulation of maintained hypothesis, testing of maintained hypothesis, evaluation of estimates and evaluation of the model’s forecasting validity. A cursory glance suggests the pre-eminence of hypothesis testing within the overall framework of regression analysis. Koutsoyiannis extols the benefits of such; and he adds that it confers scientific status on classical economics by virtue of the very fact that it is capable of sustaining rigorous testing. Many authors are critical of such claims. Hypothesis testing essentially involves what Koop (2005: 80) has referred to as ‘knocking down the straw man’, i.e. rejecting the null hypothesis and thereby establishing statistical significance. However, such a process is riven with a variety of interrelated problems.

Firstly, a finding of statistical significance does not necessarily denote scientific significance. Popper (1959: 23) defines a scientifically significant effect ‘as that which can be regularly reproduced by anyone who carries out the appropriate experiment in the way prescribed’. Much research has highlighted the remarkably high incidence of inability to replicate empirical studies in economics (Dewald et al., 1986). Hence, an econometric finding of statistical significance cannot be considered scientifically significant in any meaningful way. The problem is that the nature of data in a non-experimental discipline such as economics makes reproducibility impossible. This, in turn makes testability and falsifiability impractical, thereby rendering the whole process *de facto* unscientific. Kennedy (2003: 8) describes economic data as being ‘weak’ which refers to the fact that many of the forces governing economic behaviour are unquantifiable, being neither numerical nor measurable. O’Dea (2005: 40) even contends that they cannot be truly considered ‘economic’ and argues that the unwillingness of economists to consider such forces ‘flies in the face of its claim to be scientific’.

Some of the desirable features of any science are those of objectivity and precision. Regarding the latter, the fact that ‘outcomes are only probable to a given level of confidence, places econometrics and hence economics into a realm which is too imprecise to be deemed science’ (Nash, 2007: 57). As it is very often human behaviour that is being modelled, exact or deterministic relationships are impossible. Researchers compensate for this implicit uncertainty through the use of inferential statistics based on probability distributions. Consequently, levels of significance are assigned to outcomes. When one carries out a hypothesis test it is always at a given level of significance. It should also be noted that this imprecision is captured in the linguistic register of the terminology employed - it is best practice never to say that one rejects a null hypothesis, instead one employs the term 'fails to accept'. This highlights the fact that econometrics is 'a language for communicating results as well as a set of methods of analysis' (Krueger, 2001: 10). At an alternative level of significance, a previously statistically insignificant regression coefficient may become statistically significant. This arbitrary use of significance levels raises the interrelated question of objectivity.

Scientific credibility demands objectivity. Keuzenkamp and Magnus (1995) took issue with such an arbitrary use of significance levels whilst Berkson (1938) noted that for asymptotic samples, any null hypothesis was likely to be rejected and suggested that the choice of level should be decided by such pragmatic considerations. Unfortunately, in practice, choice is usually determined by the subjective needs of the econometrician; Keuzenkamp and Magnus (1995: 16) note that 'the choice of significance level seems arbitrary and depends more on convention and, occasionally, on the desire of an investigator to reject or accept a hypothesis rather than on a well-defined evaluation of conceivable losses that might result from incorrect decisions'. That being the case, the objectivity of the econometric process is severely compromised. This leads to a related problem extensively observed in econometrics, that of data-mining.

Leontief (1971: 390) once presciently commented on the state of econometrics describing it as 'an attempt to compensate for the glaring weakness of the data base available to us by the widest possible use of more and more sophisticated statistical techniques'. This emphasis on statistical analysis has lead to the problem of data-mining which has been frequently cited as a major source of evidence against econometrics' claim to scientific status. Data-mining consists of 'moulding or selecting models based only on an ability to pass desired statistical tests rather than underlying theory' (Hansen, 1996: 1408). The mainstay of regression analysis is the linear regression model. Over the years this model has been subjected to a dizzying array of statistical tests (heteroscedasticity, autocorrelation etc.) while undergoing significant refinement in order to deal with more complex data sources such as panel data. Such statistical myopia has led to an increased incidence of spurious regressions. For example, Hendry (1980) in his seminal paper quotes the reported case in which researchers found a higher correlation between annual inflation and cases of dysentery than that between annual inflation and the rate of change of excess money supply. This undermines scientific credibility as it becomes 'meaningless to talk about confirming theories when spurious results are so easily obtained' (Hendry 1980: 395).

The fourth step in Koutsoyiannis' econometric methodological framework (evaluation of model's forecasting validity) emphasises the importance of prediction for regression analysis. In fact, Friedman (1940: 658) argues that 'the real test of a theory' lies in its predictive ability. It is in this very area that many critics have based their refusal to accept the scientific claim of economics. Prediction implies a causal relationship and to establish this requires the use of a series of statistical assumptions (Gauss-Markov assumptions) regarding the disturbance or stochastic error term. The principle assumption is that of *zero conditional mean* which establishes a *ceteris paribus* relationship between regressor and regressand, without which a causal relationship cannot be established. In essence it is 'an attempt to control statistically what ideally would be controlled experimentally' (O'Dea, 2005: 38). Nash (2007: 56) contends that such a strong assumption nullifies the scientific process by 'invalidating the scientific status of the underlying theory'. This also impinges on the falsification issue *vis-à-vis* the Duhem-Quine problem. The Duhem-Quine problem raises the potential unreliability of a falsity finding based on hypothesis testing as a rejection of a hypothesis 'could well be due to any number of flaws with the buried assumptions rather than the falsity of the hypothesis under examination' (O'Dea, 2005: 39).

Conclusion

'An ability to use the laws of statistics/mathematics to test and potentially reject hypotheses from theory is a necessary but insufficient condition to justify economics as a science' (O'Dea, 2005: 40). This essay has shown that not only does classical economics not deserve the 'science' epithet but neither does neo-classical economics as the empirical testing process known as econometrics is not sufficiently robust to be considered a true scientific process. Although econometric methodology has added a degree of testability to the discipline, econometrics is in essence a statistical process not a scientific one. However, this does not mean that it is not a valid study in its own right as it may lead to the development of a satisfactory scientific methodology in the future and as such can be considered a milestone in economic history.

Bibliography

Berkson, J. 1938. 'Some Difficulties of Interpretation Encountered in the Application of the Chi-Squared Test'. *Journal of the American Statistical Association* 33:526-536.

Dewald, W., Thursby, J. & Anderson, R. 1986. 'Replication in Empirical Economics: The Journal of Money, Credit and Banking Project'. *American Economic Review* 76:587-603.

Friedman, M. 1940. 'Review of Jan Tinbergen. Statistical Testing of Business Cycle Theories, II: Business Cycles in the United States of America'. *American Economic Review* 30:657-661.

Hansen, B.E. 1996. 'Review Article, Methodology: Alchemy or Science?'. *The Economic Journal* 106:9:1398-1413.

Hendry, D. 1980. 'Econometrics: Alchemy or Science?'. *Economica* 47:188:387-406.

Kennedy, P. 2003. *A Guide to Econometrics* (5^{th} ed.). Oxford: Blackwell Publishing.

Keuzenkamp, H. 2000. *Probability, Econometrics and Truth: The Methodology of Econometrics*. Cambridge: Cambridge University Press.

Keuzenkamp, H.A. & Magnus, J.R. 1995. 'On Tests and Significance in Econometrics'. *Journal of Econometrics* 67:5-24.

Koop, G. 2004. *Analysis of Economic Data* (2^{nd} ed.). London: Macmillan.

Koutsoyiannis, A. 1973 *Theory of Econometrics: Ann Introductory Exposition of Econometric Methods*. London: Macmillan.

Krueger, A.B. 2001. 'Symposium on Econometric Tools'. *The Journal of Economic Perspectives* 15:4:3-10.

Leontief, W. 1971. 'Theoretical Assumptions and Nonobserved Facts'. *The American Economic Review* 61:1:1-7.

Nash, I.G. 2007. 'The Scientific Status of Economics and Econometric Methodology'. *Student Economic Review* 21:53-58.

O'Dea, C. 2005. 'Econometric Methodology and the Status of Economics'. *Student Economic Review* 19:37-42.

Pesaran, M. & Smith, R. 1992. 'The Interaction between Theory and Observation in Economics'. *Cambridge Working Papers in Economics* 9223.

Pearson, K. 1938 [1892]. *The Grammar of Science*. London: Everyman Edition.

Popper, K.R. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Ritchie, A.D. 1923. *Scientific Method: An Inquiry into the Character and Validity of Natural Laws*. London: Routledge and Kegan Paul Ltd.

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