When John Maynard Keynes likened Jan Tinbergen’s early work in econometrics to black magic and alchemy, he was expressing a widely held view of a new discipline. However, even after half a century of practical work and theorizing by some of the most accomplished social scientists, Keynes’ comments are still repeated today.

This book assesses the foundations and development of econometrics and sets out a basis for the reconstruction of the foundations of econometric inference by examining the various interpretations of probability theory which underlie econometrics. Keuzenkamp contends that the probabilistic foundations of econometrics are weak, and, although econometric inferences may yield interesting knowledge, claims to be able to falsify or verify economic theories are unwarranted. Methodological falsificationism in econometrics is an illusion. Instead, it is argued, econometrics should locate itself in the tradition of positivism.

**Hugo Keuzenkamp** is professor in applied economics at the University of Amsterdam and director of the Foundation for Economic Research in Amsterdam. Until recently he was the editor-in-chief of *Economisch Statistische Berichten*, a Dutch weekly periodical on economic policy. He has published widely in journals, including *Economic Journal, Journal of Econometrics* and *Journal of Economic Surveys*. 
Probability, Econometrics and Truth

The methodology of econometrics

Hugo A. Keuzenkamp
## Contents

*Introduction*  

1. The philosophy of induction  
2. Probability and indifference  

*Intermezzo: a formal scheme of reference*  

3. Reactive frequency and induction  
4. Probability and belief  
5. The theory of simplicity  
6. From probability to econometrics  
7. Econometric modelling  
8. In search of homogeneity  
9. Positivism and the aims of econometrics  
10. Probability, econometrics and truth  

*Personalia*  

*References*  

*Name Index*  

*Subject Index*

---

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>vii</td>
</tr>
<tr>
<td>1 The philosophy of induction</td>
<td>1</td>
</tr>
<tr>
<td>2 Probability and indifference</td>
<td>21</td>
</tr>
<tr>
<td><em>Intermezzo: a formal scheme of reference</em></td>
<td>28</td>
</tr>
<tr>
<td>3 Reactive frequency and induction</td>
<td>33</td>
</tr>
<tr>
<td>4 Probability and belief</td>
<td>67</td>
</tr>
<tr>
<td>5 The theory of simplicity</td>
<td>98</td>
</tr>
<tr>
<td>6 From probability to econometrics</td>
<td>119</td>
</tr>
<tr>
<td>7 Econometric modelling</td>
<td>142</td>
</tr>
<tr>
<td>8 In search of homogeneity</td>
<td>180</td>
</tr>
<tr>
<td>9 Positivism and the aims of econometrics</td>
<td>213</td>
</tr>
<tr>
<td>10 Probability, econometrics and truth</td>
<td>262</td>
</tr>
<tr>
<td><em>Personalia</em></td>
<td>276</td>
</tr>
<tr>
<td><em>References</em></td>
<td>281</td>
</tr>
<tr>
<td><em>Name Index</em></td>
<td>299</td>
</tr>
<tr>
<td><em>Subject Index</em></td>
<td>307</td>
</tr>
</tbody>
</table>
Introduction

Probability begins and ends with probability.

Keynes ([1921] CW VIII, p. 356)

When John Maynard Keynes accused Jan Tinbergen of practising black magic and alchemy, econometrics was still in its infancy. A critical attitude to econometrics was legitimate, as it would have been for any novel enterprise. Stubborn perseverance on behalf of the pioneers of econometrics is natural as well. However, after more than half a century of development by some of the most brilliant social scientists, and much practical experience, Keynes’ comments are repeated today by respected authorities. Has it all been in vain?

Not quite. It is true that the aspirations (or pretences) of econometrics and the accomplishments still tend to be divided by a gap, which, in turn, tends to damage the credibility of the whole discipline. Many of econometrics’ results remain controversial. Some critics claim that even the most basic aim, the measurement and quantitative description of economic phenomena, has not been accomplished. Econometric evidence has been compared with the evidence of miracles in Lourdes. Some deplore the waste of electricity used for econometric computer calculations. But a fair appraisal of contemporary econometrics cannot deny that a number of interesting empirical lessons have been learnt. The verdict that the econometric exploration was all in vain can only result from a wrong interpretation of econometric aims.

This book is a methodological investigation of this exploration. It confronts the aims with the methods and with the philosophical as well as the probabilistic foundations of those methods. It concludes that the achievements of econometrics can be found where its aspirations are put in the positivist tradition. Positivism is a philosophy which has been declared dead by many. It should be resurrected.

Positivism has an ancestor in David Hume, one of the founders of British empiricism (the forerunner of positivism). Hume ([1748] 1977, p. 114) encouraged his readers to ask about any book in their libraries,
‘Science is Measurement’, the original motto of the Cowles Commission for Research in Economics (which had a leading role in shaping formal econometrics), put econometrics clearly in the positivist tradition. Twenty years later, in 1952, this motto was changed to ‘Theory and Measurement’, reflecting the ambitions of a younger generation of researchers (headed by Trygve Haavelmo and Tjalling Koopmans) to integrate econometrics with (neoclassical) economic theory and formal probability theory. The new tradition diverted econometrics to Neyman–Pearson testing procedures, away from the positivist tradition of Karl Pearson (father of Neyman’s companion Egon), Fisher and Jeffreys. Simultaneously, in the philosophy of science positivism came under attack and was replaced by methodological falsificationism. Chapter 1 discusses this philosophy, chapters 2 to 4 deal with different approaches in probability theory. I claim that the Cowles programme in econometrics, with its Neyman–Pearson foundation and a philosophical sauce of methodological falsificationism, has done the reputation of econometrics much harm. This claim is elaborated in the chapters which follow the discussion of the various probability theories. The transition from probability theory to econometrics is shaky, as chapters 6 and 7 demonstrate. Chapter 8, which presents a case study in one of the best episodes of applied econometric inference, shows that the sampling and testing metaphors which dominated econometrics can lead to serious self-deceit. Chapters 9 and 10 bring various arguments together and recommend the positivist tradition, in which econometrics was born and to which it should be brought back again.

Finally, what about truth? Does econometrics, based on the right kind of probability, yield ‘true knowledge’? Not so. The quest for truth, which dominates much of contemporary econometrics, should be abandoned. If econometricians are able to deliver useful approximations to empirical data, they achieve a major accomplishment. What defines ‘useful’ is an intricate matter, which can only be clarified on a case-by-case basis. A model which is good for one purpose, may be inappropriate for another.

I hope that the reader will allow the author a few disclaimers. First, the focus on econometrics does not mean that there are no other ways of doing empirical economics, neither is it intended to suggest that purely theoretical work is not interesting. This book intends to provide a com-
implement to books on economic methodology which tend to ignore the strengths but also weaknesses of econometric inference.

Secondly, even though the book focuses on econometrics, I neglected some approaches that might warrant discussion. For example, there is hardly discussion of non-parametric inference, bootstraps, or even many specific cross-section topics. Many of the fundamental themes discussed here apply equally to econometric approaches which are beyond the scope of this book.

Thirdly, I have assumed that the readers of this book will be econometricians who are interested in the philosophical and statistical roots of their activities, and economic methodologists who have an interest in the scope and limits of empirical econometrics. This brings the risk that econometricians may find the econometric discussions not always satisfactory, while methodologists might complain that I have not done justice to all philosophical theories and subtleties that they can think of. I hope that both types of readers are willing to search for added value rather than concentrate on what they already know.

After the disclaimers finally some words of thanks. The book grew out of a PhD thesis, which was defended at the CentER for Economic Research of Tilburg University (The Netherlands). I also was able to discuss various parts with colleagues during visits to the London School of Economics, Duke University, the Eidgenössische Technische Hochschule Zürich, the University of Western Australia and at many seminars and conferences. For their financial support, I would like to thank in particular the Foreign and Commonwealth Office, the Fulbright Commission and NWO (Netherlands Organisation for Scientific Research).

In addition, I would like to thank many persons for inspiration, comments and collaboration. They are my fellow authors Ton Barten, Jan Magnus and Michael McAleer; many sparring partners among whom Mark Blaug, Nancy Cartwright, James Davidson, Tony Edwards, David Hendry, Colin Howson, Rudolf Kalman, the late Joop Klant, Ed Leamer, Neil de Marchi, Mary Morgan, Stephen Pollock, Ron Smith, Mark Steel, Paul Vitányi and Arnold Zellner. Michiel de Nooij provided highly appreciated research assistance.
1 The philosophy of induction

[Some other scientists are liable to say that a hypothesis is definitely proved by observation, which is certainly a logical fallacy; most statisticians appear to regard observations as a basis for possibly rejecting hypotheses, but in no case for supporting them. The latter attitude, if adopted consistently, would reduce all inductive inference to guesswork.


1 Introduction

Occasionally, the aspirations of econometrics are frustrated by technical difficulties which lead to increasing technical sophistication. More often, however, deeper problems hamper econometrics. These are the problems of scientific inference – the logical, cognitive and empirical limitations to induction. There is an escapist tendency in econometrics, which is to seek salvation in higher technical sophistication and to avoid deeper philosophical problems. This is reflected by the erosion of an early foothold of empirical econometrics, *Econometrica*. The share of empirical papers has declined from a third in the first (1933) volume to a fifth in recent volumes. This is not because most empirical values for economic variables or parameters have been settled. Despite the ‘econometric revolution’, there is no well established numerical value for the price elasticity of bananas. If *Econometrica* were to publish an issue with well established econometric facts, it might be very thin indeed. The factual knowledge of the economy remains far from perfect, as are the ability to predict its performance, and the understanding of its underlying processes. Basic economic phenomena, such as the consumption and saving patterns of agents, remain enigmatic. After many years of econometric investigation, there is no agreement on whether money causes output or not. Rival theories flourish. Hence, one may wonder what the added-value of econometrics is. Can we learn from experience in economics, and, if so, does econometrics itself serve this purpose? Or, were the aspirations too high after all, and does the sceptical attitude of Keynes half a century ago remain justified today?
An important issue in the philosophy of science is how (empirical) knowledge can be obtained. This issue has a long history, dating back (at least) to the days of the Greek Academy, in particular to the philosopher Pyrrho of Elis (c. 365–275 BC), the first and most radical sceptic. Academic scepticism, represented for example by Cicero (106–43 BC), is more moderate than Pyrrho’s. The ideas of Pyrrho (who did not write books, ‘wisely’ as Russell, 1946, p. 256, remarks) are known via his pupil Timon of Phlius (c. 320–230 BC) and his follower Sextus Empiricus (second century AD), whose work was translated into Latin in 1569. A few earlier translations are known but they have probably only been read by their translators. The 1569 translation was widely studied in the sixteenth and seventeenth centuries. All major philosophers of this period referred to scepticism. René Descartes, for example, claimed to be the first philosopher to refute scepticism.

One of the themes of the early sceptics is that only deductive inference is valid (by which they mean: logically acceptable) for a demonstrative proof, while induction is invalid as a means for obtaining knowledge. Perception does not lead to general knowledge. According to Russell (1946, p. 257),

Scepticism naturally made an appeal to many unphilosophic minds. People observed the diversity of schools and the acerbity of their disputes, and decided that all alike were pretending to knowledge which was in fact unattainable. Scepticism was a lazy man’s consolation, since it showed the ignorant to be as wise as the reputed men of learning.

Still, there was much interest in scepticism since the publication of the translation of Sextus Empiricus’ work, not only by ‘unphilosophic minds’. Scepticism has been hard to refute. Hume contributed to the sceptical doctrine (although he did not end up as a Pyrrhonian, i.e. radical sceptic). The result, ‘Humean scepticism’, is so powerful, that many philosophers still consider it to be a death blow to induction, the ‘scandal of philosophy’.

Hume ([1739] 1962) argues that the empirical sciences cannot deliver causal knowledge. There are no rational grounds for understanding the causes of events. One may observe a sequence of events and call them cause and effect, but the connection between the two remains hidden. Generalizations deserve scepticism. Hume (Book i, Part iii, section 12, p. 189) summarizes this in two principles:

that there is nothing in any object, considered in itself, which can afford us a reason for drawing a conclusion beyond it; and, that even after the observation of the
frequent or constant conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience.

The 'scandal of philosophy' is fundamental to empirical scientific inference (not just econometrics). It has wider implications (as Hume indicates) than denying causal inference. For example, does past experience justify the expectation of a sunrise tomorrow? The question was raised in discussing the merits of Pierre Simon de Laplace's 'rule of succession', a statistical device for induction (see chapter 2). Another example, popular in philosophy, deals with extrapolation to a population instead of the future: if only white swans have been observed, may we infer that all swans are white? (This is the classic example of an affirmative universal statement.)

The sceptical answer to these questions is negative. The rules of deductive logic prohibit drawing a general conclusion if this conclusion is not entailed by its propositions. There is no logical reason why the next swan should be white. Of course, swans can be defined to be white (like statisticians who define a fair die to be unbiased), making black swans a contradiction in terms. An alternative strategy is to conclude that all known swans are white. The conclusion is conditional on the observed sample. Hence, the choice is between formulating definitions or making conditional enumerations. But most empirical scientists want to make generalizations. This is impossible if the induction problem proves insurmountable. Therefore, an understanding of induction is essential.

The logical form of the induction problem is that all observed $X$ are $\Phi$ does not entail that all $X$ are $\Phi$. The next three chapters, dealing with probabilistic inference, consider a more delicate, probabilistic form of the induction problem: given that most observed $X$ are $\Phi$, what can be said about $X$ in general? The source of Humean scepticism follows from the conjunction of three propositions (Watkins, 1984, p. 3):

(i) there are no synthetic a priori truths about the external world;
(ii) any genuine knowledge we have of the external world must ultimately be derived from perceptual experience;
(iii) only deductive derivations are valid.

The conjunction of (i), (ii) and (iii) does not allow for inferring knowledge beyond the initial premises. In this sense, inductive inference is impossible.

John Watkins (p. 12) argues that a philosophical, or 'rational' answer to scepticism is needed, because otherwise it is likely to encourage irrationality. Watkins holds that Hume himself regarded philosophical scepticism as an academic joke. Indeed, Hume uses the expression jeux d'esprit (in A letter From a Gentleman to his Friend in Edinburgh, included
as an appendix to Hume [1748] 1977, p. 116). Describing the person who is afflicted by Pyrrhonism, Hume (p. 111) concludes:

When he awakes from his dream, he will be the first to join in the laugh against himself, and to confess, that all his objections are mere amusement.

Amusement. Watkins (1984, p. 12) argues, does not qualify as a rational answer to scepticism. In fact, Hume’s response is more elaborate than the quotation suggests. It relies on conventionalism (see below). I agree with Watkins that, formally, conventionalism is not very appealing (although conventions have much practical merit). Fortunately, there are alternatives. Once the source of Hume’s problem (the threefold conjunction just mentioned) is clarified, the merits of those alternative responses to scepticism can be appraised.

Watkins (pp. 4–5) discusses a number of strategies as responses to Hume’s problem. The most interesting ones are:

- the naturalist (ignoring the conjunction of propositions (i)–(iii));
- the apriorist (denying proposition (i));
- the conjecturalist (amending proposition (ii)); and
- the probabilist strategy (which takes odds with proposition (iii)).

A more detailed discussion of the probabilist strategy will be given in the next three chapters, while the remainder of this book considers how well this strategy may work in econometrics.

3 Naturalism and pragmatism

Descartes argued that one should distrust sensations. Insight in causal relations results from mere reasoning. Hume, criticizing Cartesian ‘dogmatic rationalism’, argues that such plain reasoning does not suffice to obtain unique answers to scientific questions. Cartesian doubt, ‘were it ever possible to be attained by any human creature (as it plainly is not) would be entirely incurable’ (Hume [1748] 1977, p. 103). It would not yield true knowledge either: ‘reasoning a priori, any thing might appear able to produce anything’ (Letter From a Gentleman, p. 119). Cartesian doubt is unacceptable to Hume ([1739] 1962, p. 318). It gave him a headache:

The intense view of these manifold contradictions and imperfections in human reason has so wrought upon me, and heated my brain, that I am ready to reject all belief and reasoning, and can look upon no opinion even as more probable or likely than another.

But this does not make Hume a Pyrrhonian or radical sceptic. He is rescued from this philosophical ‘melancholy and delirium’ by nature.
His naturalist strategy is to concede that there is no epistemological answer to scepticism, but to deny its importance. It is human nature to make generalizing inferences, the fact that inference is not warranted from a logical point of view has no practical implications. Hume (An Abstract of a Book Lately Published, Entitled, A Treatise of Human Nature, Etc., in Hume [1739] 1962, p. 348) concludes, that we assent to our faculties, and employ our reason only because we cannot help it. Philosophy would render us entirely Pyrrhonian, were not nature too strong for it.

The great subverter of Pyrrhonism, Hume ([1748] 1977, p. 109) writes, is 'action, and employment, and the occupations of common life'. Not reasoning, but custom and habit, based on the awareness of constant conjunctions of objects, make human beings draw inferences (p. 28). This response is known as conventionalism. According to Hume (p. 29), custom is the 'great guide of human life', and without custom or habit, those who are guided only by Pyrrhonian doubt will 'remain in a total lethargy, till the necessities of nature, unsatisfied, put an end to their miserable existence' (p. 110). Reason is the slave of our passions.

A pinch of Pyrrhonian doubt remains useful, because it makes investigators aware of their fallibility (p. 112). The fact that one cannot obtain absolute certainty by human reasoning does not imply universal doubt, but only suggests that researchers should be modest (Letter From a Gentleman, p. 116). But many scientists will feel embarrassed by the conclusion that custom is the ultimate foundation of scientific inference. Watkins, for example, rejects it. However, conventionalism may be rationally justified. This has been attempted by some adherents of the probabilistic approach. Other strategies related to Hume's conventionalism are instrumentalism (developed by John Dewey) and pragmatism, or pragmaticism, as Charles Peirce christened it. These hold that hypotheses may be accepted and rejected on rational grounds, on the basis of utility or effectiveness. The pragmatic approach can be combined with the probabilistic strategy. But it is not free of problems. Most importantly, it is an invitation to scientific obscurantism (should a theory be useful to the learned – who qualifies? – or to the mighty?). A problem with conventionalism is to give an answer to the question 'where do these conventions come from?' and to provide a rational justification for the conventions (evolutionary game theory has been directed to this question). Lawrence Boland (1982; also 1989, p. 33) argues that neoclassical economists deal with the induction problem by adopting a conventionalist strategy. Econometricians base much of their work on another convention concerning the size of a test: the well known 5% significance level. This
convention has its roots in a quarrel between Karl Pearson and R. A. Fisher, two founders of modern statistics (see chapter 3, section 3.2).

4 Apriorism

The apriorist strategy to the problem of scepticism denies proposition (i), concerning the absence of synthetic a priori truth. Immanuel Kant invented this notion of a priori synthetic truth, true knowledge that is both empirical and based on reasoning. It is neither analytic nor synthetic. The canonical example of an a priori synthetic truth is Kant’s Principle of Universal Causation, which is his response to Humean scepticism. Kant argued that everything must have a cause: ‘Everything that happens presupposes something upon which it follows in accordance with a rule’ (translated from Kritik der reinen Vernunft, Kant’s most important work, published in 1781; in Krüger, 1987, p. 72). This doctrine is also known as causal determinism, or simply as causality (Bunge [1959] 1979, p. 4).


- the method of agreement;
- the method of difference;
- the method of residues;
- the method of concomitant variations.

These methods are based on the ‘principle of uniformity of nature’, which holds that the future will resemble the past: the same events will happen again if the conditions are sufficiently similar. The method of difference starts from the premise that all events have a cause. The next step is to give an exhaustive list of possible causes, and select the one(s) which always occurs in common with the event, and does not occur if the event does not occur. A problem is to select this exhaustive list of possible causes.

Keynes ([1921] CW VIII, p. 252) refers to the principle of uniformity of nature in his discussion of reasoning by analogy, and suggests that differences in position in time and space should be irrelevant for the validity of inductions. If this principle forms the basis for induction, it cannot itself be founded upon inductive arguments. Furthermore, it is doubtful that experience validates such a strong principle. Nature seems much more erratic and surprising than the principle of uniformity of nature suggests. Still, the late philosopher Karl Popper ([1935] 1968, p. 252) explicitly argues that ‘scientific method presupposes the immutability of natural processes, or the “principle of the uniformity of nature”’. 
Likewise, Bernt Stigum (1990, p. 542) argues that this principle is a necessary postulate of epistemology. Some probability theorists advocate a statistical version of this synthetic a priori truth: the stability of mass phenomena (see in particular the discussion of von Mises in chapter 3, section 2).

In the social sciences, it is not the uniformity of nature which is of interest, but the relative stability of human behaviour. A more apt terminology for the principle would then be the ‘principle of stable behaviour’. Consider the axioms of consumer behaviour. If one assumes that preferences are stable (Hahn, 1985, argues this is all the axioms really say), then accepting these axioms as a priori truths warrants inductive generalizations. This principle solves, or rather, sidesteps, the Humean problem. If it is accepted, generalizations from human experience are admissible. But again this postulate is doubtful. Too frequently, humans behave erratically, and on a deeper level, reflexivity (self-fulfilling prophecies) may undermine uniform regularities in the social sciences. It suffers from the same problems as the principle of uniformity of nature: either it is false, or its justification involves infinite regress. But a weaker principle of stable behaviour may be accepted, by giving a probabilistic interpretation to the generalization. There should be an appreciable (non-zero) probability that stable behaviour may be expected. This is the basis for rational behaviour. A fair amount of stability is also necessary (not sufficient) for scientific inference: otherwise, it is impossible to ‘discover’ laws, or regularities.

It is hard to imagine interesting a priori synthetic truths specific to economics. The axioms of consumer behaviour are not generally accepted as true. An investigation of their validity cannot start by casting them beyond doubt (chapter 8 provides a case history of ‘testing’ consumer demand theory). Bruce Caldwell (1982, p. 121) discusses praxeological axioms of Austrian economists as an example of Kant’s a priori synthetic propositions. The Austrian Friedrich von Wieser argued that a cumbersome sequence of induction is not needed to establish laws in economics. He claimed (cited in Hutchison, 1981, p. 206) that we can ‘hear the law pronounced by an unmistakable inner voice’. Ludwig von Mises made apriorism the cornerstone of his methodology. The problem of this line of thought is that inner voices may conflict. If so, how are we to decide which voice to listen to?

5 Conjecturalism

The conjecturalist strategy denies Watkins’ proposition (ii) and instead holds that scientific knowledge is only negatively controlled by experi-
ence: through falsification. Popper provided the basic insights of the conjecturalist philosophy (also known as methodological falsificationism) in his *Logik der Forschung* in 1934 (translated as Popper, [1935] 1968). This nearly coincides with one of the first efforts to test economic theory with econometric means (Tinbergen, 1939b). Followers of Popper are, among others, Imre Lakatos and Watkins. I will first discuss Popper’s views on inference, then Lakatos’ modified conjecturalism.

5.1 Popper’s conjecturalism

Popper’s impact on economic methodology has been strong. Two pronounced Popperians in economics are Mark Blaug (1980) and Terence Hutchison (1981). Moreover, statisticians and econometricians frequently make favourable references to Popper (Box, 1980, p. 383, n.; Hendry, 1980; Spanos, 1986) or believe that Popper’s is ‘the widely accepted methodological philosophy as to the nature of scientific progress’ (Bowden, 1989, p. 3). Critics claim that the real impact of Popperian thought on economic inference is more limited (see also De Marchi, 1988; Caldwell, 1991).

5.1.1 Falsification and verification

Scientific statements are those which can be refuted by empirical observation. Scientists should make bold conjectures and try to falsify them. This is the conjecturalist view in a nutshell. More precisely, theories are thought of as mere guesses, conjectures, which have to be falsifiable in order to earn the predicate scientific. The *modus tollens* (if *p*, then *q*. But not-*q*. Therefore, not-*p*) applies to scientific inference – if a prediction which can be deduced from a generalization (theory) is falsified, then that generalization itself is false. The rules of deductive logic provide a basis for scientific rationality and, therefore, make it possible to overcome the problems of Humean scepticism. Falsifiability distinguishes science from non-science (the demarcation criterion). The growth of knowledge follows from an enduring sequence of conjectures and refutations. Theories are replaced by better, but still fallible, theories. Scientists should remain critical of their work.

So far, there seems not much controversial about the conjecturalist approach. The tentative nature of science is a commonplace. Popper went beyond the commonplace by constructing a philosophy of science on it, methodological falsificationism. A source of controversy is Popper’s critique of logical positivism, the philosophy associated with the *Wiener Kreis*. A related source is his obnoxious rejection of induction.