Thesis

Epistemic Virtues and Concept Formation in Economics

by

Julian Reiss

from the

London School of Economics and Political Science

submitted in 2002 to the University of London

for the completion of the degree of a

Doctor of Philosophy

To my Mom, Peter
and my father *in memoriam*

Abstract

The aim of this Thesis is to defend the following two main claims: (a) A "Baconian" science of economics is desirable and possible; (b) Two of Bacon's central insights are of particular relevance for modern economics: the importance of concept formation as a part of the scientific endeavour and the collaboration of theory construction and measurement.

Chapter 1 introduces the topic by way of juxtaposing and contrasting Francis Bacon's scientific method with Gustav Schmoller's philosophy of economics and showing that they share a number of crucial aspects, which are significant for modern methodological debates. It is argued that the three epistemic virtues of *phenomenal adequacy*, *explanatory power* and *exactness* are as relevant for contemporary economics as they were for Schmoller and Bacon. Chapters 2-4 critically examine various strands in the contemporary economic literature. It is claimed that methods of concept formation dominant in this literature can allow us to obtain either of the three virtues severally but not all three simultaneously.

In response to this criticism, Chapter 5 develops an alternative method of economic concept formation. In particular, the idea of *Natural Economic Quantities* (NEQs) is introduced. Essentially, an economic quantity is natural if and only if it figures in a tested causal model and it is measurable in the appropriate way. NEQs are supposed to help in building models which achieve the three epistemic virtues simultaneously, and thus allow economics to be a "Baconian" science. The theory of NEQs and its ability to help in realising all three epistemic virtues is illustrated with a case study from William Stanley Jevons's work on index numbers.

Contents

List of Figures and Tables

Prolegomena and Acknowledgements

The hero of my story is the Lord Chancellor of Verulam. Famous under the name of Francis Bacon, he invented modern experimental science, inspired the foundation of the Royal Society and provided a starting point for some versions of empiricism. Originally trained as a barrister, Bacon's career accelerated when James I ascended to the throne. In 1613 he was appointed Attorney General, in 1617 Lord Keeper and in 1618 Lord Chancellor. But Bacon took bribes, got caught, was deprived of office and had to pay a £40,000 fine. Never having been able to return to favour, Bacon died in 1626, allegedly after a cold that he had contracted from an experiment of stuffing a chicken with snow.

The history of Bacon's influence on science and philosophy is mottled. Initially almost ignored in Britain, France's empiricists and Descartes studied his works with care. Slightly later, Baconianism was the official philosophy of the Royal Society but Newton appears not to have owned either of his main works, *Novum Organum* and the *Advancement of Learning*. In the 18th century, Voltaire and Diderot expressed the greatest respect, and Kant uses an excerpt from the Preface to the *Great Instauration* as a motto for his *Critique of Pure Reason*. Mill and Whewell are indebted to Bacon in an obvious way but disagree with much of what he says in his writings. Towards the end of the 19th century, Bacon's reputation began to decline and in the 20th, his ideas were criticised heavily and fundamentally by influential thinkers such as Karl Popper and Imré Lakatos. Only in the past twenty-odd years, along with the rise of experimentalism in the philosophy of science, has it been possible again to refer to Bacon favourably without expecting the reader's scorn.

Why take such a lawyer-turned-philosopher with an unstable life and a volatile reputation to be the intellectual mentor of a PhD thesis in the philosophy of economics? Bacon is famous for criticising both the empiricist as well as the rationalist schools that he held to be dominant in the science of his time. Empiricists, so Bacon, only gather—they do not process information. Rationalists, by contrast, "spin webs out of themselves"—they don't take experimental evidence seriously. As a remedy Bacon prescribed a methodology, the "Interpretation of Nature", which would combine elements of both. Evidence would be gathered not passively but actively. It would be re-arranged, classified and its quality improved. Hypothesised generalisations would motivate new experiments and the latter's results may lead to a modification of the generalisations.

With Ian Hacking (1983, p. 248f.) I believe that at least some parts of economics are still in the era of Bacon's empiricists and rationalists. There is no shortage of analyses of empirical data nor of theoretical speculation. What sometimes seems to be missing is an effective combination of the two. Consequently, one of the objectives of this Thesis is to provide an interpretative framework in which certain methodological facets of economics can be understood along the lines of Bacon's empiricists, others along the lines of the rationalists and in which finally ways are indicated following which we may achieve a fruitful combination of the two strands.

This sounds like an ambitious project indeed, and thus in order not to disappoint the reader with what follows I must mention a number of important limitations in advance. First, I do not think that there is such a thing as the philosophy of economics *simpliciter*. What really is there is at best a philosophy of general equilibrium theory, a philosophy of post-Keynesianism, a philosophy of Marxism, a feminist philosophy of economics, a philosophy of experimental economics, a philosophy of econometrics and so on. Most case studies I analyse are drawn from only the past thirty or so years and fall into the category one broadly calls "mainstream economics". But even within that I almost completely neglect economic *theory* (general equilibrium, REM *etc.*) and focus on those models that may be thought of as being directly descriptive or explanatory of economic phenomena. I do include a number of works by historical economists, most notably by Gustav Schmoller, Carl Menger and William Stanley Jevons, but mainly (with the exception of a case study from Jevons in Chapter 5) as contrast cases that are supposed to illuminate others.

Second, there is a huge number of issues that are related and relevant to this overall objective but I lacked space and time to go into (*e.g.* idealisation and approximation, learning from false models, formalisation, mathematisation, falsification, realism *etc. etc.*). What I have tried to do is to stick to those topics that I believe have a

"Baconian" flavour to them, that is, abstraction, measurement and the search for properties that can be thought of as "natural properties".

The strategy I use is the following. In the first Chapter, these Baconian topics are introduced by means of a comparison of the respective methodologies of Bacon and Schmoller—which I believe share a number of important aspects that are germane to this Thesis. That Chapter also identifies three epistemic virtues that, so I shall argue, economic models must have in order to achieve Bacon's aim of integrating empiricist and rationalist elements in science. Chapters 2 – 4 measure a number of methodological characteristics that we can find in parts of economic practice against those epistemic virtues. As a consequence, these Chapters will be mainly critical. Chapter 5 will present the outline of a theory of "natural economic quantity" that I believe should enable the construction of models that can in principle realise the epistemic virtues demanded of economics. How this may be possible is illustrated with a case study.

An important preliminary for the understanding of the succeeding Chapters is to decide whether the enterprise should be regarded as mainly normative or as mainly descriptive about economics. My own stance in this matter is to strike a healthy balance between the "excessive" normativism of, say, early positivist or Popperian economic methodologists and the "excessive" descriptivism of some historicist methodologists. According to the view defended in this Thesis, no philosopher of economics can impose his or her own values on the ongoing process of scientific investigation. Not only would that idea be presumptuous, it would also be unlikely that economists pay great attention to such a writer. At the end of the day methodology is done for the benefit of the science it is the methodology of.

On the other hand, "mere" description and interpretation seems not sufficient as a methodologist's goal. After all, certain epistemic values do exist and it is a legitimate question to ask whether methodological practices do or do not help in realising these values. The approach taken in this Thesis is, then, to try to extract what will be called epistemic virtues from economists' own methodological and substantive reflections and measure their practices against these virtues. It goes without saying that I do not want to say those factors that I will focus on and that are consistent with a Baconian understanding of science are the only ones we can find in economics, that they are endorsed by all economists and that they must be regarded as equally important.

However, as I hope to show in the course of this Thesis, I believe that they are strong enough to tell an insightful story about some strands in modern economic thought, they are shared by at least some practitioners but general enough not to provoke immediate and obvious rejection by others.

In the course of reading for my PhD I benefited intellectually from the help of a great number of people. Most prominently of course I want to mention my thesis supervisors Nancy Cartwright (5 years) and Mary Morgan (2 years). I thank Nancy for her almost unbounded energy to make my thoughts clearer and more exact and Mary for a like energy to ground my ideas in the literature and for helpful suggestions with case studies. Responsibility for a failure on any of these accounts remains, of course, solely with me.

Dan Hausman was my second supervisor for the first two years and Jossi Berkovitz for the third. Many thanks to both of them for stimulating discussions and invaluable comments on my work. Parts of this Thesis have been presented at meetings of the London-LSE Measurement and Modelling in Physics and Economics Project. I thank all its members, especially (besides Nancy and Mary) Carl Hoefer, Hasok Chang, Roman Frigg, Till Grüne, Michela Massimi, Harro Maas, Marcel Boumans, Hsian-Ke Chao and Peter Rodenburg. One or more of the meetings were also attended by Bruce Caldwell, Max Steuer, Valeria Mosini, Lisa Lloyd, Pascal Riviere and Sonja Amadae, to all of whom I owe my thanks for very useful suggestions.

An earlier version of Chapter 1 was presented as a conference paper for the History of Economics Society Meeting 2001 at Wake Forest University, Winston-Salem, North Carolina. I thank all its participants, especially D. Wade Hands, who was my discussant at that session, and Yuichi Shionoya. An earlier version of Chapter 5 was presented as a conference paper for the International Network for Economic Methodology 2000 biannual conference at the University of British Columbia, Vancouver. Many thanks go to its participants, especially Kevin Hoover, Marcel Boumans and Steven Rappaport. A revised version of that paper appeared in the conference proceedings in the *Journal of Economic Methodology*. I am indebted to four anonymous referees for their useful comments.

This PhD thesis has been supported by a scholarship from the Friedrich-Naumann-Stiftung, which is funded by the Bundesministerium für Bildung and Forschung (the

German Federal Ministry for Education and Research) as well as funds from the Measurement and Modelling Project. Many thanks to these institutions for enabling me to carry out the studies for this Thesis.

Chapter 1

Bacon, Schmoller and Epistemic Virtues in Economics

Chapter 1

Bacon, Schmoller and Epistemic Virtues in Economics[1]

Those who have handled the sciences have been either Empiricists or Rationalists. Empiricists, like ants, merely collect things and use them. The Rationalists, like spiders, spin webs out of themselves. The middle way is that of the bee, which gathers its material from the flowers of the garden and field, but then transforms and digests it by a power of its own.
And the true business of philosophy is much the same, for it does not rely only or chiefly on the powers of the mind, nor does it store the material supplied by natural history and practical experiments untouched in its memory, but lays it up in the understanding changed and refined. Thus from a close and purer alliance of the two faculties—the experimental and the rational, such as has never yet been made—we have good reason for hope.

Francis Bacon—NO I 95[2]

1 Introduction

This is a Thesis in Baconian topics. More specifically, it is an attempt to analyse selected methodological facets of economic science from a Baconian point of view. There are two related rationales why Francis Bacon is chosen as a starting point. The first is that I have a kind of "end run" view about Bacon's scientific method. I believe that Bacon accomplished a methodology that avoids several difficulties which are the artefacts of false dichotomies and accentuations. Examples include the distinctions between impressions and ideas and that between observable and unobservable, the inductivism-deductivism dichotomy and the emphasis on meaning

[1] Earlier versions of this Chapter were presented at a meeting of the LSE-Amsterdam Measurement in Physics and Economics Group and at the 2001 HES conference at Wake-Forest University, North Carolina. Many thanks to all participants of these meetings for a helpful and stimulating discussion for the issues raised here, in particular my supervisors Nancy Cartwright and Mary Morgan, Hasok Chang, Bruce Caldwell, Lisa Lloyd, Valeria Mosini and Till Gruene in London and my discussant D. Wade Hands and Yuichi Shionoya at Wake Forest University. The responsibility for mistakes and misunderstandings remains, as usual, with me.
[2] All quotes from Bacon's *Novum Organum* (NO) refer to the Urbach and Gibson 1994 edition. The Roman numeral denotes the part of NO, the Arabic numeral the number of the aphorism.

rather than method. In very recent philosophy of science a number of these distinctions have been given up by at least some methodologists, and emphases have been shifted. Bacon's work, I believe, then, is a rich source of methodological ideas which circumvents problems associated with those dichotomies and at the same time draws our attention to topics that really matter. So this is the second rationale for choosing Bacon as a starting point. Through the lens of the Baconian looking-glass we can see clearly the significance of a number of methodological questions that I believe have great relevance for concerns of contemporary economic methodology.

One of the Baconian topics emphasised here is that of the importance of scientific *concept formation*. Bacon often called attention to the fact that adequate concepts are the key to successful science. For example, the third of his famous four groups of idols is called the *Idols of the Market-place* and refers to the observation that[3]

> ... speech is the means of association among men; but words are applied according to common understanding. And in consequence, a wrong and inappropriate application of words obstructs the mind to a remarkable extent. Nor do the definitions or explanations with which the learned men have sometimes been accustomed to defend and vindicate themselves in any way remedy the situation. Indeed, words plainly do violence to the understanding and throw everything into confusion, and lead men into innumerable empty controversies and fiction.

And further,[4]

> The *Idols of the Market-place* are the most troublesome of all; these are idols that have crept into the understanding though the allegiance of words and names. For while men believe their reason governs words, in fact, words turn back and reflect their power upon the understanding, and so render philosophy and science sophistical and inactive.

Bacon then continues to distinguish two sub-species of that idol. On the one hand, there are names for which there is no correspondent in nature. His examples include *fortune*, *prime mover* and *planetary orbs*. On the other hand, there are words that refer but in an obscure way: words that denote real objects or their features, but in a vague and muddled way. Bacon's example is *moist*: this word does not pick out objects that form a *natural* class—moist objects have not too much in common.

Bacon emphasises the correct use of our words similarly in his *Advancement of Learning*:[5]

[3] NO I 43
[4] NO I 59
[5] quoted from Hacking 1975, p. 5

> Although we think we govern our words, ... certain it is that words, as a Tartar's bow, do shoot back upon the understanding of the wisest, and mightily entangle and pervert the judgment. So that it is almost necessary, in all controversies and disputations, to imitate the wisdom of the mathematicians, in setting down in the very beginning the definitions of our words and terms, that others may know how we accept and understand them, and whether they concur with us or no. For it cometh to pass, for want of this, that we are sure to end there where we ought to have begun—in questions and differences about words.

In a third context, Bacon links his method of true induction to concept (or in his terminology, "notion") formation:[6]

> The foundations of true induction lie in the process of exclusion, which however is not completed until it arrives at an affirmative. Nor, however, is the exclusive part itself at all complete, and indeed it cannot possibly be so at first. For exclusion is obviously a rejection of simple natures, and *if we do not yet have good and true notions of simple natures*, how can exclusion be rectified?

This Thesis takes Bacon's concerns about concept formation as a vital ingredient of scientific theorising seriously. But what exactly do we understand by concept formation? Momentarily I would like to characterise concept formation as *the process or method by which a concept's rules of application are determined.*[7] "Rules of application" should not be understood too narrowly here. If we define a desk as "a kind of table equipped with drawers, compartments, *etc.*, and a flat or sloping top for writing, drawing, or reading"[8], I suppose that we have a relatively clear understanding of how to determine whether a given object is a desk or not. In this case, presumably, the rules of application are empirical rules of verification. We simply look at objects that potentially fall under the concept at hand and decide by its shape, structure *etc.* But consider concepts such as *God* or *immortality of the soul* or *element of fire*. There are, it would seem, no *empirical* rules of applying these concepts to actual objects. However, they do relate to other concepts, and the rules of application may be given by those relations. In this case we can say that they apply to *potential* objects.

---

[6] NO II 19. Emphasis added. Simple natures are essentially those for which there exist true "forms" or laws. I shall say more about this concept further below.

[7] *Cf.* above quote from Bacon in which he says that "a wrong and inappropriate *application of words* obstructs the mind..." (NO I 43), emphasis added.

[8] Webster's New World Dictionary, 2nd College Ed., entry "desk", no 1

There may be very different processes or methods by which a concept's rules of application are determined. Just consider a passage from Locke's *Essay*:[9]

> There is nothing more evident than that the ideas of the persons children converse with... are, like the persons themselves, only particular. The ideas of the nurse and the mother are well framed in their minds; and, like pictures of them there, represent only those individuals. The names they first gave to them are confined to these individuals; and the names of "nurse" and "mamma" the child uses, determine themselves to those persons. Afterwards, when time and a larger acquaintance has made them observe that there are a great many other things in the world, that, in some common agreements of shape and several other qualities, resemble their father and mother, and those persons they have been used to, they frame an idea which they find those many particulars partake in and to that they give, with others, the name "man," for example. And thus they come to have a general, and a general idea. Wherein they make nothing new, but only leave out of the complex idea they of Peter and James, Mary and Jane, that which is peculiar to each, and retain only what is common to them all.
>
> 8. By the same way that they come by the general name and idea of "man," they easily advance to more general names and notions.

In Locke's rendition mental representations of external objects are compared, similar representations are grouped together, the particularities of each representations are excluded and the similar aspects are given a common name. Compare this method of concept formation with the following passage from James Brown's *Philosophy of Mathematics* in which he introduces "original" Platonism:[10]

> A dog is a dog in so far as it 'participates' in *the form of a dog*, and an action is morally just in so far as it participates in *the form of justice*.
>
> How do we know about the forms? Our immortal souls once resided in heaven and in this earlier life gazed directly upon the forms. But being born into this world was hard on our memories; we forgot everything. Thus, according to Plato, what we call learning is actually recollection. And so, the proper way to teach is the so-called Socratic method of questioning, which does not simply state the facts to us, but instead helps us to remember what we already know.

The application rules for concepts are given by their relation to Platonic forms. This relation is sometimes called "participation".[11] A concept is adequately applied to those objects that participate in the form represented by that concept. The quoted

---

[9] II.3.7-8

[10] Brown 1999, pp. 8-9, original emphasis

[11] In other passages that relation seems to be one of causation rather than participation; see *Phaedo* 99c-102a. Brown's account of natural science comes very close to thus understanding; see his 1994 and the last section of Chapter 3 of this Thesis.

passage also nicely demonstrates a distinction often drawn by realists: the distinction between what the rules of application *are*, and how we *learn* or *know* about them. The rules are given by the relation of participation. But we learn about them by the Socratic method of questioning (the Platonic idea that we once lived among the forms provides the *reason* why we can learn about the forms by that method). Empiricists tend to blur this distinction: for Locke, the rules are given by similarities, and we know about them by observing similarities. I will follow the empiricist line and not make too much of the distinction unless otherwise noted.

Another topic Bacon is famous for is that of the collaboration of different scientific faculties. As exemplified in the quote at the beginning of this Chapter, Bacon criticised the use of both the experimental as well as the rational faculty—if employed in separation. Collaboration, though, for Bacon did not mean merely use of both faculties together. Rather, the scientist should, like the bee, transform what he or she has found in the world and add value to it according to the right scientific method.

The present-day scientific equivalent of Bacon's distinction between the rational and the experimental faculty may be a distinction between theoretical and experimental or applied science. That Bacon's concern about the collaboration of these faculties is still germane to the economics of our time can be seen from the various debates about "measurement without theory"[12] as well as the various criticisms of the lack of an empirical base of much of mainstream economics.[13] Towards the end of this Thesis I will try to show how, using various Baconian ideas, we can achieve a collaboration of the faculties that Bacon demanded for science.

In the course of researching for other Chapters of this Thesis I noticed that there is a remarkable similarity between the structure of Bacon's methodology and that of Gustav Schmoller. Those features on which they agree, are, interestingly, also features of more general and contemporary relevance. I thus place this Chapter, whose main part consists of a comparison between Bacon's and Schmoller's philosophies of science, at the beginning. It is meant to be an introduction to the issues of interest and to the Baconian topics that, I believe, are significant from today's point of view.

[12] Koopmans 1947
[13] See for example Albert 1998, ch. IV and passim.

The Chapter is organised as follows. Section 2 compares Schmoller's scientific method to Bacon's and thus identifies it as "Baconian". This section is consequently mainly historical. Section 3 translates some noteworthy requirements of economics that follow from the historical discussion into the modern language of "epistemic virtues demanded of economic models" to provide a link with the following Chapters, which are mostly concerned with the economics of the present day. Section 4, finally, looks at the issue of "collaboration of faculties", and asks what we can learn from the discussion about the relation between theory and application in economics.

2 Schmoller as a Baconian Philosopher

Francis Bacon (1561 – 1626) is—along with Descartes—sometimes regarded as the most original and influential philosophical thinker of the scientific revolution. He was deeply dissatisfied with the state of the art of science of his time and sought to reform it by way of a new, revolutionary method. With respect to their *ambition* and the *dissatisfaction* with current state of science, one can thus see remarkable parallels with Descartes. But whereas Descartes wanted to provide secure foundations for the sciences by initially sweeping away all beliefs and readmitting only what has passed his method of doubt, Bacon started with what was available and wanted to improve matters piece by piece by his method of true induction. And whereas Descartes aimed at what one might call moral certainty, Bacon, probably prompted by his almost life-long engagement in political matters, aimed at a practical exploitation of the knowledge thus gained, at "power over nature", having the improvement of human welfare in view.

Gustav Schmoller (1838 – 1917) was the most important economist in Germany at his time and the mentor of what became called the "younger German historical school". Schmoller is probably best known as one of the protagonists—and, at least to some people—the inferior party of two *Methodenstreits*: a debate between him and other historicists on one side and Carl Menger and the marginalists on the other, and a debate between him and Max Weber. The first *Methodenstreit* is often regarded to be about the correct mode of inference in economics: with Schmoller advocating induction and Menger preaching deduction. The second *Methodenstreit*, about 20 years later, was about the value-ladenness of economic principles.

Apart from having fought and in the eyes of many commentators lost these battles on method, Schmoller is well-known for his active contribution to social policy making in the German Empire. Indeed, Schmoller was one of the founders of the *Verein für Socialpolitik* in 1872, which was meant to be a forum for discussion of the acute social problems of the time—brought about by the rapid industrialisation and its social consequences: development of an industrial proletariat, urbanisation, migration from the land into the city, development of new industrial sites—and at the same time a body proposing a reformist policy.[14]

Schmoller's own social-political views were influenced by both liberal and socialist ideas. One the one hand, he advocated freedom of trade and cheered liberalism for its fight for individual rights, a constitution and autonomy of the electorate,[15] but on the other hand, he emphasised the great role the state must play for social integration.[16] For Schmoller, one purpose of reformist policy of the state was to fight back the claims of radical socialists.[17]

His own economics was means to the fulfilment of this programme. The idea of justice was the regulative principle for any social reform. By studying the laws and causal structures of the economy, politicians should be enabled to approach this regulative ideal by way of institutional reform.[18] In order to study the economy's laws and causal structures properly, however, the economists must have a method at their disposal that makes these investigations possible. Schmoller's empirical-historical method was supposed to do exactly that.

In this section I want to identify Schmoller's philosophy of economics as "Baconian" by arguing that it shares crucial aspects with Bacon's scientific method. Although I think that parallels stretch even further, especially regarding the ultimate aims of their respective philosophies—gaining power over nature or the economy in order to improve the human condition—I focus on the methodological aspects of their

[14] Schmoller 1998/1872
[15] Schmoller 1998/1881, p. 107f.
[16] *ibid.*, p. 105ff.
[17] Schmoller 1998/1872
[18] Schmoller 1998/1881

writings.[19] In particular, I want to show that the two philosophers' methodologies are identical with respect to the following five characteristics:

(1) *Empiricism*. Both philosophies can be characterised as methodologically empiricist, that is, both are premised on the claim that knowledge is gained by experiential methods.

(2) *Abstraction*. Both philosophies form abstract concepts by a three-stage process of *observation*, *classification* and *nomic/causal inference*.

(3) *Concepts*. Both philosophies regard at least some terms in science as *concepts* in Norman Robert Campbell's sense, that is, as terms whose meaning reflects knowledge about laws (or causal relations).

(4) *Induction*. Both philosophies are inductivist of a kind that makes much more complicated inferences than "inductive generalisations".

(5) *Observability*. Both philosophies use a conception of "observation" that is much broader than the one we sometimes find in contemporary empiricisms.

2.1 Empiricism

Empiricism is often thought to be either a doctrine about the sources of our knowledge or a doctrine about the meanings of our concepts.[20] The first doctrine may be called epistemic or knowledge-empiricism. In its strongest form it says that *all our (non-logical) knowledge comes from experience*. In some versions it can be a claim about the very *possibility* of knowledge: all we can know (except logic) are the contents of our experience—whatever form that may take. The second doctrine may be called semantic or concept-empiricism. It says that the meanings of words are certain aspects of the contents of our experiences.

[19] The main reason is that this is supposed to be a study of method rather than metaphysics or morality. Metaphysical issues will be striven on occasion but only in so far as they are required to understand the methods discussed. Issues of morality will be completely omitted but certainly not because I think we cannot learn from Bacon or Schmoller about these matters.

[20] See Grayling 1995, ch. 9, for a historical introduction to empiricism and a related distinction.

Bacon was not a pure epistemic empiricist. First, he held that there are two fundamental kinds of knowledge: knowledge by divine revelation and knowledge by human learning.[21] Only the source of the latter kind of knowledge is sense experience. Second, as the metaphor of ants, spiders and bees suggests, there is an active contribution of the mind to even empirical knowledge. Knowledge is not passively *received from nature* but actively *forced out of nature*, by "twisting the lion's tail": asking the right questions and conducting experiments.[22]

Neither did Bacon have the conceptual resources to be a semantic empiricist. Although he distinguished between a word and the thing it stands for (and this distinction will be important in the context of his theory of abstraction), he did not distinguish between a thing and a mental representation of it. He simply did not have the metaphysics of impressions and ideas nor the more recent conception of a verifiability criterion. And it didn't bother him. Yet his name is usually mentioned along those of Hobbes, Locke, Berkeley and Hume. So what did his empiricism consist in?

I think an apt name for Bacon's doctrine is *methodological* empiricism.[23] The route to success is method, and the right methods (to gain knowledge about *nature*) are experiential. In order to see what the "experientiality" of his methodology consists in a comparison with Descartes, another very methodologically inclined philosopher, is instructive. The contrast between the two is present in both the negative or destructive as well as the positive or constructive parts of their philosophies. Consider the destructive part first. Descartes famously begins his *Meditations* with the methodological precept of universal doubt. His aim is certainty, and the route to

---

[21] *Cf.* Kusukawa 1996.

[22] According to the interpretation outlined here it is, then, no surprise that Kant, in the Preface to the *Critique of Pure Reason*, cheers Bacon for having helped to find the right method of natural science: "Natural science was much longer in entering upon the highway of science. It is indeed, only about a century and a half since Bacon, by his ingenious proposals, partly initiated this discovery [of the 'right' method], partly inspired fresh vigour in those who were already on the way to it" (Kant 1929/1787). Kant also uses a quotation from Bacon's Preface to the *Great Instauration* as motto for the *Critique*: "[F]or myself, I ask nothing, but for the matter in hand I urge men to think of it not as an opinion but as a task to be done; and to be well assured that I am laying down foundations, not for any one sect or teaching but for the advantage and enlargement of mankind. Next I ask that they be fair to their own interests... take counsel together; ... and take part themselves. I ask, moreover, that they be of good hope and not imagine that my *Instauration* is something infinite and beyond the reach of man, when the truth is that it is the proper end and termination of infinite error" (Kant uses the original Latin text; the translation used here is that of Urbach and Gibson 1994, pp. 15f.).

[23] *Cf.* Woolhouse 1998, p. 12: "[Bacon's] empiricism is methodological. It consists in his being a propagandist for empirical and observational knowledge, and in his provision of a systematic method for increasing it".

certainty is to build one's edifice of knowledge only on principles that withstand the strongest possible doubt. The method is a rational one: hold a belief before your faculty of reason and throw it overboard in case it is dubitable. One does not need to refer to experience in order to decide whether a belief is dubitable. Reason alone is able to decide.

Bacon begins his *Novum Organum* with a destructive part as well. But his criticism consists in an enumeration of "fallacies of the mind": ways in which the mind has gone wrong in the past and the guideline to "be on the guard". These fallacies are codified in Bacon's four groups of "idols of the mind". The *Idols of the Market-place*, which arise from confusions due to the false use of language, have already been mentioned. The other three are the *Idols of the Tribe*, which arise from inclinations of the mind due to human nature, the *Idols of the Cave*, which arise from prejudices due to an individual's nature given by sex, race, profession *etc.* and the *Idols of the Theatre*, which arise from false philosophies. Nothing in Bacon's writings suggest that this list should be exhaustive. It is an enumeration of traps into which the human mind might fall. Unlike the Cartesian doubt, the method of the idols is not conclusive: even knowing that there are the various idols, one might fall into one prejudice or another time and again. And the source of the method is itself an experiential one: idols are fallacies that we have experienced people falling into.[24]

The same contrast can be seen in the constructive part of the respective philosophies. Descartes extracts knowledge of the external world from the knowledge of his own existence via his knowledge of God's existence. The metaphor to describe the Cartesian image of science is that of the tree: its roots are metaphysics, the trunk, physics and the branches, the special sciences medicine, mechanics and morals.[25] You get from bottom to top via a series of intuitively certain steps of deductive inference. The metaphor to describe the Baconian image of science is that of the pyramid:[26] "natural histories" at its base, rough regularities towards the middle and the most fundamental laws of nature or what Bacon calls "forms" at the apex. Bacon's aim, too, is certainty.[27] But his route there is via a series of steps of "true induction", each of which is guided by a long list of methodological guidelines. For

[24] or rather: being aware of the idols has proved to improve thinking
[25] See the Preface to Descartes' *Principles of Philosophy* (AT IX B 14).
[26] See for instance Losee 1993, p. 68.

example, Bacon lists no less than twenty-seven "Prerogative Instances", examples of kinds of pieces of evidence that support the inductive process (for example measurements, observations with telescope or microscope and "crucial" experiments).

To get a sense of what true induction and the methodological guidelines look like, let me summarise their crucial features here. briefly here. The three basic steps of Bacon's method of the *Interpretation of Nature* are (1) Collection of Natural Histories, (2) Arrangement in Tables and (3) Eliminative Induction. In step (1) evidence is gathered. We can call this step *observation*. In step (2) the evidence is classified according to a number of rules. Accordingly, we can call this step *classification*. In step (3) the existence of a law or "form" is inferred from the observations. Since a form can be regarded as a set of causal conditions for some feature of a phenomenon, it seems not unfair to call this step *nomic/causal inference* (though Bacon does not use this terminology explicitly).[28]

There are three kinds of tables for the second step. The first one is called the *Table of Existence and Presence.*[29] It summarises different instances of the phenomenon of interest. In Bacon's example, *heat* is the phenomenon or "nature" of interest. In this table we find a variety of hot things as diverse as sun rays, fiery meteors, horse dung and "strong vinegar, and all acids, on all parts of the body where there is no epidermis". The second kind of table is called *Table of Deviation*, or *Absence in Proximity*. Here instances are listed which as similar as possible to the instances of the first table but where the phenomenon of interest is absent. The corresponding instances to sun rays, for example, are the rays of the moon, stars and comets. The third kind is called *Table of Degrees* or *Table of Comparison*. Here, as the name suggests, instances are listed where the phenomenon of interest varies in degrees. The nature of heat is, according to Bacon, present in animals to different degrees. And, importantly, it increases "by movement and exercise, wine, and feasting, copulation, burning fevers and pain".[30] From these tables, in the third step, Bacon

[27] I shall rather say, in order to make clear the contrast with Descartes, "infallibility", because his concept has nothing to do with the mental feature of "clearness and distinctness".
[28] For Bacon, the goal of true induction is to find a nature's form. A "nature" of an object's is roughly one of its properties such as its colour or temperature. According to Mary Horton (Horton 1973, see below), a nature's form is the necessary and sufficient causal conditions for its presence. In the subsection on abstraction I discuss this issue in more detail.
[29] NO II 11
[30] NO II 13

infers a law that is governing the phenomenon of interest. He argues that the law must be a factor which is always present when the phenomenon is present, always absent when the phenomenon is absent and vary in degrees with the phenomenon. Hence, by eliminating all factors which (a) "are not found in any instance where the given nature [the phenomenon of interest] is present", (b) "are found in any instance where the given nature is absent" and (c) "are found to increase in any instance when the given nature decreases, or to decrease when the given nature increases"[31], we can infer that the sole remaining factor—if there is any such—is the factor responsible for the occurrence of the phenomenon of interest.

To this broad methodological framework of nomic/causal inference, the second part of the *Novum Organum* adds methods that are *aids* to establishing nomic or causal claims. As mentioned above, Bacon calls these "Prerogatives of Instances".[32] Some interesting examples are the following. *Solitary Instances*[33] are those in which either the phenomenon of interest is present in a circumstance that has nothing in common with other circumstances except the phenomenon, or that are similar in all respects to another circumstance except that the phenomenon is absent. *Revealing Instances*[34] are those in which the phenomenon of interest occurs "naked" or "standing on its own". We will see further below that a causal process that occurs in isolation is in contemporary discussions often accepted as making causal inference easier. *Crucial Instances*[35], of course, have become to be known as "crucial experiments" although Bacon does not presuppose that there must be an intervention (which I take as characteristic of experiments). Bacon calls an instance "crucial" when two causal hypotheses compete in their explanation of a phenomenon and there are situations that show that one but not the other cause must be responsible for the phenomenon. A nice example Bacon gives is that of the cause of weight. According to the Aristotelians, a body's weight is caused by a tendency of that body to move towards its natural place; for a heavy body that would be the centre of the Earth. A competing hypothesis was that the Earth's mass attracts the body. The Aristotelian hypothesis implies, according to Bacon, that bodies are equally heavy no matter how far from

[31] All three quotes are from NO II 16.
[32] Instances are of course situations and not methods. But one can easily translate situations into methodological precepts by demanding to either seek or produce the prerogative instances.
[33] NO II 22
[34] NO II 24
[35] NO II 36

the centre of the Earth they are. The competing hypothesis, by contrast, implies that the Earth's pull diminished with the distance to its centre. A crucial instance would thus be to compare two calibrated weights-operated clocks, one of which is located on the tower of a church, and the other, deep down in a mine. If the clock in the mine is faster than the other one, the Aristotelian hypothesis would be ruled out. Further, there is a group of five classes of instances, which Bacon jointly calls *Instances of the Lamp.* They regard aids of the senses such as microscopes and telescopes, removing causes that obstruct perception of the phenomenon and detection instruments. An interesting subclass is what Bacon calls *Supplementary* or *Substitutive Instances.*[36] They comprise a method of analogical reasoning, which I will discuss in subsection 2.3. The last seven instances, finally, concern matters of practical investigation and include mainly measurement instruments of different kinds and guidelines for experimental set ups.

Browsing through part two of the *Novum Organum*, one can easily see that the methodological precepts themselves are drawn from evidence of what has been useful in the past and what is defensible from an empiricist point of view. A good statement of this kind of methodological empiricism can be found in the Preface to the *Great Instauration* (*Novum Organum* was supposed to be part two of this):[37]

> And our journey has always to be made by the uncertain light of the sense, now shining forth, now hidden, through the forests of experience and particulars... In such a difficult pass there is no hope to be had from human judgement acting by its own power alone, nor from some lucky turn of chance. For no excellence of wit, however great, nor repeated throws of the dice of experiment can overcome these obstacles. Our steps must be guided by a thread, and the whole way from the very first perceptions of the senses must be laid down on a sure plan... but before we can reach the more remote and hidden parts of Nature, it is essential to introduce a better and more perfect method of using the human mind and understanding.

I want to argue that Schmoller's philosophy is methodologically empiricist in just this sense: the route to success is via the right method; the method is experiential; and finding methodological guidelines itself isn't an *a priori* business but a matter of experience.[38]

[36] NO II 42
[37] Urbach and Gibson 1996, pp. 12-13
[38] Which does not contradict that he might be empiricist in other senses, too.

Four features of Schmoller's philosophy point in this direction. First, there is his own emphasis on method. The fundamental principles of economics (if there are any), are not knowable *a priori* by everyone who thinks in the right way but instead may result from a tedious and continuous process of observation, classification and concept formation and of causal explanation. His image of science is very similar to Bacon's. He always emphasised the need for a broad empirical basis before any general conclusions could be drawn. And the road up towards more inclusive principles is via the correct method.

Second, Schmoller's methods are themselves experiential. The first of his three most important tasks of scientific method is observation. But observation is a broad term for Schmoller that includes observation in the narrow sense, statistics, surveys, measurements, historiography and experiments.

How does Schmoller ascertain that his methods are *experiential* methods? Why are certain kinds of measurement or surveys acceptable to the (methodological) empiricist and not others? Part of the answer is that in applying each of the methods Schmoller endorses a physical interaction of the observer with the system under study enters at some point. The examples include reading off a price tag, counting cows, registering suicides, interrogating an official. This distinguishes the empiricist's method from, say, Descartes'. In order to apply the method of doubt one does not need physical interaction with the environment. To the contrary, one can use that method to get rid of the external world altogether.

The other part of the answer is that the method must be justified in a particular way. Justification can be regarded as related to the aim of science, which for Schmoller was (among others) to find out about the laws that govern economic phenomena. A method can be justified by showing that it promotes that aim. One of Schmoller's examples of successful methods is from population statistics. Schmoller remarks,[39]

> A particularly fruitful consideration for statistics was to think of the population of a country as being partitioned, that is, to count separately according to generations, *i.e.* according to age, as well as civilian status, *i.e.* whether they are single, married, widowed or divorced. [...] Again, a wonderful stability shows up here—at least within one nation.

The method of measurement is justified (in Schmoller's eyes) because it leads to stable regularities. For instance, we find a relatively stable proportion of single, married, widowed and divorced men over 18 in the population each year. These regularities, in turn, may be used for causal explanation. In another example Schmoller talks about the causes of suicide. He lists as its "essential" causes: character, instincts and the moral forces of each individual. The hypothesis that these are the causes can be confirmed by seeing that prisoners, servants and soldiers make the greatest contribution to suicide statistics and that more singles than married, more widowed than married, and finally more divorced than married people commit suicide.[40] And these, too, are stable empirical regularities.

Compare this feature of Schmoller's methodology with its equivalent in Carl Menger's methodology of exact economics. For Menger, clearly, a method must be justified as well. But the *aim* is different: Menger seeks *strict* types (types whose instantiations are exactly identical) and *exact* laws (laws that hold by necessity and which are exceptionless). In other words, he seeks a kind of Cartesian certainty. The road to it is by mental abstraction ("breaking phenomena into their simplest parts") and applying the laws of thought (*e.g.* "It is impossible to conceive of a change of one's person from one state to another in any way other than one subject to the law of causality. If, therefore, one passes from a state of need to a state in which the need is satisfied, sufficient causes for this change must exist.")[41]. In Schmoller's methodology (as in Menger's) the means is justified by its end. But Schmoller's end is to find empirical regularities (and causal explanations) whereas Menger's is certainty. This is what makes Schmoller's methodology empiricist and Menger's non-empiricist.

The third feature pointing towards a methodological empiricism in Schmoller's work is that knowledge about the methods themselves is gathered from experience of what has worked well, and one can use some methods to correct others. For example, about observation Schmoller says,[42]

> And even today we have to approach any observation with doubt if it is correct, or if not subjective error, imperfect sight, hasty sanguine conduct, bad training, prejudice and interest present us false images.

---

[39] Schmoller 1871, p. 10. All translations of Schmoller's writings from the German are by the author.
[40] *ibid.*, p. 22-3
[41] Menger 1976/1871
[42] Schmoller 1998/1911, p. 276

And he notes about statistics,[43]

> When we ask where observation could first strip off subjective error and arrive at general truths, then it is in areas where it has subjected certain phenomena to number and measurement.

Schmoller's precept, too, is to "be on the guard". But whatever has worked in the past is worth a try in the future.

Fourth, Schmoller criticised those methodologists whose methodological considerations were not grounded in a deep and comprehensive knowledge of the subject matter (and *vice versa,* those economists who weren't trained in methodological questions). He says,[44]

> ... for the methodology of economics and its development there remains the difficulty that those professional philosophers which take care of epistemology are so remote from the particularities of our science that, all their good will notwithstanding, in incorporating our teachings they cannot do justice to its requirements, and vice versa, that most colleagues, even many that write about methodological questions, do not have the appropriate philosophical education. Among the German colleagues of the last two generations only G. Rümelin, Neuman and Hasbach, lately M. Weber and Eulenberg were properly endowed in order to work on methodological questions; one can even doubt this of J. St. Mill, more of Cairness [*sic*], C. Menger, remaining silent about others. Of all later philosophers, Wundt understands something, Dilthey much of the kind of political and economic work, Windelband and Rickert very little.

There is nothing *a priori* about the methods to be used in political economy. There is, for Schmoller, an aim of method. That is, to subject economic phenomena to the comparative and distinguishing thinking and making them intelligible (by, among other things, finding empirical laws and their causal explanations),[45] but scientists themselves determine the methods that are most useful for their endeavours.

2.2 Abstraction and Concepts

For Bacon, it seems, abstraction was more of a term to denigrate his opponents rather than one with which he would have described his own philosophy. But I think his

[43] *ibid.*, p. 284
[44] *ibid.*, p. 275
[45] *Cf. ibid.* p. 229.

concepts are abstract in an important sense, and his method is a method to arrive at concepts that are abstract in exactly this sense.

In the modern philosophical discussion by "abstract" is usually meant "that which exists outside space and time".[46] In this meaning abstract objects cannot be perceived since they are causally disconnected from the world we live in. Mathematical objects, according to some philosophies of mathematics, are abstract in this sense. Let us call this sense of abstract Platonist-abstract. This meaning of abstract is largely irrelevant for the present concern.

However, there is another sense of abstract which is more relevant here. This sense is orthogonal to the first. It defines as abstract "that which results from a process of abstraction". The process often consists of stripping off certain features of a concrete, real experience. For example, the spot on my skin has many features: it is raised, red at the base and whitish at the top, dry, round, three millimetres in diameter *etc.* But one can regard some of its features by itself, for example its "redness". This logic of abstraction implies that abstract things *cannot* exist independently. We can call this Aristotelian-abstract.[47] I will show that the sense in which Bacon's and Schmoller's concepts are abstract follows this latter tradition.[48]

As mentioned above, Bacon mostly used the term abstraction in a derogatory way. For example, he notes,[49]

> All the other notions which men have adopted up to now are aberrations, improperly abstracted and derived from things.

And in the next aphorism,[50]

> There is as much capriciousness and aberration in the construction of axioms as in the abstracting of notions.

But Bacon ridicules these abstractions because they are derived hastily and following the wrong *method* of abstraction such as the syllogism or ordinary (=enumerative) induction. This suggests that there might be a right method of abstraction for Bacon.

[46] *e.g.* Hale 1987
[47] *Cf.* Lear 1982 on Aristotle's philosophy of mathematics, which discusses the concept of Aristotelian abstraction.
[48] Brown 1999, pp. 12f. makes exactly this distinction between the two senses of "abstract". He calls Aristotelian-abstract "the older sense" and Platonist-abstract "more current usage" and identifies numbers and other mathematical entities as Platonist-abstract objects.
[49] NO I 16
[50] NO I 17

To develop what this is, I first note what Bacon means by "concrete". In various aphorisms, Bacon identifies what he calls "concrete natures" or "concrete bodies" with Aristotelian substances (a term that Bacon himself doesn't like very much) or "natures which are conjoined in a structure":

> Thus these inquiries [regarding certain motions and operations of nature] also consider natures that are concrete, or conjoined in a structure. (NO II 5)
>
> For the time to deal with these [compound forms] will be when we come to *latent processes* and *latent schematisms*[51] and their discovery, as they are found in substances (as they are called) or concrete natures. (NO II 17)

Ordinary, macroscopic objects (Bacon lists "a lion, an eagle, a rose, and gold": NO II 17) are thus *concrete* bodies.[52] Now, following the Aristotelian tradition, we can suspect that abstract natures, then, are certain features of concrete bodies. And indeed, Bacon says,[53]

> Now the precept or axiom concerning the transformation of bodies is of two kinds. The first regards a body as a troop or collection of simple natures; thus in gold the following occur together: that which is yellow; that which is heavy, up to a certain weight; that which is malleable or ductile, to a certain extent; that which is not volatile, and is not consumed by fire; that which becomes fluid, to a certain degree; that which can be separated and dissolved by certain means; and so on, through all the natures that are united in gold.

Natures are characteristics or features of concrete bodies. Bacon calls certain natures "simple". Two related questions thus arise, one more metaphysical, the other more methodological. The first question is what distinguishes "simple" natures from other natures. It asks what those characteristics are that a nature must have in order to count as a *simple* nature. The second question regards the right process or method of abstraction. It asks how we learn about simple natures.

Bacon's answer to the second question is related to his method of true induction:[54]

> Therefore—and this is the heart of the matter—if the notions themselves are muddled and carelessly derived from things, the whole superstructure is shaky. The one hope, therefore, lies in true *induction*.

[51] "Latent processes" are essentially the processes that bring about macroscopic objects, for example the processes that make a tree from an acorn. "Latent schematisms" are the hidden structures of, again, macroscopic objects. But see also Urbach and Gibson 1996, p. 133, fn 113.

[52] In order to make the terminology clearer I call these objects concrete *bodies*, and reserve the word "nature" for its characteristics.

[53] NO II 5

[54] NO I 14

Aspects of the method of true induction have already been discussed.[55] In this subsection I want to emphasise its role as a method of concept formation. I said above that Bacon, in contrast to Descartes, wanted to start from what was available. Thus my picture regarding concept formation is this. Start with the notions "muddled and carelessly derived from things", perform the inductive process using them and see what it yields. If we fail to find the fundamental law governing each nature described by our notions, adjust the conceptual or notional scheme and run it again (in the meantime, the principles about forms already established will suggest new experiments and observations *etc. etc.*). Continue this process until you have arrived at the farthest reaching laws governing the behaviour of our natures. Natures that are described by the conceptual scheme at this end-state are the "simple" natures from above. This, in turn, answers the first of the two questions.

What I have called a fundamental law thus far is Bacon's conception of a "form". That concept, however, is often chided as obscure and it is not essential to my point later on to give a convincing account of what he might have meant. But I note a couple of things here. One is that for Bacon forms and laws are the same thing:[56]

> For when I speak of forms I mean nothing but those laws and definitions of pure actuality which govern and constitute any simple nature such as heat, light, weight, in every kind of material and subject that is capable of receiving them. Therefore the form of heat or the form of light are the same thing as the law of heat or the law of light.

If we follow Mary Horton's[57] idea that Bacon's forms are necessary and sufficient conditions for the presence of a nature, we can say that a form or law has the structure:

$$L\colon C_1C_2\ldots C_N \Leftrightarrow N,$$

where the C's are causal conditions and N is a simple nature. For example, Bacon's "First Vintage" yields as the form of heat:[58]

[55] For secondary literature on that topic, see *e.g.* Urbach 1987, Malherbe 1996 and Milton 2000.
[56] NO II 17
[57] Horton 1973
[58] NO II 20, original emphasis

> *Heat is an expansive motion, checked, and exerting itself through the smaller parts of bodies.* But the extension is modified, in that *while it expands towards the circumference, it yet has some tendency to go upwards.* And this exertion through the parts is also qualified in that *it is not sluggish at all, but hurried and somewhat violent.*

Thus whenever motion of the circumscribed kind is present, heat is present. And whenever heat is present, motion of this kind is present.

My second point regards a remark about laws Bacon makes. He says, they must be "certain, free, and inclining to or having relation to action"[59]. Bacon was very interested in intervention and "gaining power over nature". His aim was the improvement of human welfare. Now, it seems that he almost defines his concept of law (or rather, that of "form") with respect to this aim.[60] At least, given knowledge of the laws of simple natures, we do have power over them:[61]

> For whoever knows the forms of yellowness, weight... and so on, and the means of superinducing them, and their degrees and measures, will see and ensure that these natures may be combined in a certain body, and from this, transformation into gold would follow... It has to be said, however, that this method of operation (that looks at simple natures, albeit in a concrete body) proceeds from those things that in Nature are immutable and eternal and universal, and opens up for human power broad paths, such as the comprehension of man (as things now stand) can scarcely grasp or imagine.

Let's summarise the points about abstraction. For Bacon, admissible abstraction is intrinsically linked to laws or forms. Only those abstract concepts are admissible concepts that pick out what he calls simple natures, and the latter are natures which are governed by a law. And one arrives at knowledge of these laws by Bacon's method of true induction.

Schmoller's ideas about abstraction are analogous. He also (implicitly) distinguishes between admissible and inadmissible abstractions, and criticises his opponents for engaging in the latter. Schmoller remarks about Karl Marx:[62]

> The type of speculating literary scholar without making own observations, without knowledge of the world and humans is Karl Marx; mathematical games were his favourite occupation; they connect with very abstract concepts and with general historic-philosophic images in his work. Because of this characteristic and despite all studies in the English bluebooks, he is

[59] NO II 4; with "free" Bacon means "unconditional".
[60] Antonio Pérez-Ramos, for example, interprets Bacon along the lines that for him *all knowledge* is operational. See his 1988, p. 109 and his 1996.
[61] NO II 5
[62] Schmoller 1998/1911, p. 281

> probably more remote from the requirement of empirically reliable research—as it is demanded today—than any other important economic thinker.

For Schmoller, abstractions that aren't grounded in observations are illicit. But where there are illicit abstractions, there are also permissible ones. Schmoller notes that all observations is based on some kind of abstraction:[63]

> All observation of nature isolates a single process from the chaos of phenomena in order to investigate it in itself. It is always based on abstraction; it analyses a partial content. The smaller the latter is, and the more isolated it presents itself, the simpler is the task.

Any observation, in order to be accessible to research, must be described in some language. But description presupposes a conceptual scheme, and concepts always classify phenomena in one way or other. Thus, one important task of scientific method, according to Schmoller, is classification and concept formation. "All concept formation", Schmoller says, "is an attempt to classify phenomena by combination of the same or similar".[64] However, classification is only preparation[65] for the real important task: causal explanation, and causal explanation proceeds by citing claims about laws and causal relations.

Thus, let us look more closely at what Schmoller has to say about concept formation. All science takes its terms from ordinary language. Scientific concept formation, according to Schmoller, is the continuation of the process of ordinary language formation by ordinary people. Ordinary language formation proceeds by conjoining mental representations with a word.[66] Concrete representations of identical or similar phenomena are thus denoted by a word in ordinary language. This is a continuous process, and therefore meanings are in constant flux.

Science aims at classification and explanation of phenomena, and thus requires a certain constancy of meaning. This is achieved by definition. A definition converts a word or name into a *concept*.[67] Concepts, in turn, classify phenomena, as has been said above. Defining something as an *economy*, says Schmoller, classifies some

---

[63] *ibid.*, p. 277
[64] *ibid.*, p. 296
[65] Schmoller writes: "Observing and describing, defining and classifying are preparatory activities. But what we want to achieve with it is the knowledge of the nexus of economic phenomena; ...", *ibid.* p. 304.
[66] *Cf.* Locke's account of concept formation as sketched above.
[67] Schmoller *op. cit.* p. 297

phenomena as an economy and others as a non-economy. According to which criteria, then, shall we classify phenomena?

We start by simple similarities we can find in our observations. Schmoller distinguishes between analytic and genetic classifications. An example of an analytic classification is Wagner's division of all economic phenomena into *private*, *public* and *charitable systems*. Here the similarities between the individual systems are their ways of financing. An example for a genetic classification is Schmoller's own of distinctions among *village*, *town*, *territorial* and *national economies*. Here the similarities obtain between the different stages of development of individual economies. But similarities between certain observable aspects of the phenomena aren't everything. In an interesting remark regarding the "ideal" end-state of concept formation, Schmoller notes,[68]

> I think we can simply say: the more straightforward the objects characteristic of a science are and the further that science is advanced in its results, the more perfected concepts it has and *the easier it can incorporate its laws and highest principles into its concepts and definitions* and deduce more from them. This is the case in parts of natural science. However, the more complicated the object of a science, the more remote it is from this ideal.

Schmoller goes on to remark that political economy is quite far away from the ideal, but he nonetheless regards a conceptual scheme that incorporates knowledge about laws into its terms as an ideal. The same point is made more clearly by Schmoller in his *Grundriß*:[69]

> It is furthermore correct that the more advanced a science is, the more it is able to lay its achieved truths and causal connections into the definitions of its highest concepts; because the latter belong to the most essential characteristics, to the representations that are essential for the word.

Claims about laws and causal connections, thus, belong to the defining characteristics of a concept.

Again, we can thus say that admissible abstractions are those that give us concepts whose referents are governed by laws. These nomic/causal claims, and Schmoller emphasises this time and again, are however not "contained" in the concepts in the

[68] *ibid.*, p. 302, emphasis added. According to the usage in this Thesis, a law is a pattern or structure in the world rather than a statement about such a pattern or structure. Thus we will say that a concept may contain knowledge or claims about laws/causal connections rather than the laws/causal connections themselves.

[69] Schmoller 1900, p. 105

way a realist about definitions has it. We cannot know about the laws by merely analysing our concepts. Rather, at the end of process of observation, classification, explanation, observation, ... we will have refined our concepts such that they are connected with nomic/causal claims which are themselves established by experience.

As an aside, I would like to mention that both Bacon and Schmoller's theories of abstraction imply that the notions created in this way are *concepts* in Norman Robert Campbell's sense. In a famous remark, Campbell defines "concepts" in his *Physics, the Elements*,[70]

> A concept is a word denoting an idea which depends for its meaning or significance on the truth of some law. The conclusion at which we have arrived is that most, if not all, of the recognised laws of physics state relations between concepts, and not between simple judgements of sensation which remain significant even if no relation between them is known.

It should be clear from the above discussion that I take it that at least some of Bacon's and Schmoller's terms are concepts in this sense.[71] To be sure, Schmoller begins the scientific process with terms that are empirical notions in the more traditional empiricist sense of the word, *i.e.* they do denote mental representations of phenomena. However, as science advances and we find out more about the laws governing phenomena, our classificatory system or conceptual scheme more and more reflects these laws and their meanings depend on claims about the latter. Eventually, our conceptual scheme will perfectly mirror the (causal) structure of the world, and thus, to say it with Spinoza, "The order and connection of ideas is the same as the order and connection of things".[72]

Therefore, in the ideal (and probably not reachable) end-state, notes Schmoller, science operates purely deductively. I want to add that in this ideal end-state all law-claims are analytic because their truth is built into the concepts that they relate themselves. The only synthetic statements would be those of the form $\exists x\, Cx$, where C can stand in for any scientific concept such as *lever*, *territorial economy* or *hot*.

[70] Campbell 1957/1922, p. 45

[71] However, recall the remark made in footnote 68. For Campbell, a law is a relation between concepts, *i.e.* itself a linguistic entity. By contrast, throughout the Thesis I will understand a law as something that occurs in reality, *i.e.* as an extra-linguistic entity. Therefore, the meaning of a concept will at best depend on a *claim* about laws/causal relations rather than on the "truth of some law" as Campbell says.

[72] *Ethics* IIP7

A kind of wholism is implied by this understanding of concepts—even before the unattainable end-state: if the significance of a concept depends on the truth of some law-claim, the whole conceptual framework might change whenever new law-claims supersede old ones, measurements are made more accurate and new phenomena are created. This wholism was beautifully described by Bacon in the *Advancement of Learning*: "Out of all the words we have to extract the sense in whose light each single word is to be interpreted".[73]

2.3 Induction and Observability

Both Bacon as well as Schmoller are usually regarded as inductivists, and it is quite evident from the above discussion why this should be so. Recently, Laura Snyder claimed in an interesting contribution[74] that the particular kind of Baconian induction is different from what is normally understood by the term, it is not limited to inductive generalisation and it incorporates both inductive and deductive elements.

Snyder compares Bacon's to Whewell's inductivism in order to make certain points about Whewell. I am comparing Bacon's to Schmoller's inductivism in order to make certain points about economics, but I think her article is instructive with respect to the nature of the Baconian inductivism. Of her five points of agreement between Bacon and Whewell, the last three are important for us:[75]

> (3) This inferential process, according to Bacon and Whewell, is called 'induction' but is not limited to inductive generalization.
>
> (4) On both their methods, this process is intended to reach hypotheses referring to unobservables.
>
> (5) Finally, Bacon and Whewell agree that an inductively obtained hypothesis be tested by its empirical consequences (*i.e.* they deny the claim that inductive generation is sufficient for confirmation).

[73] quoted from Popper 1972, p. 187, emphasis added
[74] Snyder 1999
[75] See Snyder 1999, p. 531-2. The other two points are that both Bacon and Whewell require inference from data to hypothesis and that this inference requires a process that is gradual. I hope I have covered these points sufficiently above.

It is well-known and acknowledged[76] that Bacon was himself sceptical about enumerative induction and a critic of it. In the Preface to the *Great Instauration*, for example, he writes,[77]

> But far the greatest change I make is in the very form of induction, and the judgement made from it. For the induction of which the logicians talk, which proceeds by simple enumeration, is a childish affair, unsafe in its conclusions, in danger from a contradictory instance, taking account only of what is familiar, and leading to no result.

In this quote we can already see the two points of criticism Bacon makes:[78] first, the results of simple induction are, for the most part, not "certain". We have already seen that Bacon aims at the discovery of *forms*, and a form of a given nature is "such that when it is there, the given nature infallibly follows".[79] Thus simple induction isn't of much help in discovering forms.

But another aspect of Bacon's concept of form—related to his second criticism—is more interesting: forms are usually unobservable.[80] The criticism is that ordinary induction can take "account only of what is familiar", it cannot reach the level of unobservables. But as the knowledge of forms is at least one aim of Baconian science, we cannot reach this aim by ordinary induction.

But how did Bacon think that one could? Snyder's answer is: by analogical reasoning. Among Bacon's "Prerogatives of Instances" we find two kinds of "Supplementary" or "Substitutive Instances", *viz.* substitution by degree or analogy.[81] Bacon explains:

---

[76] *Cf. e.g.* Mill 1874, p. 227

[77] Gibson and Urbach 1996, p. 21. In a footnote to NO I 17, Gibson and Urbach explain ordinary/simple/ enumerative induction by means of an example from an influential textbook of logic from 1551: "Rhenyshe wine heateth, Malmesey heateth, Frenchewine weateth, neither is there any wyne that doth the contrary: Ergo all wine heateth" (see *ibid.*, p. 47).

[78] For a discussion, see Snyder 1999, pp. 533ff.

[79] NO II 4

[80] See for example Snyder 1999, p. 534, Peltonen 1996, p. 17 and Quinton 1980, pp. 45-6. I have not been able to find a statement from Bacon himself to the effect that forms for the most part are or must be unobservable, but there are many implicit hints in his writings. First, from his own example of the form of heat as a certain kind of motion, one can see that forms are at least sometimes unobservable since not in every hot body Bacon lists one can observe motion. Second, two other conceptions, that of the latent process and that of the latent schematism already refer to unobservables. Since the investigation into forms is supposed to be deeper (see *e.g.* NO II 1), it would be surprising if they were observable. Third, the investigation into forms is classified under "metaphysics" by Bacon (see NO II 9). Again, if physical investigations already transcend the level of the observable, then metaphysical investigations certainly will, too.

[81] NO II 42

> It comes about when the non-sensible is conveyed to the sense, not by perceptible operations of the non-sensible body itself..., but by studying some related body that is sensible.

It is by this mode of reasoning that Bacon can infer that motion is present in certain cases where it does occur at the observable level (*e.g.* boiling water and flames), that it is present also in cases where it is does not occur at that level. Though more sceptical than Snyder, Mary Hesse[82] also sees analogical or more generally, hypothetical, reasoning as a solution to this problem:

> ... and he also admits a certain amount of reasoning from observed to unobserved natures, as for example when the motion which is the form of heat is said to be motion of small (not directly observable) particles. The arguments by which he arrives at this specification of the form of heat are not inductive after his own recipe, but hypothetical and analogical; but it must be remembered that they are only arguments leading to the first vintage, and elsewhere Bacon warns against injudicious use of the method of analogy for eliciting "things not directly perceptible". It cannot be said that he deals adequately with the difficulty inherent in explanations in terms of hidden natures, but given the presuppositions of his method it is impossible to see how he could have done better, for hidden natures demand hypothetical arguments.

The last of Snyder's claims is that Bacon's "induction" requires confirmation by instances or experiments deduced from statements about laws or causal relations that have been inductively established in order to be complete. Indeed, Bacon says in NO I 106,

> Now in establishing axioms by means of this induction, we must also examine and check whether the axiom so established is only fitted to and made to the measure of those particulars from which it is derived, or whether it is larger and wider. And if it is larger or wider, we must look to see whether it confirms its largeness and wideness by indicating new particulars, as a kind of collateral security; lest we either stick fast in things already known, or perhaps weakly grasp at shadows and abstract forms[83], not solid and actual material things.

Thus Bacon's inductive method contains a "whiff of deductivism". Inductively established fundamental laws cannot be trusted unless they are further confirmed by suggested new experiments and particulars.

From these considerations Snyder infers that Bacon's (and Whewell's) method is inductive in a sense that throws an interesting light on some discussions in 20th century philosophy of science. Some logical positivists thought that inductive

[82] Hesse 1964, p. 147
[83] In this context Bacon of course refers to Platonist-abstract forms in the above sense.

generalisations can never come up with theories because they can only infer from observable to unobserv*ed* but observ*able*.[84] But this means that inferences to theories or theoretical hypotheses could never be rational. "Rational inference" in this context means either deduction or inductive generalisation. Therefore, we end up in a dilemma: either we must remain in the realm of the observable or we must admit non-rational elements into our science. Snyder argues that Baconian induction makes the dilemma obsolete: because his method allows us to transcend the boundaries of the observable and the method of inference is rational (analogical reasoning), we can have theoretical hypotheses and do not need irrational elements in our science.[85]

Snyder's main point seems correct to me: that Baconian induction is a much more complex activity than mechanical enumeration and elimination. In my view, it differs also at another level, and this is—not surprisingly—at the level of concept formation.

Let us look at John Stuart Mill's classical definition of inductive inference. Mill says,[86]

> Induction, then, is that operation of the mind, by which we infer that what we know to be true in a particular case or cases, will be true in all cases which resemble the former in certain assignable respects. In other words, *Induction is the process by which we conclude that what is true of certain individuals is true of the whole class…*

My point is that because concept formation is part of the inductive process, the assignment of individuals to classes is itself part of the matter. And the classes are defined, at least partially, with respect to the relations that hold between different aspects of the individuals of that class. We do not simply infer: A_1 is B, A_2 is B, A_3 is B *etc.*, therefore all As are B (or some elaborate version of this), but the meaning of

---

[84] *Cf.* for instance Ernst Nagel in his *Structure of Science*: "An immediate corollary to the difference between experimental laws and theories just discussed is that while the former could, in principle, be proposed and asserted as inductive generalizations based on relations found to hold in observed data, this can never be the case for the latter" (Nagel 1960, p. 85).

[85] It seems to me that the question itself is rather obsolete, for we have known for at least forty years that science is not a mechanical or simple rule-following business, and if "rationality" demands it to be mechanical or simple rule-following, most philosophers would happily admit that parts of the scientific enterprise are irrational. I am also not too sure whether analogical reasoning can be mechanical-rational in a sense that Snyder needs.

[86] Mill 1849, p. 210

the predicate "*A*" is itself a result of the process (and thus the assignment to classes), and the fact that all *A*s are *B* is part of the meaning of "*A*".[87]

Of course, I believe that Schmoller's inductivism has exactly the features that distinguish Bacon's inductivism from simple inductive generalisation. To repeat these features:

(1) it allows inferences about unobservables

(2) it has deductive elements

(3) concept formation is part of the process.

The point about unobservables is trivial, in part, in the context of Schmoller's investigations because a great share of economic phenomena are unobservable mass phenomena.[88] "Observing" in economics always includes counting, measuring, surveying *etc.* as well as observing in the narrower sense. But in at least one context Schmoller uses an inference from something "observable" (in a sense to be specified) to something unobservable—similar to Bacon's use of analogy. That is when we infer about the motives of other people. Schmoller argues,[89]

> To observe economic phenomena means to establish the motives of the respective economic actions and their results, their course and effect in the external world. We recognise the motives of our actions directly by observing our own mental life; *from us we infer to others.* What happens in the world we know through our sense impressions, which we interpret and understand as objective events. All our experience stems thus from these to sources of perception.

Thus Schmoller did admit of introspection as an important source of knowledge, and he counted our own "mental life" as observable. Schmoller nonetheless emphasised time and again the importance for economics of an objective psychology.[90] But this psychology would make use of introspection (and thus inferences about mental states, which are unobservable for most people), and Schmoller was a harsh critic of behaviourism.[91]

---

[87] I don't want to presuppose a regularity view of laws here. One can read my "all *A*s are *B*" in any empiricist way one wants including regularity/tendency/*c.p.* law/capacity.

[88] I use the concept of observability here in van Fraassen's sense of "imperceptible to the unaided senses". See his 1980.

[89] Schmoller 1998/1911, p. 276, emphasis added.

[90] See *e.g. ibid.* p. 311ff.

[91] See for example Backhaus and Hansen 2000.

Very clearly, Schmoller's inductivism has deductive elements. Exactly parallel to Snyder's point (5) from above, Schmoller says that any inductively established rule must be tested by consequences:[92]

> Also the last test of every inductively established principle is that it proves true when it is continuously used deductively. From this one can follow how closely related induction and deduction are... For years I've told the students that like the left and right foot to walking, induction and deduction to the same extent belong to scientific thinking.

This latter remark resembles very nicely Bacon's suggestion that "Such a road [of scientific investigation] is not level, but rises and falls; first ascending to axioms, then descending to works" (NO I 103). I like both metaphors of the walking and the road because they suggest (a) that the scientific process is gradual and (b) that it is continuous, not ending once we've reached stage three of the process.

That, finally, concept formation is part of the process is an explicit feature of Schmoller's methodology. The second stage of his method is classification and concept formation, and as we have seen, concepts are formed in such a way as to reflect knowledge about stable empirical regularities.

Let us take stock here. I have tried to show that Bacon's empiricism has a number of features which I believe are instructive for economics: its emphasis on *method*; the central role of *concept formation*; that *the meaning of concepts reflects nomic/causal knowledge*; the insignificance of *observability*; the joint use of *induction and deduction*. I have also tried to establish that Schmoller's economic methodology reflects these features. The reason for including the comparison with Schmoller in this Chapter is to give some plausibility to the claim that Bacon's ideas may be relevant for economics. I will not argue further for this claim *directly*. However, in the next section I will present an *indirect* further argument for it. For that purpose I will hypothesise what kinds of epistemic virtues Bacon and Schmoller would accept in the light of their methodologies. But because these are the virtues that are endorsed also by a number of contemporary economists (as I will try to show), these methodologies are of relevance for economics.

[92] Schmoller 1998/1911, pp. 321f. See also Schmoller 1998/1881, p. 102

3 Epistemic Virtues in Economics

I thus leave the historical discussion in order to make way for an investigation of the relevance of Bacon's and Schmoller's ideas today. Because the ultimate aim of this Thesis is to learn about current economics I have to digress to draw some inferences from the discussion of Bacon and Schmoller so far for modern theorising techniques. (Contemporary) economists construct models. But except by a few nonconformists model construction is not conducted for its own sake. Models are vehicles to gain knowledge. They are the spectacles through which economists see their world. They construct models to *learn*: about the economy, but also about our theories.[93] I believe that this feature of modelling is true for the anti-realist and the realist alike. Whether one is an instrumentalist of Friedmanian breed, a Samuelsonian operationalist or a realist à la Mäki, one builds and uses a model to investigate (describe? predict? explain? understand?) aspects of the economy.

If this is true, there may be qualitative differences between the models constructed. We might be able to learn better using one kind of model rather than a different kind, and learn better in some situations with this kind of model and in other situations with that kind of model. Features of models that pick out such qualitative differences I shall call *epistemic virtues*. What epistemic virtues would Bacon and Schmoller suggest?

It follows directly from the discussion so far that Bacon and Schmoller regard the adequate description of phenomena, their classification into kinds and causal explanation as the central aims of science. I propose to translate these into the language of epistemic virtues of models as *phenomenal adequacy* and *explanatory power*. As I will explain shortly, a phenomenally adequate model will not only describe features of reality correctly, it will also pick out features that occur repeatedly and are well behaved (and hence we can speak of "kinds"). A model that has explanatory power will also help us to gain understanding of this stable feature.

I would like to insinuate that at least Schmoller would also have regarded *exactness* as a third salient epistemic virtue of economic models.[94] By exactness Schmoller

[93] *Cf.* Morgan and Morrison 1999.

[94] This case is probably harder to make for Bacon—he hardly ever uses the concepts of "exactness" and "precision" *etc*. However, I think that implicit in his criticism of "carelessly abstracted notions" is the requirement that objects to which our notions are applicable should form distinct classes and that

means more or less accurate measurability. An exact law, for him and as we shall see for William Stanley Jevons as well, is a law whose quantities are measurable (or countable). In his lecture on moral statistics, for example, Schmoller notes:[95]

> The most obvious entity, which turns a researcher into a statistician is the number in itself, or more accurately, the *exactness* of observation. [...] Only *exact* mass observation has revealed to us that a regular mathematical rhythm governs the colourful chaos of life and that a lawfulness is apparent in the phenomena of personal and social life, a lawfulness which is probably different in its nature from the lawfulness observed in the planetary orbits and the attractions of chemical molecules, but which manifests the almost same certainty, the same relentlessness in its measurable results or at least appears to do so.

And throughout his work Schmoller cheers methods of statistics and measurement. In the passage I used above already, Schmoller notes:[96]

> When we ask where observation could first strip off subjective error and arrive at general truths, then it is in areas where it has subjected certain phenomena to number and measurement.

And further,[97]

> The importance of the statistical method for the progress of all knowledge in the realm of state, society and economy was immense, nonetheless. The perfection of it was one of the most significant advances in the area of the social sciences for 150 years. Statistics has in many ways substituted the experiment that is lacking here; only it has created a sense of exactness and precision in this area of knowledge; it has replaced many vague images with fixed quantitative representations; it for the first time has allowed to subject the mass phenomena, which hitherto had been accessible to a vague estimation, to precise observation, and to use their countable characteristics for an absolutely certain characteristic; it has noticed the changes in the development through its use of tables, visual representations and other aids of comparison, pointed towards recognition of causes and allowed to measure the influence of certain essential and accidental causes.

However, exactness is no end in itself for Schmoller, and we have to be aware of the potential dangers of pseudo-exactness which is not approved by the phenomena. For example, discussing the mathematised political economy of his time, he notes,[98]

---

all objects of one class should bear objective similarities to one another. This is more or less the concept of exactness I shall use below, in Chapter 3.

[95] Schmoller 1871, p. 5. See also Schmoller 1998/1911, p. 332.

[96] *ibid.*, p. 284

[97] *ibid.*, pp. 285f.

[98] *ibid.*, p. 320

> One will not be able to deny that in [the mathematical] form the results of abstract theory can be represented cleanly and precisely, that its mode of inference is often more certain than the one found in the accustomed writings, and that it increases the intuitive understanding of certain processes, at least for the mathematically trained person. [...] The constructions and formulæ, however, employ elements which in fact cannot be determined, they are not amenable to measurement, and, by filling in fictitious magnitudes for psychological causes and immeasurable market relations, they give an impression of exactness that does not exist.

Using exact mathematical concepts is a good, but only as long as the phenomena represented lend themselves to measurement and thus concepts can be given empirical meaning.

Next I would like to demonstrate that *phenomenal adequacy*, *explanatory power* and *exactness* are demanded also of contemporary economics. This discussion, in addition, should make the conceptions as they are used in this Thesis somewhat clearer.

A model is phenomenally adequate essentially if it describes or predicts the phenomena correctly. I propose to use "phenomenal" adequacy rather than, say "empirical" or "descriptive" or "predictive" adequacy for two reasons. First, "phenomenon" seems to be a conception which is used and understood by many contemporary economists. For example, in an article I will discuss in greater detail in Chapter 2, George Akerlof says:[99]

> The example of used cars captures the essence of the problem. From time to time one hears either mention of or surprise at the large price difference between new cars and those which have just left the showroom. The usual lunch table justification for this *phenomenon* is the pure joy of owning a "new" car. We offer a different explanation.

In a more recent article in political economy, its author Colin Campbell summarises his topic of interest as follows:[100]

> The ability of small but zealous groups of individual citizens to secure their preferred political outcomes is the topic of this paper. Of particular interest are cases in which such a group is able to do so even if the proportion of all citizens who share its preferences is small. There are many contemporary examples of this *phenomenon*. [...]
>
> Positive economic analysis of this "decisive minority" *phenomenon* has yielded two explanations. One is that despite the one citizen-one vote structure of American elections, avenues exist for an individual to express the intensity of her political preferences in an

[99] Akerlof 1970, p. 489, emphasis added

addition to the direction of those preferences. [...] The second explanation is realized by framing political participation as a public-good problem, in which it is in each citizen's interest to free-ride on the actions of others. [...]

In these two quotes three important characteristics of the conception of phenomenon that I use in this Thesis are apparent. First, a phenomenon is a feature of interest. It is an occurrence or a process that entices the attention of economists. That in Singapore on March 3, 1971, a rubbish bin fell over is not a phenomenon. But the rise of the dot.com economy is one. And so is the fact that asset bubbles mostly burst. Second, a phenomenon is usually a stable, recurring feature of the (economic) world. As such, it is general; a type rather than a token. The same (or a similar) phenomenon can be instantiated in many different historical and regional contexts. However, in some cases individual occurrences may be called phenomena, if they are remarkable enough (such as the 1929 crash, or the aforementioned rise of the internet industry). Third, phenomena are the objects of economic *explanations*. To the extent that an economic model can be explanatory, it will explain a phenomenon. The aim of economic model building is to demonstrate that an instance of the phenomenon occurs under the assumptions of the model. Consequently, phenomenal adequacy will mean, until further notice, reaching this aim.

The second reason to use the conception of a phenomenon is that it is an key element in the analyses of scientific practice in some recent contributions to experimentalist philosophy of science. Ian Hacking, in his *Representing and Intervening*, for example, emphasises the first two characteristics of this conception:[101]

A phenomenon is *noteworthy*. A phenomenon is *discernible*. A phenomenon is commonly an event or process of a certain type that occurs regularly under definite circumstances. The word can also denote a unique event that we single out as particularly important.

James Woodward, on the other hand, emphasises the second and third characteristics in his *Data and Phenomena*:[102]

Phenomena, as I shall use the term, are relatively stable and general features of the world which are potential objects of explanation and prediction by general theory.

---

[100] Campbell 1999, p. 1200, emphasis added and original emphasis removed
[101] Hacking 1983, p. 221, original emphasis
[102] Woodward 1989, p. 393. *Cf.* also Bogen and Woodward 1988.

The usage of the word, then, by economists and experimentalist philosophers of science coincides enough to solicit our acceptance, and thus I shall employ the term in just this sense.

Whether or not models that describe features of the world can also be explanatory, and if so in what sense, is of course at the heart of the realism-antirealism debate. Although I shall defend a moderate realist point of view in this Thesis, I do not intend the realism-antirealism issue to be the its main focus.

There are a number of uses of the term "explanation" or "explains" in economics. Some of them are entirely consistent with an instrumentalist philosophy of economics. In theoretical economics, for instance, one often finds a sense of explanation in which, roughly, a model explains a phenomenon if an instance of the phenomenon can be derived from the model's assumptions. This sense of explanation, obviously, coincides with my usage of the term "phenomenal adequacy". On the other hand, in econometrics "explaining the data" often means explaining its variance in the statisticians' sense.

But there are other senses of explanation which are stronger. Some of these are incompatible with an instrumentalist philosophy of science. We can summarise these senses with the equation explanation = phenomenal adequacy (of a certain kind) + *X*, where *X* may be one of the following:

- compatibility with our intuitions[103]
- simplicity[104]
- the phenomenally adequate model is a causal model[105]
- the phenomenally adequate model incorporates knowledge about laws[106]
- the phenomenally adequate model incorporates necessary relations between universals[107]

[103] See Kreps 1990a, p. 12, who seems to hold such a view. See also Chapter 5 of this Thesis which briefly discusses Kreps's view.
[104] Aumann 1985 says that simplicity is an important feature of modelling in game theory.
[105] Rappaport 1998 defends an account of causal explanation in economics.
[106] Lawson 1997 ascribes a DN model of explanation to mainstream economics.

- the phenomenally adequate model (or its concepts) can be used to derive a large number of different phenomena[108]
- *etc.*

Because of the ubiquity of the usage of "explanation" or "explains" I think it is fair to assume that explanatory power *in some sense* is an epistemic virtue demanded of models in economics. Below I will argue that the unification aspect of "good" economic models or concepts is the sense of explanation that is really relevant in contemporary economics. In the last Chapter, finally, I will argue that a "good" economic explanation is a causal explanation, which also achieves some unification of phenomena.

Exactness is much harder to get a grip on. This is because the virtue of exactness is rarely explicitly targeted, neither by economists nor by economic methodologists nor by philosophers. There are notable exceptions of course, including Menger and Jevons whose views I shall discuss below. But the main reason to hold that exactness is a significant epistemic virtue of contemporary economics is that it is often cited as a defence of the fact that economics has become increasingly mathematised over the past sixty years or so. Consider for example David Kreps's justification of game theory:[109]

> (3) Game theory comprises formal mathematical models of 'games' that are examined deductively. Just as in more traditional economic theory, the advantages that are meant to ensue from formal, mathematical models examined deductively are (at least) three: (*a*) It gives us a clear and precise language for communicating insights and notions. In particular, it provides us with general categories of assumptions so that insights and intuitions can be transferred from one context to another and can be cross-checked between different contexts. (*b*) It allows us to subject particular insights and intuitions to the test of logical consistency. (*c*) It helps us to trace back from 'observations' to underlying assumptions; to see what assumptions are really at the heart of particular conclusions.

We can learn from Kreps that exactness is not an end in itself but a presupposition for other important features such as objectivity and consistency. Again, I will say

[107] Menger 1960/1871 may hold such a view. See Mäki 1997. To be fair, though, for Menger an explanatory ("exact") model, or rather, law, in his terminology, cannot be phenomenally adequate. See Chapter 5 for a discussion of this claim.
[108] See Chapter 2 for a defence of a version of this claim.

more about what exactly I mean by exactness but so far I submit that exactness is not a primary aim but a derivative one, albeit one which is a vital ingredient in the realisation of other virtues economic models might or might not have. At this point, however, we can already distinguish two important senses of exactness: extensional and intensional exactness.[110] A concept *C* is extensionally exact if and only if for any object or process or structure *a* one can unambiguously determine whether *a* is or is not *C*. Concepts that are inexact in this sense may lead to the famous sorites paradoxes.[111] For example, there is no way to determine unambiguously whether certain men are bald or not or whether certain objects are heaps. By contrast, "having 13,500 hairs" or "being 146 pebbles" is exact.[112]

A concept *C* is intensionally exact if and only if it has a determinate meaning. By "determinate meaning" I mean that it has either a unique definition or, in case of multiple definitions, the definitions are equivalent (*i.e.* they must apply to the same set of actual or potential objects or processes or structures). Being a brother in this sense is intensionally exact, because it can be defined as "male sibling". A rival definition ("son of the same parents") is equivalent as it applies to the same objects.[113] Let us take an example from Bacon to understand what an intensionally inexact concept is. "Moist" is intensionally inexact because it may mean for example[114]

> something that readily surrounds another body, but also something with no definite boundaries and unable to become solid; something which yields easily in every direction; something which easily subdivides and scatters itself; or easily coalesces and becomes one; easily flows and is set in motion, easily adheres to another body and makes it wet; and which easily liquefies, or melts, when it was previously solid.

In this case extensional and intensional inexactness coincide because there are actual objects for which we cannot decide whether they are nor are not moist. "Fortune", on the other hand, is only intensionally inexact because for Bacon there is no such thing but its various definitions apply to different potential things.

---

[109] Kreps 1990b, pp. 5-6
[110] *Cf.* Roy Sorensen's characterisation of "vagueness" (Sorensen 1992, ch. 7).
[111] See Sainsbury 1988 on these paradoxes.
[112] if we know how to determine what a hair is and a pebble
[113] This is Sorensen's example.
[114] Urbach and Gibson 1994, p. 65

Given this distinction, we can see that even if two economists agree that exactness is an important epistemic virtue, they may disagree about the form of exactness a model should take. Let us further distinguish between intensional exactness and mere intensional exactness. Mere intensional exactness is intensional but not extensional exactness, *i.e.* it applies to objects which are only potential but not actual.[115] This in addition to the observation that Kreps stresses a model's role in understanding over its role in providing an accurate representation of phenomena[116], may lead one to the conclusion that Kreps regards mere intensional exactness as a virtue, whereas clearly for Schmoller it is a vice.

So far the conceptions of our epistemic virtues of phenomenal adequacy, explanatory power and exactness are clearly not precise enough and in need of specification in the light of actual scientific practice. The next three Chapters will mainly fill in these gaps. Chapter 2 discusses the process of abstraction by which economists form concepts which supposedly represent phenomena, as well as a version of explanatory power which is consistent with a great deal modelling practice. Chapter 3 interprets models that are constructed using concepts abstracted by the process of the previous Chapter and points to a trade-off that obtains between explanatory power and exactness. Chapter 4 analyses models that are abstracted by measurement and provides an interpretation. It, too, points to a trade-off between explanatory power and exactness. Chapter 5 takes those ideas together and provides a theory of concept formation that aims at a simultaneous realisation of the three epistemic virtues. The theory will be illustrated by means of a case study.

4 Conclusions

I began this Chapter by noting that this is a Thesis in Baconian topics. The introduction mentioned concept formation and the collaboration of faculties as two Baconian topics of interest. Then I tried to give some content to these terms. A key

[115] Bacon's "prime mover" may thus be merely intensionally exact. Let us assume the term is precisely and uniquely defined. As, according to Bacon, there is no referent for it, it cannot be extensionally exact. Hence, it is *merely* intensionally exact.
[116] See his 1990a/b.

element in Bacon's philosophy is his emphasis on *method*. Concept formation, in turn, is a vital ingredient of Bacon's scientific method. The concepts that the method of concept formation generates will reflect nomic or causal knowledge. And that method is a broadly inductive method, which nonetheless uses some deductive elements and does not shy away from unobservables.

In the next three Chapters I will look at a couple of methods of concept formation that one can find in contemporary economics. It will be shown that the nature of the method of concept formation is crucial for the realisation of epistemic virtues. Not all methods of concept formation will result in economic models that are phenomenally adequate, explanatory and exact at the same time. Quite to the contrary, we will see that the methods examined imply trade-offs between these virtues.

Ignoring exactness for the time being, we can say that models that are merely phenomenally adequate but not explanatory (*i.e.*, they describe or predict features of reality but do not further our understanding of it) are the models of Bacon's "empiricists". "Empiricists", says Bacon, "like ants, merely collect things and use them". On the other hand, models that are merely explanatory (in *some* sense![117]) but not phenomenally adequate are the models of Bacon's "rationalists": "Rationalists, like spiders, spin webs out of themselves". The middle way would be a way that combines empiricist and rationalist elements—and realises the two virtues at once.

Looking at contemporary economics, it seems that Bacon's urge still has some relevance. Ian Hacking, for example, in his chapter on Bacon in *Representing and Intervening*, notes,[118]

> Hence we can diagnose doubts some of us share about the social sciences. Those fields are still in a world of dogmatics [rationalists] and empirics. There is no end of 'experimentation' but it as yet elicits almost no stable phenomena. There is plenty of speculation. There is even plenty of mathematical psychology or mathematical economics, pure sciences which have nothing much to do with either speculation or experimentation. Far be it from me to offer any evaluation of this state of affairs. Maybe all these people are creating a new kind of human activity. But many of us experience a sort of nostalgia, a feeling of sadness, when we survey

[117] What I mean here is that we can have models that further our understanding by showing that certain characteristics follow from accepted first principles but without being predictively successful or directly relevant for the description of real phenomena. Parts of game theory have sometimes been interpreted in this way.
[118] Hacking 1983, pp. 248-9

> social science. Perhaps this is because it lacks what is so great about fairly recent physical science. Social scientists don't lack experiment; they don't lack calculation; they don't lack speculation; they lack the collaboration of the three. Nor, I suspect, will they collaborate until they have real theoretical entities about which to speculate—not just postulated 'constructs' and 'concepts', but entities we can use, entities which are part of the deliberate creation of stable new phenomena.

Maybe this statement is just the result of the smug self-confidence of a philosopher of natural science who does not understand much of the work of social scientists. But his feeling of nostalgia is felt by at least some of the scientists working in the field. In an article surveying work of 30 years in the area of the natural rate of unemployment, Olivier Blanchard and Lawrence Katz remark,[119]

> Nevertheless, we feel that a divide has grown between macroeconomists and labor economists, at least on this side of the Atlantic. Too much theoretical work on the natural rate of unemployment by macroeconomists is divorced from microeconomic evidence, and too much microeconometric work seems in search of a broader theoretical framework for interpretation. This is unhealthy. We thus end with a plea for more joint efforts by macro and labor economists to better integrate theoretical and empirical work on wage determination and unemployment.

Maybe again Blanchard and Katz's views are idiosyncratic or the natural rate case is unrepresentative or both. But I guess many people have the feeling that there is something right in what Bacon, Schmoller,[120] Hacking and Blanchard and Katz have to say about their respective fields. Now inasmuch as these considerations are true of contemporary economics, I think there is much to be learned from Bacon and Schmoller.

What I think can be taken from Bacon and Schmoller is a framework of thinking about how the empirical and rational faculties might co-operate. In both cases, the kind of empiricism defended does not render the mind inactive in a way suggested for example by Karl Popper's bucket metaphor. A wrongly understood empiricism, according to Popper, interprets the mind as a bucket that, initially empty, is gradually filled with information from the external world which the mind keeps essentially

[119] Blanchard and Katz 1997, p. 70

[120] So far, I haven't mentioned Schmoller's equivalent to the ant and the spider. But one finds many passages in which Schmoller criticises the classicals' and neo-classicals abstract, speculative reasoning, and others in which he criticises some historicists' "thoughtless poly-historics".

unadulterated.[121] By contrast, in Bacon's and Schmoller's philosophies (and so too in Popper's) the mind is active in many ways: it seeks, gathers and classifies information. It deliberately *creates* new phenomena (Hacking). It invents measurement procedures, experimental set ups and social survey techniques. It devises analogies and hypothesises explanations. Certainly, the level of analysis at which I have presented their thoughts is too abstract, too general to make claims about real methodological improvements: these have to be made at the level of concrete methods of investigation. But it is a framework nonetheless.

One of the purposes of this Thesis is to fill in the obvious gaps in this framework. I shall do so from two sides. On one side, by analysing detailed case studies from economics I want to give the distinctions drawn here and the terms employed a meaning relevant for contemporary economics. On another side, I want to use a number of conceptions and tools of analysis provided by modern experimentalist philosophy of science, which I take to be the contemporary equivalent of Bacon's and Schmoller's projects, and see whether and how they can apply to economics. In so doing, I will adopt an unrepentantly naturalist stance throughout the four succeeding Chapters. That is, I deny that there is any principled difference in the methodology of investigating natural and the social sciences. What I do not deny is that there may be a variety of differences in the concrete circumstances of our epistemic enterprises. Certain kinds of experiment that are characteristic of physical science may not be possible in the social and life sciences for ethical or practical reasons. It is possible that in the realm of the social sciences we more frequently encounter situations of causal complexity. Standard methodological principles of, say, physically isolating a process of interest can perhaps not be applied in other areas. But none of these differences implies that methodological insights cannot be transferred from one area to another, and that the sciences cannot mutually profit from one another. In this way I think that we can at least consider whether we can make fruitful loans from other contexts and see what they can teach us about economics—which is the ultimate focus of this thesis.[122]

[121] Popper 1972, ch. 2. Popper in fact ascribes an empiricism of this kind to Bacon, see Appendix of that work.
[122] It may also be the case that claims made here about economics may be transferable to other sciences but whether this is the case is beyond the scope of this Thesis.

I hope that eventually there will emerge a perspective on economics which is consistent with important "experimentalist" ideas, and which also teaches us how there may be a tenable middle ground between the strict rationalism of the spiders and the strict empiricism of the ants.

Chapter 2

Unifying Intuited Concepts: Abstraction and Explanatory Power in Theoretical Economics

Chapter 2

Unifying Intuited Concepts: Abstraction and Explanatory Power in Theoretical Economics

It cannot be that axioms established by argumentation should have any value for discovering new works; for the subtlety of nature is far greater than that of argument. But axioms properly and methodically derived from particulars can very well point to and indicate new particulars again, and so render the sciences active.

Francis Bacon—NO I 24

1 Introduction

John Stuart Mill argued that political economy must be an abstract science. For Mill, this meant that political economists 1) isolate a small set of causal factors which are well understood from the chaos of concrete economic circumstances and 2) deduce results from the principles governing this set. Truths derived in this way are "abstract truths": they are not directly informative about real, concrete circumstances but only about what the known factors (or factors of interest) contribute to the situation.

To a large extent, modern economics seems to follow Mill's scheme. Economic phenomena, understood in the sense of the previous Chapter as "stable, noteworthy and recurrent features of the economy", are isolated from their concrete historical and local contexts and are explained by deriving representations of them from a thought experiment or model that focuses on a small set of causal factors such as *asymmetric information*, *quality differences*, *transportation costs* and *spatial distance*. Any failure to predict an economic outcome correctly can always be assigned to what Mill called a "disturbing cause": a factor operative in the concrete situation that was omitted in the abstract thought experiment or model.

Three important questions arise in the present context. First, in what sense is whatever we call abstract abstract? We have already discussed two senses of abstract. In the previous Chapter we distinguished between Platonist-abstract and Aristotelian-abstract. Platonist-abstract is what exists outside space and time. Aristotelian-abstract is whatever is formed by a process of abstraction. Mill's sense of abstract belongs to the broad Aristotelian category. His method of abstraction is *analysis*: breaking a complex phenomenon into parts each of which is governed by a law. Hence for Mill, individual causal factors are isolated from a complex and their behaviours are examined severally.

The second question, which is related to the first in an obvious way, is how the concepts that denote the abstract things are formed. This is the question of what is the method or process of abstraction. Again, we have already discussed three important methods of forming abstract concepts.[1] According to the Lockean view, we form abstract concepts by extracting those features of our ideas that resemble each other and giving them a common name. According to the Platonic view, abstract concepts are formed by a process of purifying an idea in a rational dialogue. And finally according to the Bacon-Schmoller view, they are formed by finding features of reality which are law-governed (and incorporating knowledge about these features into our concepts).

The third question, eventually, is how the abstract relates to the concrete. According to Mill, we can get from an abstract truth to a concrete one by a process he called "synthesis":[2]

> The method of the practical philosopher consists, therefore, of two processes; the one analytical, the other synthetical. He must *analyze* the existing state of society into its elements, not dropping and losing any of them by the way. After referring to the experience of individual man to learn the *law* of each of these elements, that is, to learn what are its natural effects, and how much of the effect follows from so much of the cause when not counteracted by any other cause, there remains an operation of *synthesis*; to put all these effects together, and, from what they are separately, to collect what would be the effect of all the causes acting at once.

There are many other ways to get from the abstract to the concrete, however, often discussed in the context of de-idealisation and concretisation.[3] Although this third question is intriguing too, I shall focus almost exclusively on the first two in this Chapter. More specifically, I shall argue that there is a tradition in modern

[1] An abstract concept is a concept that denotes an abstract thing.
[2] Mill 1948/1830, p. 159, original emphasis
[3] See for example Nowak 1980, Cartwright 1989, ch. 5 and Cartwright 1999, ch. 2.

economics in which abstract concepts are formed by a process that combines intuition and casual observation in such a way that a thought experiment fitted with these concepts yields a representation of the phenomenon of interest.[4]

Abstract concepts are not formed for their own sake but as a means of constructing thought experiments and models that help us in realising certain epistemic virtues. In this Chapter I focus on "explanation". That is, I will examine how abstract concepts are used to construct models that are thought to have explanatory power. In particular, I shall argue that it is consistent with a great deal of economic practice and commentary that economic models or concepts have explanatory power to the extent that they help to unify and systematise our thinking about economic phenomena. I will argue for this claim in section 3 after having discussed economists' method of abstraction.

2 Abstract Concepts

George Akerlof begins his seminal article *A Market for "Lemons": Quality Uncertainty and the Market Mechanism*[5] with the words: "This paper relates quality and uncertainty". Here two factors of interest are denoted by two abstract concepts: quality and uncertainty. Clearly, these two concepts are abstract. They do not denote any concrete objects or situations we could point to or see or touch. But what exactly is the sense in which they are abstract, and how do we form these concepts?

2.1 Forming Abstract Concepts By Thought Experiments

The phenomenon of interest in Akerlof's article is that in markets where the quality of a good matters, average quality tends to be relatively poor and market size small.[6]

---

[4] The formulation about the combination of intuition and casual observation is taken almost verbatim from Kreps 1990, p. 12. See Chapter 5 for the full quote and a discussion.

[5] Akerlof 1970. Akerlof has just been awarded the Nobel Prize for this work.

[6] I want to draw a distinction here between a stable, recurrent feature of the economy in the sense Akerlof uses the term, and a stronger sense. For Akerlof, the phenomenon of interest is known by casual observation. For Bogen and Woodward, on the other hand, in order to call something a phenomenon, it must have undergone systematic empirical investigation of the kind described in their article. I shall say about this kind of empirical investigation much more in Chapter 4. I will equivocate

An example he uses to demonstrate the existence of this phenomenon is that of the automobile market. An aspect of this example is that there is a large price differential between new cars and those which have just left the showroom. Akerlof notes that the "usual lunchtable justification" of this aspect is that people enjoy having new cars. But he wants to offer a different explanation. Akerlof present both an intuitive explanation as well as a mathematical model. The intuitive explanation is the following:[7]

> Suppose... that there are just four kinds of cars. There are new cars and used cars. There are good cars and bad cars (which in America are known as "lemons"). A new car may be a good car or a lemon, and of course the same is true of used cars.
>
> The individuals in this market buy a new automobile without knowing whether the car they buy will be good or a lemon. But they do know that with probability q it is a good car and with probability $(1 - q)$ it is a lemon; by assumption, q is the proportion of good cars produced and $(1 - q)$ is the proportion of lemons.
>
> After owning a specific car, however, for a length of time, the car owner can form a good idea of the quality of this machine; *i.e.*, the owner assigns a new probability to the event that his car is a lemon. This estimate is more accurate than the original estimate. An asymmetry in available information has developed: for the sellers now have more knowledge about the quality of a car than the buyers. But good cars and bad cars must still sell at the same price—since it is impossible for a buyer to tell the difference between a good car and a bad car. It is apparent that a used car cannot have the same valuation, it would clearly be advantageous to trade a lemon at the price of a new car, and buy another new car, at a higher probability q of being good... Thus the owner of a good machine must be locked in. Not only is it true that he cannot receive the true value of his car, but he cannot even obtain the expected value of a new car.[8]

An intuitive explanation of a phenomenon such as this I would like to call a *thought experiment* in Roy Sorensen's sense. Although in principle one could with as much justification say that Akerlof provides a *model* of the situation, I prefer the terminology of the thought experiment in this context because it is more specific and emphasises the non-material character of the representation. I shall reserve the term

---

on these to senses of the term but I want to note that it is always possible that systematic empirical investigation will show that the "phenomenon" is not a genuine phenomenon after all—just as in the case of Bacon's "moist" objects.

[7] *ibid.*, p. 489

[8] As an aside, this story obviously cannot explain that aspect of the phenomenon which Akerlof stressed above: that the car's price drops immediately after sale. For if the buyer needs to verify the quality of the car, most certainly he needs some time to do so (Akerlof says, he owns the car "for a

"model" for the more formal or mathematical representations of the kind discussed below and use "thought experiment" for both the more intuitive representations such as Akerlof's as well as mathematical models. I note, however, that there is a continuum between the two kinds of representation (Akerlof uses some mathematical notation in his thought experiment), although there are probably paradigms of each.[9]

Sorensen defines a thought experiment as[10]

> an experiment... that purports to achieve its aim without the benefit of execution. The *aim* of any experiment is to answer or raise its question rationally[,]

and experiment as[11]

> a procedure for answering or raising a question about the relationship between variables by varying one (or more) of them and tracking any response by the other or others.

The variables of interest are the subjective probabilities of the market participants about the cars' quality, the cars' price and market size. The varied variables are the subjective probabilities of the buyers. Responses of the price and, implicitly, of the quantity variables are tracked.[12] The question Akerlof is interested in is apparently what kinds of factors may bring about the phenomenon that in markets where the quality of a good matters, average quality tends to be relatively poor and market size small. He shows that if the subjective probabilities of the market participants are varied in such a way that an informational asymmetry between owners and potential buyers results, or more precisely such that owners *know more* than buyers, the phenomenon of interest arises in this thought-experimental setting.

Using this thought experiment, Akerlof can demonstrate the plausibility of his thesis that it is an asymmetry of information which triggers the decrease in market prices and quantities. Given our intuitions about the economic behaviour of individuals and certain simplifying assumptions about the market (there are only four kinds of cars *etc.*), Akerlof shows that it is conceivable that asymmetric information causes the drop in prices and quantities.

---

length of time"). But the aspect of the phenomenon, which matters, is that the price drops *immediately*.

[9] On the continuum character of thought or *virtual* experiments and mathematical models, see Morgan 2001.

[10] Sorensen 1992, p. 205, original emphasis

[11] *ibid.*, p. 186

But Akerlof does more than that. He also provides a mathematical model of the situation. One can reconstruct his model on the basis of the following premisses:[13]

(1) The demand for used automobiles depends "most strongly"[14] upon two variables—the price of the automobile p and the average quality of used cars traded, μ, that is: $Q^d = D(p, \mu)$.

(2) Supply and average quality of cars depends only on price: $\mu = \mu(p)$ and $S = S(p)$.

(3) There are just two groups of traders with utility functions $U_1 = M + \Sigma x_i$ (where M is consumption of goods other than automobiles and x_i is the quality of the i^{th} car) and $U_2 = M + \Sigma 3/2x_i$.

(4) Both types of traders are von Neumann-Morgenstern maximisers of expected utility.

(5) Group one has N cars with uniformly distributed quality x, $0 \leq x \leq 2$; type two traders have no cars.

(6) The price of other goods is 1.

(7) The traders have incomes Y_1 and Y_2, respectively.

(8) All goods are infinitely divisible.

(9). The average quality of cars μ is known to all participants but the individual car's quality x_i is known only by its owner.

From these premisses, Akerlof derives the following conclusions.

(C1) Type-1 traders demand automobiles according to the schedule $D_1 = Y_1/p$ if $\mu/p > 1$ and $D_1 = 0$ if $\mu/p < 1$.

(C2) Type-1 traders supply automobiles according to the schedule $S_1 = pN/2$ if $p \leq 2$.

(C3) Type-2 traders demand automobiles according to the schedule $D_1 = Y_2/p$ if $3\mu/2p > 1$ and $D_1 = 0$ if $3\mu/2p < 1$.

(C4) Type-2 traders *ex hypothesi* do not have cars (premiss 5): $S_2 = 0$.

(C5) Adding demand schedules yields: $D(p, \mu) = (Y_2 + Y_1)/p$ if $p < \mu$, $D(p, \mu) = Y_2/p$ if $\mu < p < 3\mu/2$ and $D(p, \mu) = 0$ if $p > 3\mu/2$.

(C6) From (C5) it follows that demand is always zero (as the price is p and the average quality is $\mu = p/2$, we are always in the bottom section of the schedule).

---

[12] A completion of Akerlof's story would be along the following lines: if the price drops significantly, owners of good cars will be less likely to sell their cars as the expected market price is below their estimation of the value of the car. Hence quantities exchanged drop along with prices.

[13] *Cf. ibid.*, pp. 490ff.

[14] says Akerlof. In fact, it depends only on these factors.

Thus Akerlof can demonstrate that a representation of the phenomenon of interest must occur under the assumptions made. The difference between the intuitive thought experiment and the mathematical model is as follows. Whereas the *intuitive* thought experiment described above could demonstrate that it is *conceivable* or *plausible* that an instance of the phenomenon will occur when asymmetric information is present in a market structure described by it, the mathematical model shows that it *must* occur in a highly specific structure.

How do the intuitive thought experiment and the mathematical model relate to the issue of interest here, *i.e.*, concept formation? A comparison with Bacon's and Schmoller's theories of concept formation will be instructive. Take Bacon's theory first.

According to the interpretation of Bacon offered in Chapter 1 above, a "good" abstraction is one where the abstract concept picks out a simple nature. A simple nature is a nature which has a unique form, that is, it has necessary and sufficient causal conditions for its occurrence. Why is "moist" a badly abstracted concept on this reading? Because it is not the case that the same necessary and sufficient causal conditions are present in every object that we call "moist". Why is "heat" a good abstraction? Because whenever heat is present, motion of a specific kind is present.

For Schmoller, of course the concept of a law is more complex than the simple interpretation of a Baconian form that I have sketched. For him a law, in the sense of an occurrent regularity, is always structure dependent. That is, the regularity happens on account of a causal structure, and it happens only as long as the causal structure is present. We may observe a fixed number of suicides in Paris each year. But this "law" will change in line with the moral and socio-economic factors that bring it about. The general idea to connect accounts of concept formation with laws is, however, present in Schmoller's writings as it is in Bacon's. "Good" or "admissible" abstractions are concepts that help us in the formulation of law-statements and causal explanations. The *territorial economy* is an admissible abstract concept because it represents a phenomenon that is part of a law of the development of economic systems from the family to the national economy, and that law dictates, among other things, that all economies at a certain stage share a number of fixed properties.

There is a positive and a negative analogy between the concept formation theories of Bacon-Schmoller and the practice of modern economics. The positive analogy is that in all three cases "good" abstractions are those that help us formulating *law-claims*. In the theory that has been ascribed to Bacon, there is even a one-to-one-to-one correspondence between an abstract concept, a simple nature and a law. In Schmoller's case the situation is more complex but here too abstract concepts pick out features of economic phenomena that are law governed. In the same manner Akerlof's thought experiment as well as his model describe a situation in which asymmetric information regularly brings about low market prices and quantities.

The seminal character of Akerlof's contribution is in part determined by the fact that he invented the term "asymmetry in information", or as it has come to be known, "asymmetric information". Naturally, it is not the term itself that has been decisive but the fact that it denotes a factor that under certain conditions regularly brings about certain market situations: a factor that can be used for explaining economic phenomena.[15] The point is that the phenomena Akerlof is interested in have a great number of characteristics. His contribution consists in having isolated in thought one of them which is able to bring about market situations of interest in a regular way under specific conditions.

So much for the positive analogy. The negative analogy is the mode of finding out about the laws in which the phenomena represented by the abstracted concepts play a part. In Bacon's methodology we have seen there is an iterated three stage process of observation, classification and eliminative induction, which is supplemented by Bacon's twenty-seven "Prerogatives of Instances". In Schmoller's methodology there is a similar iterated three stage process of observation and description, classification and concept formation, and causal explanation, and a number of tools that help during one stage or the other such as the "historical method" and "statistics and surveys".

The formulation of law-claims in modern economic science, by contrast, proceeds to a great extent by thought experiments of the kind described above. Instead of a

[15] The market outcome relevant here is now called "adverse selection". According to Milgrom and Roberts 1992, adverse selection refers to "the kind of precontractual opportunism that arises when one party to a bargain has private information about the something that affects the other's net benefit from the contract and when only those whose private information implies the contract will be especially disadvantageous for the other party to agree to a contract" (p. 595).

method which, like Bacon's and Schmoller's if we want to follow the arguments of Chapter 1, combines induction and deduction in the process of forming concepts and establishing nomic claims, a kind of hypothetico-deductive method is used almost exclusively in this process. Empirical work is frequently conducted but as a means of *testing* certain implications of a thought experiment rather than as a basis for concept formation and inductive inference.

2.2 The Empirics of Asymmetric Information

Let us try to find evidence for the claims made in the previous paragraph. It is difficult to substantiate claims of this kind with respect to actual economic practice because they involve a statement of the form $\neg\exists x\ (Ix)$ (where the I could signify inductive methods and a model of the theory could be the set of economic articles), and statements of this form are never verifiable in a finite time if the set of objects to which I could apply is at least potentially infinite.[16] In this case I am claiming that there are virtually no inductive methods in modern economic science but I cannot verify this hypothesis as one can always claim that the next investigated article (or teaching practice or university corridor chat) will involve inductive methods. However, I have tried to find at least some evidence for the assertion. Among other things I performed an ABI Inform search for the terms "asymmetric information", "informational asymmetry" and "informational asymmetries" for the years until 1985. The search landed 79 hits.[17] In most cases, Akerlof's results were extended to other phenomena. That is, it was demonstrated that asymmetric information is an important factor that contributes to the production of a great variety of market results, including *brain drain, low employment, litigation, organisational slack* and *bidding cartels.*

Of the 79 articles mentioning one of the three terms in citation or abstract, 72 or 91.14 per cent were purely theoretical. What I mean here is that these articles provided thought experiments in the sense employed in this Thesis and present no reports of either material economic experiments, nor analyses of real data. Seven

[16] This is just Popper's argument that statements of the form $\exists x\ (Ix)$ are not falsifiable upside down. One cannot falsify the hypothesis that there is a unicorn because no empirical evidence would be sufficient to contradict it. The next object investigated could always be a unicorn.

[17] In fact, the result was 85. However, of these five appeared twice for some reason and one was an article not on asymmetric information but mentioned the term only as a suggestion for future research.

articles, or 8.86 per cent, were at least partly empirical. In most of these, implications of the postulated thought experiments were derived and "tested" by means of an econometric model. In one article, however, an experiment was conducted to study the influence of asymmetric information on budgetary slack. Because of this article's uniqueness I will quote the abstract in full:[18]

> A single-period experiment is conducted to test empirically the effects of private information about productive capabilities, risk preferences, and participation on budgetary slack. Five hypotheses related to budgetary slack are formulated and tested. The experiment used 40 full-time MBA students who are randomly assigned either to an Information Asymmetry group or a No Information Asymmetry group. The results confirm the hypotheses that a subordinate who participates in the budgetary process builds budgetary slack and that slack is in part attributable to a subordinate's risk preferences. In addition, while the possession of information gives a subordinate greater opportunity to misrepresent productive capability, this opportunity is mitigated by social pressure to provide truthful information. Greater variation in slack production was exhibited by subordinates having private information about productive capability than by those who lacked this information.

In this last article, an attempt is made to empirically investigate whether, among other things, there is a causal relation between asymmetric information and budgetary slack. The point I want to emphasise is not whether its author, S. Mark Young, successfully established the causal link. The point to emphasise is rather that an empirical investigation about hypothesised causal connections is comparatively rare. In the great majority of cases thought experiments instead of material experiments are supposed to do that job.

This case is supposed to confirm that there are above mentioned positive and negative analogies with Bacon's and Schmoller's theories of concept formation. On the one hand, to know what a quantity of interest *is*, is to know what it *does*, and what the quantity does is determined by the laws that govern its behaviour. On the other hand, the laws governing the quantity's behaviour are, by and large, not investigated by inductive methods of causal/nomic inference of the kind describe in Chapter 1. Rather, they are determined by the thought experiments and models that we usually find in the economic literature.

In order to get some insight into the distinction I am trying to draw consider the following famous thought experiment. According to the Aristotelian theory of

[18] Young 1985

motion, velocity equals force divided by resistance. This theory, however, makes the motion of projectiles baffling because there appears to be nothing pushing the moving object, *i.e.*, no force that moves the object. One solution to the puzzle was that the air displaced by the front of the flying object hastens to its back in order to prevent the creation of a vacuum. Jean Buridan criticised this theory with the following *reductio*. Imagine Aristotle's theory were true. This would imply that if one reduced the surface area of the moving object exposed to air pushing from behind, the object would move more slowly (*e.g.* by sharpening a flying arrow's back its flight would be shortened). But for all we know, this is absurd. Hence Aristotle's theory must be false.[19]

Compare this with one of Bacon's *Crucial Instances*. He discusses the same theoretical puzzle and argues that there are two competing causal hypotheses. One is the Aristotelian idea that air particles cause the projectile to move. The other one seems to be Buridan's impetus theory: that the "mover" imparts an enduring quality called *impetus* to the projectile.[20] Bacon argues,[21]

> But the following, among others, could be a Crucial Instance concerning this question: that an iron sheet, or a fairly stiff iron wire, or even reed or quill sliced down the middle, after being pressed into a curve between finger and thumb, springs up. For it is obvious that this cannot be ascribed to the air collecting behind the body, because the source of motion is in the middle of the sheet or reed, not in the extremities.

The difference between the two cases is that Buridan held his argument to be convincing without material execution of the experiment, whereas for Bacon, only a real experiment can confirm Buridan's hypothesis. One could paraphrase Bacon by saying that *if* the iron sheet, wire, reed or quill were to spring up after being pressed into a curve, *then* Aristotle's theory would be ruled out. Buridan, by contrast, is persuaded without actually sharpening an arrow's back. But Buridan's argument would not be convincing were it not for our intuitions about flying objects, *i.e.* our causal background knowledge. *A priori* there seems nothing absurd about the claim that sharpening an arrow's end would slow down its velocity. It is only absurd given all we know about projectiles.

---

[19] For a brief discussion of this thought experiment, see Sorensen 1992, p. 197.

[20] Bacon uses the following words to describe the second hypothesis: "[motion comes about] from the parts of the body itself not enduring the pressure, but moving forwards in succession to relax it" (NO

It is, I believe, in a very similar way that thought experiments in economics are held to be convincing. A thought experiment (or model of course) in economics is convincing if it accords with our intuitions about the behaviour of people. A difference is that the intuitions do not concern real people and their real motives and actions but rather idealised agents that act only on the economic motive. If we enquire whether a thought experiment or model is acceptable or not, we might ask ourselves, "assuming the agent is acting on the economic motive alone, what *would* his or her behaviour be like?" If the thought experiment uses kinds of behaviour that are described in the answer to that question, the thought experiment is acceptable.

This idea about thought experiments implies also that it is not the case that *no* observations or data are involved in writing down models or conducting thought experiments in order to establish new claims about laws. To the contrary, very frequently articles begin with a description of an observed phenomenon of interest. Akerlof's example of the car market is a case in point. Akerlof writes:[22]

> The example of used cars captures the essence of the problem. From time to time one hears either mention of or surprise at the large price difference between new cars and those which have just left the showroom.

This is clearly an observation. But in devising his thought experiment, Akerlof does not create any *new* data or make *new* observations. He uses what is known already and hypothesises a factor, asymmetric information, which could, under certain conditions, produce an instance of the phenomenon. In the vast majority of cases, and certainly in the literature on asymmetric information until 1985, laws governing the behaviour of factors of interest are established in exactly this way.

I have not the space here to develop thoughts about the intriguing question of how thought experiments as I understand them in this Chapter can be informative about real economic situations.[23] Nor can I talk about implications regarding the relation between real and thought experiments. My point is rather a descriptive one: to find an answer to the question of what is the method of abstraction used in contemporary

---

II 36). I can only guess that this is a version of the impetus theory, but as that theory was Aristotle's main competitor at the time, this seems a reasonable guess.

[21] NO II 36

[22] Akerlof, *op. cit.*, p. 489

[23] For accounts dealing with this question, see Brown 1991, Sorensen 1992, Horowitz and Massey 1991, Kuhn 1981/1964, Häggqvist 1996 and Mach 1920, pp. 183-201. Mach's account (he coined the term *Gedankenexperiment*) is probably the most similar one to my story about Buridan.

theoretical economics. I hope, thus, to have verified the first thesis of this Chapter, *viz.* that there is a tradition in modern economics in which abstract concepts are formed by a process that combines intuition and casual observation in such a way that a thought experiment fitted with these concepts yields a representation of the phenomenon of interest.

3 Economic Explanation

Why does Akerlof think his model *explains* the phenomenon of the large price differential between new and second-hand cars? First, note, as I have done in the previous section, that Akerlof uses the automobile market as an *example*. This suggests that the phenomenon he is really interested in is of a more general nature and the price differential in the automobile market is a feature of a subclass this more general phenomenon.

This impression is confirmed by section III "Examples and Applications" of his paper. In this section he discusses no less than four applications of his model to markets as diverse as insurance, labour, business where honesty matters and credit. In the insurance market, the phenomenon is that for high risks insurance premiums are higher and the level of coverage is lower than their equilibrium levels. In the labour market it is that workers from ethnic minorities tend to get employed less frequently. In the third context, Akerlof claims that his model can shed light on the cost of dishonesty. The latter consists, he says, not only in the amount by which the purchaser of the bad good is cheated but also in the loss of legitimate business that is "driven out" by bad business. Last, there are two "phenomena" in credit markets of underdeveloped countries that Akerlof describes. One is that in India so-called managing agencies, which are dominated by communal groups, control a major fraction of the industrial enterprise. The other is that moneylenders charge their clients "extortionate" rates in India.

3.1 Lemons and DN-Explanation

We have said above that Akerlof's thought experiment and model yield a representation of the phenomenon of interest. Economists often call this an explanation. After the passage just quoted from Akerlof's article, he says that the usual lunch table justification for the phenomenon is the "pure joy of owning a new car", but that he offers a different *explanation*.[24] One could thus postulate a kind of deductive-nomological (DN) model of explanation in economics and say something along the lines:

(E) *Condition E.* A phenomenon of interest P is explained if a representation of P can be demonstrated within a thought experiment or mathematical model.

It seems, though, that it is not enough to say that *any* model that is able to demonstrate a representation of a phenomenon of interest also explains it. The logical positivists had four "adequacy conditions" that were supposed to guarantee that the premisses of the argument (the "explanans") that supposedly explains some observational statement (the "explanandum") are appropriate. The four conditions were:[25]

(a) The set of propositions constituting explanans and the explanandum must be a valid deductive argument.

(b) There must be at least one general law-statement among the premisses of the explanans.

(c) The explanans-propositions must be testable.

(d) The explanans must be true.

Let us assume, for the sake of the argument, that economic thought experiments and models by and large fulfil these criteria, or rather, that they fulfil them in as much as any other good scientific explanation. Would our condition (E) be a theory of economic *explanation* in this case? Many philosophers and economists would probably deny this. As an example, look back at the opening of Akerlof's article.

[24] *ibid.*
[25] *Cf.* Salmon 1992

Here he describes the phenomenon of the price differential between new and almost new cars, and then says, "The usual lunch table *justification* for this phenomenon is the pure joy of owning a 'new' car". That is, he says is that there is a thought experiment or model that involves the concept of *pure joy of owning a new car* from which a representation of the phenomenon of interest can be derived. Let us assume that this "explanation" fulfils the condition (E) and the adequacy conditions. But the "new car joy model" is not an explanation, it is only a "lunch table justification". Akerlof's own thought experiment and model, on the other hand, do genuinely explain the "phenomenon" (in his view at least). So what is the difference between the lunch table justification and Akerlof's approach? Among the differences between a genuine and a false explanation relevant in the context of economics, the following four come to mind.

First, the "lunch table justification" is *ad hoc.* There is an empirical puzzle that cries out for explanation. We "explain" the phenomenon by writing the result into people's preferences. This is certainly, and not only for the Popperians among economists, bad scientific style. Almost *any* phenomenon could be explained in such a way.

Second, the model makes a substantive assumption about human behaviour. Though not necessarily *ad hoc* in the first sense, many substantive assumptions about human behaviour are considered by economists to be *ad hoc* in a different sense. Consider Dan Hausman's reconstruction of the methodological rules that govern general equilibrium theory and a further comment regarding Keynes's theory about people's propensity to consume:[26]

> 1 Generalizations about choice or other economic phenomena are *ad hoc* and should be avoided unless they are derivable from equilibrium theory and further legitimate generalisations about preferences, beliefs, and non-economic constraints on choices.
>
> 2 Further generalizations about preferences, beliefs, and constraints are legitimate and may be incorporated into economic theories only if they do not threaten the central place of rational greed, the possibility of equilibrium, or the universal scope of economics. [...]
>
> One sees these methodological rules at work especially in the reactions of economists to macroeconomic theories that lack explicit microfoundations. Keynes' claim that the marginal propensity to consume is less than one is by itself regarded as "*ad hoc*"... It is legitimate only

[26] Hausman 1992, pp. 95-6

> if it can be shown to follow from equilibrium theory such as Modigliani's life-cycle hypothesis or Friedman's permanent income hypothesis... Modigliani's and Friedman's hypotheses about beliefs and preferences are not *ad hoc,* because they do not threaten the explanatory unity of equilibrium theory. Generalizations about wage or price stickiness have been criticized as *ad hoc* on the same grounds...

This kind of ad hocness is different from the first kind as the assumption of a marginal propensity to consume smaller than one cannot explain every economic phenomenon. But it is an assumption of substance which is not derivable from general equilibrium theory.

Hausman's last remark suggests a third criticism of the lunch table justification, one which I consider to be the most important. The assumption of a strong preference of new over old cars is entirely local: it does not unify and systematise our thinking about economic phenomena. All one can "explain" with the concept of *pure joy of owning a new car* is one single phenomenon—that under investigation. By contrast, Akerlof's lemons model is (so he claims) applicable to a wide range of cases, cases as diverse as minority employment and credit markets in underdeveloped countries.

The ability of Akerlof's model to unify diverse phenomena is related to a fourth difference between genuine and false explanations. It is sometimes stressed by philosophers of science that a genuine explanation should further our understanding of the explained phenomenon or phenomena.[27] This seems to be a requirement made by at least some economists too. Close to a passage I already quoted in Chapter 1, David Kreps says,[28]

> *the point of game theory is to help economists understand and predict what will happen in economic contexts.* [...]
>
> (2) Game theory *by itself* is not meant to improve anyone's understanding of economic phenomena. Game theory... is a tool of economic analysis, and the proper test is whether economic analyses that use the concepts and language of game theory have improved our understanding.

The lunch table justification does not help much with understanding the phenomenon of interest. New cars are more expensive than old cars because people enjoy having new cars. After this "explanation" we know exactly as much as before. But showing that the explanandum-phenomenon systematically depends on the operation of a

[27] See *e.g.* Schurz 1988
[28] Kreps 1990, p. 5f., original emphasis

factor which can be used in explaining a great number of different phenomena reduces the complexity of the economic world, and via this path—according to some philosophers—furthers understanding.[29]

Following the idea that a genuine economic explanation unifies and systematises thinking about economic phenomena, we may say that a thought experiment or model is considered explanatory if (a) a representation of the explanandum-phenomenon can be demonstrated within it; and (b) a great range of other, dissimilar phenomena can be demonstrated by it or by a similar thought experiment or model. Let us look at a second case study to try to make this idea more precise. Whereas the first case study was from micro economics and thirty-odd years old, the second one is from macro economics and very recent.

3.2 Lemons, International Trade and Unification

The case I want to look at is an NBER working paper by Maurice Obstfeld and Kenneth Rogoff, entitled "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?". I here reproduce their Abstract in full because it very nicely summarises the project of the paper and illustrates the point that I want to make.[30]

> The central claim in this paper is that by explicitly introducing costs of international trade (narrowly, transport costs but more broadly, tariffs, nontariff barriers and other trade costs), one can go far toward *explaining* a great number of the main empirical puzzles that international macroeconomists have struggled with over twenty-five years. Our approach elucidates J. McCallum's home bias in trade puzzle, the Feldstein-Horioka saving-investment puzzle, the French-Poterba equity home bias puzzle, and the Backus-Kehoe-Kydland consumption correlations puzzle. That one simple alteration to an otherwise canonical international macroeconomic model can help substantially to *explain* such a broad arrange of empirical puzzles, including some that previously seemed intractable, suggests a rich area for future research. We also address a variety of international pricing puzzles, including the purchasing power parity puzzle emphasized by Rogoff, and what we term the "exchange-rate disconnect puzzle." The latter category of riddles includes both the Meese-Rogoff exchange rate forecasting puzzle and the Baxter-Stockman neutrality of exchange rate regime puzzle. Here, although many elements need to be added to our extremely simple model, we can still show that trade costs play an essential role.

[29] On this point, see Friedman 1974 and Kitcher 1981.
[30] Obstfeld and Rogoff 2000, Abstract, emphasis added

I clearly cannot go through all six "puzzles", and there is no need to do so. But by means of a couple of examples I want to illustrate (a) how their model can demonstrate the puzzling phenomena and (b) why they consider their models to be explanatorily more satisfactory than alternative models that also capture those phenomena. The first phenomenon Obstfeld and Rogoff discuss is that of the home bias in trade puzzle.

The "puzzle" consists in the fact that econometric evidence suggests that international goods markets are far more segmented than is predicted by free trade models. It has been shown, for example, that trade among Canadian provinces was twenty times greater than trade between Canadian provinces and U.S. states.[31] Obstfeld and Rogoff derive this result by adding an assumption of so-called iceberg shipping costs to an otherwise standard macroeconomic model. Iceberg shipping costs are such that only a fraction $1-\tau$ of the goods shipped arrive at the opposite shore. Their model is a two-country two goods endowment economy, in which the utility function of a representative home consumer is given by

$$C=\left(C_H^{\frac{\theta-1}{\theta}}+C_F^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}, \tag{2.1}$$

where C_H is home consumption of the home good, C_F is home consumption of the foreign good and θ is the elasticity of import demand. Consumers in the foreign country have identical preferences with consumption C_H^* and C_F^*, respectively. The endowments with local goods in each country is Y_H at home and Y_F in the foreign country. There are four prices:

Good	**produced at home**	**produced abroad**
consumed at home	P_H	P_F
consumed abroad	P_H^*	P_F^*

Assuming competition and iceberg shipping costs, arbitrage implies:

[31] This appears to be a puzzle mainly for economists. I expect the fact that, say, the British are more likely to trade with themselves rather than with the French to be hardly surprising for at least some others.

$$P_F = P_F^*/(1-\tau), \tag{2.2}$$

$$P_H = (1-\tau)P_H^*. \tag{2.3}$$

Defining $p \equiv P_F/P_H$ and $p^* \equiv P_F^*/P_H$ yields:

$$p^* = p(1-\tau)^2. \tag{2.4}$$

Maximising (2.1) yields the first-order conditions:

$$\frac{C_H}{C_F} = p^{\theta}, \quad \frac{C_H^*}{C_F^*} = (p^*)^{\theta}. \tag{2.5}$$

Combining (2.4) and (2.5) gives

$$\frac{C_H}{C_F} = (1-\tau)^{-2\theta}\left(\frac{C_H^*}{C_F^*}\right). \tag{2.6}$$

Under the simplifying assumption that $Y_F = Y_H$, it can be shown that home consumption of the home good is related to home consumption of the foreign good in the following way:[32]

$$C_H = (1-\tau)^{1-\theta} pC_F. \tag{2.7}^*$$

With this formula we can produce a representation of the phenomenon. If trade costs τ were zero, consumption of the home and foreign goods would be identical in monetary terms. With trade costs of 0.25 and an elasticity of import demand of 6, the ratio of home to foreign consumption is 4.2:1. Obstfeld and Rogoff then claim that these values ($\tau = 0.25$, $\theta = 6$, $C_H/pC_F = 4.2$) are all very plausible in the light of recent empirical findings.

Obstfeld and Rogoff' capture the phenomenon within a mathematical model. Formula $(2.7)^*$ is deduced mathematically from the assumptions, and one can show that there is a home bias in trade by filling in the appropriate numbers that can be estimated empirically. There is a problem, however, with the estimation of τ. As stated in the Abstract, which was quoted above, τ may include "transport costs, tariffs, nontariff barriers and other trade costs". First, there are the usual measurement problems that are ubiquitous in empirical economics. Tariffs can be relatively straightforwardly measured, but non-tariff barriers only indirectly using an

[32] Formula numbers with no asterisk follow Obstfeld and Rogoff's numbering, the ones with asterisk are not explicitly their's.

economic model—a measure which of course is only as good as the assumptions that go into the model. Transport costs again are measurable but they tend to be underestimated due to what is called substitution bias. Deriving the extent of the home bias with a formula such as (2.7)* would require the estimation of transport costs of all goods that *might* be traded. However, we can measure transportation costs only for goods that are *actually* traded. It is likely, now, that goods with lower transportation costs are substituted for goods with very high transportation costs. Thus, an empirical measure tends to underestimate τ. Other trade costs, finally, include for example currency conversion costs and exchange rate uncertainty, the latter of which, again, is not easily measurable. Again, we can see that a model is thought to explain a phenomenon despite the fact that it employs concepts—even essential concepts—the quantities represented by which are not readily measurable (a fact which in turn makes it hard to establish a law-claim that employs these concepts inductively). This fact should further confirm the suggestion made in the previous Section, *viz.* that law-claims are established by thought experiments rather than experiential methods.

There is a second with the model from the point of view of the authors. The way the model is formulated (see equ. 2 and 3), "trade costs" comprise everything that may raise the price of foreign goods. Now, it is possible, as Obstfeld and Rogoff mention, to derive the exact same results we have in this model with a different model that introduces a home bias in *preferences* ω. Equation (2.1) would then read:

$$U = \left(C_H^{\frac{\theta-1}{\theta}} + \omega C_F^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}. \tag{2.8}$$

Obstfeld and Rogoff say about this, "One can easily show that the effects of home bias in preferences ($\omega < 1$) are isomorphic to the effects of trade costs τ".[33] The situation is almost exactly equivalent to the Akerlof case above: the phenomenon can be captured both by a model that writes the solution into the agents' preferences and by one that introduces a novel concept. The authors claim that "it is more illuminating to derive trade biases from other frictions [than biased preferences]".[34] What are *their* reasons to think that this is the case?

[33] *ibid.*, p. 9
[34] *ibid.*

In the context of discussing their solution to the "home bias in equity portfolios puzzle", Obstfeld and Rogoff mention "potential explanations" other than their own trade cost explanation. These include models describing factors such as human capital, non-traded consumption goods, asymmetric information and legal restrictions.[35] After applying their own trade-cost model to this puzzle by enhancing it with uncertainty and capital markets and showing that the phenomenon can be captured with the model, Obstfeld and Rogoff summarise their view about the advantages of their approach:[36]

> As we have noted, our explanation not only has the merit of (extreme) simplicity, but it is also more convincing because the same basic approach seems to help explain such a diverse range of puzzles. Finally we note that our results are consistent with recent empirical work...

In order to formalise the idea of the unification aspect of economic models let us define the following.

(1) *To capture a phenomenon.* A model or thought experiment M^i *captures* a phenomenon P if and only if a representation of P can be demonstrated within M^i. (Above, for example, we have seen that a representation of the phenomenon that the consumption of home goods tends to be higher than that of foreign goods can be demonstrated within Obstfeld and Rogoff's model.)

(2) *Quantities and concepts.* In M^i concepts represent quantities from the set $\mathbf{Q}^{Mi}$. (For example, *consumption of home and foreign goods, trade costs* and *elasticity of demand.*)

(3) *Condition E'.* For any given set of independent phenomena $\mathbf{P}$, if the model or thought experiment $M^i \in \mathbf{M}$ (which is the set of models or thought experiments that capture $\mathbf{P}$) is a member of a set $\mathbf{K}$ of thought experiments or models such that a subset of $\mathbf{Q}^{Mi}$, the quantities represented by the concepts of M^i, for each $M^k \in \mathbf{K}$ is identical, it has explanatory power.[37]

---

[35] *ibid.*, p. 21 and 26

[36] *ibid.*, p. 26-7

[37] Two remarks about this condition. First, I take the ability to unify and systematise thinking about economic phenomena as formalised here as a *sufficient* condition rather than a definition of explanatory power. The reason is that there may be different ways for a model to have explanatory power, unification being only one of them. I will discuss an alternative in Chapter 5. Second, the formulation is slightly inaccurate in that I really believe that unification is at best an INUS (insufficient but necessary part of an unnecessary but sufficient) condition of explanatory power. The

(4) *Explanatory power.* Explanatory power is proportional to the cardinality of **K**, that is, the more models or thought experiments there are that represent the same quantities and capture a great range of independent phenomena, the more explanatory power does $M^i \in \mathbf{K}$ have.

In order to understand condition E′, let us set **P** = {Obstfeld and Rogoff's six puzzles}; M^1 = {the model that captures the first trade puzzle}, M^2 = {the model that captures the second trade puzzle}, *etc.*; **M** = {the set of all models that jointly capture **P**}; $\mathbf{Q}^{M1}$ = {*consumption of home goods, consumption of foreign goods, prices of home goods, prices of foreign goods, trade costs...*}, $\mathbf{Q}^{M2}$ = {*savings, investments, trade costs...*}, *etc.*; **K** = {M^1, M^2, ..., M^6}. In this case, then, any member of **K** has explanatory power because among the quantities represented by its concepts we find *trade costs*, which is the single member of the subset of each $\mathbf{Q}^{Mi}$ that is identical for each $M^k \in \mathbf{K}$. Explanatory power of each M^k is the greater, the more phenomena there are in **P**.

In order to make the theory fully consistent with the economic practice in so far as I have examined it, we need another condition. Although, as I have pointed out in the previous section, empirical work regarding the behaviour of the quantities represented in the thought experiments and models is comparatively rare in the process of *establishing* that some quantity is related to another in a systematic way, there is much empirical testing of implications of thought experiments and models of the kind discussed above. Hence, it is unlikely that *any* model that captures some phenomenon of interest and that uses quantities that can be employed to build other models that capture a great range of phenomena will be regarded as explanatory. If for example a quantity is used in a large set of models that capture many phenomena, but there is evidence that the quantity is not present in many real instances of the phenomenon, these models will not be regarded as having explanatory power. Thus we must amend (E′) and say that

model must have "something more" than merely the ability to unify. What this "something more" is, will be discussed in Chapter 4. Many of the articles building upon the work of the Friedman and Kitcher papers on unification are concerned with exactly this question. See for example Barnes 1992, Sabatés 1994, Morrison 2000 or Schurz and Lambert 1994. Thanks to Roman Frigg for pointing out to me the link with this literature.

(5) *Condition E″.* Condition E′ is fulfilled and there is virtually no significant evidence that in situations instantiating members of **P**, the quantities of the subset of $\mathbf{Q}^{Mi}$ that is identical across the members of **K** are not present.

This theory of economic explanation can be easily shown to agree with the cases discussed above. For example, although the "lunch table justification" captures the phenomenon of the price differential, its quantity: "the pure joy of owning a new car" is represented in only one model, *viz.* exactly this one.[38] Furthermore, the result is demonstrated only verbally. Akerlof's thought experiment and model have explanatory power as the quantities represented in them such as "asymmetric information" figure in models that can be used to capture a great range of phenomena.

The theory proposed here also suggests why the "home bias in preferences" is a worse explanation than the trade costs model. Remember that the two models are isomorphic, and note that Obstfeld and Rogoff even claim that "it is important to recognize that a home bias in demand for goods can serve much the same function as trade costs in our analysis throughout this paper". That is, they suggest that the two models certainly agree on the first condition (E′). However, the alternative fares worse with respect to the known evidence: "Helpman (1999) argues that once one controls for income, there is no clear evidence of home bias in preferences…". That is, their model but not the alternative is consistent with available evidence, and thus the trade costs model is the better explanation.

4 Conclusions

Francis Bacon laid out his philosophy of the interpretation of nature as an alternative to at least two prevailing schools, the rational and the empirical.[39] This Chapter was

[38] We may of course write down any number of models using this quantity but they, presumably, would not capture independent phenomena.

[39] Bacon mentions also a third school, the "superstitious", which "out of faith and piety mix theology and tradition with their philosophy; among these, the vanity of some has led them astray to look for and derive science from spirits and supernatural beings" (NO I 62). For obvious reasons, this school is

and the next Chapter will be concerned with a tradition of concept formation in economics that shares important characteristics with Bacon's rationalists.

According to Bacon, rationalist philosophers (in particular, Aristotle) subject their scientific investigation in an exaggerated way to formal or logical considerations (NO I 54), examine only a small number of particulars if any (NO I 25) and jump from claims about these to claims of highest generality (NO I 19). The tradition I have been trying to sketch here is similar in so far as 1) concerns of primarily formal appeal such as consistency, simplicity, explicitness and mathematical expressibility are regarded as being of great weight, 2) empirical investigation is not suppressed altogether but its function is more appropriately rendered as one of motivation illustration and testing rather than one of a experiential fundament from which the investigation has to start, and 3) from formalising some illustrative cases principles of great generality are derived.

One point of criticism Bacon made when discussing what he called the rationalist philosophers was that the method of scientific investigation he ascribed to them—the "anticipation of nature"—would not yield concepts that categorise things according to the natural classes into which they fall. That is, concepts formed by the method of anticipating nature pick out features of things that do not have much in common. I have tried to show in the first part of this Chapter that a crucial difference between the Bacon-Schmoller method of concept formation and that of modern theoretical economics as analysed in this Chapter is the use of real experiments, observations and measurements in the one tradition, and that of thought experiments in the other. Since the "phenomena" of the latter tradition are not investigated empirically in any systematic way, its models cannot be phenomenally adequate in the stronger sense, *i.e.* adequate to a *real* phenomenon.

However, they have the virtue of being explanatory in the sense that they help systematising and organising our thinking about economic phenomena (in the weaker sense of casually observed stable features of the economic world).

less relevant in a modern context, and hence I shall omit it here. For Bacon's own words, see for example *Novum Organum* Part I (esp. aphorisms 61-5 and 95, which is the motto of Chapter 1). For secondary literature, see various essays in Peltonen 1996b and Urbach 1987, esp. ch. 4 on the idols and ch. 2 on the interpretation of nature.

The next Chapter continues the examination of the "rationalist" tradition in economics and in what way its models may or may not achieve the remaining epistemic virtue of exactness.

Chapter 3

Concepts Functionally Defined: Exactness Versus Explanatory Power

Chapter 3

Concepts Functionally Defined: Exactness Versus Explanatory Power

Our discussion will be adequate if it has as much clearness as the subject-matter admits of, for precision is not to be sought for alike in all discussions, any more than in all the products of the crafts.
[... F]or it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician scientific proofs.

Aristotle—Metaphysics K3

1 Introduction

Chapter 1 defined concept formation as a method or process by which the rules of application of scientific concepts are determined. Chapter 2 discussed what admissible abstract concepts in theoretical economics are. The aim of this Chapter is to shed light on the question of how concepts that are abstract in the sense of the previous Chapter may be thought to *apply* to objects, processes or systems.[1] In terms of the three questions regarding abstraction that were raised in that Chapter (concerning the *kind* of abstract concept, the *process* of abstraction and the *relation* between the abstract and concrete), the interest of this Chapter falls in the domain of the third question.

An issue such as the application of abstract concepts to objects, processes or systems is most fruitfully discussed in the light of the particular cognitive aims or epistemic virtues that scientists want their concepts to achieve. We have identified phenomenal adequacy, explanatory power and exactness as important epistemic virtues endorsed by economists. In this Chapter we shall therefore investigate a number of theories of concept application in their relation to these epistemic virtues.

[1] I will keep referring to "objects, processes or systems" in order to express a metaphysical neutrality between substantivalist ontologies such as David Armstrong's, process ontologies such as Whitehead's and ontologies that understand the world as made up of "entities and their activities" such as Machamer, Darden and Craver 2000.

There are two main contenders for a theory of the application of concepts in theoretical science: operationalism and functionalism.[2] According to operationalism, the rules of application are fixed by the operations that measure the quantities that are represented by concepts. The answer to the question of how the concept of length applies, for example, is given by the operation of aligning a measuring rod to an elongated object in a specified way and counting the number of basic units from one end of the object to the other.

Functionalism, by contrast, regards the application rules for concepts as being given by the concepts' role or function (hence functionalism) in a system of claims. In science, this usually means "law-claims", and these, in turn, may be given by a system of axioms. According to functionalism, for example, the rules for applying the concept of point are given by the fundamental principles of Hilbert's axiomatisation of geometry.

In this Chapter we will argue that those concepts in theoretical economics that are abstract in the sense discussed in the previous Chapter are best regarded as being characterised functionally. Subsequently we will distinguish two variants of functionalism and claim that according to one position we can get exact concepts that lack explanatory power, and according to the other, we get explanatory concepts that are inexact.

2 Abstract Economic Concepts and their Meaning

We continue the discussion of those economic concepts that are formed by a process that combines intuition and casual observation in such a way that a thought experiment fitted with these concepts yields a representation of the phenomenon of interest. We will regard any economic concept referred to in this Chapter as abstract in precisely this sense.

In order to get an understanding of how abstract economic concepts can be thought to apply to objects, processes or systems, it may be useful to analyse a set of

[2] For this distinction and a criticism of both positions with regard to the concept of mass in the

examples that we regard as paradigmatic. Compare the following definitions, each taken from a standard economics textbook.

1. *Moral Hazard* "is the form of postcontractual opportunism that arises because actions that have efficiency consequences are not freely observable and so the person taking them may choose to pursue his or her private interests at others' expense", Milgrom and Roberts 1992, p. 167.

2. By *efficient choices* or *options* "we mean ones for which *there is no available alternative that is universally preferred in terms of the goals and preferences of the people involved.* More precisely, if individuals are sometimes indifferent about some of the available options, then a choice is efficient if there is no other available option that everyone in the relevant group likes at least as much and at least one person strictly prefers", Milgrom and Roberts 1992, p. 22 (original emphasis).

3. *Postcontractual opportunism* "arises because contracts are incomplete and imperfectly specified, so that the parties to the contract can exploit loopholes to gain an advantage over one another", Milgrom and Roberts 1992, p. 137-8.

4. An "*externality* is present whenever the well-being of a consumer or the production possibilities of a firm are directly affected by the actions of another agent in the economy", Mas-Colell, Whinston and Green 1995, p. 352 (original emphasis).

5. Situations of *asymmetric information* are "situations where one economic agent knows something that another economic agent doesn't", Varian 1992, p. 440.

6. *Players* "are the individuals who make decisions. Each player's goal is to maximise his utility by choice of actions", Rasmusen 1994, p. 10.

In all these cases we see that the abstract definiendum-concept is characterised in terms of other abstract concepts. Moral hazard, for example, is characterised (among other things) in terms of postcontractual opportunism, efficiency and observability; efficiency in terms of preferences, and preferences, in turn, are often characterised

Schwarzschild solution in relativity theory, see Bartels 1990.

by a set of mathematical properties of a relation.[3] It seems, thus, that at least some of the abstract concepts of theoretical economics are characterised by their relations to other abstract concepts. An unambiguous example is the following:

1. *Market power* is "the ability to alter profitably prices away from competitive levels", Mas-Colell, Whinston and Green 1995, p. 383, and
2. In a *competitive economy* "a market exists for each of the L goods, and all consumers and producers act as price takers. The idea behind the price-taking assumptions is that if consumers and producers are small relative to the size of the market, they will regard market prices as unaffected by their own actions", Mas-Colell, Whinston and Green 1995, p. 314.

In these passages market power is characterised in terms of competition, and competition in terms of price-taking behaviour, that is, the absence of market power.

The textbook examples suggest that *operational* definitions do not play a prominent role in theoretical economics. This image is confirmed by the investigation mentioned in Chapter 2 which found that less that ten per cent of the work on asymmetric information had any empirical content at all, and in no case was "asymmetric information" actually operationalised (*i.e.*, there were no descriptions of procedures which aim at measuring the quantity represented by the concept *asymmetric information* in the articles surveyed). It is further confirmed by that Chapter's observation that sometimes even essential concepts used in theoretical models cannot be operationalised. I am far from wanting to assert, however, that it in all cases is not *possible* to operationalise at least some of the abstract concepts. At most I can claim that I have found little endeavour to do so. For the purpose of this Chapter, then, I will focus on theories that regard application rules for concepts as given by the role they play within some kind of theoretical framework. The next Chapter will deal with a different kind of tradition within economics, a tradition in which concepts are more naturally thought to be operationalisable and sometimes operationalised.

One theory which regards application rules being given by the theoretical framework in which they appear is functionalism. Although according to Willard van Quine[4],

[3] See Varian 1992, ch. 7 or Mas-Colell, Whinston and Green 1995, ch. 1 for an introduction. I have not included the characterisation of preferences or the preference relation merely because of its

functionalism dates back to at least 1818, the most famous application is David Hilbert's *Foundations of Geometry*[5]. Hilbert asked the question, "How can one introduce the fundamental concepts of a theory in a way that guarantees that the axioms of the theory are sound?" He answers: by postulating that the concepts are *defined* in a way that they satisfy the axioms.[6] (I take a definition of a concept to be an expression that demarcates the objects or processes or systems to which it applies from those to which it does not apply. The demarcation may involve actual as well as fictional objects. A *brother*, defined as "male sibling" applies to all people who are male and siblings and not to any other people. In this example at least some of the objects to which the concept applies are actual. A *mermaid*, defined as a "being, half woman-half fish" applies to all objects which are half woman and half fish. This, presumably, includes only fictional objects.[7])

In the second of Hilbert's "Mathematical Problems"[8], he summarises his idea as follows: "When we are engaged in investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the relation subsisting between the elementary ideas of that science. The axioms set up are at the same time the definitions of those elementary ideas..."[9] Definitions thus understood have become to be called "implicit", "functional" or "contextual" definitions, or definitions "by postulation".[10]

Moritz Schlick extends this idea in his *General Theory of Knowledge*[11] to make it applicable to all sciences, not just mathematics. The general problem Schlick sought to solve was "How is objective science possible?". Gaining "objective knowledge" for Schlick in this context means not only assuming the point of "view from nowhere", that is, independent of any personal tastes or interests of the knower. It also implied for him that knowledge be absolutely exact and definite.[12]

---

clumsiness.

[4] Quine 1964, p. 71

[5] Hilbert 1959/1899

[6] This reading follows Schlick 1925, part I, §7.

[7] I owe a lot to Sorensen's 1991 distinction between intrinsic and extrinsic vagueness and their characterisations in my thinking about definitions, rules of application, exactness *etc.* The two examples are also due to Sorensen.

[8] Hilbert 1902/1900

[9] quoted from Shapiro 2000, pp. 152f.

[10] See Shapiro 2000, p. 155, Brown 1999, ch. 7, Schlick 1925, part I, §7.

[11] Schlick 1925

[12] See for instance *ibid.*, part I, §5.

According to Schlick, knowledge *by concepts* is able to achieve this goal. However, not every theory of concept formation will be equally suitable. *One* alternative is concept formation by "concrete definition"[13]. Schlick discusses the advantages of "knowledge by concepts concretely defined" by means of an example:[14]

> If one hands me a piece of metal, I will not be able to recognise whether it is pure silver or not as long as I am dependent on perceptions that I gain from merely seeing or touching it. For the memories of the representations that I have of the silver are not exact enough in order to be distinguished from the representations of similar metals such as tin or specific alloys. The situation is entirely different if I make use of the scientific *concept* of silver. In this case it is defined as an element with a specific weight 10.5, an atomic weight 108, of a specific electric ductility *etc.* and I only need to test whether the metal at hand has these properties in order to decide with all desirable precision whether I have silver in front of me or another substance.

Schlick, however, quickly notices that this solution is only a temporary one. For testing the metal for the properties that form the content of the scientific concept depends ultimately on sense impressions such as reading a scale or a galvanometer. Hence inexactness sneaks in again, though, as Schlick remarks, "that difficulty can be moved to that area where error can be excluded with a certainty which is sufficient for all purposes of special science".[15]

It is important to notice that uncertainty is eliminated only for the *practical* purposes of special science, not in principle: "[Forming concepts by concrete definition] the eventual retreat to the immediately given is inevitable, and therefore the formation of absolutely exact concepts seems altogether impossible"[16].

Schlick's solution to the problem of objectivity and exactness is to adopt Hilbert's theory of implicit definitions and extend its use to all science. Schlick thought that one could eliminate partiality and vagueness by postulating that concepts are *defined* by the fundamental axioms of the formal system in which they figure. The subjective and vague element of concrete definitions could be avoided exactly because there are

[13] *ibid.*, §5. Although Schlick discusses this approach to concept formation in the whole section, he introduces the term only on p. 28.
[14] *ibid.*, p. 25. Translated from the German by the author. Emphasis is original.
[15] *ibid.*, p. 26
[16] *ibid.*, p. 27

no connections between concepts and perception: "Implicitly defined concepts are rigorous because they are entirely detached from the given intuition".[17]

Let us call this position "axiomatic functionalism". It says that the fundamental axioms *define* the concepts of a deductive system. That is, the objects, processes or systems to which a concept applies are delineated by the relations to the other concepts of the axioms. The set of axioms, in turn, applies to all domains in which the axioms are true.[18] A point, for instance, is any object that bears the same relations to the other objects described by Hilbert's axiomatisation of geometry in any domain where these axioms are true. Let us contrast Schlick's 1925 theory of concept formation with a related but different position that we may call "nomic functionalism".

Norman Robert Campbell, in his *Physics The Elements*, argues that the application rules of physical concepts depend on truths about scientific laws:[19]

> Thus our first conclusion is that many of the words used in expressing scientific laws denote ideas which depend for their significance on the truth of certain other laws and would lose all meaning if those laws were not true. [...] It will be convenient to have a name for such words and they will in future be called "concepts." A concept is a word denoting an idea which depends for its meaning or significance on the truth of some law.

Claims about laws are propositions that state relations between concepts. The meaning of concepts depends on the truth of a nomic claim. These two suggestions, imply a holism about nomic systems: all (or most) nomic claims are interconnected, and changing parts of the system in the light of new experimental findings will

---

[17] Coffa 1991, p. 176. "Intuition" here appears in the Kantian sense of being the faculty of the mind responsible for representations or imaginations, or which is capable of being affected by the external world.

[18] This position may appear to be a non-starter. It holds that axioms or assumptions (see below) apply to those and only those domains in which they are true. But that obviously does not solve the problem of application since we usually don't readily know how to interpret axioms or assumptions empirically. However, here "application" may include application to abstract or fictitious objects, processes or systems. The real issue for the purpose of the arguments presented in this Chapter is which theories of application realise which epistemic virtues. Through this issue the problem of application—more narrowly defined as the problem of how to apply axioms or assumptions to *empirical* objects, processes or systems—re-enters as the problem of finding a model which is phenomenally adequate. Looking at the problem from this point of view allows us to find positions that result in models that may realise some epistemic virtues at the expense of others—as we shall see below.

[19] Campbell 1957/1922, p. 45. Recall from Chapter 2 that Campbell understands laws as claims rather than as patterns in the world.

propagate through the whole system. One example Campbell uses to illustrate this idea concerns Hooke's law:[20]

> Thus we saw that Hooke's Law, because it involves the concept *force*, depends for its significance on the laws of dynamics, and in particular on the law of the equality of action and reaction. Now this law is asserted to be true for all forms of force and applies therefore, let us say, to the force which we call the pressure of radiation; in its application to radiation the law involves again the laws of radiation and, though them, the laws of electrodynamics.

Like Schlick, Campbell also uses silver as an example. He suggests that if the truth of all nomic claims describing silver was known, statements associating silver with its properties would be analytic. Since the *concept* of silver reflects the "truth" that any lump of stuff to which it applies melts at 960° and that its density is 10.5 the nomic propositions claiming that "silver melts at 960°" and "the density of silver is 10.5" are analytic.[21] The only experimental truth left is whether or not silver, characterised in this way, does actually exist.

Although there are obvious similarities between Schlick's conception of an implicit or functional definition and Campbell's views, there are two important differences. First, Campbell rejects the view that any set of law-claims describing the behaviour of the referent of some concept *defines* that concept. He argues that no sustainable distinction between "defining" and "non-defining" properties can be made, but also that the absence of a sharp distinction does not matter. This is because there is a fundamental difference between logic and science. Logic is the study of accurate and systematic thought, and it depends for its power on the sharp definition of words. Science, by contrast, is in large part illogical, and thus we would "be led into nothing but error if we try to force scientific reasoning into the forms prescribed by logical canons".[22] The contrast between Campbell and Schlick could not be greater than when Campbell says,[23]

> Our words then are not instruments by means of which the process of thought is conducted, but merely convenient means of recalling to our minds thoughts which have once passed through them or of calling up in the minds of others thoughts which are passing through our own.

The second and related difference is that Campbell is not worried by the following kind of inexactness and openness implied by his account of concept formation. For

[20] *ibid.*, p. 50, emphasis added
[21] *ibid.*, p. 54
[22] *ibid.*, p. 52

Campbell the application rules for concepts are given by the law-statements in which they figure. However, the fact that the set of law-statements in which the concept *silver* plays a part may be different for different people or at different times is no predicament for the proper use of the concept.[24] The important fact, though, is that there is a set of law-statements which as a whole determines the application rules for the concepts which figure in it:[25]

> A law is a single whole, or at least, if it is capable of analysis, the parts into which it can be analysed are not those into which it can be divided grammatically. These considerations apply to "force" as much as to "silver." Because we state that the force on a certain body is 1 dyne or that the extension is proportional to force, it does not follow that we can state significantly that force "is" something or other. Though laws state relations between concepts, the significance of those concepts can hardly be separated from that of the laws they are used to state.

We can thus characterise *nomic functionalism* as follows. The rules of application of scientific concepts are given by the nomic claims in which they occur. The concepts apply to all objects, processes or systems of which these claims are true. The difference between axiomatic and nomic functionalism is one of degree rather than one of kind. If a certain physical domain is closed in the sense of being exclusively and exhaustively described by a set of nomic claims, and this set is completely axiomatised, they are identical with respect to the concepts applicable to this domain. But as we shall see shortly, "complete" axiomatisation of a set of claims involves strict formal requirements which may not always be satisfied (and even satisfi*able*) by a set of nomic claims even though the latter may be true, or true for the most part, or approximately true for certain objects, processes or systems. Furthermore, a set of nomic claims may be open in a sense that an axiomatic system cannot be open. Adding a new nomic claim to an existing set may expand our picture of the natural world. An example would be that we experimentally establish a new property of silver which has hitherto not been known. By contrast, adding a new axiom to a complete axiomatised system will either be superfluous or render the system inconsistent.[26] The doctrine that deals with concepts whose rules of application are

[23] *ibid.*

[24] For Campbell laws are propositions rather than patterns of behaviour in the real world. Most conceptions of law regard it as a condition of lawfulness that the law-proposition must be true. But even if we grant this, it is possible that different people know different subsets of all the true nomic facts about an object, process or system, and that in the course of time people learn more nomic facts.

[25] *ibid.*, p. 55

[26] The difficulties associated with Gödel's incompleteness results are obviously ignored here.

characterised by a system of the latter kind I call "axiomatic functionalism", and the doctrine that deals with concepts whose rules of application are characterised by a set of law-claims of the former kind, "nomic functionalism".

Let us now turn to economics. We need one more preparatory step for our analysis of abstract economic concepts. The reason is that we need to mould the vocabulary of the two approaches into a language that is applicable to economics. For *prima facie* there is only a small amount of scientific work in economics which resembles the formal systems of mathematical logic, and only very few economic relationships carry the name "law".

Take axiomatisation first. It is a truism that the use of mathematics in economics has experienced a great rise during the past century or so. Roger Backhouse[27] estimates that the proportion of articles in the *Economic Journal* and the *American Economic Review* that use algebra has increased from around 10% in 1930 to around 75% in 1980. But mathematisation is not always the same as axiomatisation. According to Backhouse, axiomatisation "involves reducing a body of knowledge to a set of independent axioms, with all propositions being derived from those axioms using well-defined logical rules".[28] Standard logic texts would normally add that an axiomatised system requires, besides the two elements Backhouse mentions (axioms and inference rules), a number of further elements, *viz.* primitive symbols, definitions, and rules for the formation of correct expressions.[29] The advantages of axiomatisation (among others, explicitness, generality, objectivity, minimalism[30]) are most pertinent if a system is not only axiomatised, but if its axioms are simple and not too great in number, consistent (that is, it is impossible to deduce a contradiction from the axioms), complete (that is, every truth of the system is deducible from the axioms) and independent (that is, no axiom is a logical consequence of the others).[31]

Even if it is clear that even formal systems cannot realise the above mentioned virtues all at once, economic models have a long way to go to come anywhere near the ideals of formal logic. Just look at the following—I think not untypical—

[27] Backhouse 1998
[28] *ibid.*, p. 1848
[29] See for instance Stigum 1990, ch. 2 or Machover 1996 passim.
[30] See Suppes 1968.
[31] See for example Courant, Robbins and Stewart 1996, p. 215.

example. In a model of local governance (from Tiebout 1972), we have, among others, the following "axioms":[32]

> (A1) Consumer-voters living under *local governments* are fully mobile and move to the community whose tax and expenditure package best satisfies their preferences for local pure public goods.
>
> (A2) Consumer-voters have full knowledge of the tax and expenditure packages of the different local governments.
>
> ...
>
> (A4) The source of consumer-voter income provides no obstacle to their mobility—for example they might derive all their income from dividends on common stock.
>
> ...

In this example, depending on the sense of "fully mobile", (A1) may imply both (A2) and (A4). For if informational asymmetries and sources of income provide obstacles to mobility, consumer-voters are not "fully mobile". We may of course understand the "fully mobile" either as "fully mobile except for possible informational and income-related obstacles" or as "fully mobile in the everyday sense of the word" (*e.g.*, "they have cars"), which would make the axioms independent, but then one would equivocate on two senses of "mobility" in (A1) and (A4)). Because in many cases, economic models use a natural rather than formal language ambiguities such as this are almost unavoidable. Hence we cannot speak of economic models as formal systems.

To cut a long story short, economic models, usually, are not axiomatised but they make informal use of mathematical techniques.[33] This view is shared by Backhouse who notes that:[34]

> More general [than axiomatisation] is mathematisation: the use of mathematical techniques (geometry, algebra, set theory, topology) in economic arguments. This definition is informal, but it has the advantage of corresponding exactly to what people have in mind when they talk of the mathematisation of economics in the last fifty to sixty years.

Therefore, I think it is appropriate to modify the characterisation of axiomatic functionalism and define: *axiomatic economic functionalism* regards economic

[32] This model is discussed by Steven Rappaport in his 2001. See pp. 280f., original emphasis.
[33] For this distinction, see also Chick 1998, p. 1861, and Backhouse 1998, p. 1848.
[34] *ibid*,

concepts as being defined by the assumptions (informal "axioms") of an economic model.[35]

A fundamental concept of the second approach was that of a scientific law. There are indeed some regularities in economics that carry the name of a law. Charles Kindleberger, in his *Economic Laws and Economic History*[36], mentions four: Engel's Law, The Iron Law of Wages, Gresham's Law and The Law of One Price. We might add Say's Law, the law of comparative advantage, and maybe the law of the quantity theory of money.

However, the unit of analysis for both economists as well as philosophers of economics is usually the model rather than the law. In the usage of this Thesis (and contrary to Campbell's usage) a law is a pattern of behaviour in the real world.[37] This implies that a model—an abstract or linguistic entity—can be thought of as (at best) *representing* a law or laws. A law, on the other hand, in economics can be thought of as a relation between economic quantities. For example, Akerlof's model from the previous Chapter may be said to represent the law that asymmetric information causes adverse selection; Hotelling's model, which I will discuss in a moment, may be said to represent the law that transportation costs cause minimal product differentiation.

Therefore, the definition of nomic functionalism should be modified in order to capture this particularity. I shall thus say that according to *nomic economic functionalism* concepts apply to those systems of economic quantities of which the nomic claims expressed by economic models[38], are true.

Let us now turn to the analysis of the two theories of concept application in the light of epistemic virtues.

---

[35] We may add here, "... or the assumptions of an economic theory" in order to capture systems such as revealed preference theory, which—arguably—is not a model. The definitions presented here are supposed to apply to cases such as revealed preference theory too.
[36] Kindleberger 1989
[37] "Pattern of behaviour" is supposed to capture anything from regularity, *ceteris paribus* regularity, tendency, capacity, relation among universals *etc*. The emphasis is on the "in the real world".
[38] The same caveat applies again in this definition: economic models *or theories*.

3 Functionalism and Explanation

What, then, do the two different approaches imply for the epistemic virtues of phenomenal adequacy, explanatory power and exactness? The advantage of axiomatic functionalism is that concepts understood in this way are very exact and unambiguous. *Asymmetric information* in the automobile market, for example, refers to assumption (9) of Akerlof's model as discussed in the last Chapter and means, "The average quality of cars μ is known to all participants but the individual car's quality x_i is known only by its owner". There is no ambiguity in this claim.[39]

Axiomatic functionalism implies that each model defines its own concepts—a feature that it shares with definitional operationalism as we shall see in the next Chapter. A standard objection to definitional operationalism—the proliferation of concepts—therefore applies to axiomatic functionalism as well. Each economic model defines its own conceptual framework and the concepts from different models are mutually incommensurable.

If the unification account of explanation given above is correct, functionally defined concepts are not useful for explanation. This is because a main feature of the account of explanation given was that explanatory power of a concept is proportional to its ability to unify and systematise our thinking about economic phenomena. But if each model defines its own concepts, its concepts can figure in only one model, which can at best capture one phenomenon(-type). Axiomatically functionally defined concepts have little explanatory power.

But even if one did not follow the unificatory account of explanation, functionally defined concepts can do little to explain features of the actual world. This is because

[39] There is no ambiguity if it is possible to tell, for any given domain, whether the assumptions are true of it or not. But there is a particularity of economic concepts which makes it easier to find rules of how to *construct* domains in which assumptions are true. Many economic concepts are homonymous with everyday concepts—take Akerlof's *cars*, their *quality*, and *traders* as examples. Now a rule for the construction of a system in which Akerlof's assumptions are true could read as follows. Take *actual* exemplars of the objects Akerlof is talking about, say actual cars which have a good or bad quality and actual traders. Now transform these objects such that they satisfy Akerlof's assumptions. That is, remove all the cars' properties except their *quality*, and arrange the quality distribution such that it is uniform between zero and two. Take actual traders and suppress all their motives of action except that of expected utility maximisation, and arrange the latter such that only the cars' quality matters in a way prescribed by the assumption. Not all of these transformation procedures will be feasible in the actual world. But still it seems that there is a relatively clear rule of what the result must look like in order to satisfy the assumptions. This is why I would say that these concepts are *exact*. But this kind of exactness may be bought at the expense that only fictional objects, processes or systems are described by the assumptions, which is the case when at least some of the transformation procedures are not feasible in the actual world.

of a trade-off that was very clearly seen by Schlick in his original account of axiomatic functionalism. Schlick notes that we have the choice either of understanding the concepts being characterised merely formally, by the axioms of a theory, or of understanding them to refer to real things (real cars in Akerlof's example, say). But if we do the latter, exactness of the concepts is not guaranteed:[40]

> In such an abstract science as number theory, for example, we erect the edifice for the sake of the pleasure obtained from the play of concepts. But in geometry, and even more in the empirical sciences, the motive for putting together the network of concepts is above all our interest in certain intuitive [as above, "intuitive" is to be understood in the Kantian sense] or real objects. Here the interest attaches not so much to the abstract interconnections as to the examples that run parallel to the conceptual relations. In general, we concern ourselves with the abstract only in order to apply it to the intuitive. But — and it is to this point that our consideration returns again and again — the moment we carry over a conceptual relation to intuitive examples, we are no longer assured of complete rigor. When real objects are given to us, how can we know with absolute certainty that they stand in just the relations to one another that are laid down in the postulates through which we are able to define the concepts?

And Schlick concludes about this matter:[41]

> It is therefore all the more important that in implicit definition we have found an instrument that enables us to determine concepts completely and thus to attain strict precision in thinking. To achieve this end, however, we have had to effect a radical separation between concept and intuition, thought and reality. While we do relate the two spheres to one another, they seem not to be joined together at all. The bridges between them are down.[42]

Thus we can have concepts whose meaning is entirely precise and objective but it lacks connection to reality. In so far as explanation is thought to concern reality (and I will defend that it should in Chapter 4), implicitly or functionally defined concepts have no explanatory power.

As one might expect, matters with nomic functionalism are just the reverse. Following nomic functionalism, concepts may help in building explanatory power, because it is possible that they figure in many models that capture a variety of independent phenomena. But are they exact?

It is not trivial to say what exactly is meant by "exactness". Above, we used it in a sense of axiomatic exactness, and thus the fact that implicitly defined concepts are

[40] Schlick 1925, §7
[41] *ibid.*

exact amounts to no more than a tautology. Exactness is never an end in itself. Schlick, for example, as we have seen, was aiming at a kind of objectivity, which was not to be had unless one has got rid of all intuitive elements of his theory of concept formation, that is, all elements that concern the "immediately given" or direct conscious experience.

I would thus argue that there is a hierarchy in the epistemic virtues discussed here. Exactness is secondary; it helps to promote the primary virtues of phenomenal adequacy and explanatory power. I have two reasons to believe in this hierarchy. One is that with Aristotle I believe that one should never be more exact than the subject matter at hand allows. Conceptual exactness is sometimes bought at the cost of being only an exactness of fictional objects, processes or systems. The second reason is that exactness is frequently thought to be a virtue that is supposed to promote others.

Consider the two senses of exactness that we find in the works of William Stanley Jevons and those of Carl Menger. Jevons contrasts mathematical and exact science. Any science whose characteristic properties admit of degrees is a mathematical science:[43] "To me it seems that *our science must be mathematical, simply because it deals with quantities*. Wherever the things treated are capable of being *greater or less*, there the laws and relations must be mathematical in nature". An exact science, by contrast, is one whose quantities are measurable with a high degree of precision. Many people are notoriously wary of the mathematical method in political economy. This is because, according to Jevons, they mistake a mathematical science for an exact science. But even in the absence of precise measurement, the quantities characteristic of a science must be represented mathematically.

For Jevons, absolute exactness is an illusion. All mathematical sciences are only *relatively* exact, and they are ordered by the precision with which their quantities can be measured. Thus, astronomy is a more exact science than political economy, but both are mathematical sciences.[44] According to Jevons, then, exactness is a function of precision in measurement. I take it for granted that precise measurement is not an

[42] Schlick of course implies in this passage that it is never the case that all transformation procedures of the kind sketched above are feasible in the actual world.
[43] Jevons 1871, p. 3. Original emphasis.
[44] *Cf.* Jevons 1871, pp. 3-16 and Jevons 1874, ch. 21.

end in itself but serves to describe phenomena (which ultimately provides the basis for a causal explanation of them).[45]

Contrast Jevons's ideas about exactness with Menger's—laid out in the same year. Menger distinguishes two orientations of theoretical science, *viz.* the exact and the empirical-realistic branches.[46] Both are concerned with the general as opposed to the particular aspect of phenomena. The method of the latter is induction: by classifying phenomena according to similarities and recording regularities of co-occurrence and succession, the investigator establishes claims about empirical laws. The method of the former is deduction: phenomena are broken into their simplest parts such that exact identity between the tokens of the same phenomenon-type can be warranted. The laws of thought allow us to infer the exact laws governing the behaviour of exact types. They hold by necessity and are thus exceptionless. Exactness for Menger means (a) strict identity (as opposed to mere similarity) of the tokens of a phenomenon-type and (b) necessary and exceptionless co-occurrence or succession of phenomenon-types in laws.

Although Menger does not explicitly endorse this view I am presenting momentarily, I think the best available defence of the distinction Menger draws between the two orientations of theoretical science is along the following lines. Nomic claims according to the exact orientation carry additional epistemic virtues. Explaining a phenomenon, for Menger, means to subsume it under a strict regularity. Since an empirical law holds, at best, "for the most part" (it is not exceptionless), explanatory power is impaired, at least according to some views of explanation, including Menger's. Consider one of Nancy Cartwright's nice examples:[47]

> Sometimes super laws [*i.e.*, empirical regularities], even when they are available to cover a case, may not be very explanatory. This is an old complaint against the cover-law model of explanation: 'Why does the quail in the garden bob its head up and down in that funny way whenever it walks?' ... 'Because they all do.'

Cartwright's point is that mere regularities are often considered not to be explanatory. And this is *a fortiori* the case for regularities which hold only for the

[45] *Cf.* Jevons 1874, ch. 13
[46] Menger 1976/1871
[47] Cartwright 1983, p. 70

most part. For Menger (as for a number of recent accounts[48]), the additional bit a regularity needs in order to be explanatory is that of necessity. A phenomenon is considered to be explained only when it is subsumed under a regularity that cannot fail to hold. Again, we see that exactness is subordinate to another epistemic virtue, here explanatory power. We thus may demand that any sense of "exact" as a property of concepts must help to promote either phenomenal adequacy or explanatory power or both.

Inexactness of the concepts that are nomicly functionally characterised according to the view defended here arises because model results are rarely very stable across different sets of assumptions. It is a truism that most law-statements, physical and social alike, have a *ceteris paribus* clause in front, which says that the law holds only under certain conditions. Because of the way models are often built in economics, these *ceteris paribus* clauses tend to be very restrictive.[49] Let us look at a number of examples to illustrate this claim.

The first example, which I am going to discuss in greater detail than the other examples, is a version of Hotelling's famous model of the spatial duopoly.[50] Hotelling's problem was to determine location and supply functions of two duopolists that compete for business along a line segment. An example for an application would be two competing filling stations or restaurants along a limited segment of a motorway. The version presented here follows the analysis of the Hotelling model that I was taught in my postgraduate microeconomics course.[51]

The two duopolists' cost function is given by

$$C^i(q_i) = c_i q_i + f_i \quad i = A,B, \tag{3.1}$$

where $c_i > 0$ denotes marginal costs, $f_i \geq 0$ fixed costs and q_i the quantity. Transport costs (which the consumer has to bear) are linear:

$$T(\delta) = t\delta, \tag{3.2}$$

[48] See in particular Dretske 1977 who stresses this point. Armstrong's 1978 and Tooley's 1978 views are similar to Dretske's.
[49] For a defence of this view, see Pemberton forthcoming.
[50] Hotelling 1990/1929.
[51] Guyer 1994, ch. 22. I follow Guyer's treatment exactly with the exception of some small notational changes.

with δ denoting the distance from the supplier and $t > 0$ is a parameter. Every consumer demands exactly one unit of the homogeneous good:

$$Q^k = 1, \tag{3.3}$$

for all K consumers ($k = 1, 2, \ldots K$). The two duopolists play a dynamic game which consists of two subgames. In the first phase, they determine their locations on a line segment with fixed length L, in the second phase they determine their supply prices (quantities are then given by the demand function). The optimal strategies are derived by backward induction, that is, first the pricing strategies are determined (given the duopolists' respective location), and then given their pricing strategies, locations are determined.

If L denotes the total length of the segment, a the distance of duopolist A from the left, b the distance of duopolist B from the right, x_i the subsegment between a and b served by A, and x_j the subsegement served by B (*cf.* Figure 3.1), then

$$L = a + b + x_i + x_j. \tag{3.4}$$

The division of the subsegment between a and b is dependent on the equilibrium condition

$$p_i + tx_i = p_j + tx_j, \tag{3.5}$$

where p_i is the price A charges, and p_j the price B charges. This equation defines the location of the marginal consumer who is indifferent between the two suppliers. Solving (3.4) for x_j and substituting into (3.5) yields:

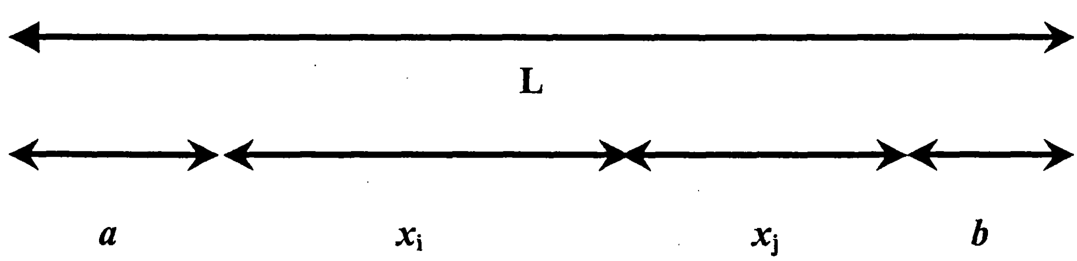

Figure 3.1 Distances in Hotelling's Model

$$x_i = X^i(p_i, p_j) = \frac{L-a-b}{2} + \frac{p_j - p_i}{2t}, \tag{3.6}$$

for $i = A, B, j = A, B$ and $i \neq j$.

Consumers are equally distributed along L. This implies (with (3.3)) the market share of duopolist A:

$$q_A = Q^A(p_A, p_B) = X^A(p_A, p_B) + a \tag{3.7}$$

(and analogously for q_B). The profit is defined as turnover minus cost:

$$\Pi^i = (p_i - c_i)Q^i - f_i. \tag{3.8}$$

The first order optimality conditions are accordingly:

$$\frac{\partial \Pi^i}{\partial p_i} = q_i - \frac{p_i - c_i}{2t} = 0, \tag{3.9}$$

and solving for p_i yields:

$$p_i^* = c_i + 2tq_i. \tag{3.10}$$

The reaction function p_A^* is calculated by substituting for q_i from (3.6) and (3.7):

$$p_A^* = R^A(p_B) = \frac{c_A t(L-a-b) + p_B + 2ta}{2}. \tag{3.11}$$

A (Bertrand-Nash) equilibrium is defined as $p_i^* = R^i(p_j^*)$. Substituting accordingly yields:

$$\tilde{p}_A = \frac{t(3L+a-b) + 2c_A + c_B}{3}. \tag{3.12}$$

Supplied quantities are given by (3.6) and (3.7):

$$\tilde{q}_A = \frac{3L+a-b}{6} \quad \text{and} \quad \tilde{q}_B = \frac{3L+b-a}{6}, \tag{3.13}$$

whereas the profits are given by (for $c_A = c_B$):

$$\tilde{\pi}_A = \frac{t(3L+a-b)^2}{18} - f_A \quad \text{and} \quad \tilde{\pi}_B = \frac{t(3L+b-a)^2}{18} - f_B. \tag{3.14}$$

The interesting part now follows. I have said above that the solution to the game is determined by backward induction, by determining the pricing strategy given the location first and then determining the location. In his original article, Hotelling derives from profit formulæ that are very similar to (3.14) what has been termed his agglomeration theorem. Because the profits of both producers increase with the distance to the end of the line segment, Hotelling has argued that producers will tend to move closer together. Since he regarded the distance only as a "figurative term for a great congeries of qualities"[52], he concluded that there is "an undue tendency for competitors to imitate each other in quality of goods, in location, and in other essential ways".[53]

However, Hotelling's result is dependent on a number of factors. First, Hotelling did not take into account the possibility of the non-existence of a duopolistic equilibrium. But it is always possible that a producer will want to decrease his price below his competitor's price and thus take over the whole line segment. He will do so once the profit from monopoly exceeds the profit from the duopolistic equilibrium. Thus, the condition for the existence of a duopoly is given by (assuming $c_A = c_B = c$ and $f_i = 0$):

$$L(p_A^{MPLY} - c) \leq \frac{t(3L + a - b)^2}{18}. \tag{3.15}$$

The monopoly price is given by consideration that once the price for a good from supplier A for the consumer (*i.e.* producer price and transportation costs) is lower along the whole subsegment between a and b, the supplier B is driven out of the market. Hence,

$$p_A^{MPLY} = \tilde{p}_B - t(L - a - b). \tag{3.16}$$

Substituting for the monopoly price in (15) and the optimal price for B from (3.12) yields:

$$L(a + 2b) \leq \frac{(3L + a - b)^2}{12}. \tag{3.17}$$

For symmetric duopolists ($a = b$) we get the condition:

[52] Hotelling 1990/1929, p. 61
[53] *ibid.*, p. 50

$$a = b \leq L/4, \tag{3.18}$$

which means that both suppliers will have to be located in the outer quarters of the line segment. That is, the principle of minimal differentiation (another name for Hotelling's result) holds *only if the two suppliers do not move to closely together.*

This discontinuity is an artefact of the assumption of linear transportation costs. If one replaces equation (3.2) with the following:

$$T(\delta) = t\delta^2, \tag{3.2'}$$

the equilibrium expressions for prices, quantities and profits read (again, assuming $c_A = c_B = c$):

$$\tilde{p}_A = t(L-a-b)\frac{3L+a-b}{3} + c, \tag{3.12'}$$

$$\tilde{q}_A = \frac{3L+a-b}{6} \tag{3.13'}$$

and

$$\tilde{\pi}_A = t(L-a-b)\frac{(3L+a-b)^2}{18} - f_A. \tag{3.14'}$$

The most interesting result is, however, the following. Differentiating (3.14') with respect to a, the location parameter of supplier A, we get:

$$\frac{\partial \tilde{\Pi}_A}{\partial a} = -t(3L+a-b)\frac{L+3a+b}{18} < 0. \tag{3.19}$$

The principle of minimal differentiation has turned into a principle of *maximal* differentiation in the presence of quadratic transportation costs.

But now let us assume that for institutional or other reasons, the prices are exogenously fixed. In this case competition can only proceed via location. Producers will determine their location so as to maximise demand. The demand equation becomes (from equations (3.6) and (3.7) with $p_i = p_j$):

$$q_A = Q^A(a) = \frac{L-a-b}{2} + a, \tag{3.20}$$

which increases in a. Thus, supplier A will move as closely together with B as possible, and vice versa (which is easily seen when a and b are interchanged in

(3.20)). Again, in markets with no price competition the principle of minimal differentiation results![54]

I think Hotelling's model of spatial competition is a good example of how in some cases the results of economics models are very sensitive to the assumptions one makes at the outset. An important feature of this sensitivity is that it is not easily known what kinds of changes the models are sensitive to. In this example we have seen that minimal, though restricted differentiation results for linear transportation costs and price competition, maximal differentiation for quadratic costs and competition and minimal differentiation in the absence of price competition. But there are more results possible. It has been shown[55] that if one assumes transportation costs are $t\delta^{\alpha}$, with α in [1, 2], there exist equilibria for $\alpha > 1.26$, differentiation is maximal in [1.67, 2], but some differentiation, though not maximal exists between 1.26 and 1.67. I do not dare predictions for α outside the interval [1, 2].[56] The results of course also depend on the possibility of price competition. But before one has produced a model that is robust with respect to some kinds of changes, these findings suggest that one should rather be careful as to the generality of a model's results.

One might object that the Hotelling model is a very special case, and that it is not representative of all economics. I would certainly agree that it is not representative of *all economics*. The model is micro economic, and the modelling techniques it uses are found mainly in micro economics—with the exception of macro with micro foundations of course. However, I do think that the Hotelling model is typical of a large class of economic models. In order to give a bit more flesh to the bones of this claim just consider the following abstract, which is taken from an article by Exeter economists David de Meza and Ben Lockwood that discusses the Grossman-Hart-Moore theory of the firm:[57]

> This paper studies the Gossman-Hart-Moore (GHM) "property rights" approach to the theory of the firm under alternating-offers bargaining. When managers can pursue other occupations while negotiating over the division of the gains from cooperation, the GHM results obtain. If

[54] These results are well known. *Cf.* Tirole 1992, section 7.1.
[55] Economides 1986. *Cf.* Tirole 1992, p. 286, note 12.
[56] Although it is known that for no transportation costs ($t = 0$) the Bertrand solution results, *i.e.*, goods are priced at marginal costs c—which is identical to the solution with competition. See Tirole 1992, p. 280.
[57] De Meza and Lockwood 1998, p. 361, original emphasis

> taking the best alternative job terminates bargaining, outcomes are very different. Sometimes an agent with an important investment decision should *not* own the assets he works with; sometimes independent assets should be owned together; sometimes strictly complementary assets should be owned separately.

That is, under the original GHM assumption that managers can pursue other occupations during the bargaining process, we get one set of results but if accepting an alternative occupation terminates the bargaining process, a very different set of results obtains.

I want to mention a third example. The Black-Scholes option pricing formula is derived from seven important assumptions.[58] Since most of them are false, at least for the most part, in the past almost thirty years a great number of models have appeared in which one assumption or the other has been relaxed or changed. In one such article its author concludes:[59]

> Option prices are sensitive to the stochastic processes that determine underlying stock prices… Relaxation of these assumptions can produce large percentage changes in option prices.

My last example is slightly differently pitched. It does not concern a demonstrated sensitivity to the assumptions made but an expression of the caution economists should take as long as model results are not demonstrated for all relevant situations. It comes from a commentary on the trade cost model by Obstfeld and Rogoff, which was discussed in Chapter 2. Recall that Obstfeld and Rogoff's aim is to demonstrate that a number of international trade puzzles can be solved by focusing mainly on one factor: trade costs. One of the assumptions the authors make in order to derive their results is that of complete asset markets. This assumption implies that relative consumption levels internationally are perfectly correlated with real exchange rates.[60] However, there is vast empirical evidence against this correlation. Obstfeld and Rogoff remark about this state of affairs:[61]

> We do not take this as too damning, since for us the complete markets assumption was only a useful device for calibration, and not a religious conviction. Trade costs would play essentially the same role in a world with, say, trade in debt and equities but not a complete set of Arrow-Debreu securities.

[58] See Black and Scholes 1973.
[59] Beenstock 1982, p. 40. Quoted from Pemberton forthcoming, p. 10
[60] Obstfeld and Rogoff 2000, p. 30
[61] *ibid.*, pp. 30f.

Fair enough, one might think, but a commentator, Charles Engel, warns:[62]

> That [remark] may be true, but it needs to be demonstrated. Can trade costs play a quantitatively significant role in resolving the puzzles in such a model? At this stage, this seems not much more than a conjecture. The models that are presented in this paper all have the implication that relative consumption levels are perfectly correlated with real exchange rates. OR [the authors] provide us with no evidence about models in which this link is broken.

I take Engel's point to be that trade costs may be a significant factor in the explanation of all sorts of phenomena in international trade. Because models are often assumption-sensitive, however, as long as a particular result is not demonstrated within a mathematical model, we have no good reason to believe in it. Or, to turn this claim the other way round, we have no reason to believe in the result of any mathematical model outside the domain of validity of its assumptions.

Assumption-sensitivity in economics is a truism. But it presents a problem for economic concept formation according to nomic functionalism. This theory says that concepts apply to those objects, processes or systems of which the nomic claims that are expressed by the models in which these concepts play a part are true. There seems to be no way of demarcating those objects, processes or systems of which the nomic claims are true from those of which they are false. Of course, they are true of those objects, processes or systems of which the model's assumptions are true. But then we are back at axiomatic functionalism with its associated problems. On the other hand, we may want to say that the nomic claims are true of all objects, processes or systems which instantiate the quantities represented by the concepts of the model. For example, the nomic claims in which *transportation cost* plays a part is true of all systems with transportation cost. But in this case we will fall into contradictions: transportation cost can either produce a market with minimal product differentiation or a market with maximal product differentiation but not both.

We have therefore two alternative readings of the *ceteris paribus* clause. According to the first, the nomic claims are true only of those objects, processes or systems of which the assumptions of an economic models are true. This is a very restrictive interpretation. We have seen that this interpretation brings us back to axiomatic functionalism. According to the second reading, the *ceteris paribus* clause is empty.

---

[62] Engel 2000, p. 8

All nomic claims apply universally. But in this case we have seen that they will often contradict each other.

Hence we would like to find an interpretation of *ceteris paribus* clauses that allows us to determine to which objects, processes or systems the nomic claims apply without being overly restrictive or overly liberal. Marcel Boumans and Mary Morgan[63] provide an insightful discussion of the nature of *c.p.* clauses. They distinguish three general meanings of *c.p.* clauses in the context of material and virtual experiments:[64]

(1) other things being held constant ("*ceteris paribus*")

(2) other things being absent ("*ceteris absentibus*")

(3) other things being negligible ("*ceteris neglectis*").

Their interpretation is causal. "Other things being held constant", for them, means that the nomic claim one seeks to establish with the experiment holds only in a causally homogenous background; "other things being absent" that causal factors other than the ones modelled are absent; and "other things being negligible" that other causal factors are present, but their influence on the result is either nil or "within the limits of experimental error" (the gravitational influence of distant galaxies on the tides, say).

Using the method of causal inference of, say, isolating a causal system from all possible confounding factors justifies the formation of a claim "*ceteris absentibus*". The claim applies to all actual systems in which the factors described in this clause are indeed not operative. Since, however, the use of these empirical methods of causal inference is largely absent in the formation of nomic claims and concepts in the tradition I am analysing here, we cannot use this strategy. Applying a nomic claim that has been established by a thought experiment or mathematical model to actual economic systems will involve a great leap of faith. This leap of faith, in turn, will make our concepts *inexact*. Because there are no strict rules in the application of the concepts and nomic claims to real systems, application will to some extent become arbitrary and results ambiguous. Note that his arbitrariness obtains both on a

[63] Boumans and Morgan 2001

[64] Material experiments are real experiments conducted in a physical/social environment. Virtual experiments are the thought experiments and mathematical models of the kind discussed here. There are also hybrid forms such as computer simulations.

semantic as well as on an epistemic reading of "concept application".[65] A vague *ceteris paribus* clause *in principle* obstructs the determination of what objects, processes or systems the model is applicable to. Hence, the meaning of the concepts that figure in the nomic claims expressed by economic models is inexact. But, as a consequence, it will also be hard to find empirical principles that are informative about how to apply the models to actual economic phenomena. This threat is clearly seen by Boumans and Morgan too:[66]

> In this research tradition [mathematical modelling], *ceteris paribus* conditions are flexible friends for theory construction but offer pitfalls for the unwary economist who applies those theories to the world.

It seems, then, that we are facing a trade-off here. We may interpret concepts as characterised axiomatically functionally. In this case, they are, in some sense very exact: they apply to those and only those systems of which the assumptions of the model which is definitional of the concept are true. But these concepts cannot explain. They do not unify and systematise thinking about economic phenomena because each model defines its own concepts.

On the other hand, we may interpret concepts as characterised nomicly functionally. In this case, concepts may be used for building models that have explanatory power: they figure in many models that may capture a great range of independent phenomena. But they will be very inexact. Because of the problems associated with finding an appropriate understanding of the *ceteris paribus* clause characteristic of so many scientific law-claims, their rules of application are notoriously vague.

4 Conclusions

This Chapter has analysed economic concepts that may be thought of as defined by their relations to other concepts in economic models. Two positions were distinguished, axiomatic and nomic functionalism. Eventually a trade-off result emerged. On the one hand, one can either define one's concepts axiomatically

---

[65] *Cf.* the distinction drawn in Chapter 1, section 1.
[66] *ibid.*, p. 14

functionally, that is, by the assumptions of a particular model. One will then have very exact concepts: they clearly delineate which objects, processes or systems fall under them and which do not. However, these concepts will not be very informative of reality, neither descriptively nor explanatorily. They cannot describe *actual* objects, processes or systems because of the fictional character of the assumptions. They cannot explain much because each model defines its own concepts. On the other hand, one can characterise one's concepts nomicly functionally. This may allow one's concepts to have explanatory power, but they will not be exact because the rules of applying them to objects, processes or systems are indeterminate.

There is a way, though, according to which models thus understood may be thought of as descriptive (and, in fact, explanatory) of real phenomena. Let us suppose that the assumptions of a model whose concepts we understand as axiomatically functionally defined characterise *fictional objects*. Let us also suppose that these fictional objects are real—not actual but real in the Platonist-abstract sense. James Robert Brown holds a view that enables us to interpret relations between fictional objects to be responsible for phenomena that obtain in the actual world. He summarises his view with the following statements:[67]

1 Laws are relations between abstract entities.

2 They are real, but exist outside of space and time.

3 Laws are causally responsible for the regularities that do obtain in the physical world.

Given Brown's notion of phenomena which is very similar to the one held in this Thesis, it is possible that models that describe purely fictional objects and their relations should be phenomenally adequate as well as have explanatory power. But the claim that (Platonist-) abstract entities or purely fictional objects which exist outside space and time should be causally responsible for actual phenomena seems contentious to say the least. I shall not refute it here; I simply note that it is incompatible with the general Baconian theme that defines this Thesis.

Let us, then, turn to concepts that may be thought of as being defined operationally. The problem of application should be solved much more readily for these concepts. But can these concepts help us in building models that are phenomenally adequate?

[67] Brown 1994

Chapter 4

Artefacts, Multiple Operationalism and the NAIRU

Chapter 4

Artefacts, Multiple Operationalism and the NAIRU

Once he was inside her, fear was derailed and biology took over.
The cost of living climbed to unaffordable heights;
though later, Baby Kochamma would say it was a Small Price to Pay.
Was it?
Two lives. Two children's childhoods.
And a history lesson for future offenders.

Arundhati Roy—The God of Small Things

1 Introduction

The NAIRU[1] is an intriguing concept. It is a part of the standard parcel of taught macroeconomics; it is a topic of considerable significance in theoretical and macro econometric research; and it is a key indicator for policy making. And yet, there is widespread disagreement among economists about virtually all aspects of it, including:

- the adequacy of its very name[2]
- the relation between the concepts "NAIRU" and "natural rate of unemployment"[3]
- the determinants of the NAIRU[4]
- whether the causal structure behind the NAIRU concept is right
- the speed of adjustment of unemployment towards the NAIRU

[1] *N*on-*A*ccelerating *I*nflation *R*ate of *U*nemployment
[2] It is often noted that the name gets a derivative wrong and should be called Non-*Increasing* Inflation Rate of Unemployment. See for example Mellis and Webb 1997, p. 4.
[3] I have found many different answers to this question in the literature, including (a) they are different names for the same concept (*e.g.* Stiglitz 1997); (b) they are "twin concepts", (c) they are conceptually different but empirically indistinguishable (*e.g.* Thirlwall 1983), (d) the NAIRU is a proxy for the natural rate (*e.g.* McAdam and Mc Morrow 1999); (e) the NAIRU is a measure or operationalisation of the natural rate (*e.g.* Dornbusch and Fischer 1990); (f) the natural rate is a long-term whereas the NAIRU is a more short or medium-term concept (*e.g.* Richardson *et al.* 2000, p. 32) and (g) they are really different concepts because they are based on incompatible theoretical frameworks (*e.g.* Tobin 1997).

- whether the NAIRU is even in principle measurable with a reasonable degree of accuracy[5]
- the value of NAIRU for country x and period $t_n - t_{n+1}$[6]
- whether there is a NAIRU for country x and period $t_n - t_{n+1}$[7]
- whether the NAIRU is stable or time varying[8]
- how useful the NAIRU is for policy recommendations.[9]

This apparent tension between the NAIRU's pervasiveness in both economic theory and practice and the problems surrounding its concept indicate that the NAIRU presents an instructive case study. In this Chapter I shall focus on one of its many aspects, *viz.* the measurement of the NAIRU, and address the other issues only as necessary to deal with measurement accurately. Because a great deal of controversy is involved and disagreement about measurement issues is far from settled, the case provides a rich opportunity to examine the multifaceted epistemological strategies employed by economists.

The second subspecies of Bacon's *Idols of the Market-place*—the fallacies of the mind that relate to the inadequate use of language—is the custom of employing terms as if they signify something real when in fact they don't. Bacon's examples include *fortune*, *planetary orbs* and the *prime mover*. Prompted by this idea of Bacon's, in this Chapter I want to investigate the question of whether the methods we can find in the more empirical branch of contemporary economics can establish that its concepts refer—exemplified by the methods surrounding the NAIRU concept.

In an important contribution, Kevin Hoover[10] has asked "Is Macroeconomics For Real?" His answer is triumphantly in the affirmative: macro economic entities exist independently of any individual human mind, and objectively, that is, not constituted

[4] This point and the following two a mentioned as controversial issues in Cromb 1993, p. 27.
[5] McAdam and Mc Morrow 1999
[6] This, I believe, is implied by the fact that at least for the US and the UK there exist many different estimates, none of which is accepted to be the "correct" one. For the US, see for example Staiger *et al.* 1997, for the UK Coulton and Cromb 1994.
[7] This I base on the fact that some authors reject the concept completely such as Galbraith 1997, or at least much of the theoretical framework behind it such as Eisner 1997 (whose reasoning implies that the Phillips curve relationship does not obtain in a way the advocates of the NAIRU understand it).
[8] See *e.g.* Gordon 1997.
[9] For an advocate, see Stiglitz 1997, and for a critic, Galbraith 1997.

by "the representations of macroeconomic theory".[11] Hoover employs two main arguments to defend this thesis: Hacking's argument from manipulability, and an argument from explanatory success. According to the first argument, at least some irreducibly macro economic aggregates are real because we can reliably manipulate them in order to achieve certain macro economic outcomes. According to the second argument, some macro economic aggregates are real because of the explanatory success of real[12] business cycle models which use these aggregates for testing and estimation. Hoover writes, "Thus, to the degree that such models are successful in explaining empirical phenomena, they point to the ontological centrality of macroeconomic and not microeconomic entities".[13]

My project here is comparable in flavour but less ambitious. Using arguments similar to Hoover's I focus on one example in detail and want to ask whether methods of measurement currently in use can establish whether the NAIRU "is for real" and by that token question the relevance of Bacon's idols for contemporary applied economics.

Considerations about the ability of a measurement, testing or experimental procedure to establish a genuine phenomenon rather than an artefact of that procedure have frequently been addressed with reference to the *reliability* of the procedure in question or the *robustness* of its results. In order to see what kinds of questions are asked when the reliability of a procedure or the robustness of its results is at stake, let us look briefly at four examples. The first two are more naturally classified under the topic of reliability of a procedure, the second two under the robustness of a procedure's results.

Bogen and Woodward: Control of Possible Confounding Factors

One of James Bogen and James Woodward's objectives in their 1988 article[14] is to draw the readers' attention to the distinction between what they call *phenomena* and *data*. As we have seen in Chapter 1, phenomena are stable and general features of the world which are of theoretical interest. Data, on the other hand, are the

---

[10] Hoover 1995
[11] *ibid.*, p. 236
[12] "Real" in this context is of course opposed to "nominal" rather than "unreal".
[13] *ibid.*, p. 253

observable outcomes of measurement or experimental procedures. They can play the role of providing evidence for the existence of phenomena. But when, we may ask, are we in a position to argue convincingly that a specific set of data in fact does constitute reliable evidence for the existence of a phenomenon of interest? One type of consideration, according to Bogen and Woodward, is whether possible confounding factors and sources of error have been adequately controlled. "Confounding factors", the authors define,[15]

> are factors which can produce data similar to that which would be produced by the phenomenon of interest and thus yield spurious candidates for that phenomenon. We also include under this heading factors which introduce so much noise into the data that it becomes unusable as evidence. Controlling for such factors is central to establishing reliability.

In their example, the phenomenon of interest is the existence of neutral currents, which are weak interactions mediated by the neutral Z° particle. This phenomenon is predicted by the Weinberg-Salam theory, which unifies the weak and electromagnetic forces. The difficulty with finding reliable evidence for the existence of neutral currents is that it can only be inferred from tracks in bubble chambers, which can be produced by other processes than the neutral current.

In particular, it was known that when neutrinos strike the wall of the bubble chamber and the surrounding apparatus, they produce a large and unknown number of neutrons, which in turn may produce a shower of hadrons.[16] Events of this kind are indistinguishable from genuine neutral current events. The strategy to establish that at least some of the events were genuine neutral current events and not produced by confounders, which was used by a number of CERN experimenters when they investigated neutral currents, had a number of different elements. One was to calculate an upper limit for the size of the neutron background with various methods. If the observed number of events that could have been produced by the phenomenon of interest (genuine and pseudo) exceeds this limit, one could argue at least that some of those events were due to neutral currents. On the other hand, for theoretical reasons one could argue that pseudo events were more likely to occur in the vicinity of the wall of the bubble chamber. If many of the observed events were to occur near

---

[14] Bogen and Woodward 1988
[15] *ibid.*, pp. 327f.
[16] *ibid.*, p. 328ff.

its centre, or indeed they would be distributed uniformly within the chamber, one again could argue that at least some events were due to neutral currents.

In this case, a possible confounding factor has been controlled for by estimating an upper limit if its size and subtracting this value it from the number of all events that could have been produced by either the genuine (neutral current) or the pseudo process (neutron background). Let us call this method of controlling for confounders *calculation*. Other methods of controlling confounders include *crafted isolation* (for example by conducting an experiment far beneath the surface of the earth in order to shield the apparatus from background radiation), *natural isolation* (where shielding occurs naturally) and explicit *modelling* of the factor's contribution.[17]

Hudson: Reliable Process Reasoning

In a second case, relevant empirical knowledge gained independently in a different area can be used to argue for the preference of one procedure over another one in the establishment of a genuine phenomenon. The phenomenon of interest in this example is a small cell organelle called the mesosome. The mesosomes case is of particular interest in this context because the phenomenon was first believed to be an artefact, then for about fifteen years mesosomes were generally believed to exist until at the beginning of the eighties doubts about their existence gained the upper hand.

There are various accounts of these events, including Sylvia Culp's[18] reconstruction of them as a story where the robustness of the results of measurement procedures in a sense similar to the one defended below was decisive in the acceptance and rejection of the phenomenon's genuineness and Nicolas Rasmussen's[19] study which argued that evidential considerations did not play a decisive role. In one contribution, however, Robert Hudson[20] reasons that the case for the non-existence of mesosomes was made on account of the fact that certain specimen preparation methods are known independently to produce artefacts but others are known not to. Since mesosomes are found when the bacteria are fixed using osmium (a method that often

[17] See Woodward 1989 for a detailed discussion of these and other methods.
[18] Culp 1994
[19] Rasmussen 1993
[20] Hudson 1999

leads to artefacts) but not when using uranyl acetate or freeze-substitution (methods that do not tend to produce artefacts), they are judged to be artificial.[21] In this case, we believe a certain measurement procedure to be more reliable than others because it is known empirically to work for a number of well-understood phenomena.

Chang: Minimalist Overdetermination

Hasok Chang's[22] account of the "real scale of temperature" can illustrate the idea of the robustness of measurement results as an argument for a phenomenon's reality. In the 18th and 19th centuries thermometry experienced an intense debate concerning the choice of thermometric fluids. There were three main contenders: air, methyl alcohol and mercury. Since different fluids gave rise to different readings except at the fixed points, the question arose which type of thermometer would give the "correct" value—if there is any. After his discussion of Jean André De Luc's and Pierre-Simon Laplace's experimental and theoretical reasoning, Chang analyses Henri Victor Regnault's series of experiments as employing a strategy of "minimalist overdetermination". According to Chang, Regnault's strategy was "minimalist" in that he managed to avoid all possible theoretical assumptions except what Chang reconstructs as a "basic ontological conjecture", *viz.* that there is an objectively existing property, "temperature", which can be measured correctly by some thermometer-type.

Regnault used a criterion of "comparability" to justify his choice of the air thermometer as the type which gives the closest approximation to the real temperature. Chang paraphrases Regnault's argumentation: "if a thermometer is to give us the true temperature, it must at least always give us the same reading under the same circumstance; similarly, if a type of thermometer is to be an accurate instrument, all thermometers of that type must at least agree with each other in their reading".[23] Therefore, this minimalist strategy is also one of overdetermination: all thermometer-tokens of the same type must agree in their readings if the ontological conjecture is to be true.

[21] *ibid.*, pp. 304ff.
[22] Chang 2001
[23] *ibid.*, p. 41

In this case the overdetermination consists in the fact that we can vary certain parameters that do not have a significant influence on measurement results. Here the range of variation is limited by the *type* of thermometer. Two thermometers filled with air at different densities, and made of different kinds of glass are of the same kind. But a thermometer filled with air is not of the same type as one filled with sulphuric acid.[24]

Hacking: The Argument from Coincidence

Ian Hacking[25] argues that one concern of microbiologists is to distinguish artefacts from real objects in their microscopic images of biological specimen. One way to test whether a given structure on a representation is real or artificial, according to Hacking, is to subject the same specimen to a different kind of microscopic observation, one which uses unrelated physical processes. In his example, low-powered electron microscopes detect "dense bodies" in red blood platelets. These dense bodies may be either a real feature of red blood cells or an artefact of electron microscopy. One test is to determine whether dense bodies are also revealed by light microscopes. This is possible in the given case because a low resolution electron microscope has roughly the same power as a high resolution light microscope.

Dense bodies are not detected by every kind of light microscopic technique, but they are by preparing the specimen and subsequent observation with a fluorescent microscope. Hacking concludes:[26] "It would be a preposterous coincidence if, time and again, two completely different physical processes produced identical visual configurations which were, however, artifacts of the physical processes rather than real structures in the cell".

Here again a parameter of the measurement procedure is varied. In this case the range of variation is and must be across different kinds of physical processes, described by "different chunks of physics", as Hacking puts it. The fact of the agreement of results of these different physical processes can be used to give credence to each method.

---

[24] *ibid.*, p. 50
[25] Hacking 1983, ch. 11
[26] *ibid.*, p. 201`

The four strategies described above are very different strategies, but they are all empiricist strategies. The empiricism of minimalism is probably the strongest one: according to Chang, Regnault attempted to discard all theoretical assumptions except the "basic ontological conjecture". In Hacking's example, microbiologists do not try to avoid theoretical assumptions but mutually corroborate alternative methods because measurement procedures give the same results, or very nearly same results, in the relevant aspects, and those procedures are described by different physical theories. In Hudson's case, low-level empirical knowledge about the properties of specimen and preparation methods gained in contexts different from the one at stake is used to assess the reliability of that method. Finally, in Bogen and Woodward's case (where probably the greatest use of theory is made) all care is taken that at least some of the observable events must be due to the phenomenon of interest rather than due to alternative processes which are observationally indistinguishable.

Let us circumscribe reliability and robustness in the following way. First, reliability pertains to a measurement *procedure* itself whereas robustness pertains to its *results*. Second, we will say that a measurement procedure is reliable if we have good reason to believe that it ensures that its results represent an aspect of the phenomenon of interest. We will say that a procedure's results are robust if they do not change significantly when parameters that we believe should not affect the results are varied.

One of the main insights from the analysis of the NAIRU case presented below is that considerations of reliability and robustness in this sense play a comparatively insignificant role in economists' justification of particular NAIRU measures. Sometimes economists do test for the robustness of their results, but this occurs usually at the very far end of the measurement procedure, that is, in the selection of the data set used (*e.g.* it is tested whether NAIRU point estimates are similar when the GDP deflator is used as proxy for inflation and when the CPI-U-X1 is used). Reliability of measurement procedures is a factor which is often considered but in most cases it relates to either theory or policy considerations rather than evidential considerations. Measurement procedures are justified with criteria that will be classified into three groups below—theoretical, statistical and pragmatic criteria. Among these criteria one finds some which seem similar to those of reliability and robustness but these are of comparatively low prominence.

The principal conclusion I shall draw from this observation is that the concept of the NAIRU that is associated with the measurement procedures discussed below can either be regarded as being *defined* by the measurement procedure used (*i.e.*, as being defined operationally), or as being defined within a theoretical framework, and the measurement results giving empirical meaning to it. Operational definition implies that each NAIRU measurement procedure is constitutive of a different concept. But this in turn implies that NAIRU concepts are "peninsular" concepts (using Harriet Margenau's phrase[27]), which are not very readily associated with theoretical, explanatory concepts. On the other hand, an independent, theoretical definition, as will be seen, is associated with great measurement uncertainty. Therefore, NAIRU measures either have theoretical significance and are inexact to a large extent, or they are exact but lack theoretical import. Before engaging in a philosophical discussion of the methods used to justify particular measures of the NAIRU I shall briefly introduce the concepts and the various methods employed to measure it.

2 NAIRU—Concept and Measurement

The NAIRU—or Non-Accelerating Inflation Rate of Unemployment—is a concept with a long and varied history and pre-history.[28] One can regard its theoretical background as a kind of synthesis of (neo-)classical/monetarist and Keynesian elements. But let us begin with a fragment of pre-history from the time just before Keynes.

2.1 Theory Behind the NAIRU

In the neo-classical Walrasian general equilibrium model there existed no involuntary unemployment. Unemployment, in so far as it occurred, was a result of the economic agents' rational deliberation of the trade-off between labour and

[27] See Margenau 1950, p. 87. Quoted from Wimsatt 1981, p. 137.
[28] The first half of this section draws considerably, but not exclusively, from Espinosa-Vega and Russell 1997.

leisure. However, in the light of the Great Depression and the happenings in its aftermath the Walrasian model did not appear to be an adequate representation of the actual US economy. Unemployment rose to five million in 1930 and up to thirteen million in 1932, not all of which could readily be explained with reference to the general equilibrium model.

John Maynard Keynes[29] in 1936 provided an alternative explanation of the events. In the Keynesian model unemployment is caused by a failure of aggregate demand to equal the equilibrium level. Low aggregate demand, in turn, may be caused by *money hoarding*, a *low propensity to consume* and, likewise, a *low propensity to invest*. In 1958, LSE economists Alban Phillips[30] hoped to find empirical support for the Keynesian idea that wage pressure depends on the tightness of the labour market. He investigated "whether statistical evidence supports the hypothesis that the rate of change of money wage rates in the United Kingdom can be explained by the level of unemployment and the rate of change of unemployment".[31] Phillips found that they were negatively correlated.[32] He thus discovered a mere *statistical* regularity, but Keynes's followers integrated it within the broader Keynesian framework as a kind of law-like relationship. The reasoning behind it was that in situations of tight labour market conditions employers must bid up wages in order to attract workers. Higher wages then would lead to additional actual hires, and thus the unemployment rate would fall. Although Phillips's investigation was into the relationship between unemployment and wage inflation, the Phillips curve was soon used to describe the relationship between unemployment and inflation *simpliciter*, which could mean wage or price inflation. Since the bulk of firms' costs are indeed wages, movements in the level of wages would sooner or later affect their price setting behaviour too.

A nice aspect of the Phillips curve mechanism was that it appeared to make economic policy possible. Despite the frustration of the original Keynesian hope that aggregate demand could be stimulated and thus involuntary unemployment reduced *at no cost*, demand management still seemed viable though now a reduced unemployment rate had to be bought at the cost of higher inflation.

---

[29] Keynes 1936
[30] Phillips 1958
[31] *ibid.*, p. 284. Quoted from Espinosa-Vega and Russell 1997, p. 6.
[32] Phillips was actually not the first to discover a relationship of that sort. As early as 1926 Yale's Irving Fisher found that a negative correlation obtains between unemployment and inflation, though he examined *price* rather than *wage* inflation. See Fisher 1926, reprinted as Fisher 1973.

In the late 1960s Milton Friedman[33] and Edmund Phelps[34] claimed for theoretical reasons that the Phillips relationship would not be stable in the long run. Their reasoning was essentially that workers and firms, being able to predict movements of the price level, could not be misled systematically in their decisions to seek employment and to hire, respectively. The hypothesised Phillips relationship obtains only subject to the validity of the assumptions that price expectations are rigid and that workers do not resist wage cuts if they are caused by inflation. But these assumptions seem implausible in the long run. Only in the short run could unemployment deviate from what Friedman called the *natural rate*; in the long run, unemployment would always tend towards that quantity. In other words, the long-run Phillips curve is vertical. Friedman defined the natural rate as follows:[35]

> The natural rate of unemployment is the level which would be ground out by the Walrasian system of equilibrium equations, provided that there is imbedded in them the actual structural characteristics of the labor and commodity markets, including market imperfections, stochastic variability in demands and supplies, the cost of gathering information about job vacancies and labor availabilities, the costs of mobility, and so on.

The *natural rate* is thus an equilibrium concept. Unemployment at the natural rate will frequently be positive rather than zero because of imperfections in the market characteristics such as search costs and other causes of "friction" and "structural mismatch".[36] According to the natural rate hypothesis, there is a fixed rate of unemployment at which the labour market is in equilibrium in the sense that workers' and employers' plans are consistent. Wages rise at the growth rate of productivity. Greater employment can only temporarily be bought by driving workers into what Friedman called "money illusion". That is, a change in the inflation rate, brought about by a sudden monetary expansion, is able to reduce workers' real wages and hence unemployment as long as the change is not predicted. A stable inflation rate, however, will soon be incorporated into the workers' plans and therefore cannot induce greater employment. One of the upshots of this monetarist programme was the comparative incapacity of an interventionist policy to

---

[33] Friedman 1968
[34] Phelps 1967
[35] *ibid.*
[36] Friction is the time it takes workers to find a new job. Structural mismatch occurs whenever the characteristics of the labour sought in a particular market do not equal the characteristics of the labour offered, when for example certain qualifications are in high demand but job-seeking workers have different qualifications.

affect the labour market in a beneficial way. Any demand policy, in the long run, would only cause inflation and leave employment at its natural rate.

Indeed, Friedman's and Phelps's predictions regarding the instability of the Phillips curve were verified in the seventies and after. The data then obtained were incompatible with the alleged simple inflation-unemployment trade-off. As a consequence, Keynesians were prompted to look for ways to reconcile their views about the possibility of beneficial intervention with the new empirical situation and the monetarist insights that gained credence as a result. In 1975, Franco Modigliani and Lucas Papademos introduced a concept labelled *noninflationary rate of unemployment* (NIRU). They defined it "as a rate such that, as long as unemployment is above it, inflation can be expected to decline—except perhaps from an initially low rate" and claimed that its existence "is implied by both the 'vertical' and the 'nonvertical' schools of the Phillips curve".[37] Although taking important monetarist elements into account in this way, the authors used their idea of a NIRU to argue in favour of monetary expansion. They write,[38]

> We conclude that the monetary authority should be prepared to accommodate the temporary rapid rate of growth of the money supply required for the strong recovery we advocate, which we believe is consistent with a gradual abatement of inflation. By contrast, holding to monetary growth targets of the 1974 magnitude would very likely make for a sluggish recovery with rising unemployment, and might even produce a new downturn.

The main differences between the "monetarist-amended" Keynesian view of the Phillips curve relationship and the more purely monetarist view can be said to be the following. Friedman posited his "natural rate" in a context in which he claimed the ineffectiveness of monetary policy. Keynesians regarded that rate as a mere constraint on policy. Interventionist policy is possible for the latter—as long as it does not drive the unemployment rate too far below its long run value. The reason for this difference is twofold. First, monetarists believed that the actual unemployment rate would almost always be near the natural rate whereas Keynesians regarded an unemployment rate which is persistently much above that rate as not only conceivable but as actual most of the time. Second, for the monetarists the Phillips curve was steep in the vicinity of the natural rate, *i.e.* additional employment could only be bought at considerable extra inflation. But for

[37] Modigliani and Papademos 1975, p. 142

the Keynesians, who believed that the actual rate was nowhere near the natural rate anyway, the Phillips curve was flat. That is, reductions of unemployment would lead to a mere incremental additional pressure on prices. The second main difference was that the monetarist mechanism regarded monetary inflation as the *cause* of unemployment, whereas according to the Keynesians the influence on the price level is channelled through the labour market mechanism. Despite these differences James Tobin, in 1980, pronounced that by 1970 there had emerged a "consensus view" among macro economists. It consisted of the following five elements:[39]

1. The nonagricultural business sector plays the central role in determining the economy's rate of inflation. [...]
2. Variations in aggregate monetary demand, whether the consequences of policies or of other events, affect the course of prices and output, and wages and employment, by altering the tightness of labor and product markets, and in no other way. [...]
3. The tightness of markets can be related to the utilization of productive resources, reported or adjusted unemployment rates, and capacity operating rates. [...]
4. Inflation accelerates at high employment rates because tight markets systematically and repeatedly generate wage and price increases in addition to those already incorporated in expectations and historical patterns. At low utilization rates, inflation decelerates, but probably at an asymmetrically slow pace. At the Phelps-Friedman "natural rate of unemployment," the degrees of resource utilization and market tightness generate no net wage and price pressures up or down and are consistent with accustomed and expected paths, whether stable prices or any other inflation rate. The consensus view accepted the notion of a nonaccelerating inflation rate of unemployment (NAIRU) as a practical constraint on policy, even though some of its adherents would not identify NAIRU as full, equilibrium, or optimum employment.
5. On the instruments of demand management themselves, there was less consensus. [...]

It has been suggested[40] that Tobin gave birth to the term *NAIRU* in this passage, but this is impossible since he uses it already in 1978 in a joint paper with Martin Neil Baily[41], and Arthur Okun mentions the acronym in a paper published in the same year without repeating its full name, which suggests that its was well-known at the

[38] *ibid.*, p. 141
[39] Tobin 1980, pp. 23f.
[40] *e.g.* by King 1998
[41] Baily and Tobin 1978

time.[42] But whoever coined the term, let us call *NAIRU* that concept which is part of the macro economic model of the "consensus view" described in the passage above.

It is possible to conceptually distinguish between the NAIRU and the natural rate if we stick closely to the Keynesian and monetarist framework, respectively. Deviations from the NAIRU could thus be said to cause inflation in one framework, whereas money causes both inflation and deviations from the natural rate in the other. Although Tobin in earlier publications used the two terms interchangeably, in a 1997 contribution, he says:[43]

> NAIRU and NATURAL RATE are not synonymous. NAIRU is a macro outcome of an economy with many labor markets in diverse states of excess demand and excess supply. NAIRU represents an overall balance between the inflation-increasing pressures from excess-demand markets and the inflation-decreasing pressures from excess-supply markets. The natural rate, as described by Friedman, is a feature of Walrasian market-clearing general equilibrium. While the NAIRU fits into a Keynesian model, the natural rate is an aspect of a New Classical model. The determinants of the two are theoretically different, and so are their implications for policy. The NAIRU varies from time to time as the relationships between unemployment, vacancies, and wage changes vary, and as the dispersion of excess demands and supplies across markets changes. In this decade, these developments appear to be reducing the NAIRU, in contrast to the unfavorable circumstances of the 1970s.

However, despite these differences I will follow standard usage[44] and employ the two terms interchangeably unless the context makes the distinction significant.

One of the differences between the Keynesian and the monetarist interpretations of the NAIRU/natural rate was that Keynesians believed that the actual unemployment rate would tend to be fairly remote from the natural rate whereas monetarists believed that it is almost always close to it. This fact suggests that finding a reliable measure associated with that concept is an issue of considerable substance.[45] In the

---

[42] Okun 1978

[43] Tobin 1997. Many thanks to James Tobin for the permission to reproduce this abstract.

[44] In early writings Tobin himself used the terms interchangeably but preferred NAIRU because it has a less normative flavour, see *e.g.* Baily and Tobin 1978; for a similar point of view, see Samuelson and Nordhaus 1992, p. 614 . Okun 1978 uses both as if they were the same, and so do most contributions to the 1997 *JEP* symposium on the NAIRU. But there are many other views, see footnote 2 to this Chapter.

[45] It is of course of real importance only for Keynesians. Friedman introduced his concept in a context in which he argued against intervention. From his point of view, thus, determination of the level of the natural rate is secondary. Why should one measure the natural rate if one does not base one's policy decisions on it? In an article for the *Wall Street Journal* in 1996, Friedman said, "The natural rate is a concept that does have a numerical counterpart—but that counterpart is not easy to estimate and will

following I shall then briefly introduce measurement methods that have been discussed in the literature.

2.2 Measuring the NAIRU

There are three main approaches to measuring the NAIRU: the univariate, the reduced-form or Phillips curve and the structural approach. Let us begin with the simplest method, the *univariate* approach.

The Univariate Approach

As the name suggests, a time series for only one variable is needed to estimate the NAIRU following this approach, and this is the unemployment rate. The idea behind the univariate approach is that theory implies that the actual unemployment rate will eventually tend to revert to the NAIRU, and hence one can calculate it by decomposing a time series into a trend and a cyclical component.[46] The trend component then represents the NAIRU, and the cyclical component transitory deviations from it. Indeed, a number of macro economics textbooks define the natural rate as the structural and frictional component of the actual rate, as opposed to the cyclical component.[47]

There are various ways to extract the trend component from a time series, including the calculation of a moving average and the usage of a filtering method such as the Hodrick-Prescott (HP) filter. The Bank of England[48], for example, uses the HP filter to generate a smooth estimate of the trend component or NAIRU U_t^* from the following minimisation problem:

$$\min_{U_t^*}\left\{\sum_{t=1}^{T}(U_t - U_t^*)^2 + \lambda\sum_{t=2}^{T-1}(\Delta U_{t+1}^* - \Delta U_t^*)^2\right\}. \tag{4.1}$$

depend on particular circumstances of time and place. More important, an accurate estimate is not necessary for proper monetary policy." Quoted from Espinosa-Vega and Russell 1997, p. 13.

[46] Distinguishing the two concepts, this procedure appears to measures the natural rate rather than the NAIRU. But as just said, we do not follow these intricacies here because they are not essential to the argument made further below.

[47] See *e.g.* Mankiw 1994.

[48] Bank of England 1999, p. 85

The smoothness of the trend component depends on the choice of the parameter λ There are various ways to determine that parameter. One can for example choose it in such a way as to maximise the statistical fit of the resulting Phillips curve—as suggested by the Bank of England[49], or adopt a "smoothness prior" such as Robert Gordon who argues that the resulting time series must have certain statistical properties in order to be economically sound.[50]

The Reduced-Form Approach

The second, and from what I have seen, most common approach is the *reduced form* or *Phillips curve* approach. A widely discussed example of this approach is Gordon's "Triangle model", in which the inflation rate depends on three kinds of factors, *viz.* expected or past inflation, demand conditions and supply shocks. The demand conditions are often proxied by the "output gap", *i.e.* the difference between actual and potential output, and that is in turn often proxied by the "unemployment gap", *i.e.* the difference between actual unemployment and the NAIRU, such that Gordon's equation becomes:

$$\pi_t = a(L)\pi_{t-1} + b(L)(U_t - U^*) + c(L)z_t + e_t, \tag{4.2}$$

where π represents inflation, z supply shocks and L the lag operator. There is a stable and a time-varying version of this model. If, for example, we simplify the equation by using only one lag of the inflation rate (and defining $\Delta\pi = \pi_t - \pi_{t-1}$), and using lagged unemployment instead of the contemporaneous rate and no supply shock lags, we have:

$$\Delta\pi_t = \alpha + \beta(U_{t-1} - U^*) + \gamma z_t + \varepsilon_t, \tag{4.3}$$

---

[49] *ibid.*

and can calculate the NAIRU in the absence of supply shocks as α/β.[51] If, on the other hand, we would like to estimate a time-varying NAIRU, we time-index the NAIRU and assume in addition to (4.2) a so-called transition equation

$$U_t^* = U_{t-1}^* + \varepsilon_t \tag{4.4}$$

and calculate the NAIRU by jointly estimating (4.2) and (4.4), whereby the variance of the error term ϵ must be chosen independently. This is Gordon's 1997 version.[52] In this approach the NAIRU is indirectly measured by assuming the Phillips curve relationship and reading off the value the NAIRU must have had in order to make that relationship true. In the Phillips curve relationship the unemployment figures as a *cause* of inflation but its own causes are left unspecified.[53] In the structural approach, on the other hand, the *determinants* of the NAIRU are explicitly modelled as well.

The Structural Approach

One version of the structural approach uses the so-called bargaining framework. Here the labour market is represented by three equations: one for price setting, one for wage setting and a third for labour supply. The price-setting equation summarises the aggregate demand for labour as a function of the marginal product of labour. In one version it reads:[54]

$$p - w = a_0 + a_1 n + a_2 \Delta n - a_3(p - p^e) - q + ZL_p + ZT_p, \tag{4.5}$$

where p, w, n and q represent the logs of prices, wages, employment and "trend labour efficiency", respectively, and ZL_p and ZT_p represent long-lasting and

[50] Gordon 1997. He uses a multivariate approach but his reasoning can *mutatis mutandis* be applied to the univariate case.
[51] See McAdam and Mc Morrow 1999 for this version.
[52] See below for a detailed discussion.
[53] In (4.3), for instance, unemployment has been said to Granger-cause (changes in) inflation. See McAdam and Mc Morrow 1999, p. 9.

temporary factors affecting the price formation of firms, respectively. The long-term factors may for instance be degree of competitiveness or the cost of capital, the short-term factors oil-price shocks or other supply shocks.

The wage-setting equation, on the other hand, models real wages as a decreasing function of unemployment and is often obtained from explicit micro economic models.[55] In one version it reads:[56]

$$w - p = b_0 - b_1 U - b_2 \Delta U - b_3(w - w^e) + q + ZL_w + ZT_w, \tag{4.6}$$

where U is the level and ΔU the change of the unemployment rate, and ZL_w and ZT_w represent long-lasting and temporary "wage push" factors such as the strength of unions, income support measures, the degree of mismatch between skills and geographical location of workers *etc.*

Finally, the labour supply equation may read as follows:[57]

$$l = c_0 - c_1 U + ZL_l, \tag{4.7}$$

where l is the log of the labour force and ZL_l represents factors influencing labour market participation. Given these equations, the NAIRU can be calculated as follows:

$$U^* = \frac{d_0 + a_1 ZL_l + ZL_p + ZL_w}{d_1}, \tag{4.8}$$

where $d_0 = a_0 + a_1 c_0 + b_0$ and $d_1 = b_1 + a_1 c_1$.[58] Empirically, one can estimate the parameters of (4.8) by specifying the various shock variables and jointly estimating econometric versions of (4.5)-(4.8).

---

[54] Richardson *et al.* 2000, p. 30

[55] for instance in Layard *et al.* 1991

In each of these three approaches an array of questions need to be answered before one can obtain values for the NAIRU concerning among other things the basic form of the equation, the exact specification, which time series to use, the value of certain parameter values such as the λ of the HP filter, how expectations are modelled *etc.* In the next section I want to discuss various strategies that have been employed to answer these questions. The discussion will show that most of these in use cannot be utilised to show that the NAIRU is real.

3 Justifying Measurement Procedures

Prior to discussing the various strategies that have been offered to justify NAIRU measurements, a number of conceptual clarifications are in order. I propose to define a *measurement procedure* **MP** as an ordered quadruple $<\mathbf{A_1}, \mathbf{A_2}, \mathbf{D}, \mathbf{R}>$ of a set $\mathbf{A_1}$ of assumptions about the *specification* of the procedure; a set $\mathbf{A_2}$ of assumptions about the data series; a set **D** of *data*; and a set **R** of *results*. In many cases the NAIRU is measured by running a regression. *Assumptions about the specification*, then, concern mainly the variables that enter the regression equation(s) and the functional form of the equation(s). The assumptions about the data make the assumptions about the variables more precise. Among other things, they specify the period and area of interest and where the data are taken from. In the Phillips curve specification, for example, changes in the inflation rate are regressed on past changes in the inflation rate, the output gap and supply shocks. This selection of variables and the exact functional form is given by the assumptions about the specification. *Assumptions about the data* concern the period of interest (*e.g.* "1953 – 1996"), the area ("USA" or "OECD countries", "individual Euro zone countries" or "pooled cross-country data"), and which exact measure of the variables is taken ("CPI-U" or "GDP deflator", "married males unemployment rates from the BLS", "relative food and

[56] Richardson *et al.* 2000, p. 31

[57] *ibid.*

[58] The NAIRU is calculated under the assumption that price and wage expectations are met ($p - p^e = w - w^e = 0$), that short-term shocks are absent and that n and U are stable ($\Delta n = \Delta U = 0$). Under these assumptions, substitute (4.7) in (4.5) (labour demand equals labour supply) and set the resulting equation equal to minus (4.5). Solving for U yields the NAIRU.

energy prices from the CPI series"). Naturally, the *data* are the data series themselves. The *results* are the output of the regression, sometimes only a number or a time-series or confidence intervals, *t*-statistics, residuals, standard deviations *etc. etc.*

A justification *criterion* is a principle according to which the adequacy of an aspect of a measurement procedure is assessed. The criteria are classified into three *groups*. A justification *strategy* is a cluster of criteria drawn from the three groups that assesses the adequacy of the whole measurement procedure.

In his contribution to the *Journal of Economic Perspectives* symposium on the NAIRU, Joseph Stiglitz argues that NAIRU is a useful concept for academics and policy makers according to three criteria. He asks, first, whether deviations from the NAIRU (the "unemployment gap") are a *useful predictor* of changes in the inflation rate. Second, he asks whether economists can *explain* the evolution of NAIRU. Finally, he asks whether the NAIRU is a useful concept to *frame policy decisions*.[59] Stiglitz's criteria for the usefulness of the NAIRU concept are thus predictive success, capability of being explained and policy relevance. In the classification of criteria economists use to justify their measurement procedure for the NAIRU, which I present below, I want to follow Stiglitz roughly, and distinguish three groups of criteria. The first group (including Stiglitz's *second* criterion) embraces various theoretical criteria. They concern questions about explanation and understanding and the embedding of measurement procedures in economic theory. The second group (including Stiglitz's *first* criterion) regards statistical criteria. They concern issues of predictive capacity, statistical fit, confidence intervals *etc.* The third group (including Stiglitz's *third* criterion), finally, regards pragmatic criteria. They concern policy issues and issues that relate to simplicity (of an equation) and ease of computation.[60]

Within each group, each specific criterion relates to either of two levels, the level of *assumptions* and the level of *results*. For example, a measurement procedure can be justified because some of the assumptions about the specification follow from

[59] Stiglitz 1997, p. 4. Stiglitz's three criteria are reflected in the titles of many articles discussing NAIRU, for example "The United Kingdom NAIRU: Concepts, Measurement, and Policy Implications" (Mellis and Webb 1997), "The Concept, Policy Use and Measurement of Structural Employment" (Richardson *et al.* 2000) and "The NAIRU Concept—Measurement Uncertainties, Hysteresis and Economic Policy Role" (McAdam and Mc Morrow 1999).

economic theory. But it can also be justified because economic theory predicts that two variables are related, and the results have shown them to be significantly correlated. For the aid of the understanding the following table summarises the groups of criteria.

	Theoret. Criteria - conceptual - economic - explanatory	Statistical Criteria - statistical fit - predictive accuracy - sensitivity	Pragmatic Criteria - policy-related - simplicity
Assumptions			
Results			

3.1 Theoretical Criteria

The group subdivides into three categories: conceptual, economic and explanatory criteria. Conceptual criteria demand that the measurement procedure be in line with the general idea of the NAIRU, that is, with its definition as the "non-accelerating inflation rate of unemployment" and the essential Friedmanian or Phillips curve reasoning behind it. Economic criteria are more stringent. They demand that the specification of the measurement procedure derives from (or sometimes, is consistent with) an economic model that is built of (a) economic agents and their constraints, (b) an equilibrium definition/conception and (c) optimising behaviour of the economic agents.[61] Thus, conceptual and economic criteria differ in that the requirements to fulfil a conceptual criterion are more vague, often implicit, and can relate to both micro and macro economic reasoning, whereas the model referred to in an economic criterion is an explicit micro economic model.

Explanatory criteria, finally, ask whether the NAIRU measurement can be used to *understand* features of an economic system. NAIRU can here figure as either a cause or an effect: movements in the NAIRU given the unemployment rate or, vice versa,

[60] In my language then we can say that Stiglitz's justification *strategy* consists in employing the three *criteria* of predictive success, capability to be explained and policy relevance, each of which is drawn from a different *group*.
[61] For such an understanding of an acceptable economic model, see Rogerson 1997, p. 77.

movements in the unemployment rate given the NAIRU are sometimes said to *cause* changes in the inflation rate. On the other hand, various factors are, in other models, said to be responsible for the evolution of NAIRU (or cross-country differences), models in which NAIRU figures as an *effect*.

The structural approach models are often cheered for doing well with respect to these theoretical criteria.[62] On the one hand, they are said to be explicitly derived from micro economic theory.[63] This is an economic criterion: the specification derives from an explicit economic model that satisfies the three requirements above.[64] On the other hand, they can be used to *explain* the evolution and cross-country differences of the NAIRU.[65]

An interesting point to note is that economic theory determines aspects of the measurement procedure at a very general and abstract level. For instance, in Layard *et al.*'s model only a general variable z_w, called "long-run wage pressure", is derived from theory (this is the equivalent to ZL_w in the model of (4.5)-(4.8)). The real explanatory work is done by specifying what this variable amounts to in a concrete setting. Cromb 1993, for instance, analyses the three factors: *generosity of benefits*, *mismatch* and *union strength*.[66]

I have not found any study which claims that there is a real factor z_w that explains the movement of NAIRU over time. Rather, z_w is thought to represent wage pressure "elements", and these elements themselves explain the NAIRU.[67]

What, in turn, determines the "elements" and exactly what data are taken to measure them, must be determined by means other than theory. We can say that theory underdetermines an economic explanation. In order for an explanation to be acceptable, theory must be supplemented with features drawn from historical or statistical analysis.

---

[62] Richardson *et al.* 2000, Cromb 1993, McAdam and Mc Morrow 1999, Pichelmann and Schuh 1997

[63] Cromb 1993, p. 27. For a monograph-length study of the NAIRU and its micro foundations, see Layard *et al.* 1991.

[64] See Layard *et al.* 1991, chs 2-7 for the development of various such models.

[65] "Empirical results obtained from these models thus allow for causal interpretations of the NAIRU estimates. In contrast to time series models, the movements of the NAIRU are 'explained' by various labour market variables... which are inserted into the empirical models", Pichelmann and Schuh 1997, pp. 13-4.

[66] See Cromb 1993, pp. 29ff. *Cf.* also Mellis and Webb 1997, pp. 16ff.

[67] See for example Pichelmann and Schuh 1997, p. 14.

The reduced-form or Phillips curve approach is usually judged to be *conceptually* sound.[68] Within this approach the NAIRU is measured to be that rate of unemployment that is consistent with non-accelerating inflation. Gordon goes as far as claiming that "The NAIRU is meaningful *only* within a well-specified model of the inflation process"[69]—which his "Triangle model" of course provides. Here one can well see the difference between a conceptual and an economic criterion. Gordon demands a well-specified model of the *inflation process*. His model explains (in a statistical sense) the variation of the change of inflation in terms of past changes of inflation, the output (unemployment) gap and supply side shocks. But there are no explicit agents, constraints, optimising behaviour or equilibria in his model. Although it is true that the Phillips curve approach is consistent with many structural models (and thus, in principle derivable from a microeconomic framework), most authors do not use this fact in their justification of this approach, and if it is mentioned at all, it appears only as an additional virtue.[70]

In most versions, the Phillips curve is read causally: deviations of the unemployment rate from the NAIRU *cause* changes in the inflation rate. Therefore, the reduced-form approach has explanatory virtues as well. However, the NAIRU itself remains unexplained according to this approach. Whether in a constant[71] or a time-varying version[72], the level or evolution of the NAIRU are not explained by factors responsible for it.

The discussion in the literature of the univariate approach reflects the importance of theoretical criteria. The univariate approach is often called "atheoretical" and it is criticised for being conceptually poorly defined.[73] The main idea behind it, though, is consistent with the concept of NAIRU: the idea that unemployment always reverts to its mean or trend over time. Furthermore, the trend can be interpreted as reflecting factors such as hysteresis, *i.e.* a causal interpretation is possible. However, since the causal factors at work are not modelled explicitly, the univariate approach tends to be regarded as not useful in this respect.

---

[68] Gordon 1997, Richardson *et al.* 2000
[69] Gordon 1997, p. 13, emphasis added
[70] Richardson *et al.* 2000, p. 37
[71] See *e.g.* Gordon 1977.
[72] Gordon 1997
[73] See *e.g.* Richardson *et al.* 2000, p. 36 and McAdam and Mc Morrow 1997, p. 5.

So far, all criteria have concerned the assumptions of the specification. But theoretical criteria can also come into play in the interpretation of the results of a measurement procedure. Gordon discusses the results of the contribution of Staiger *et al.*[74] to the same symposium and argues that they make "no economic sense".[75] Staiger *et al.* estimate confidence intervals for their time series of the NAIRU, and their estimate of a 95 per cent confidence interval for the year 1990 for the US is 5.1 – 7.7 per cent. Gordon now suggests that this result is not consistent with the concept of NAIRU. This is because "[i]f the NAIRU had been 5.1 percent since 1987, inflation would not have accelerated during 1987 – 1990, since the actual unemployment rate never fell below 5.1 percent in any calendar quarter. If the NAIRU had been 7.7 percent in the period since 1987, inflation would not have decelerated during 1990 – 93, since the actual unemployment rate never rose above 7.7 percent in any calendar quarter".[76] The relation between deviations from the NAIRU and the inflation rate are thus treated as analytic. That actual unemployment deviates from NAIRU just means that inflation must accelerate or decelerate.

A similar criticism is levelled against the univariate approach. Conceptually speaking, Richardson *et al.* 2000 argue, NAIRU measured according to the univariate methods is likely to be biased relative to the "true" NAIRU when the inflation rate is falling. When inflation decreases this is due to either an increase in the actual unemployment rate or a decrease in the NAIRU. Because univariate methods define the NAIRU basically as a steady or moving average of the actual unemployment rate, the NAIRU is inert. Therefore, univariate methods will produce a measure which is biased upwards.

To repeat a point alluded to above: despite their importance, theoretical criteria are in no case stringent enough to suggest a unique measurement procedure. In fact, at most they suggest or imply a subset of the first set of assumptions, those that regard the specification. Theory informs us about *some* of the variables that enter a specification (wages, unemployment and wage pressure in the structural approach, say; inflation, unemployment and supply shocks in the reduced form approach and unemployment in the univariate approach), but not all of them. Which variables account for the "wage pressure", and which for the "supply shocks" is not implied by

[74] Staiger *et al.* 1997
[75] Gordon 1997, p. 29

micro economic theory or by the concept of the NAIRU. Except some constraints theory might impose on the form of the specification, it is in general silent about it (*e.g.* whether the output gap should be linear or non-linear). And theory has nothing to say about the actual choice of the time series (though in some cases, theory may contain information about for example the choice of the CPI or the GDP deflator as indicator of inflation), and obviously nothing about the period and area of interest.

3.2 Statistical Criteria

Naturally, statistical criteria abound in assessing the adequacy of particular NAIRU measures. They mostly pertain to measurement *results*. The structural model of Layard *et al.* 1991 is able to "explain" the evolution of unemployment in 19 OECD countries, and estimated parameters have the correct signs and are statistically significant:[77]

> The two estimated equations for all 19 OECD countries from 1956 to 1988 are set out in Table 12. Overall, the equations look satisfactory, with the relevant variables being correctly signed and significant. But the crux of the matter lies in their ability to explain the data. Recall that we are using only four time-varying variables and a mere 12 parameters (excluding the country dummies) to explain the dramatic fluctuations in unemployment in 19 countries over some 33 years.

In this context, "explaining the data" of course means that the R^2 value of the regressions is high. If this ability to "explain the data" is a point in favour of the structural approach, a point against it is that results are known to be very specification-sensitive[78]. This poses a problem if either the exact causal structure responsible for NAIRU is not known or if not all of the variables influencing it are measurable or measured. If the estimate is very sensitive to the specification but (a) the true causal structure is not known and thus it is not clear which variables must enter the equations; (b) not all variables are measured or measurable or (c) both, the estimate will tend to be biased. Layard *et al.* for example appear to give more credence to the NAIRU measured by reduced-form approach than the one measured by their own structural approach. In their calculation of the breakdown of the NAIRU change with respect to its causes, they add the influence of *unmeasured*

---

[76] Gordon 1997, p. 29
[77] Layard *et al.* 1991, p. 433

causes to the *measured* causes (*e.g.* unemployment benefits, mismatch, union power). The influence of unmeasured causes is calculated by taking the difference between the NAIRU levels estimated by the reduced form and structural approaches.[79] This line of reasoning implies that Layard *et al.* assume a reduced-form estimate to give the more accurate result.

The reduced-form approach does similarly well with respect to statistical criteria regarding the explanation of the data, correctness of signs and significance of the relevant parameters.[80] In addition, it is sometimes said to be *a priori* likely to be more robust to specification errors than the structural approach.[81]

Another statistical criterion is predictive capacity. Gordon 1997 uses a recursive procedure to test whether his Triangle model can accurately track the inflation process in the US of the 1990s. He finds that it can.[82] A remarkable feature of this criterion, or the way Gordon uses it, is that it cannot distinguish various rather different forms of the Triangle model. Gordon's 1997 time-varying specification fares just as well in tracking the inflation process as his older fixed-point estimate. Staiger *et al.* 1997 find in addition to this that the predictive capacity of the triangle model is relatively insensitive to variations in the point estimate of the NAIRU and aspects of the data specification. For example, if the core PCE (private consumption expenditure) inflation measure is used, an estimated NAIRU of 4.5% predicts an inflation in 1997 of 2.1%, 5.5% predicts 2.2% and 6.5% predicts 2.7%. The predictions are similar when GDP inflation or CPI inflation is used.

Tests of sensitivity to specification changes also fall into the category of statistical criteria. Gordon, for example, tests his model with respect to different smoothness parameters, different time series for the inflation variable and with and without supply shocks.[83] Staiger *et al.* 1996 and 1997 test a huge number of specifications, the changes occurring mainly in the time series for the inflation variables, the in- or exclusion of supply shocks, expectations formations and constancy of the NAIRU. Richardson *et al.*, among other things, test their specifications with respect to the

[78] Cromb 1993
[79] See Mellis and Webb 1997.
[80] See for example Staiger *et al.* 1997.
[81] Richardson *et al.* 2000, p. 37
[82] Gordon 1997, pp. 25f.
[83] *ibid.*, pp. 21ff.

filtering procedure used. As has been said above, the reduced-form approach is usually said to succeed with respect to the sensitivity criterion.

The univariate approach also does well with respect to at least one statistical criterion. According to the Bank of England, a Hodrick-Prescott filtered NAIRU is significantly correlated with movements in inflation.[84]

3.3 Pragmatic Criteria

The last group concerns criteria that relate to the *use* of the measurement procedures. Roughly, they divide into two subgroups, one regarding the complexity of the actual production of a NAIRU estimate and the other regarding policy relevance.

A drawback of the structural approach is its relative mathematical complexity (*e.g.* regarding the number of variables and parameters) and the fact that a number of causes of changes in the NAIRU are unmeasurable or unmeasured. Layard *et al.* reason about their structural model:[85]

> However, Table 18 reveals that in the last two periods, but particularly in the 1980s, we are unable to provide complete explanations for the rise in equilibrium unemployment as estimated by our earlier method of removing the inflation, trade balance, and hysteresis (Δu) effects from the actual unemployment rate. This is surely the result of our inability to capture all the relevant exogenous factors at work... For example, there is a certain amount of evidence that skill mismatch has been a more serious problem in recent years than it was in earlier decades.

Richardson *et al.* add:[86]

> However, a more general specification problem with structural modelling concerns the number and identity of explanatory variables, which is potentially large, and the sensitivity of results to the particular subset of variables chosen for inclusion in the model. This is, itself, an important limitation when the objective is to apply the same specification across many countries.

Here we can see that sometimes criteria mutually constrain and interact. Because structural models use a large number of variables, they tend to be specification-sensitive (which is a statistical criterion). In different countries variables are often differently measured and the availability of data varies. Therefore, the structural

84 Bank of England 2000, p. 85
85 Layard *et al.* 1991, p. 446
86 Richardson *et al.* 2000, p. 35

approach is of limited use when one aims at cross-country comparisons (which is a pragmatic criterion).

The reduced-form approach is generally said to succeed with respect to pragmatic criteria. It uses a relatively small number of variables, all of which are measured in many countries. It is comparatively simple mathematically. These points concern the specification of reduced-form equations at a fairly general level. More specifically, some aspects of the specification are often chosen with its usefulness for policy making in view. Gordon 1997, for instance, argues that one reason for his model to relate the output gap to *price* rather than to *wage* inflation is that the Fed targets that variable rather than the other one.

Again an interaction between two criteria is presented by the recent tendency to provide confidence intervals with the point estimates of the NAIRU.[87] On the one hand, from a statistical point of view it is advantageous to know the probability that the true value of an uncertain variable lies within a specific confidence interval. On the other hand, if the confidence interval is large, the usefulness of an estimate for policy considerations may be limited. The clearest case for such a limitation occurs when the actual unemployment rate falls into the confidence interval. In this case the actual rate may be at, above or below the NAIRU, a situation in which it is practically impossible to base one's decisions on that measure.

The univariate approach is very simple mathematically and estimates are easy to construct.[88] Data availability is an issue of considerably less importance as data from only unemployment series are used. However, as argued by Pichelmann and Schuh 1997, it cannot provide a basis for interventions because the NAIRU is not causally explained, and the univariate approach does not model interactions between economic variables.

To summarise briefly I want to highlight the following points. First, theory underdetermines the measurement procedure. No theory is strong enough to derive the details of the assumptions needed about specification and data. Where theory comes in is usually at the upper end of the measurement procedure, that is, it determines very general features of the specification.

[87] See for example Staiger *et al.* 1996, 1997, Richardson *et al.* 2000 and Pichelmann and Schuh 1997.
[88] See among others Pichelmann and Schuh 1997.

Second, sensitivity tests usually are conducted at the lower end of the procedure, that is, concerning the data series or the filtering procedures employed.[89] Third, theoretical, statistical and pragmatic criteria are often interrelated and they sometimes mutually constrain each other. At best, only if many criteria are used in conjunction it is plausible to hope that they are strong enough to yield all assumptions of the entire measurement procedure.

4 Reality and the NAIRU

How about considerations of reliability and robustness of the kind discussed in the first section of this Chapter? Can the justification strategies discussed so far establish that the NAIRU is a genuine phenomenon?

Let us first briefly discuss Hoover's first argument for the reality of macro economic aggregates because I think it can easily be shown not to be applicable to the NAIRU case.[90] Recall that the argument was that macro economic aggregates may be thought to be real because they can reliably be used to manipulate other macro economic aggregates. If Hacking's "If you can spray them, then they are real" idea should have any force, it must be thought to be valid in the macro economics case too. Hoover's examples include the Federal funds rate which can be used to shift the yield curve and the general price level that can be used to lower real interest rates (at least if changes are unanticipated).

Compelling as Hoover's argument may be in the context of his examples, I do not think it can work for the NAIRU. I know of no attempt to manipulate the NAIRU in order to change the inflation rate. But suppose on an archipelago in the Pacific people have a preference for an inflation rate of ten per cent. Call this archipelago Inflatonia. Although Inflatonia is a constitutional monarchy, people are otherwise very egalitarian. Currently, the people of Inflatonia are discontented because, for some time now, inflation has been well below the target rate. And this is the case

[89] Although Staiger *et al.* 1996 and 1997 consider "literally hundreds of specifications" with different functional forms, models of the expectation process *etc.* Gordon 1997 also tests for different functional forms.

despite the fact that very few workers in Inflatonia are without job. Happily, Amanaki Filimoeiki, the queen's son, returns from a three-year study trip to London, England (it is not known at which college he studied, though), in order to become the islands' minister of finance. During his stay abroad, Amanaki learned about the relationship between the output gap and the inflation rate. Let us further suppose that he missed the classes on monetarism because they were scheduled at the same time as the all-London undergraduate meetings in the philosophy of science, which he preferred to attend.

Urged by his mother to make the people of Inflatonia happy, Amanaki thus sets out to raise the NAIRU to a level above the actual rate of unemployment such that through the Phillips curve mechanism inflation would increase. He would for example increase the generosity of benefits, strengthen the unions, make it more difficult for workers to change the location and the kind of their jobs, complicate commuting between the islands and take other measures to raise the NAIRU and hope that in this way the inflation rate would accelerate. Even in Inflatonia, such an economic policy would be absurd. After all, it would only work in case the NAIRU *given the actual rate* increases. But I do not see how in Inflatonia the mentioned measures should affect the NAIRU without affecting the actual unemployment rate.

Of course, many of the labour market policies that have been conducted in *e.g.* the US, the UK, the Netherlands, parts of Scandinavia and New Zealand in the past ten or so years can be interpreted as attempts to lower the NAIRU. They have been, however, clearly aimed at reducing the actual unemployment rate rather than at using the NAIRU mechanism to influence inflation.

Let us, then, examine how important in the NAIRU case are strategies similar to what we called reliability and robustness. For the sake of brevity, in what follows I focus almost exclusively on the reduced-form approach as an exemplar for the others. I believe that most of the arguments given will apply *mutatis mutandis* to the other two approaches as well.

[90] Hoover's second argument will be analysed below.

Control of Possible Confounding Factors

Confounding factors are causal factors that can either mimic the operation of the phenomenon of interest and thus create "spurious events" or bathe the signal in so much noise that it is impossible to detect it. In Gordon's Triangle model the phenomenon of interest is of course the causal link between the unemployment gap and changes in the inflation rate. Now we may ask whether we can interpret Gordon's study as an attempt to control systematically for confounders. What are possible confounding factors in the case of the NAIRU? The following determinants of inflation are frequently discussed in the literature:[91]

- demand ("demand-pull inflation")
- supply shocks ("cost-push inflation")
- inertia
- the budget deficit
- monetary policy (or "money")
- the prices of import goods in local currency
- wage inflation
- regulation
- central bank independence
- openness to trade
- *etc.*

Of course, at least some of these were suggested by different, and sometimes incompatible theoretical frameworks. But all have somewhere been cited as contributing factors to or causes of inflation. Gordon's model essentially recognises the first three factors as causes: inertia, excess demand, and supply shocks. He operationalises them as follows. Excess demand is proxied by the "output gap" (actual minus potential GDP), and the latter by the unemployment gap which in turn is measured by the deviation of the actual unemployment rate from the NAIRU. Gordon uses the standard Bureau of Labor Statistics all workers time series as a measure of unemployment. Supply shocks are measured by three variables: excess productivity growth, relative import prices and the relative prices of food and energy. Inertia is naturally measured by lags of the inflation rate. Furthermore, dummies for the Nixon price controls are included, which we may classify under the heading "regulation".

The point I want to emphasise is that Gordon's study seems to presuppose that all other possible causal factors are either absent or operate through the variables that he does include. He does not engage in a systematic investigation of whether his model is an adequate representation of the actual causal structure. With respect to the role of money, Gordon recognises this:[92]

> [G]rowth in the money supply is not a unique cause of inflation. What matters is excess nominal GDP growth, which depends not just on the rate of monetary growth but also in [*sic*] the growth in the velocity of money. In a literal sense, the triangle model predicts inflation without using information on the money stock. In an economic sense, this implies that any long-term effect of money growth on inflation operates through channels that are captured by the real excess demand variables.

Compare this statement with Milton Friedman's famous dictum that "Inflation is always and everywhere a monetary phenomenon". There are of course worlds in which these two statements are compatible but in many they are not. I cannot take a stance on the issue here but Gordon seems to *presuppose* that money operates only through the real excess demand variables rather than engaging in a systematic investigation of that question. To be more precise, he presupposes that the contribution of money and its velocity operate through the unemployment gap rather than through excess demand, since he uses a measure of the former and not the latter.

The same point holds for the other variables. The Nixon price controls are not the only factor that can influence inflation under the rubric of "regulation". The budget deficit may influence prices via routes other than the unemployment gap (*e.g.* through financial markets). Furthermore, many changes took place in international trade arrangements in the period Gordon investigates. Surely, all of these factors may have only a negligible influence on prices. But if we want to regard the NAIRU as a genuine phenomenon, it would be good to have an argument that makes this claim

[91] For the first three, see Mankiw 1994, pp. 305-6, for the fifth, see Friedman 1968, for the others, see IMF 2001.
[92] Gordon 1997, p. 18

convincingly. Many of the factors cited here may not be measurable. But again, it could be possible that one can calculate an upper limit—such as Bogen and Woodward's CERN experimenters did—and if a contribution of the unemployment gap remains, we could have an argument that it is real.

To summarise, Gordon controls for *some* confounding factors, mainly those that are accepted by Neo-Keynesian theory. But he does not appear to engage in a systematic attempt to control for all of them. And the problem is not that these other factors are not known. The problem is that these other factors are not liked by a certain theory—and that theory, in turn, does not enjoy widespread agreement among economists.

Reliable Processes

The gist of Richard Hudson's reconstruction of the rise and fall of bacterial mesosomes is the following:[93]

> Again, the strategy of these experimenters is to use empirical considerations wherever possible not only in justifying theoretical pronouncements, but in supporting the experimental procedures used in such justifications.

I think it should be clear from the analysis in the previous section that empirical considerations—though not entirely absent—play a relatively minor role in the overall justification of a measurement procedure.

An example of a kind of empirical consideration I have found is the use of an unemployment series that counts only registrations of married males rather than all unemployed persons as conducted by CBO 1994. The reasoning behind this is that this measure is less affected by changes in demographics because married males have a stronger than average attachment to the labour force.[94] It is known independently and empirically that married males have a strong attachment to the labour force, and therefore, a procedure that uses these data is judged to be more reliable than one that uses more inclusive data.

Again, let us look at Gordon's study as an exemplar. Consider the following of Gordon's remarks:[95]

[93] Hudson 1999, p. 306
[94] Staiger *et al.* 1997, p. 40
[95] Gordon 1997, p. 15

> The equations estimated in this paper use current and lagged values of the unemployment gap as a proxy for the excess demand parameter D_t...

But do we have good reason to believe that the unemployment gap is a reliable proxy for excess demand? Gordon does not let us know. He continues:[96]

> Although the focus here is on using the unemployment gap to predict inflation, the ultimate exogenous demand factor in this model is "excess nominal GDP growth," which is the extent of which growth of the nominal GDP exceeds the growth of potential output. [...] By treating excess nominal GDP growth as exogenous, the triangle model focuses on the inflation process without the distraction of building a model of the determinants of aggregate demand. Admittedly, this simplification sweeps two-thirds of macroeconomics under the rug. Moreover, it ignores channels by which inflation feeds back into the determination of nominal GDP.

Most other aspects of the specification and the choice of data are also justified with respect to criteria other than empirical. Expectations are not modelled because price inertia is compatible with rational expectations. One reason wages are omitted is that the Fed targets price and not wage inflation. The "smoothness parameter" (see above) is determined by using a "smoothness prior": "the NAIRU can move around as much as it likes, subject to the qualification that sharp quarter-to-quarter zig-zags are ruled out".[97] *Etc. Etc.*

To repeat myself: I am far from asserting that empirical considerations are absent from economists' justification strategies. But there appear to be many questions which could be answered with reference to empirical considerations but aren't. If we believe in Hudson's arguments, we might say that whether or not price inertia is compatible with rational expectations is a relatively lacklustre question. The interesting question would be how we can find out empirically how the relevant market participants form their expectations; if they are inert we can use Gordon's specification, but it is certainly possible that they are formed in a way incompatible with the Triangle model.

So much for reliability. And how about robustness? Certainly, economists occasionally do test aspects of their measurement procedure for sensitivity with respect to the data series used. For example, Gordon estimates his time-varying NAIRU using three alternative price indices, the GDP deflator, the personal

---

[96] *ibid.*
[97] *ibid.*, p. 22

consumption expenditures index (PCE) and one of the consumer-price indices, called CPI-U-X1[98]. He summarises the results as follows:[99]

> The time-varying NAIRU series for the PCE deflator and the CPI-U-X1 are quite close to each other prior to 1980; the CPI-U-X1 series for the NAIRU is lower from 1980 to 1990 and higher after 1990. By mid-1996 a substantial gap had opened up between the NAIRU for CPI-U-X1 (5.8 percent) and for the PCE deflator (5.4 percent), with the NAIRU for the GDP deflator in between (5.6 percent). Prior to 1980, the NAIRU for the GDP deflator was generally lower than that for the two consumption price indexes, by as much as half a percentage point in the mid-1970s.

But Gordon neither indicates whether he does or does not think that these results are robust over changes in the inflation series used, nor which of these series is the appropriate one to use.[100] One interesting and relevant question for example would be whether the data series have been collected/constructed independently. Do they come from the same statistical offices, and have they been constructed according to the same principles? Gordon does not engage in questions such as these. But let us turn to the strategy of minimalist overdetermination and the argument from coincidence in more detail.

Minimalism

One important point to mention in this context is that the measurement procedures of both the reduced-form and the structural approach do not have a stable NAIRU as a *result*; rather the existence of a stable NAIRU is *built into them.* Consider the reduced-form approach first. I repeat here the two equations of Gordon's Triangle model:

$$\pi_t = a(L)\pi_{t-1} + b(L)(U_t - U_t^*) + c(L)z_t + e_t, \tag{4.2'}$$

$$U_t^* = U_{t-1}^* + \varepsilon_t, \tag{4.4}$$

[98] For completeness, the CPI-U-X1 is the same as the consumer-price index for urban consumers (CPI-U) except a difference in 1967-1983 in the treatment of the shelter component. See Gordon 1997, p. 21 fn. 13 for an explanation.
[99] *ibid.*, p. 24
[100] The latter point is a fact that might imply that he thinks that they are interchangeably usable. But this seems unlikely given the—in Gordon's judgement—"substantial" difference between the series in some periods.

where π is inflation, U unemployment, U^* the NAIRU, z a vector of supply shocks, L is a lag operator and e and ϵ are error terms. The variance of ϵ determines the smoothness of the NAIRU schedule, the smaller the variance, the smoother the NAIRU. The NAIRU is estimated as that value of U^* which maximises the likelihood of the coefficients, given (a) the "smoothness prior"—the variance σ_ϵ—and (b) a constraint on the coefficients of the inflation lags that assures their sum to be unity. Employing this method of estimating the NAIRU obviously makes sure that there exists such a thing as a "non-accelerating inflation rate of unemployment". U^* is calculated to be that rate of unemployment at which inflation does not change (*i.e.*, the sum of the coefficients on lagged inflation is one; it is not *estimated* to be one but *constrained* to be one).

The same feature can be observed looking at structural models. As we have seen, in this case the NAIRU is calculated as that rate of unemployment where short-term shocks are absent, markets clear and expectations are fulfilled. Thus again, the relation between the NAIRU and inflation is built into the measurement procedure. The variables *change in inflation*, *trade balance* and *change in unemployment* are simply eliminated from the long-run supply-side constraint in order to have an equation for the equilibrium level of unemployment. The equation for the equilibrium level of unemployment and the supply-side constraint are then combined to obtain (4.8), from which the values for the NAIRU can be derived.

Now we may say that the circularity of the kind found in the measurement of the NAIRU is ubiquitous in the sciences. In Chang's account of the "real scale of temperature" the researchers faced a circularity very comparable to the one discussed here. The temperature case belongs to a class of intricate cases which face what Chang has called "the problem of nomic measurement". His account of that problem is as follows:[101]

(1) We want to measure quantity X.

(2) Quantity X is not directly observable, so we infer it from another quantity Y, which is directly observable.

(3) For this inference we need a law which expresses X as a function of Y, as follows: $X = f(Y)$.

[101] Chang 2001, p. 5

(4) The form of this function f cannot be discovered or tested empirically, because that would involve knowing the values of both Y and X, and X is the unknown variable that we are trying to measure.

Our case is exactly parallel to Chang's. The quantity to be measured, X, is the NAIRU. The NAIRU is not directly observable, but another one is: inflation (Y).[102] In order to infer the NAIRU from inflation, we need a law-like relationship such as the Phillips curve (*e.g.* (4.2') & (4.4)). And the form of the Phillips curve cannot be tested empirically because that would involve knowing the NAIRU.

Chang suggests minimalism as a default strategy if there are no reliable auxiliary conjectures.[103] Now we could say that the existence of the NAIRU is our "basic ontological conjecture". Would minimalism work in this case?

At least *prima facie*, there is a case for minimalism in the measurement of the NAIRU. There is one basic ontological conjecture: there is an objective property of certain kinds of economic systems, which is called the "non-accelerating inflation rate of unemployment", *i.e.* in these kinds of economic systems, inflation will not accelerate when the actual unemployment rate is at the NAIRU (providing supply shocks are absent). Different tokens of the same type of measurement procedure must give the same results, that is, they must satisfy the "comparability" test. For the time being I want to take the three different approaches as procedure-types, different specifications within one approach as tokens of that type.

Let us again focus on the reduced-form approach. The univariate approach has for a long time lived only a shadow existence.[104] Comments on the results of the structural approach suggest that it fails the comparability test: results are very sensitive to the exact specifications, that is, different "tokens" (specifications) of the same "type" (kind of measurement procedure) do not give the same results. By contrast, results of

[102] We would probably call inflation measurable rather than observable. But from what I have seen in the economics literature, Chang's terminology is consistent with economists' practice.
[103] *ibid.*, p. 61
[104] The reasons were discussed above. I think it is safe to disregard the univariate approach here because it plays virtually no role in contemporary research. In the literature from 1990 onwards, I have been able to find only one favourable reference, in Bank of England 1999: "Work at the Bank, using different values of the smoothing parameter, has found that HP-[Hodrick-Prescott]filtered NAIRU (based solely on actual unemployment) has been significantly correlated with movements in inflation, suggesting that even a simple approach like this yields reasonable results", p. 85.

the reduced-form approach are often said to be robust under different specifications, or at least more robust than the structural approach.[105]

A number of considerations indicate, however, that minimalism does not work as a strategy for the NAIRU. First, only *some* of the results are *relatively* robust under different specifications. Robustness considerations mostly pertain to some statistical properties of estimation results, for example sign and significance of the relevant parameters or usefulness as a predictor. The focus for the purpose of this paper is, however, on the point estimate of the NAIRU itself, and this seems to be specification-sensitive to a degree comparable to the sensitivity one finds in the structural approach. In Table 1 in the Appendix I reproduced point estimates of the US-NAIRU for 1970, 80 and 90 calculated under different specifications. They all follow more or less the same basic approach, that is, the Phillips curve specification (except the comparison with the survey of UK structural models by Coulton and Cromb 1994, of course). They differ with respect to inflation and unemployment data series, expectations formation and filters used. Below the point estimates the range into which the estimates fall is given for each respective year. I state the ranges including and excluding results from specifications using series for married male unemployed because it has been argued that the NAIRU is different for this group and unemployment rates are lower.

It can be seen from the table that the range of variation is considerable.[106] I do not have a precise criterion that determines how much variation is "within usual measurement error" or "acceptable". But the ranges displayed intuitively seem to be very large. One may ask how useful an indicator with a large error band is for policy considerations. The relevant policy question is, how large is the unemployment gap? If it is large and positive, demand policy may be activated in order to reduce unemployment. If by contrast the actual unemployment rate is near or below the NAIRU cautious or restrictive fiscal policy should be implemented. If there is a large variation in the measurement of the NAIRU, then, policy makers will not be able to

[105] See for example Richardson *et al.* 2000 who state that the reduced-form approach "is *a priori* likely to be more robust to specification errors than the corresponding structural approach" (p. 37) and Staiger *et al.* 1997 who discuss the robustness issue and find that most basic conclusions remain unchanged under different specifications.

[106] I have not stated confidence intervals because (a) they are not always given in the literature, (b) they presuppose the existence of a true value for the estimated variable and (c) they presuppose a reliable measurement procedure. See for example Blalock 1979 on the latter two points. (b) and (c), however, are exactly the issues that are questioned here.

receive much guidance from this indicator. The range of the NAIRU including married males comprises the actual rate in all three years, that is, the output gap may be positive, nearly or exactly zero, or negative. Excluding married male unemployment series, this still holds for 1980. And for the other years, the output gap may still be large or negligible, and each will recommend different policy measures. In this sense, the variation in the measurement of the NAIRU is also large.

The second reason why the strategy of minimalist overdetermination is not applicable to measures of the NAIRU is that it is not the case that one makes only one or a small number of theoretical commitments. Although a great many specifications are tested, especially in Staiger *et al.* 1996, they all follow the same basic theoretical framework, which is broadly Keynesian. As pointed out above, one controversial matter is, for instance, the role of money. According to monetarism, money—and only money—causes inflation in the long run. The measurement procedures of the reduced-form approach, however, suggest that money is not the unique cause of prices. This is clear from the fact that the output or unemployment gap figures as a cause of inflation. Money is not explicitly modelled as a determinant of inflation. Any effect money might have on inflation must be channelled through the real demand variables.[107] But this means that it uses additional theoretical assumptions, which of course may be false. And this in turn means that the strategy is not minimalist.

The third reason is that the "basic ontological conjecture" were one to attempt a strategy of minimalism—the assumption of the existence of a NAIRU—is itself not uncontroversial. Although Gordon 1997 argues that "the triangle is resolutely Keynesian"[108], there seem to be views that are more or differently Keynesian. Galbraith for example argues that for a "real Keynesian" the notion of an aggregate labour market does not even make sense.[109] Eisner 1997, in a similar vein, finds that "the NAIRU has never had any sound base in theory"[110] and gives a number of reasons why the concept is theoretically unsatisfactory from a Keynesian point of view. The point here is of course not to argue for or against the concept, a specific understanding of it or whether and how truly Keynesian it is. The point is the rather

[107] See Gordon 1997, p. 27 on this point.
[108] *ibid.*
[109] Galbraith 1997
[110] Eisner 1997, p. 197

trivial one that the NAIRU does not seem to enjoy a degree of acceptance similar to that of the existence of an objective temperature. And if it does not, the minimalism strategy collapses.

Coincidence

Hacking's argument from coincidence cannot be readily applied to the economics case because he demands that measurement or experimental procedures which are described by "different chunks of physics" give similar results. But the different chunks of physics Hacking refers to explain the behaviour of properties that can often happily co-exist. In Hacking's case we find the different properties of photon beams and electron beams, respectively, and the properties of the different materials with which these microscopes are made.

The "different chunks of economics" often do not describe properties, or I'd rather say economic systems, that can co-exist in a similar way. Keynesian economics is to a large degree incompatible with monetarism, and different versions of Keynesianism incompatible with each other. The Keynesian and the monetarist views of the world describe in fact different worlds. And thus it appears not to make much sense to compare measurement results from different theoretical frameworks.

But we may find independent laws that govern the behaviour of the NAIRU, each of which faces Chang's "problem of nomic measurement", but because their results coincide, credence is given to the measure as well as the statements about the laws into which the concepts enter. The Phillips curve relationship is one that specifies the NAIRU and its effects on the inflation process. Now it is possible that there are laws into which the NAIRU enters as an effect. If these two results coincided regularly I think credence could be given to each method.

Something of the kind is done in Layard *et al.*'s comparison of the results of the structural with the results of the reduced-form approach. As we have seen, the results did not coincide enough to convince Layard *et al.* Furthermore, the inflation process is part of the structural model, as wage decisions are explicitly modelled (and price inflation is often regarded as a direct consequence of wage inflation). But this kind of route seems to me to be the way to go if we wanted to follow Hacking.

5 Operationalism and the NAIRU

Thus far, I hope to have established that empiricist strategies to justify particular measures of the NAIRU (a) in some cases do not play a significant role in the economic literature on the topic, and (b) in some cases cannot play that role because they are not applicable to the NAIRU case. I now want to draw some conclusions from these observations.

Chapter 3 has distinguished two main contenders for a theory of the application of concepts in theoretical science: operationalism and functionalism. Undoubtedly, the NAIRU is a concept whose empirical counterpart is at least not very readily observable. As most of this Chapter is concerned with measurement issues it is natural to try to find a kind of operationalist interpretation of the NAIRU.

5.1 Operationalism: Definitional vs Multiple

In order to do so, I would like to introduce a distinction by Donald Campbell[111] between "definitional operationism" and "multiple operationism".[112] Definitional operationism is the view according to which "operations are regarded as defining terms in a scientific theory"[113]. On the other hand, multiple operationism is a programme which recognises (a) that there is no rock-bottom criterion against which to check our measurement procedures, experiments or tests and (b) that "great inferential strength is added when each theoretical parameter is exemplified in two or more ways, each mode being as independent as possible of the other, as far as the theoretically irrelevant components are concerned".[114]

Paraphrasing Campbell in a way that parallels the language of Chapter 3, we may say that *definitional* operationalism claims that the rules of concept application are fixed by the procedure that measures a quantity represented by some concept; each measurement procedure defines its own concept. *Multiple* operationalism, by contrast, holds that the rules of application are given by the nomic claims into which the concepts enter: the concepts apply to those objects of which the nomic claims in

[111] Campbell 1988/1969
[112] I shall use his term "operationism" only with direct reference to his paper and otherwise stick to the nowadays more common "operationalism".
[113] *ibid.*, p. 31, emphasis removed
[114] *ibid.*, p. 33

which they figure are true. However, our causal and conceptual background knowledge entails that the quantities represented by the concepts of interest are measurable in various ways. Measurements, in turn, allow us to decide empirically to which objects the concepts apply.

Campbell argues that at the time of his writing, definitional operationism was seldom explicitly endorsed by scientists. However, he writes, "[i]t persists in some sense in all those studies complacently satisfied with dependence upon a single method. It persists in those studies implicitly presuming that the resulting measurements are unbiased indicators of the theoretical variable".[115]

If Campbell is right about this, it seems that there are two options in the NAIRU case, corresponding to Campbell's two "operationisms". One can either insist on the fact that there is *one* theoretical quantity[116], which is measured in multiple (theoretically underdetermined) ways. But then we have to accept that there may be a great uncertainty in the determination of that quantity. The degree of uncertainty involved in measurements of the NAIRU in fact approaches levels at which the policy relevance and theoretical significance of the concept becomes suspicious.[117]

Or one can decide that the NAIRU concept is to be interpreted operationally. But then we will face a plethora of problems that are associated with what one might call the "strong programme of operationalism". Famous among these problems is that of the proliferation of concepts. If the identity conditions of a concept are given by those of its measurement operations, concepts would multiply excessively. This can be seen from the analysis of the NAIRU case, too. Essentially, no two studies use the same operation, and thus following this line of argument, no two studies measure the same NAIRU. This, at least, poses an impediment to comparisons of different NAIRUs, be it through time or across countries. It incapacitates communication among scientists.

---

[115] *ibid.*, p. 34

[116] Identity conditions in this case are given by the theory in question, and thus a more Keynesian "Phillips-curve NAIRU" would be a different quantity than a more neoclassical "wage-bargaining approach NAIRU".

[117] See among others Staiger *et al.* 1996 and 1997, Gordon's 1997 remark that "The recent suggestion of Staiger, Stock and Watson (1996) that the NAIRU for the year 1990 could range from 5.1 to 7.7 percent *makes no economic sense*" (p. 29; my emphasis), Pichelmann and Schuh 1997 and McAdam and Mc Morrow 1999.

A related problem that has greater significance from the point of view of this Thesis is that concepts that are operationally defined in this way are difficult to relate to theory. In a discussion of Campbell's views on operationalism, William Wimsatt argues:[118]

> Against all this [the programme of multiple operationalism], suppose one did have only one means of access to a given quantity. Without another means of access, even if this means of access were not made definitional, statements about the value of that variable would not be independently testable. Effectively, they would be as defined by that means of access. And since the variable was not connected to the theory in any other way, it would be an unobservable, a fifth wheel: anything it could do could be done more directly by its operational variable.

According to definitional operationalism, then, theories become partly redundant. I say partly because they are still instructive in the construction of the measurement procedure. But once the procedure is constructed, the theoretical term seems either to reduce to the operational one or to become redundant.

Thus a gap opens between explanatory and applied work. It is not the case any more that the quantity we are aiming to measure is essentially a theoretical quantity, which can be given empirical significance in various ways. Rather, the operation exhausts the concept's meaning. Theoretical work is relevant at best only in the process of constructing the operation.

We have said in Chapter 2 that the explanatory power of a model is given by the ability of its concepts to unify and systematise thinking about phenomena. This work is usually done by means of theoretical models: models that usually represent the operation of a small number of causal factors and that can be carried from situation to situation. However, if the measurement procedures *define* their concepts, no concept will be used in theoretical models across different situations. This is because, as we have seen, different researchers use different measurement procedures.

But suppose they did not. Suppose that there was only one measurement procedure for all NAIRUs at all places and times. In this case the problem of concept proliferation would not arise because there is only one procedure and thus only one concept. However, we have also seen that the methods surrounding NAIRU

[118] Wimsatt 1981, p. 137

measurements cannot establish that the NAIRU is (part of) a genuine phenomenon, *i.e.* that the NAIRU is real. But if we want to say that a model is explanatory, it must tell us something about the world. There is no such thing as a purely instrumental model with explanatory power.

This is the reason why, in Chapter 1, explanatory power has been described as "phenomenal adequacy + *X*". Phenomenal adequacy makes sure that the alleged explanation picks out a real feature of the world. The *X* that is relevant for economics has been identified as the ability to unify and systematise thinking about phenomena.

This argument may cause an immediate objection. As I will discuss in more detail in the next Chapter, Nancy Cartwright[119] has claimed that there is a trade-off between the "facticity" of a scientific law and its explanatory power. Let us take what she calls facticity as the same as phenomenal adequacy. Her claim is then that we can have either phenomenal adequacy or explanatory power but not both.

Part of the answer to this objection is that Cartwright ascribes a structure to physics, which I don't think is true of economics. The structure is marked by the distinction between phenomenological laws and fundamental laws.[120] Phenomenological laws can be interpreted realistically. That is, they can be thought to represent actual states of affairs. By contrast, the fundamental laws of physics "lie": they do not describe facts. But fundamental laws unify and systematise. And the price we pay for this ability is their remoteness from the facts.

I do not see that the same structure is true of economics. There are different kinds of models but they do not differ in the same way Cartwright understands the laws of physics. Models of the NAIRU are not more phenomenally adequate than Akerlof's of Hotelling's models. And it does not seem that the fundamental principles of revealed preference theory, or any other theory of rational economic man, involve abstraction and idealisation to a degree that is different enough from the degree we find in Akerlof's and Hotelling's models in order to justify this distinction.

The other part of the answer to the objection is the role of concepts in the account of explanatory power offered in Chapter 2. Essentially, it is the concepts that do the systematising work. A necessary condition for a model to have explanatory power is

[119] Cartwright 1983, chs 2 and 3
[120] See *e.g.* p. 1.

that it uses concepts that are also used by many other models with which we can capture many different phenomena. But given such an understanding of explanatory power, it is possible that a phenomenally adequate model also has explanatory power. A model captures a real phenomenon. But it does so by means of employing concepts that are used in many other models—and thus allows us to unify and systematise thinking about phenomena.

The point is that Cartwright's fundamental laws of physics would not have any explanatory power were it not either (a) for the fact that there is a level of phenomenological laws which do describe facts, and this level is summarised by the fundamental laws or (b) for the fact that the fundamental laws themselves unify in a way which reflects a "natural classification"[121] of phenomena or (c) for both. It is not the case that *any* theoretical framework which unifies and systematises is also explanatory. But then it is true that if a NAIRU, operationally defined, is not for real, there is a gap between theoretical (explanatory) and applied (econometric) work.

Now recall Kevin Hoover's second argument for the reality of macro economic aggregates. The argument was that in so far as today's main contender of Keynesian macro economics, *viz.* real business cycle models, are empirically successful, they point towards the reality of macro not micro economic entities. If the story told so far is correct, namely that operationally defined concepts cannot be explanatory, Hoover's second argument cannot apply to the NAIRU case (if the NAIRU is taken to be operationally defined). One the other hand, his argument may apply for a NAIRU multiply operationally understood. But what is won with a NAIRU that is real but so fuzzy that it seems unusable in policy making?

5.2 Conventionalism to the Rescue?

One possible way to narrow the gap between empirical and theoretical work is a more thorough conventionalism. Conventionalism defines its own reality. Conventionalism would, in fact, be fairly consistent with much reasoning one finds in economic analyses. Thus one could give greater weight to criteria I have called "pragmatic" above. One pragmatic criterion was that of policy relevance. Gordon 1997 deliberately uses price rather than wage inflation in his specification, for the

[121] This is Pierre Duhem's term, see his 1954/1908.

Fed targets prices and not wages. We can extend this line of argument and say that certain price indices are of policy relevance and not others, ditto for the unemployment series, ditto for other aspects of specification and data. Once constructed, we adopt this measurement procedure *conventionally.* The problems of uncertainty and concept proliferation then do not arise. The numbers relate to theory because we define them to do so.

Indeed I think that reasoning along these lines is what one finds in a number of economic cases. Although I cannot go into details here, I assume that the measurement of inflation and other index numbers provide examples: a large number of seemingly arbitrary decisions must be taken (*e.g.* about the selection of households, goods and outlets, the treatment of quality differences and new goods, which index number formulæ to use *etc. etc.*), and these decisions are conventionally established.

However, eventually conventionalism cannot overcome the difficulties of (definitional) operationalism faces. First, conventionalism presupposes a reasonable level of consensus (all researchers and policy makers working with concepts bearing the same name must agree). Given the immense disagreement one can observe about the NAIRU it does not seem to be a realistic option now or in the near future. Second, even if there were consensus at the academic level, it is unlikely to extend to the government offices where actual decisions about at least some measurements are taken (the US *Bureau of Labor Statistics*, say, or the *Statistical Office* in the UK). As a consequence, one would have at best a standard within national borders. This again makes international comparisons difficult.[122] Third, even assuming that all practical problems could be solved, conventionalism risks that we are conventionally drawing a consistent representation of the world but increasingly the world departs more and more from our representation. At some point somebody might notice. Imagine a policy based on a botched measure of the NAIRU, one which, say, consistently overstates the "true" NAIRU (assuming there is one). Policy making based on such a measure may throw so many people out of work that, even if they do not appear in official figures (because our representation is consistently false), at some point they may create a social problem that is noticeable by means different from official statistics. It seems, then, that conventionalism is a strategy that (a) is not very likely to succeed practically and (b) even in an ideal world it risks being only a short-run success.

To summarise, analysing the justification criteria or strategies that are employed in measuring the NAIRU reveals that economists are facing a dilemma. The dilemma is that between a strategy of multiple operationalism, which solves the problem of relating empirical to theoretical work at the cost of great uncertainty of the measure, and a strategy of definitional operationalism, which provides exact[123] measures but ones that produce "peninsular" concepts that are not very readily related to theory.

6. Conclusions

In Chapter 1 three principal epistemic virtues of economic science were identified: phenomenal adequacy, explanatory power and exactness. Analysing the NAIRU shows another trade-off result between these virtues. The justification strategies found in NAIRU measurements cannot establish its reality. But this means that no model in which the NAIRU figures can be phenomenally adequate because phenomenal adequacy implies that there is a real phenomenon. But we may be looser in our requirements and say that the NAIRU measures are "somehow" robust under different specifications. In this case the NAIRU may be said to be a genuine phenomenon, but it is a very fuzzy quantity.

The same kind of trade-off is present with respect to explanatory power. A concept will only have explanatory power if it helps us to unify and systematise our thinking about phenomena. But in order to do so, concepts must be embedded in some kind of theoretical framework. We have seen that if the NAIRU is understood as being operationally *defined*, a gap opens between theoretical and applied work. NAIRU concepts thus conceived not only proliferate with intemperance, they can also not be brought in touch with other theoretical concepts. On the other hand if we understand

[122] I believe that inflation measurement faces problems of this kind.

[123] For the purpose of the argument I ignore the fact that a large extent of uncertainty already enters a single point estimate, often measured by confidence intervals. It remains true, though, that a single measure with a certain confidence interval is more exact than a wide range of measures, each of which comes with a confidence interval.

the NAIRU multiply operationally, it becomes inexact: it will be difficult to tell, for any given economic system, whether it is in a situation of a positive, zero or negative output gap.

Appendix

Table 4.1

Year		Basic Form	Inflation	UE data	Expectatations	const./time var.	1970	1980	1990
Study	OECD*	Phillips Curve	CPI-U?	all?	adaptive	Kalman, recursive	5.6	6.2	5.7
		Phillips Curve	CPI-U?	all?	adaptive	HP, recursive	4.7	7.2	6.1
	SSW**	Phillips Curve	CPI-U all	all	rec AR(12) fcast	constant	6.41	6.41	6.41
		Phillips Curve	CPI-U all	all	adaptive	2 breaks	5.12	8.81	6.18
		Phillips Curve	CPI-U all	all	rec AR(12) fcast	2 breaks	8.4	8.4	6.02
		Phillips Curve	CPI-U-food	married male	adaptive	constant	3.54	3.54	3.54
		Phillips Curve	CPI-U-food	married male	rec AR(12) fcast	spline, 3 knots	2.25	5.19	3.65
		Phillips Curve	CPI-U, p in levels	all	n/a	spline, 3 knots	7.01	10.78	7.6
		Phillips Curve	CPI-U, lags chosen	all	adaptive	spline, 3 knots	9.35	5.25	5.71
	KSW***	Phillips Curve	CPI-U all	all	adaptive	Kalman, var=0%	6.26	6.26	6.26
		Phillips Curve	CPI-U all	all	adaptive	Kalman, var=5%	6.27	6.55	6.22
		Phillips Curve	CPI-U all	all	adaptive	Kalman, var=10%	6.18	7.13	6.06
		Phillips Curve	CPI-U all	all	adaptive	Kalman, var=15%	6.09	7.87	5.88
	Gordon****	Phillips Curve	GDP	all	adaptive	Kalman, var=0%	6.05	6.05	6.05
		Phillips Curve	GDP	all	adaptive	Kalman, var=20%	6.2	6.4	6.3
		Phillips Curve	GDP	all	adaptive	Kalman, var=40%	6.3	6.55	6.55
		Phillips Curve	PCE	all	adaptive	Kalman, var=20%	6.45	6.45	6.3
		Phillips Curve	CPI-U-X1	all	adaptive	Kalman, var=20%	6.4	6.6	6.25
	Range						[2.25:9.35]	[3.54:10.78]	[3.54:7.6]
	Range (without married males)						[4.7:9.35]	[6.05:10.78]	[5.7:7.6]
UNEMPLOYMENT: January							3.9	6.3	5.3
UNEMPLOYMENT: First quarter							4.17	6.3	5.3
UNEMPLOYMENT: annual average							4.73	6.9	5.62

for a comparison:									
	Coulton and	structural	UK data			**1969-73**	**1974-80**	**1981-87**	**1988-90**
	Cromb 1994	(survey)			**Range**	[1.6:5.6]	[4.5:7.3]	[5.2:9.9]	[3.5:8.1]

*Richardson et al. 2000; figures may be slightly inaccurate as they have been read off from graphs

**Staiger et al. 1996

***King et al. 1995

****Gordon 1997; figures may be slightly inaccurate as they have been read off from graphs

Chapter 5

Natural Economic Quantities and Their Measurement

Chapter 5

Natural Economic Quantities and Their Measurement[1]

Concepts which have no foundation in nature may be compared to those Northern forests where the trees have no roots. It needs nothing more than a gust of wind, or some trivial event, to bring down a whole forest of trees—and of ideas.

Denis Diderot—Thoughts on the Interpretation of Nature

1 Introduction

In Chapter 1 I claimed that this was a Thesis in Baconian topics. As the Chapter argued, an important distinction Bacon drew was that between "right" and "false" abstractions. There are two cases of false abstractions. In the first, the abstract notions refer to something real, but they are vague and obscure. In the second, the abstract notions refer to nothing real at all. We can say that the abstract concepts of Chapters 2 and 3 are "false" abstractions in the first sense. There is no doubt that such things as "transportation costs" or "asymmetric information" really exist. But the concepts are muddled and vague, as we do not know what the quantities they describe they do under what circumstances in reality. Their behaviour is not investigated empirically in a systematic way. I argued in Chapter 4, on the other hand, that the methods I found in the more empirical branch of economics cannot establish that concepts such as the "NAIRU" refer to something real.

In this Chapter I want to take the various threads spun in the previous chapters together and see whether we can fill Bacon's distinction with some content relevant

[1] Earlier versions of this Chapter were presented at a meeting of the LSE-Amsterdam Measurement in Physics and Economics Group and at the 2000 INEM conference at the University of British Columbia, Vancouver. Many thanks to all participants of these meetings for a helpful and stimulating discussion for the issues raised here, in particular my supervisors Nancy Cartwright and Mary Morgan, Hasok Chang and Carl Hoefer in London and Kevin Hoover, Marcel Boumans and Steven Rappaport in Vancouver. The paper presented in Vancouver has been published as Reiss 2001. I also thank four anonymous editors of the *Journal of Economic Methodology* for their invaluable comments. The responsibility for mistakes and misunderstandings remains, as usual, with me.

to the epistemic virtues demanded of contemporary economics. More precisely, I want to draw a distinction between "natural economic quantities" (represented by concepts that "rightly" abstract) and quantities which are not natural (represented by concepts that "falsely" abstract). Most of us probably feel that there is something right about such a distinction, but it is difficult to flesh out precisely what it amounts to. Getting the distinction clearer is the purpose of this Chapter.

My strategy is the following. There are two major traditions in the philosophical literature trying to understand this distinction. In one, it is spelled out in terms of natural laws and causation. According to this approach, a quantity (or property, kind *etc.*) is natural if and only if it figures in laws and/or causal relations in the appropriate way. This approach is also the one advocated by Bacon, Schmoller and Campbell, as we have seen in Chapter 1.

In the other tradition, the distinction is marked by measurability. According to this latter approach, a quantity is natural if and only if it is measurable in the appropriate way.

Below, I shall argue that we need a mix of both requirements. I shall do so by first discussing one version of the natural law theory, and amend and change that theory's distinction step by step in order to fit our requirements. In the course of doing that, I will introduce one version of the measurability theory and again amend and change it. Eventually these approaches will be merged and I'll argue why we need to do so. Before going into that, however, I motivate the following considerations by discussing our three epistemic virtues and their trade-offs.

2 Epistemic Virtues and Their Trade-offs

The discussion in the previous Chapters has shown that some modelling practices found in contemporary economics result in various trade-offs between the epistemic virtues that were identified in Chapter 1. Sometimes, we get concepts that are very exact but they have no explanatory power. At other times we have explanatory but inexact concepts.

It is natural to ask whether similar trade-offs obtain between exactness and phenomenal adequacy or between explanatory power and adequacy. Here I shall briefly discuss some of Carl Menger's ideas, which imply that the first trade-off is possible, and then some of Nancy Cartwright's, who urges the importance of the second trade-off.

Recall that Carl Menger distinguished two orientations of theoretical science: the empirical-realist and the exact orientation.[2] The empirical-realist orientation uses concepts that are formed by classifying observations according to similarities ("real types"). It establishes nomic claims by inductively inferring regularities of co-occurrence and succession ("empirical laws"). The exact orientation uses concepts that are formed by breaking phenomena into their simplest parts ("strict types"). It establishes nomic claims by applying the rules of thinking ("strict laws").

Menger's further reasoning regarding these two orientations and their relation implies that there is a trade-off between the epistemic virtues of economics. Following the empirical-realist orientation, we can have concepts that are phenomenally adequate but lack exactness. Two occurrences of a phenomenon that falls under a real type will never be exactly identical. By the same token, empirical laws will never be necessary and exceptionless. Following the exact orientation, on the other hand, we have exact concepts and necessary laws, but they lack empirical content, and thus can't be phenomenally adequate. Menger said that a methodological fallacy is involved in the attempt to test the claims of exact research against "full empirical reality". That is, for him it would be erroneous to look at a simultaneous realisation of the two conflicting virtues.

The second trade-off, that between explanatory power and phenomenal adequacy, is discussed by Nancy Cartwright in her "The Truth Doesn't Explain Much" and "Do the Laws of Physics State the Facts?".[3] Her argument is roughly the following. We have the choice between either subscribing to the facticity view of laws, which holds that law-claims describe facts about reality, or not to do so. Since law-claims on a literal reading at best can be true only of isolated processes, which are few and far between, they cannot have explanatory power, because explanatory power requires unification of phenomena. If, on the other hand, we want to use law-claims in

---

[2] Menger 1976/1871

different contexts, we have to give up the facticity view. In this case, we can have unifying, that is, explanatory law-claims, but they will not be true of the facts or phenomena. Cartwright summarises:[4]

> Most scientific explanations use *ceteris paribus* laws. These laws, read literally as descriptive statements, are false, not only false but deemed false even in the context of use. This is no surprise: we want laws that unify; but what happens may well be varied and diverse. We are lucky that we can organize phenomena at all. There is no reason to think that the principles that best organize will be true, nor that the principles that are true will organize much.

David Kreps seems to observe a similar trade-off in contemporary economics. He argues,[5]

> The standard acid test is that the theory should be (a) testable and (b) tested empirically, either in the real world or in the lab. But many of [Kreps's] models and theories... have not been subjected to a rigorous empirical test, and some of them may never be. Yet, I maintain, models untested rigorously may still lead to better understanding, through a process that combines casual empiricism and intuition.
>
> By casual empiricism joined with intuition I mean that the reader should look at any given model or idea and ask: Based on personal experience and intuition about how things are, does this make sense? [...]
>
> Imagine that you are trying to explain a particular phenomenon with one of two competing theories. Neither fits the data perfectly, but the first does a somewhat better job according to standard statistical measures. At the same time, the first theory is built on some hypotheses about behavior by individuals that are entirely ad hoc, whereas the second is based on a model of behavior that appeals to your intuition about how people act in this sort of situation. I assert that trying to decide which model does a better job of "explaining" is *not* simply a matter of looking at which fits better statistically. The second model should gain credence because of its greater face validity, which brings to bear, in an informal sense, other data.

In both cases, the authors subscribe to particular views of explanation: Cartwright relates explanatory power to the ability of nomic claims to unify phenomena; Kreps seems to endorse an account that relates to intuition. But the point is clear: in many cases we can have either explanatory power or phenomenal adequacy but not both.[6]

---

[3] in Cartwright 1983

[4] *ibid.*, p. 52f.

[5] Kreps 1990

[6] Hartmann 1997 makes a similar point, though regarding a trade-off between empirical adequacy and understanding. The way Hartman understands these two conceptions, interestingly, is very similar to Kreps's.

However, Cartwright's and Kreps's analyses conflict with the discussion of Chapter 1. That Chapter suggested that explanation and description could be two sides of the same coin. For Schmoller, any description uses *concepts*. But since our causal and nomological knowledge is part of the meanings of our concepts, any description will be explanatory. There are at least two strategies to handle this divergence. Either one could refute Cartwright's and Kreps's argument by showing that their trade-offs do not really obtain in the way they see them. Or one could show that they are a product of their respective accounts of explanation and/or description, and that the trade-offs can be diminished for other accounts of explanation and/or description. I will follow the latter path. With respect to the requirement of exactness, by contrast, I will try to argue that Menger's requirement is overly restrictive from the point of view of the general Baconian outlook of this Thesis.

First, then, I want to introduce a conception of *natural economic quantity* (NEQ), which is motivated by these considerations. The main idea is that if one incorporates nomological and causal knowledge into one's concepts one can realise all three epistemic virtues at once (for *some* account of explanatory power, phenomenal adequacy and exactness), though the extent to which one will achieve this will depend on how well-behaved the objects or properties of interest are. The consecutive section will show how NEQs realise the epistemic virtues.

3 Natural Economic Quantities

The distinction I want to propose in this chapter, *viz.*, the distinction between *natural economic* quantities and quantities which are not natural, follows two traditional philosophic debates. Although I believe that these two debates roughly share the same goal, they are discussed in very different contexts. The two debates are that about the nature of properties and that about the distinction between primary and secondary qualities, respectively. The goal I believe they share is the clarification of the relation between reality and our means to know it. The two different contexts in which this goal is discussed are that of metaphysics and the philosophy of language, and that of the nature of our sense impressions, respectively. In the first case the relation is the between an entity of our language—a predicate—and its referent—an aspect of reality: a property or an object. In the second, it is that between an aspect of the content of our consciousness—an impression or idea—and a characteristic of a real object—a quality.

Let us look at the debate about the nature of properties first. In what follows I will exclusively focus on one kind of theory of properties, *viz. natural property theories*. According to natural property theories some but not all predicates refer to real or "natural" properties. The details of which predicates refer to a natural property are, in turn, fleshed out in terms of causal relations and laws of nature. According to these views, then, those and only those properties are "natural" which figure in causal relations and/or laws of nature. These views are represented by, among others, David Armstrong 1997/1992, Hugh Mellor 1997/1991 and 1995, Hilary Putnam 1975 and Sidney Shoemaker 1997/1984.[7] In what follows I shall analyse Shoemaker's doctrine in more detail.

3.1 Schoemaker and Natural Properties

Shoemaker examines the view according to which there is a property corresponding to every predicate. He does not begin his discussion with the difficulty referred to above, *viz.* that this view implies that there are Platonist-abstract things, but rather with another implication of it, *viz.* that according to this view the properties of an object can change without there being genuine change in the object itself. If "being married to Socrates" counts as a property, Xanthippe ceases having it the moment Socrates dies although, presumably, nothing within Xanthippe changes.

Shoemaker introduces the terms "Cambridge property" and "mere Cambridge property". This follows a criterion Peter Geach calls the "Cambridge criterion", which reads: "The thing called '*x*' has changed if we have '*F*(*x*) at time *t*' true and '*F*(*x*) at time *t*'' false, for some interpretations of '*F*', '*t*' and '*t*''".[8] Cambridge *changes* thus comprise real changes (which according to Shoemaker occur within the changing object) as well as changes of the kind that make Xanthippe a widow. Mere Cambridge changes are those latter kind of changes. Shoemaker applies this

[7] See Armstrong 1997/1992, Mellor 1997/1991 and 1995, Putnam 1975 and Shoemaker 1997/1984.
[8] Shoemaker 1997/1980, p. 229f.

terminology to *properties* as well. Cambridge properties, accordingly, are both real and mere Cambridge properties, and mere Cambridge properties[9]

> will include such properties as being 'grue' (in Nelson Goodman's sense), historical properties like being over twenty years old and having been slept in by George Washington, relational properties like being fifty miles south of a burning barn, and such properties as being such that Jimmy Carter is President of the United States.

Shoemaker's theory of ("genuine", "real" or "natural") properties is the following:[10]

> *[P]roperties are clusters of conditional [causal] powers.*

In order to understand this idea and its details I will discuss Shoemaker's own example in which the property under question is "being knife-shaped", and during that discussion I will also illuminate Shoemaker's use of the concepts involved.

A causal power, according to Shoemaker, is a function from circumstances to effects:

$$CP =_{df} \mathrm{f} : \mathbf{S} \rightarrow \mathbf{E},$$

where *CP* is the causal power, **S** a set of circumstances and **E** a set of effects. The causal power to cut wood, thus, is a function from the circumstances of, say, the presence of a knife, a piece of wood, suitable pressure and the "right" application of the knife to the wood to the effect of the wood being cut.[11]

Things have *conditional* causal powers in case the ability to exercise the power depends on the presence of certain properties. An object *a* has the property *r* of being

[9] *ibid.*, p. 230. Although I see that both in case of changes as well as in case of properties Shoemaker wants to draw a distinction between something genuine or real or intrinsic and something which is not, I do not see the exact parallelism between changes and properties. A theory of genuine changes could cut across changes very differently than a theory of genuine properties cuts across properties. This is even true for Shoemaker's own theory. Imagine there is a property (in Shoemaker's sense a cluster of conditional causal powers), which is had by all particulars for all time. This property could never enter a Cambridge change but would still qualify as a genuine property for Shoemaker. Nonetheless I shall follow Shoemaker here and call bogus properties and changes "mere Cambridge".

[10] *ibid.*, p. 235, emphasis added

[11] For Shoemaker the relata of causation are events.

knife-shaped. This property gives it the causal power P to cut wood only in the presence of other properties $\mathbf{Q}$, *viz.*, those of being made of steel and being of an appropriate size. More formally, Shoemaker's condition reads:[12]

> a has P conditional on $\mathbf{Q}$ if a has r such that having $\mathbf{Q}$ and r are causally sufficient for having P while having $\mathbf{Q}$ only is not sufficient,

where a is an object, P a causal power, $\mathbf{Q}$ a set of properties and r a property. Hence, the object has the power to cut wood conditional on a set of properties (being made of steel, being suitably sized *etc.*) if it is knife-shaped (r) and being knife-shaped together with the properties in $\mathbf{Q}$ is causally sufficient for the power to cut wood while having the properties in $\mathbf{Q}$ only isn't.

The property of being knife-shaped, then, is identified with a cluster of such conditional causal powers, say the power to cut wood and fish bones, to crack oysters, to carve fillets and to inflict pain on my pet. It is important to note that the "cluster" includes *all* conditional causal powers for Shoemaker.[13]

This formulation is blatantly circular, and Shoemaker acknowledges this:[14]

> It will not have escaped notice that the account of properties and property identity I have offered makes free use of the notion of a property and the notion of property identity. It says, in brief, that properties are identical, whether in the same possible world or in different ones, just in case their coinstantiation with the same properties gives rise to the same powers. This is, if anything, even more circular than it looks. For it crucially involves the notion of sameness of powers, and this will have to be explained in terms of sameness of circumstances and sameness of effects, the notions of which both involve the notion of sameness of property. And of course there was essential use of the notion of a property in my explanation of the notion of a conditional power.

We need to know what a property is in order to know what a conditional causal power is, and the definition of a property makes use of the concept of conditional causal power. I believe that the circularity is vicious because it does not solve Shoemaker's original problem of demarcating intrinsic from mere-Cambridge properties. Let us look at an example. Shoemaker called "being slept in by George Washington" a mere-Cambridge property. However, we can certainly imagine that this property, in conjunction with other properties (*e.g.* "being such that people

---

[12] *Cf. ibid.*, p. 234
[13] *ibid.*, p. 240
[14] *ibid.*, pp. 242f.

believe that one was slept in by George Washington") has the causal power of achieving higher prices at auctions. Shoemaker discusses this objection: "Beds that were slept in by George Washington may command a higher price than those that lack this historical property, but presumably this is a result, not of any causal potentialities in the beds themselves, but of the historical beliefs and interests of those who buy and sell them".[15] I think this is correct, but Shoemaker begs the question with respect to the problem he set out to solve, *viz.* to find a criterion that demarcates properties intrinsic to the object of interest from those that are not. We know exactly as much as before: the crank property is not real because it is not a property of the object itself—but we need to know what it means to be an intrinsic property of an object in order to tell what causal powers the object has.

Again, Shoemaker seems to acknowledge this problem. He introduces a distinction between genuine and mere-Cambridge *powers* by means of an example which is due to Robert Boyle. In Boyle's example there is the key which, *inter alia*, has the power to open locks of a certain design. It happens that it also opens the lock on the front door to my house. But whereas the former power is genuine—one needs to change something *in the key* in order to incapacitate it—, the latter is mere-Cambridge since a change on the door or its particular lock would suffice. In order, then, to flesh out the difference between genuine and mere-Cambridge *properties*, the account may refer only to genuine *powers*, but the latter are of course identified in terms of genuine properties. However, Shoemaker is not troubled, for[16]

> the notion of a property and the notion of a causal power belong to a system of internally related concepts, no one of which can be explicated without the use of the others. [...] And it is perfectly possible for a 'circular' analysis to illuminate a net-work of internal relationships and have philosophically interesting consequences.

Shoemaker says that "There is no such thing as tracing a property through a series of changes in its causal potentialities—not if it is a genuine property, *i.e.*, one of the sort that figures in causal laws". But unless the notion of a causal law is made more precise, there might be such changes in causal potentialities. Suppose that "having the colour red" is a property.[17] Among its conditional causal powers is that whose

[15] *ibid.* p. 241

[16] *ibid.*, p. 243-4

[17] I don't see why one shouldn't. There is certainly a cluster of causal powers that could identify it—appearing red to a competent observer under normal daylight, enraging a bull, making cars stop in

effect-event is a certain quale experienced by an observer. But this quale is certainly not only an effect of the object's being red but also of the observer's perceptual apparatus. A "competent observer's" perceptual apparatus, on the other hand, might change as a consequence of evolutionary pressure they have undergone. And thus the property might have undergone a change in its causal potentialities, too. We may of course define the property red by its physical characteristics (those that Mary *did* know). But in this case we will have to find a distinction does the job of the primary-secondary distinction, and we will discuss this below.

Let us now turn to Shoemaker's motivation for holding such a theory. His argumentation is essentially the following. Imagine there were properties in the world which would be completely causally inert. Since learning about a property will always involve a causal interaction between the property and the observer—be it relatively direct in observation or mediated by a scientific instrument in measurement or an inference which is based on the knowledge of some correlation—one could never know about the existence of this property. But properties that could never be known of do not add to our understanding of the world, and thus we can safely discard them.

I share Shoemaker's general epistemological concerns, but I also think that he does not take them seriously enough. It is arguably a necessary condition for the possibility of knowing a thing that it is causally active in some sense. But this is hardly enough. First, it would be good to have strategies available which tell us how to learn about the thing and not only what real things *are*—especially if what they are is motivated by epistemological concerns. Second, it seems that although Shoemaker aims at reducing the number of properties in the world, he does not go far enough. For there are many clusters of conditional causal powers, most of which will be of absolutely no interest for us.

Shoemaker does briefly discuss these two issues. For reasons that will appear promptly, let us discuss the second point first. On Shoemaker's account it is true that all properties are clusters of conditional causal powers, but it is not true that all clusters of conditional causal powers are properties. This is because one could

---

front of a traffic light *etc.* One might object that laws describing these powers are not "proper" laws. But this is the point I want to make—unless Shoemaker presents a more precise concept of causal randomly conjoin conditional causal powers, which, intuitively, would not make genuine properties. Shoemaker's example is that of an object (a knife made of wax) that has both the power of cutting wood conditional on being knife-sized and made of steel and that of being malleable conditional on being at a temperature of 100°C.[18] But these powers are not parts of any common property. Therefore, Shoemaker adds to his theory the condition that[19]

> conditional powers X and Y belong to the same property if and only if it is a consequence of causal laws that either (1) whatever has either of them has the other, or (2) there is some third conditional power such that whatever has it has both X and Y.

Causal laws, thus, determine which conditional causal powers are joined to form properties.[20] There are, so to speak, natural clusters of conditional causal powers, and artificial ones like the cluster of the powers to cut wood and melt at 100°C. Although I have not been able to find a counter-example to Shoemaker's theory from the point of view of common objects and their properties (Shoemaker's knife, Locke's gold, the Cartesian piece of wax, say), I am slightly sceptical about the application to more scientific contexts.

Consider our model from Chapter 2. Akerlof wants to establish the claim that, *ceteris paribus*, the introduction of asymmetric information decreases market size. Let us assume he is right about that. So could we say that: there is an object a (a market? an economic system? a causal structure? a mechanism?), which has P (the causal power to decrease market size?) conditional on **Q** (the absence of interfering factors? other things being equal? other things being right?), and since r (asymmetric information?) is a cluster of such causal powers, it is a property?

It seems the analogy is very loose at best. It would require much speculation to get the metaphysics right, more than we have space for here.[21] But it appears that at least

---

law, one has no reason not to suppose that they are not proper laws. *Cf.* also p. 239, where Shoemaker suggests that "green" could be a property.

18 *ibid.*, p. 245

19 *ibid.*, p. 246

20 In fact, this is all that laws do: "causal laws can be viewed as propositions describing the causal potentialities of properties", p. 244.

21 I have a hypothesis regarding that metaphysical question, though. I think that a genuine metaphysics for economics presupposes a kind of systems or mechanisms ontology rather than either a substance or a process ontology. There are certain things characteristic of economics which are quite clearly not objects such as the stock or quantity or flow of money, and other things which are quite clearly not processes such as any ordinary economic good (although in both cases there may be ways

we would need to allow negative or relational properties if we wanted to save Shoemaker's account in the light of economics cases (the fact that causal factor *F* does not interfere with the working of the structure or mechanism to bring about the decrease in market size doesn't seem to be a causal power of that structure of mechanism). But this would run counter to Shoemaker's original starting point: genuine properties are intrinsic powers of objects. Thus I would suggest a different scheme: let us take the requirements and methods of economics seriously while at the same time keeping in mind Shoemaker's motivation and solution, *viz.* to find a distinction between real and crank properties and that distinction has to do with laws and causation.

The first point was that it would be nice to have epistemological strategies telling us how to find out about properties. Shoemaker hints at such a strategy. Since for him causal powers cluster in a unique way (see the above conditions), we can find out about properties by determining which causal powers are joined regularly, label them and then find out which other causal powers are in the cluster as well.[22] However, this seems a very crude strategy, and I think there are better strategies available. I shall discuss some below.

To summarise briefly, I think that while Shoemaker's theory makes the notion of property a great deal clearer, he fails to deliver on his primary concern, *viz.* to show how genuine (as opposed to mere Cambridge) properties add to our understanding of the world. Importantly, Shoemaker's theory shows that the concepts of properties and those of laws and causation are intimately connected. On his view, properties just *are* clusters of conditional causal powers. But laws enter at two levels. First, causal powers are in fact *ceteris paribus* laws for Shoemaker. As mentioned above, he regards them as functions from circumstances to effects, but this is one way of formalising a deterministic *c.p.* law. A thing that has the causal power to cut wood will—if the circumstances are right ("*ceteris paribus*")—cut wood. Second, laws tell causal powers how to associate into properties. This fact makes it possible to have empirical strategies to find out about properties.

to force these entities into either metaphysics). The biological mechanisms ontology of Machamer *et al.* 2000 is, I believe, a step in the right direction for economics too.

[22] *ibid.*, p. 246

Properties and laws are intimately connected concepts. However, knowing this fact is almost futile unless we have carefully worked out empirical strategies to find out about those laws and what properties figure in them. Shoemaker has sketched a crude way to understand such a strategy. Further below I want to spell out in more detail what kinds of empirical strategies might be helpful in the economics case. But first, let us take Shoemaker's account and discuss its applicability to economics.

3.2 Moulding Shoemaker's Theory Into an Economic Frame

Summarising what has been said so far, Shoemaker's theory can be stated as follows:

(NP1) *A property is natural if and only if the laws of nature fix a unique cluster of conditional causal powers, which is the natural property.*

Economists and economic methodologists typically speak of models rather than laws. In my usage of the terms, we cannot simply substitute the language of laws with that of models. A model, I take it, is at best a *representation* of a law in the sense that "$mg = m\mathrm{d}^2s/\mathrm{d}t^2$" is a model (an either linguistic or abstract entity) that represents the law of free fall (a fact or pattern of facts or power that obtains in the real world).[23] Now, the move from laws to models brings with it a move from an objective level to a subjective level: from the world to ways we talk about or represent the world. For epistemological reasons such a move towards subjectivity may be conceded. Whether or not we have a causal model is much more readily knowable than whether there is a law. Some objective content, however, we should saved by a requirement that the model has certain properties which can be objectively verified (see below).

In the language of models the condition then reads:

---

[23] The usage of "law" as referring to something *real* (as opposed to something linguistic) is not shared by all commentators. Most writers in the tradition of logical empiricism but many others as well usually mean by law a *statement* (see Campbell 1957/1922, Hempel 1966 or Goodman 1983 among many others). Shoemaker uses "law" in this sense, too. A number of realist philosophers, however, draw a distinction and mean by "law" something in the world and use for the linguistic entity "description of law" or "statement of law" or something similar. See for example Mellor 1995. I follow this latter usage.

(NP2) *A property is natural if and only if there are (good? true? adequate? tested?) models that describe a unique cluster of causal powers, which is the natural property.*

What kinds of models do we find in economics which ascribe causal powers to some of the properties whose behaviour is described by the model? This is a difficult question as many economic models either directly use causal language or may be interpreted in a causal way. Just look at a couple of examples, one from theoretical economics and one from econometrics. Both come from sources that have been discussed before.

De Meza and Lockwood summarise their contribution to the literature on the property rights approach to the theory of the firm with these words:[24]

> The *mechanism at work* is that the distribution of property rights over these assets *determines* the bargaining power of agents over the returns to investments which *enhance* the productivity of these assets, which in turn *determines* incentives to invest. The ownership pattern that maximises aggregate surplus then defines the scope of the firm.

I emphasised the terms in the quote which can be causally understood. We can paraphrase De Meza and Lockwood by saying that the distribution of property rights *causes* bargaining power over returns, the returns *cause* asset productivity to increase, and productivity *causes* investment behaviour. The whole situation is a *causal mechanism* responsible for the scope of the firm.

The example from econometrics is taken, not surprisingly, from the NAIRU case. One of the implications Robert Gordon draws from his triangle model is that[25]

> ... since excess nominal demand is the ultimate *cause* of inflation, a sensible anti-inflation policy should target this variable in a direct way.

But why should we believe these authors? The evidence presented by De Meza and Lockwood is a derivation of their results from assumptions which are, at best, strongly idealised descriptions of real relations. But there are no suggestions that even if they are true descriptions De Meza and Lockwood get the causal order right.

[24] De Meza and Lockwood 1998, p. 361, emphasis added
[25] Gordon 1997, p. 18, emphasis added

The "determines" from the quote might as well be read functionally. Turning to Gordon, he presents statistical evidence that inflation is indeed correlated with the unemployment gap (which he takes as a proxy for excess nominal demand). But correlation is not causation, and the causal order might just be the reverse, or there may be a common cause.[26]

A common way to represent causal relations in economics is by introducing structural equations. Kevin Hoover has developed such an approach.[27] His main idea is that interventions that change the data generating process should, in some cases, are reflected in the measured data in a way that is indicative of the underlying causal structure. Let me explain this by means of an idealised example.

Let us assume that there is a structure in which death penalty rate (D) and crime rate (C) are causally connected. The true data generating process is the following:

$$C = \alpha D + \epsilon \tag{5.1}$$

$$D = \beta + \eta, \tag{5.2}$$

where $\epsilon \sim N(0, \sigma_\varepsilon^2)$, $\eta \sim N(0, \sigma_\eta^2)$, $\mathrm{cov}(\varepsilon, \eta) = 0$, $E(\varepsilon_t \varepsilon_s) = 0$ and $E(\eta_t \eta_s) = 0$ for $t \neq s$.

This data generating process is unobservable. However, Hoover shows that if there interventions which upset the process of one and only one variable, changes in the measured correlations can be informative about the true causal structure.[28] If, for instance, an intervention in the data generating process of D occurs (a change in the legislation, say), a structural break in the marginal distribution of variable D, f(D), will be observable. If (5.1)-(5.2) is the true causal structure, a break will also occur in the following distributions: the distribution of D conditional upon C, f($D|C$), and the marginal distribution of C, f(C). However, the distribution of C conditional upon D will remain stable. If, on the other hand, in fact C causes D, f(C), f($C|D$) and f(D)

26 I am far from suggesting that there is anything wrong with either De Meza and Lockwood's or Gordon's models. Evidence about the correctness of their models with respect to causal structure may well be had. The point I am making here is that these models by themselves are not sufficient to make causal claims.

27 Hoover 2001

will break while $f(D|C)$ will remain stable. In a case of mutual causation all four distributions will be unstable, and in case of mutual independence the two distributions on which the intervention occurs will be unstable. Inference of causal direction according to Hoover's approach relies on what James Woodward calls *coefficient invariance*: the stability of a causal structure under changes in the coefficients.[29]

By no means I want to suggest that Hoover's approach is the only one, or even the best. But it seems that one needs a method *of this kind* if we follow Shoemaker's approach in general but substitute "laws" with "models". The reason is simply that a law-claim brings its truth with it, a model doesn't. Shoemaker, whose motivation is primarily to establish relations between important philosophical concepts, can say that there are laws that are responsible for causal powers to cluster this way and not otherwise. Our motivation, on the other hand, is methodological. Therefore, we talk about models, and among these a sub-species for which methods have been developed that allow the inference of causal relations. I call these models *causal models*.

The next two adjustments I want to make are primarily terminological. Shoemaker talks about properties. The causal models of the three above approaches use variables, which can be thought of as representing *quantities*. From now on I shall thus talk about natural quantities rather than properties.[30] The other adjustment is that in causal models variables stand for causes or effects rather than "causal powers". I also adopt this terminology. With these changes the condition reads:

(NQ1) *A quantity is natural if and only if there are (good? true? adequate? tested?) causal models in which the quantity is represented as either a cause(-variable) or an effect(-variable).*[31]

---

28 See *ibid.*, ch. 8.

29 See *ibid.* for a derivation of these results from the assumptions about the data generating process and Woodward 1997.

30 A quantity may be thought of as a property which comes in degrees. Since, however, any quality (which does not come in degrees) can be transformed into a quantity by means of a function that assigns the quantity a value of 1 just in case the quality is present and a 0 just in case it is not, in making this move there is no loss in generality.

31 In Shoemaker's account it looks as if only *causes* ("causal powers") can enter a property, not effects. For him, however, "causal powers" include effects. "Being made of wax", for example, could

I still need to defend the choice of *causal models*. What makes causal models special? Why would not any economic model do? For Shoemaker the motivation was epistemological. He sees properties and causal powers as intimately connected because properties that were causally inert would make no difference to our understanding of the world. Our motivation was to get models that have the virtues of explanatory power, phenomenal adequacy and exactness simultaneously. So what is the contribution of causal models to the realisation of these virtues?

I believe that what is special about causal models is that, under certain conditions, they can be informative about reality even if they do not get the causal structure of the situation they are describing exactly right. Let me explain how this is possible, and then what distinguishes causal models from other models.

I need to introduce a different kind of invariance in this context. Recall the relationship between death penalties and crimes:

$$C = \alpha D + \epsilon. \qquad (5.1)$$

We may find at some point that, contrary to our initial assumptions, the crime rate is actually caused not only by the death penalty rate but also by another factor, "education". Let us suppose education is measurable, and factor it in in the following way:

$$C = \gamma D + \delta E + \nu, \qquad (5.3)$$

where E is the education variable. In some cases, now, the coefficient on D will remain unaltered after a change like this, *i.e.*, $\gamma = \alpha$. In these cases we can say that

have the causal power to be left a mark on under suitable conditions. Therefore, again, there is no change in generality here.

there is a stable causal tendency of D to affect C. Let us call this kind of invariance *capacity invariance*.[32]

The remarkable characteristic of capacity invariant models is that they can sometimes tell us something significant about a real feature of the economy without getting the causal structure exactly right. This characteristic appears to be of considerable consequence in situations of causal complexity. It is very plausible to assume that economic phenomena occur in isolation only in very rare circumstances.[33] Hence a requirement of models to mirror the real causal structure exactly may be overly restrictive. It can be overly restrictive because either we are not able to include all factors in our models (for example because they are not measurable or we have not thought of them), or because we do not want to include them as they are not factors of theoretical interest. If, in such cases, we can measure the stable causal tendency of those factors that we do include, we can describe adequately at least part to the whole complex phenomenon. We can describe what the known factors or the factors of interest *contribute* to the whole situation. In the above example, we may know that there is an army of other factors responsible for the crime rate. If our model has capacity invariance, though, we can represent what the known factors contribute to the crime rate. We have seen above, in Chapters 3 and 4, that many models in theoretical economics (*e.g.* the Hotelling model) and at least some models in applied economics (*e.g.* the NAIRU models from the structural approach) are *not* capacity invariant in this sense. Results in these kinds of modelling traditions, and this in econometric models often includes coefficient estimations (which is the relevant result here), are very sensitive to the assumptions made at the outset. But these assumptions describe which causal factors to include and which to exclude. Let us call models that are capacity invariant *causal tendency models*. Our naturalness condition now reads:

(NQ2) *A quantity is natural if and only if there are (good? true? adequate? tested?) causal tendency models in which the quantity is represented as either a cause(-variable) or an effect(-variable).*

---

[32] Cartwright 1989, ch. 4, argues that capacity invariance is a feature of econometric models which is generally presupposed by economists if they are to be used for prediction and planning.

Above I said that it is important to have trustworthy strategies at one's disposal following which we can find out about our natural properties or quantities. With respect to causal tendency models, we are in the lucky position that a host of empirical methods that can be thought to help in establishing causal tendency claims have been developed and analysed philosophically in recent experimentalist philosophy.

For example, as we will see in detail below, Jevons estimates the contribution of gold to the increase of the price level by comparing the price levels of two years in which the trade activity was at the same point of the cycle. Woodward calls this method "uniformity of background". We can measure the contribution of a factor if we can control that factor while other factors operate uniformly. In Jevons's case we do not directly control the stock of gold, but we know that it has risen as a consequence of the Australia and California findings. If only gold and trade cause prices, and in the two years compared the trade activity was indeed the same, the change in prices occurred between the two years will be due to gold.[34]

Jevons's reasoning relies on the capacity invariance of his model. Unless there is the stable tendency of gold to cause prices, his argument won't succeed. In this case we can see how closely related measurement and causal modelling may be. We cannot measure monetary (or gold) inflation unless we have a capacity invariant model. But we will not be able to test whether gold actually causes prices (according to Hoover's methods, say) unless we can measure it.

Unlike Jevons, most contemporary applied economists tend to focus on data *analysis* rather than the construction of measurement procedures.[35] Typically, the approach that is taken is what one might call Aristotelian-contemplative. If the causal relation between two variables X and Y is at stake, usually economists obtain data from statistical offices such as the *Bureau of Labor Statistics* or the OECD or the IMF and

---

[33] According to Hacking 1983 and Cartwright 1983, for example, this is also true of the phenomena of physics.

[34] As we will see, Jevons in fact uses an average of his price level variable through the cycle for this calculation. This is due to the fact that in his view we cannot reliably determine where exactly in the cycle we are. My reasoning presupposes that we could determine that point with considerable accuracy.

[35] This is only regarding *direct* measurements. In case of indirect measurements (such as the NAIRU), much attention is spent on the construction of procedures.

regress X on Y or vice versa. If the data series appear not to measure variable X, say, but some other variable which is for theoretical or other non-empirical reasons believed to be correlated with X, one takes it as a proxy for X. Except in cases where the capability of a procedure to reliably measure a variable is seriously questioned—such as in the controversy around the issue of whether the consumer price index is a reliable measure of the variable *cost of living*[36]—the focus is on what one does with the data rather than how the data have come about. This may have to do with an academic division of labour between econometricians and statisticians, but it appears odd in the light of the general Baconian theme of this Thesis. One of the metaphors Bacon used (see Chapter 1) was that of science as a uneven road. He said,[37]

> Such a road [of scientific investigation] is not level, but rises and falls; first ascending to axioms, then descending to works.

Chapter 3 argued that in economics the ascent to axioms, or what we would call "economic laws" does not resemble very much the Baconian model. The claim here is that the descent to works doesn't resemble this model much either.

But accurate measurement matters if we want to test for causal relations. In order to see that, consider a case such as David Lewis's where students' philosophical abilities are measured using the examination grid for the French exam. Any model that relates philosophical abilities with some other variable but uses data obtained from this measurement procedure will not be very informative.

Or consider the simple causal structure in Figure 5.1. In this structure, an intervention I and taxes T cause government spending G, which in turn causes some quantity X. We can test our causal models only using measured variables. Let us call the variable that is supposed to measure government spending g and that which is supposed to measure taxes t. Imagine as well g does measure G, but t measures in fact X. The intervention is observable by historical rather than statistical methods.

---

[36] See *e.g.* Baker 1998, Boskin 1998, Deaton 1998, Nordhaus 1998 and Pollak 1998.
[37] NO I 103

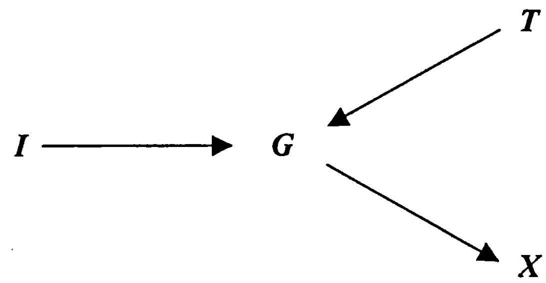

Figure 5.1 A causal structure.

Inferring causal direction according to Hoover's methods now would conclude that G causes T whereas in fact the situation is just the reverse.[38] In this case our variable t measures a real quantity but not the quantity of interest, which leads us to a wrong conclusion about causal structure. In another case we can imagine that a variable measures not the wrong quantity but no quantity at all. In these cases, we may infer that the quantity represented by t (which now supposedly measures nothing) is causally independent from G, but in fact T causes G.

Consider a third case. The quantity *cost of living* is often proxied by the consumer price index (CPI). As mentioned just a moment ago, recently doubts have been shed on whether the CPI really does measure the cost of living. Various so-called "biases" have been identified that are thought to distort that indicator, including among others the "new goods bias", the "quality bias", the "substitution bias" and the "outlet bias".[39] We can say in this case that the CPI does not measure "inflation" but a mix of at least five causal factors. A problem may occur when this mix is not constant. Say it is a law that the quantity of money has a stable causal tendency to affect the cost of living. We will not be able to detect this law, however, if we use the CPI as a

[38] This is of course not an artefact of Hoover's system. All methods of causal inference that make use of statistical methods (besides Hoover, for example Granger and Bayesian nets) presuppose that variables are appropriately measured.

[39] The new goods bias arises because the new goods are introduced whose contribution cannot be readily reflected in a price index (price indices are often averages of price changes; a new good, however, does not have a "previous period price"). The quality bias is related but here the problem is not due to the novelty of goods but due to quality changes. The substitution bias arises because people change their consumption patterns in response to price changes. Finally, the outlet bias obtains because people may change their shopping habits in a way that cannot be reflected easily in an index. For a discussion, see references in footnote 37.

proxy for it if the contribution of the other factors continuously changes (unless by chance they cancel out).

Thus we can distinguish three cases of "mismeasurement": in the first case the measured variable represents no real quantity at all; in the second case the variable represents a real quantity but not the quantity of interest; and in the third case, the measured variable does not only represent the quantity of interest but a (usually unstable) mix of various quantities.

3.3 Measurement and Natural Quantities

The criterion which makes sure that the variable represents a real quantity comes from the literature on the distinction between primary and secondary qualities.[40] Roughly speaking, the distinction is one between qualities which are independent of the sensory apparatus and those which are not. A standard example for a secondary quality is "warmth". Whether an object feels warm or cold depends among other things on the temperature of one's hands. It may also feel different to different people, or different at different times. The qualities of being triangular or square are not dependent on the sensory apparatus in the same way. Hence they are primary.

The distinction is discussed in a way which is particularly relevant for the purposes of this essay by Derk Pereboom 1991.[41] Pereboom argues that the early modern philosopher-scientists Galileo, Descartes, Boyle and Locke, important details in their arguments and positions notwithstanding, held that (a) there are no secondary qualities in the objects which resemble our ideas of them and (b) that all real qualities in the physical world are either primary qualities or are wholly constituted of primary qualities.[42] The reasons to believe in these theses were the following:[43]

(1) the overall explanatory success of a physical science which refers only to primary qualities

(2) the mathematical expressibility of primary but not secondary qualities

(3) the immunity of primary but not secondary qualities to perceptual relativity arguments.

---

[40] On the Lockean view, a quality is a feature of the external world which can be the causal basis for an idea (a mental representation of a quality or complex of qualities). See *Essay* II, viii, 8. This coincides almost exactly with the notion of property from above. I will follow the early modern terminology here but submit that one could use "qualities" and "properties" interchangeably.

[41] My thanks to Jordi Cat to draw my attention to this article.

[42] *ibid.*, p. 68

Of these three reasons for the primary-secondary distinction, Pereboom takes mathematical expressibility to be the most important one. This is because mathematical expressibility is a necessary and sufficient condition for a purported quality's reality or the existence of a real quality which is appropriately causally related to the purported quality.

Let us see what the steps of the overall argument are. It is argued first, that *if* a quality is mathematically expressible, *then* it will be immune to perceptual relativity arguments. Instead of trying to understand the early moderns' notion of mathematical or geometrical expressibility, Pereboom uses a formalisation of Rudolf Carnap's.

According Carnap, if we want to call some property to be a "quantity", five rules must be satisfied. Pereboom takes these five rules as a modern version of the Cartesian demand for geometrical representability.[44] Carnap's rules are taken from his *Introduction to the Philosophy of Science* and read:[45]

(1) The first rule states that there must be an empirical test that specifies when two magnitudes of the quality are equal...

(2) The second rule says that there must be an empirical test that tells us when one magnitude of the quality is greater than another...

(3) The third rule defines an easily (and empirically) recognizable point of reference on the scale for the quality...

(4) The fourth rule specifies a unit of measurement for the scale...

(5) The fifth rule... tells us under which empirical conditions two differences in the values of magnitudes for a quality are the same...

The reason for Pereboom to hold that mathematically expressible qualities are not subject to perceptual relativity arguments is that these rules guarantee that there are empirical procedures for determining the existence and value of a quantity; and these procedures that are available to everyone and open to intersubjective agreement. Thus mathematical expressibility is *sufficient* for absence of perceptual relativity.

---

43 Pereboom 1991, p. 63

44 *ibid.*, pp. 74ff.

45 Carnap 1966, pp. 62-69. Quoted from Pereboom, *op. cit.*, p. 76.

Second, immunity from perceptual relativity is *necessary* for a quantity to be real. Pereboom argues,[46]

> An idea of a purported physical quantity being relative to the state or nature of the perceptual mechanism constitutes strong evidence that the quality (as represented) is not real...

This is roughly analogous to Hacking's distinction between facts and artefacts of a detection device. If a biological entity appears only on images of one kind of microscope, say, there is good reason to believe that it is an artefact of this detection technique rather than a real property of the analysed specimen. In the same vein, Pereboom argues that if a quality is relative to the state or nature of the perceptual mechanism, we can safely infer that it is not real. It is of course always *possible* that the temperature felt by the right hand is always the one that is "in the object", or that felt by women over 50 or that felt by philosopher-kings. But this, arguably, would constitute *ad hoc* reasoning.[47]

On the other hand, we can argue that if a purported quantity is not observer-relative that we have good reason to believe that either the quality is real, or there is a real quality which is appropriately causally related to the purported quality (real temperature which causes felt standard warmth exactly proportionally within the natural limits of the observers perceptual mechanism, say). There is, however, a problem with this suggestion. It is possible that different physical qualities are responsible for the production of the same type of sensation in humans—*i.e.* that there can be intersubjective agreement on the sensation, but the sensation-type may be nonetheless caused not by a single physical quality but by a number of very different qualities. For example,[48]

> Sensations of unique yellow... can be produced by light of wavelength 580 nanometers, or a combination of light of 540 and 670 nanometers, and an indefinite number of other combinations.

Pereboom's way out is to demand that the purported quality be immune to perceptual relativity *across* species. Although it is always *possible* that common ancestry brings

[46] *ibid.*, p. 81

[47] Hence, in this context the early moderns' first argument plays a role. It is always possible that qualities are real but escape our perception, and "our perception" in this case means the perception of each of us under the appropriate conditions. But the existence of these qualities would not add much to our understanding of the world. Thus eventually Pereboom's argument rests on a similar foundation than Shoemaker's.

[48] *ibid.*, p. 85

with it a similar sensitivity to physical stimuli, immunity across species does, thus Pereboom, constitute fairly strong evidence for the reality of a quantity. The analogue argument can be made in the context of measurements and experiments. Hacking demands that the processes by which we try to establish the reality of some entity (or quantity) be *physically* different in order to rule out as well as can be done the possibility of a common artefact. In a similar vein, I shall argue that measurement procedures in economics should be independent of one another if the aim is to establish that the measured quantity is real rather than artefactual.

Hence, third we can say that a quality's being immune from perceptual relativity is *sufficient* for that quality's reality or the existence of a quality that is appropriately causally connected with the purported quality.

Fourth, Pereboom argues that for purported qualities which intuitively come in degrees (unlike being a neutron, being a quark or being pregnant, say), they are immune from perceptual relativity just in case they are mathematically expressible (according to Carnap's five rules). But he qualifies:[49]

> Perhaps there is some way, other than by a process of the sort Carnap describes, to distinguish a physical quality that intuitively comes in degrees and is immune to perceptual relativity arguments. If there really is such a way, [the condition saying that immunity from perceptual relativity and mathematical expressibility are co-extensive] will not be exceptionless. But I am not aware of such a method.[50]

This closes the circle. Mathematical expressibility is sufficient for immunity from perceptual relativity. The latter is necessary and sufficient for the reality of a quality. And mathematical expressibility is necessary for immunity.

Pereboom can thus say:[51]

> (d') A purported physical quality being mathematically expressible, as defined by Carnap's five rules, constitutes strong evidence that either the quality (as represented) is real or there is a type of real quality whose various degrees are appropriately causally correlated with the degrees of experiential contents of the idea of the quality.

Carnap's five rules are conditions for measurability. If one can run Carnap's empirical tests for a given quantity, that quantity is measurable. Thus, if we follow

[49] *ibid.*, p. 87

[50] Further below I shall argue not only that there is such a method, but also that we need it if there should be any real economic qualities.

Pereboom's arguments and take measurability in Carnap's sense as strong evidence for the reality of a quality (and ignore fact that there might be a "type of real quality whose various degrees are appropriately causally correlated with the degrees of experiential contents of the idea of the quality"), we might say:

(NQ3) *A quantity is natural if and only if*

(a) *there are causal tendency models, in which the quantity is represented as either a cause(-variable) or an effect(-variable)*

(b) *it is measurable (as defined by Carnap's rules).*

Although I think this condition is a great deal nearer to the truth, it is still not satisfactory. The reason is that Carnap's rules are not very helpful in the context of *economic* measurements. Most economic measurements proceed either by index numbers or by running regressions in which the quantity is measured either by an estimated parameter or by construction from estimated parameters. But then Carnap's rules are in most cases either trivially fulfilled or never fulfilled.

They are trivially fulfilled if we take the measurement result as the basis for applying Carnap's rules. If the quality of interest is unemployment (U) and IN is the index number that supposedly measures unemployment, rule (1) would then read $U_1 = U_2$ iff $IN(U_1) = IN(U_2)$, rule (2) $U_1 > U_2$ iff $IN(U_1) > IN(U_2)$ *etc.* These rules are arguably empirical: they use real data and some mathematical manipulation. The disturbing factor is that in this case almost *any* index number would do. But we certainly do not want to say that every index number is equally good (and represents the same quantity).

On the other hand, the rules will almost never be fulfilled if we take the quantities themselves as the basis for application. Unemployment is a particularly simple case, though, for which it might almost work. If Σ_1 is the set of unemployed people in country 1 and Σ_2 the set of unemployed people in country 2, rule (1) could be interpreted as $U_1 = U_2$ iff there is a one-to-one mapping between Σ_1 and Σ_2, which at least in principle can be verified empirically. Rule (2) could then read $U_1 > U_2$ iff

---

[51] *ibid.*, p. 86

there is a one-to-one mapping between Σ_1 and a proper subset of Σ_2, which also could in principle be verified empirically. Things are complicated if we want to make comparisons through time. Furthermore, I have assumed until now that it is clear what it means to count as an unemployed. But this is not at all clear. One could define as unemployed (a) anyone who is registered unemployed, (b) anyone who wants to work but has no work (c) anyone who is actively looking for a job, (d) anyone who does not have work for more than *x* days, (e) nobody, because there is no such thing as involuntary unemployment *etc. etc.* A one-to-one mapping between two countries which define unemployment very differently may not make sense at all. But unemployment is probably one of the simpler cases of economic measurement. I see no way of applying Carnap's rules for quantities such as the price level or the NAIRU.

I think it is reasonable to assume that, these difficulties notwithstanding, at least some economic quantities are measurable. If this is true, Carnap's rules are, as I have tried to show, not *necessary* as conditions of economic measurability. This follows if there are quantities which are measurable but Carnap's rules are not applicable. But neither are they *sufficient*. If we assume that the rules are applicable the unemployment case as hinted at above, and also that the two countries have very different standards that define what it means for a person to count as unemployed, "unemployment" will be measurable according to Carnap, but economically make no sense. We need a different set of rules.

To get an idea of what these rules might look like, I want to make use of some of the thoughts that were discussed in the context of analysing the NAIRU case. First, a trivial but necessary rule is that we need a measurement procedure. I want to follow the terminology of Chapter 4 and define:

Rule 1. There must be a measurement procedure ***MP****, understood as an ordered quadruple* $\langle A_1, A_2, \mathbf{D}, \mathbf{R}\rangle$ *of a set* A_1 *of assumptions about the* specification *of the procedure; a set* A_2 *of assumptions about the* data series*; a set* ***D*** *of data; and a set* ***R*** *of results.*

As stated above, we need conditions that guarantee (a) that **MP** measures the quantity of interest, and only that quantity, and (b) that that quantity exists. (a) is a pragmatic condition. It tells us whether there the appropriate kind of link between the quantity of interest and our measurement procedure. (b) is an ontological criterion. It says that there should correspond something in the economic world that we give the name of a "natural quantity".

The reason we need (a) is to exclude such cases as the second and the third examples from above where either a quantity was measured but not the quantity of interest and a mix of various quantities was measured, respectively. We could attempt to measure the NAIRU using Gordon's triangle model but feeding it with data series about American chicken production, rainfall in London and earthquake strength in San Francisco rather than inflation, unemployment and shocks. And we could attempt to measure monetary inflation with an indicator that in fact measures not only inflation due to money but also inflation due to oil price shocks.

We know that a certain measurement procedure measures the right quantity by appealing to our background knowledge. Consider a case I have discussed at greater length in an earlier paper.[52] Jevons analyses the measurement of the lunar tide. The lunar tide is the tide due to the gravitational influence of the moon. He argues that we can correctly measure it by drawing an average between the spring and the neap tides. But why would that average measure lunar tide? The reason is that we know that when the moon is new (or full) the gravitational forces of both sun and moon act such that their influences on the tides *add*. This produces the phenomenon of spring tides. On the other hand, if the moon is in quadrature, the sun's pull acts such as to reduce the moon's influence. This is why we have neap tides. Thus: spring tide = lunar tide *plus* solar tide, neap tide = lunar tide *minus* solar tide, and drawing the arithmetic average yields the lunar tide.

In another case, Staiger *et al.* 1997 argue that testing their NAIRU model for robustness by exchanging their unemployment series (all unemployed) with a series that describes the unemployment rates for married males would be fallacious. Since married males have consistently lower unemployment rates than average, the NAIRU with married males measures a different quantity than the one with all

[52] Reiss 2001

unemployed. Being married and male can be thought of as a property which is causally responsible for lower unemployment rates. Given this knowledge, we know that the two NAIRUs are different.

However, occasionally our background knowledge will be conceptual rather than causal. In another example from the NAIRU case Gordon argues that the confidence interval technique Staiger *et al.* use in part does not make economic sense. He says,[53]

> The recent suggestion of Staiger, Stock and Watson (1996) that the NAIRU for the year 1990 could range from 5.1 to 7.7 percent makes no economic sense. If the NAIRU had been 5.1 percent since 1987, inflation would not have accelerated during 1987-1990, since the actual unemployment rate never fell below 5.1 percent in any calendar quarter. If the NAIRU had been 7.7 percent in the period since 1987, inflation would not have decelerated during 1990-93, since the actual unemployment rate never rose above 7.7 percent in any calendar quarter.

Obviously Gordon implies that the NAIRU exists. But given that, there are conceptual limits to what value it can take.

Thus, calling the quantity of interest Q, we add another rule:

*Rule 2. Causal and conceptual background knowledge implies that **MP** measures Q.*

There is, however, a problem with this rule. Depending on what is meant by "knowledge", this rule can let in either too much or not enough. For many philosophers, "knowledge" implies truth.[54] If we require this, there won't be any natural quantities, or very few if we are lucky. Jevons's lunar tide will not be a natural quantity on this conception. Although we believe that Newton's model of the planetary system (on which Jevons's reasoning is based) is approximately true, it is exactly false. We also know that there are many more causal factors responsible for the determination of the tides than just moon and sun, all of which Jevons idealises away. Hence, strictly speaking we know that Jevons's procedure does not measure lunar tide.

---

[53] Gordon 1997, p. 29

[54] For an introduction to the theory of knowledge, see Stugeon 1995.

A distinction Jevons draws seems appropriate to deal with this situation. Jevons distinguishes the "precise mean result" from the "probable mean result" and characterises:[55]

> *The precise mean result*
>
> "It may give a result approximately free from disturbing quantities, which are known to affect some results in one direction, and other results equally in the opposite direction."
>
> *The probable mean result*
>
> "It may give a result more or less free from unknown and uncertain errors;..."

For Jevons, the lunar tide is a "precise mean result". For the sun affects the tide sometimes positively, sometimes negatively, and on average the influence cancels out. There are other causal influences, for instance the friction of the earth, the gravitational pull of other planets and measurement errors to mention but some. But we know that as compared to the influence of sun and moon they are negligible. Hence Jevons can argue that the result is "approximately free from disturbing quantities".

The distinction is not sharp, though. The stronger the influence of the unmodelled causes is, the more appropriate it is to say that the result will hold "only with probability" rather than "precisely".

Thus one amendment we should make is to demand that the procedure approximately measures the quantity of interest rather than "measures" it with no qualification.

Rule 2'. Causal and conceptual background knowledge implies that ***MP*** *approximately measures Q.*

This still does not solve the problem of what to count as "knowledge". There are very different requirements to be found in the relevant literature.

[55] Jevons 1877, p. 359f.

Reconsider Hasok Chang's temperature case. Here the problem is to find the right "standard" to measure temperature. All thermometers can be calibrated such that they show 0°C at the melting point of water and 100°C at the boiling point. But how about points in between? In the best of all worlds we would know one of the substances (*e.g.* air, mercury or alcohol) to expand linearly such that we can divide the interval between 0° and 100°C into 100 equal segments and each measures one degree. Unfortunately, we don't know which substance if any does expand in this way, and there is no absolute standard against which to calibrate.

One solution to that problem that was proposed by Pierre-Simon Laplace involved the invocation of a theory, *viz.* caloric theory.[56] According to this theory, matter consists of small atoms or molecules, each surrounded by an atmosphere of caloric. The latter was thought of as a self-repulsive fluid, which is attracted to matter. The atoms or molecules of matter themselves attract each other through chemical affinity.

People at the time took this theory as a reason to believe that the laws involving heat were most likely to be revealed in gases rather than liquids or solids. Thus, they thought that between the three substances air, spirit and mercury, air would be most regularly behaved. Together with some additional assumptions this theory yielded the result that air expands linearly. Thus an air thermometer is the preferred measurement instrument.

This case shows that sometimes *theories* are admissible background knowledge. But this means that knowledge cannot imply its truth since the caloric theory is false.[57]

Laplace's reasoning was later superseded by a reasoning that was more strictly empiricist. In the 1840's Henri Victor Regnault tried to get rid of as many theoretical assumptions as possible in the design of his measurement techniques. He instead used a criterion of "comparability": first, if a thermometer is supposed measure temperature correctly, it must give us the same value under the same circumstances; second, if a *type* of thermometer is supposed to measure temperature correctly, all tokens of that type should agree in their readings.[58] Regnault conducted experiments, and found that the air thermometer is the only type of thermometer for which the comparability criterion is fulfilled. However, he believed that comparability is only a

[56] Chang 2001

[57] Though at the time not believed to be false.

necessary but insufficient condition for measuring real temperature. On the other hand, there were no other candidates left. So the air thermometer is the preferred one. Regnault regarded the choice of the air thermometer as partly justified by the comparability criterion.

The important point for my story is that Regnault rejected the idea that a theory (which may always be false) should play a role in determining the choice of a measurement instrument. But I believe that Chang overlooks one aspect in his theory of minimalist overdetermination. It is not enough to have a basic ontological conjecture ("there exists an objective property called temperature") and use the comparability criterion to determine which instrument best measures that objective property. For if it was, *any instrument* that fulfils the comparability criterion measures temperature. But this is absurd. Regnault must use more background knowledge if he is to justify is choice of the preferred instrument.

The lesson from Chang's case here is that different *kinds* of knowledge may be regarded as appropriate background knowledge for the choice of a measurement instrument or procedure. The problem is that we need the *right* kind of knowledge. Otherwise we will have to accept for example the assumptions of neo-classical economics as knowledge, and this may imply that only measurement procedures that have the right kinds of micro foundations will ever measure *anything*.[59] And this would run counter to the general Baconian outlook of this Thesis.

But in order to avoid having to define what the "right kind of knowledge" is, we shall make the rule relative to a body of causal and conceptual background knowledge at time *t*, which is strong enough to imply that the measurement procedure approximately measures the quantity of interest, and let others define what good empirical knowledge is. Our rule then reads:

Rule 2''. A body of causal and conceptual background knowledge CBK_t implies that MP approximately measures Q.

---

[58] *ibid.*, p. 8

[59] For a view that comes close to this, see Rogerson 1997.

The next problem was the ontological one: does *Q* exist? I discussed a related point at greater length in the previous Chapter, but I think the best way to find evidence that a measurement result represents something real rather than an artefact is to have different measurement procedures which give similar results. James Woodward[60], for example, says:

> To say that a certain experiment or measuring procedure has detected a phenomenon is to suggest that the data produced do not just reflect causal interaction which are idiosyncratic to the particular experimental procedures employed, and that the phenomenon in question exhibits stable characteristics which are detectable via several different procedures.

Woodward uses two important conceptions in this statement: *stable* characteristics of a phenomenon and *different* procedures. About the second conception Ian Hacking has written in his 1983. He has been portrayed to be saying that "a scientific entity should be and is accepted as real in proportion to the number of independent experimental methods that can be used to show its existence".[61] For Hacking, this independence condition is met when the detection processes are *physically* different, described by entirely "different chunks of physics".[62] Thus, there is good reason to believe that an entity is real if it shows up both by observation with a light microscope as well as by observation with an electron microscope.

As I have remarked in Chapter 4, Hacking's criterion won't do for economics. But there is a sense in which measurement procedures may be independent. Let me illustrate this thought with a couple of examples. For the sake of illustration I first want to discuss two examples in which we know that the quantity of interest exists but we do not know what measurement procedure should reliably measure it. I will then show that the same reasoning can be applied if the existence of the quantity is at stake.

Candidates for a new job are sometimes asked to estimate a certain quantity about which they are more or less ignorant in order to demonstrate their reasoning skills. In such an interview I was once asked to estimate the car maker Mercedes's turnover. I didn't know the figure. But I could reason that Mercedes sells a percentage *x* of its cars in Germany, in Germany live so many people, a percentage *y* of them drive cars,

---

60 Woodward 1989, p. 401

61 Rasmussen 1993, p. 227

62 Hacking 1983

some of them Mercedes, a Mercedes on average costs so much *etc. etc.* I arrived at a certain number. But then I roughly knew Daimler-Benz's (which was the holding company at the time) turnover. And I knew that a large percentage but not all of Daimler's turnover would come from Mercedes. So I again estimated a percentage and calculated the corresponding number for Mercedes's turnover. The numbers coincided exactly. My interviewer then told be that this fact vindicates my first estimation. I was startled.

The second example comes from LSE colleague Jim Thomas. In his contribution to the *Economic Journal* controversy on the black economy, Thomas discusses several concrete measures of the black economy.[63] One of them is the so-called "Cash-Deposit Ratio" measure. It rests on three assumptions:[64]

> First, a year... must be identified in which the black economy did not exist. Secondly, transactions in the black economy are carried out exclusively using cash. Finally, the velocity of circulation of cash is the same in both the non-black and the black economies.

Essentially, the size of the black economy is then estimated by calculating a cash-deposit ratio in the year where there was no black economy, assuming that this ratio would remain constant over the years, thus deriving a measure of black economy cash and finally calculating the size of the black economy using a version of the quantity theory.

Thomas criticises this measure on the basis of the unrealisticness of its assumptions, especially the first one. Has there ever been a time with no black economy? "Perhaps in the Garden of Eden", Thomas says, "but even there we do not know what else the Serpent got up to". However, no matter how unrealistic the assumptions are, I think the measure could be supported in case there are other measures (based on very different, independent assumptions) which coincide. It seems, for instance, that a straightforward approach would be to aggregate estimates of the sizes of all parts of the economy that are unnoticed by statistical offices: drug dealing, prostitution, tax evasion *etc. etc.* There are individual measures for each of these. If the two measures coincide, there is at least some reason to believe that the measure is correct.

This reasoning is reminiscent of Jean Perrin's reasoning when he determined Avogadro's number and argued for the existence of atoms. As is well known, Perrin

---

63 Thomas 1999

found thirteen physically different ways to calculate the number. Of the first of these, he was not too convinced. In this case Avogadro's number is derived from Van der Waal's equation and the kinetic theory of gases.[65] But Perrin has doubts about this method:[66]

> [T]he probable error, for all these numbers is roughly 30 per cent, owing to approximations made in the calculations that lead to the Clausius-Maxwell and Van der Waal's equations... The Kinetic Theory justly excites our admiration. [But] it fails to carry complete conviction, because of the many hypotheses it involves.

Now this last remark is similar to Thomas's criticism. The determination of the black economy involves a great deal of dubious assumptions. But if we were in Perrin's lucky situation and had different methods whose results coincided, the reliability of each of the methods would be vindicated. Maybe it is false that all transactions in the black economy are cash transactions. But maybe it doesn't matter too much. Maybe the assumption is good enough to yield a reliable estimate. And we have good reason to believe that this is the case if we have different methods that yield comparable results. The same point is made by Cartwright in the discussion of Perrin's argument:[67]

> Here is where coincidence enters. We have thirteen phenomena from which we can calculate Avogadro's number. Any one of these phenomena—if we were sure enough about the details of how the atomic behaviour gives rise to it—would be good enough to convince us that Avogadro is right. Frequently we are not sure enough; we want further assurance that we are observing genuine results and not experimental artefacts. This is the case with Perrin. He lacks confidence in some of the models on which his calculations are based. But he can appeal to coincidence. Would it not be a coincidence if each of the observations was an artefact, and all agreed so closely about Avogadro's number?

In the third economic example we don't know whether the quantity of interest exists (as was the case for Avogadro's number), and this is the situation that matters at present. The quantity of interest is the NAIRU. I see no *a priori* reason why there must be a NAIRU. Maybe the inflation process has nothing to do with unemployment. But if there are various measurement procedures, based on

---

64 *ibid.*, p. F382
65 *Cf.* Cartwright 1983, p. 84.
66 Perrin 1916, p. 82. Quoted from *ibid.*
67 *ibid.*

independent assumptions, whose results coincide reliably[68], we have some reason to believe not only that the measurement procedure is vindicated but also that the NAIRU exists.

We can say about the reduced-form approach that it measures the NAIRU indirectly from its *effects* (on the inflation process), and about the structural approach that it measures the NAIRU indirectly from its *causes* (wage pressure *etc.*). We can also say that the two measurement procedures are to some extent independent as they use very different data series (with the exception of the unemployment rate, presumably). Now if the results of these two procedures coincide, we have evidence for the existence of a NAIRU.

Layard *et al.* 1991 compare the results from the reduced-form and the structural approach. The results do not coincide. But the strategy they employ is to take the result from the reduced-form approach to be the correct one and ascribe the failure of the structural approach to produce the same number to their inability to capture all relevant exogenous factors which are responsible for the NAIRU, among which they see skill mismatch to be the most important.[69] But imagine that the results did coincide reliably. Then, we have reason to believe that the NAIRU exists.

Let us follow the interviewer from the first example and say that measurement procedures $\mathbf{MP}_2$, $\mathbf{MP}_3$... $\mathbf{MP}_n$ *vindicate* measurement procedure $\mathbf{MP}_1$ iff $\mathbf{A}_1^1$, $\mathbf{A}_1^2$, ... $\mathbf{A}_1^n$ are mutually independent, $\mathbf{A}_2^1$, $\mathbf{A}_2^2$, ... $\mathbf{A}_2^n$ are mutually independent, $\mathbf{D}_1$, $\mathbf{D}_2$, ... $\mathbf{D}_n$ are mutually independent and $\mathbf{R}_1$, $\mathbf{R}_2$, ... $\mathbf{R}_n$ are similar. I have not much to say about the concepts of independence and similarity. In both cases I have just a minimal number of vague clues. Sets of assumptions will certainly be independent from each other if they follow from different theoretical frameworks. However, this may already be too much as different theoretical frameworks will often define different concepts. A Keynesian NAIRU may be a different concept from a neo-classical natural rate. Hence the assumptions should be different but not too different. Even in Chang's minimalism story, different methods must accept the same basic ontological conjecture.

68 By "reliable coincidence" I mean that results of measurement procedures coincide repeatedly and in different situations.

69 Layard *et al.* 1991, p. 446

With respect to the data series, it is sometimes possible that they are produced by independent physical processes. For example, inflation rates may be calculated by different statistical offices that use different baskets of goods. But again, the series cannot be too different. It has been argued that the married male NAIRU is a different concept than the all unemployed NAIRU. One will of course say that the data series should all aim at measuring the same quantity but arrive at it by different processes. The question of whether they measure the same quantity then is just repeated at this lower level, and it cannot be answered *a priori*.

The same problem enters with respect to similarity. We want to say that the diversion of Regnault's measurement results of less than 0.3° in the range from 0° to 340° (and always below 0.1 per cent of the magnitude of the measured values) is little but that the diversion of NAIRU results of 4.73 per cent for the year 1980 (and up to more than 78 per cent of the magnitude of the measured values) is not. But I do not see a general standard. Above it has been argued that two results are not similar if the difference matters for policy considerations. I think this is a good criterion, but as we have seen, it may let in a lot. Future work on both conceptions of independence and similarity is very important for the Baconian project of this Thesis.

Bearing these limitations in mind, let us define rule 3.

Rule 3. ***MP_1*** *must be vindicated by a number of measurement procedures* ***MP_2, MP_3, ... MP_n****. The greater* n, *the more the belief in* Q*'s naturalness is justified.*

I need a further argument for why I put so much faith in the argument from coincidence at the expense of all other possible arguments for the reliability of the measurement procedure. It is true, I believe, that coincidence of the kind discussed above is a *sufficient* condition for the reliability of the process under question. *If* we have thirteen relevantly different procedures, each of which yields a similar result, we can infer each method's reliability. But why make it a necessary condition?

The reason is that I think that in economics we are hardly in the situation to be entirely convinced that *any* measurement procedure is the right one and the only right one. All of the procedures I have looked at in the course of the last several years (including inflation, unemployment, the NAIRU, financial market indices, national accounts, the black economy *etc.*) involve a great deal of assumptions the confidence in whose empirical validity is often waning. In such a situation, it seems, it is a viable strategy to find methods following which the dependence on these assumptions is diminished. Chang's minimalism would be one such method. But minimalism is just a special case of what I called vindication: a case in which the sets $\mathbf{A}_1$ and $\mathbf{A}_2$ are "minimal" (but given the complexity of most economic investigations, frequently that will still involve a significant number of assumptions). And recall that Chang has recommended minimalism as the default strategy for situations in which reliable auxiliary conjectures are absent.

The problem I think economists have in, for example, trying to estimate the NAIRU is that they want to get the measure *right*. This is why they use elaborate strategies of the kind described in Chapter 4 to justify their measure. But in the absence of well entrenched empirical knowledge about the behaviour of economic quantities no single method will attract enough confidence to disregard all other possible methods. Thus, a method that does not require much firm background knowledge will be preferable to methods that do. In other words, I believe that in economics we are in a situation much like the one ascribed to Perrin in his measurement of Avogadro's number. There is a number of different ways to measure the quantity of interest. Each of them would do in we had enough confidence in the causal processes leading from the quantity to the end of our measurement procedure. But we rarely have that confidence. In this situation, then, it seems best to appeal to coincidence—if anything.

This fact is reflected in making the vindication criterion *necessary*. But by no means I want to argue that it should be so for all circumstances and at all times. It may well be that certain economic quantities are measurable by one and only one procedure or that others are so simple or so well known that we do not need to vindicate measurement procedures. But until someone has a convincing argument that there are such quantities we had better stick to Rule 3.

The naturalness condition now reads:

(NQ3) *A quantity* Q *is natural relative to* CBK_t *if and only if*

(a) *there are causal tendency models, in which the quantity is represented as either a cause(-variable) or an effect(-variable)*

(b) *it is measurable (as defined by rules 1, 2″ and 3).*

But now hang on. If the quantity is measurable (as defined by rules 1, 2″ and 3), why do we need (a)? For Pereboom it is measurement that matters. Measurability makes a quantity real. I have changed the rules of measurability to make possible that there are any *economic* quantities that are measurable. If (b) makes a quantity real already, and (b) also uses causal information to establish measurability, isn't (a) superfluous?

A central claim of this Thesis is that mere measurability does not suffice for naturalness. Quantities that are causes and/or effects in causal laws, which we are sometimes able to represent by our causal tendency models, influence other quantities, they take part in the course of events, they evolve and interact in regular ways. Knowing about them systematically reduces complexity. Hence, they are the object of our scientific interest.

On the one hand, this requirement presents an advantage, but on the other, an important limitation. The advantage is that the natural quantities approach follows the Aristotelian intuition that sciences are concerned with the general aspects of phenomena, not with what is "accidental" or particular. According to Aristotle, knowledge or understanding of things acquaintance with their causes. But these causes are the general and abstract principles that govern particular things. And in turn, whatever is merely particular will be unreal from the point of view of science.[70]

The limitation is that it may be the case that causal *laws* (generic causal relations between quantities) are rare, and hence so too are natural quantities. In economics

[70] *Cf. Physics* I.1. The characterisation is similar to Menger's definition of *theoretical* science, see his *Untersuchungen*. This view may or may not conflict with Schmoller's view on economics. According to some commentators, Schmoller's position in the *Methodenstreit* was to emphasise the importance of the historico-individualistic aspect of phenomena. In his own methodological writings, however, one can find a fairly balanced view on the topic. On the one hand, so Schmoller, we have to attempt to explain as much of an individual phenomenon as possible with reference to general causes and laws. Inasmuch as they cannot capture that phenomenon, on the other hand, we will have to attend to the peculiarities of the situation. In the latter case, however, we can try to find similar particular phenomena and begin to generalise. See his 1998/1911. But the view that only the general aspect of

there is probably an indefinite number of quantities which are measurable in the sense of rules 1-3 but do not enter any causal laws. Consider an example from financial market theory. According to the Capital Asset Pricing Model (CAPM), the expected return on a risky asset is the sum of the return on the risk-free asset and the excess return on the market times the asset's so-called beta:

$$\overline{R}_a = R_f + \beta_a(\overline{R}_m - R_f), \tag{5.4}$$

where $\overline{R}_a$ is the expected return on asset a, R_f is the return on the risk-free asset, $\overline{R}_m$ is the expected return on the market and β_a is defined as the asset's covariance with the market portfolio divided by the variance of the market portfolio. Two quantities are of interest here: $\overline{R}_a$ and $\overline{R}_m$. In both cases it is the *expected* values that are the relata in (5.4). We can, however, regard the *realised* values of these quantities as at least potentially measurable in the sense used in this Chapter. R_a is measurable in a straightforward way; the measurement of R_m, on the other hand, encounters all the difficulties commonly associated with index number faces (a number of these will be discussed in the next section). For the sake of the argument, let us assume that it is measurable. R_m is not, however, the quantity that plays a causal role in the CAPM. At best, we can use several measured values of it in order to construct an estimate of $\overline{R}_m$. But the latter quantity is the one that we focus on in the formulation of causal models and explanations; and therefore it is the one that, if any, we may call "natural".

Measurability thus is not enough for naturalness. And we have seen that the existence of causal models is not sufficient either because they can be misleading if the quantities are not appropriately measured. An additional argument that being a cause or an effect in a causal model is not enough either has to do with exactness as understood in Chapter 3. Suppose a Parmenides-like figure claims that the true causal tendency models have been revealed to him in a vision and now he is going to teach them to us. Or suppose a Moses-like figure claims that he has had a chat with a

phenomena is "real" or pragmatically interesting is certainly denied by Scotian particularists such as Cartwright 1999 or pluralists such as Dupré 1996.

supernatural being high on the mountain of Sinai, and that he now has engraved true causal tendency models on marble tables. In both cases we can easily infer all the natural quantities the world is made of. But these quantities would not be natural *for us*. Unless measurability is an ingredient of naturalness, we will not be able to determine which *actual* objects, processes or systems those true models apply to. These models may be true, but they are empirically empty.

(NQ3), then, is almost a good condition. The last change concerns (a). It is not enough that there are causal tendency models. We can write down causal tendency models for anything. We can claim that our meeting at the market is the natural quantity *Q*, posit any generic cause *P* for it and write down an equation.

In order for a quantity to be natural, then, these models must have been tested in order to establish that *Q* actually has tendencies to do or to be acted on that the causal models ascribe to it. The difference between (NQ3) and the next version is roughly that between describing an experimental set up theoretically and actually having run the experiment (and thus having established whatever claim one wanted to establish). All approaches to causal inference come along with certain methods that tell how causality is to be inferred from the data (and, sometimes, other knowledge). How this works in case of Hoover's approach to causal inference has been shown above. Thus we additionally require that the causal tendency models have been used to establish that *Q* has the tendencies to do *X*, *Y*, *Z*. We'll also introduce a minute terminological change. Since we are concerned with economics here, and the causal tendency models of interest will be economic models, we will call *Q* a natural *economic* quantity and define:

(NEQ) *An economic quantity of interest* Q *is natural relative to* CBK_t *if and only if*

(a) *there are causal tendency models in which* Q *is represented as either a cause(-variable) or an effect(-variable)*

(b) *it has been established that* Q *has the causal tendencies that the models of (a) ascribe to it*

(c) *it is measurable (as defined by rules 1, 2'' and 3).*

I think this condition is sound. It demonstrates Bacon's as well as Schmoller's claims that induction and deduction are part of the same process of scientific investigation. A plausible research strategy could look like this. Construct measurement procedures in an area of scientific interest. Use rules 1, 2″ and 3 to find *measurable* quantities. For a given measurable quantity, hypothesise a causal model that explains how the quantity is systematically related to other quantities. Test whether those quantities are measurable. If they are, then test whether the hypothesised causal link really exists. If it does, it is confirmed that the quantity is natural.

Let us now examine a concrete case study with which I hope to be able to show in what way natural economic quantities may help in building models that realise our three epistemic virtues simultaneously.

4 Naturalising the Value of Gold

The protagonist in my little case study is the Victorian polymath William Stanley Jevons. Jevons is well-known for his contributions to areas as diverse as logic[71], philosophy of science[72], meteorology[73], computing[74], economic theory[75] and statistics[76]. My focus is here on his contributions to index number theory and economic measurement. I did not pick this case study because I believe that Jevons has convincing arguments that the value of gold is a natural economic quantity in the sense outlined in section 3. In part he lacked the conceptual resources for doing so, and in part the results he achieved do not make the interpretation of his quantity as a natural economic quantity compelling. I picked it rather because I think we can readily interpret Jevons's study as a *search* for natural economic quantities. Jevons *aimed at* achieving something like what we would call establishing the naturalness of

---

71 Jevons 1991/1890 and 1965/1870. For a discussion, see Mosselmans 1998.

72 Jevons 1874

73 See Maas 1999b.

74 See Maas 1999a.

75 Jevons 1871

76 See Stigler 1971.

the value of gold.[77] In what follows, I will first try to argue that this is the case, and then show how Jevons's model of the economy, were his investigation to be successful, could achieve the three epistemic virtues of phenomenal adequacy, explanatory power and exactness all at once.

4.1 Jevons and the Value of Gold

In his 1863 (published in 1884) *A Fall in the Value of Gold Ascertained, and Its Social Effects Set Forth*[78], Jevons set out to investigate whether and to what extent the value of gold has depreciated during a certain period of time, and what the likely causes of the depreciation were. We can say that he faced an ontological problem and a metric problem.[79] Jevons's ontological problem was whether the phenomenon of monetary inflation really existed and whether the events following the gold discoveries of Australia and California constitute an instant of that phenomenon. The metric problem was to measure the *extent* to which the gold discoveries have led to a depreciation in the value of gold—if this link exists.

Corresponding to the three conditions of naturalness defined in (NEQ), three questions arise with regard to Jevons's investigations: (a) Is there a causal model in which the value of gold is represented as a cause or effect? (b) Has it been established that the value of gold has the causal tendency ascribed to it in (a)? (c) Is the value of gold measurable? Let us consider each question in turn.

The Causal Model

Jevons bases his investigation on a model of the economy with various ingredients. Its boldest claim is probably a particular version of the quantity theory. In its crudest fashion the quantity theory only says that money and prices in an economy vary proportionally. Jevons adds to this raw idea the following claims:[80] (i) only money

---

[77] In my 2001 I was obviously less cautious and claimed that Jevons did establish the naturalness of the value of gold conditional upon the adequacy of Jevons's version of the quantity theory as a model of the economy. The discrepancy arises in part because in that earlier paper the conditions for naturalness were slightly less stringent.

[78] Jevons 1884/1863

[79] See Aldrich 1992.

[80] This is not to be read as a historical claim. Most of these ideas were much older than Jevons and quite common wisdom.

that is "active" in the sense of being used in transactions is causally relevant, (ii) only gold can be active money, (iii) gold and prices do not only co-vary but gold actually *causes* prices and (iv) the relationship between gold and prices holds only in the long run.[81] For reasons I shall explain later, I want to interpret this version of the quantity theory as an ascription of a capacity. The Jevons's quantity theory then reads "the quantity of gold has the capacity to raise or lower prices"[82].

To this Jevons adds the idea that the price of each good in the economy is governed by a variety of (largely independent) causes that he calls the "conditions of supply and demand". Jevons says,[83]

> Thus, if the value of an article, A (gold, for instance), falls in comparison with several other articles, B, C, D, E..., so that the same quantity of A purchases less of each of B, C, D, E, than it used to do, this may arise either from causes affecting A only or from causes affecting each of B, C, D, E. The value of A may fall from a lessened demand or an increased supply...

Jevons draws a distinction between various "temporary" causes of the fluctuations in prices and a "permanent" cause.[84] The temporary causes are the individual conditions of supply and demand, which include also influences such as the trade cycle. The permanent cause is gold.

Establishing a Causal Tendency

Jevons's investigation is one more good example of how ontic and metric issues are related. I said above that Jevons faced one ontological and one metric problem. But in order to have good reason to believe that the phenomenon of monetary inflation exists, Jevons must be able to measure the value of gold. It will not suffice, though,

[81] For the classical economists, here including Jevons, the term "money" was intimately connected to the idea of a "means of exchange". Many things besides gold can serve this role, *e.g.* bank notes—whether convertible or not—, cheques, bills of exchange *etc.* There was a great debate in the 19th century over the issue of whether only gold and convertible notes could influence prices or also other "monies" between the so-called Currency school (who held the former view) and the Banking school (who held the latter view). Jevons clearly saw the complexities involved with an attempt to define the term "money" (see his 1875, pp. 248 ff.) but he thought that these could be safely ignored for the purpose of this study because, "*Prices temporarily may rise or fall independently of the quantity of gold in the country; ultimately they must be governed by this quantity*", Jevons 1884/1863, p. 32 (his emphasis).

[82] I should be more careful here and say, "changes in the stock of gold in the economy", and "price changes". As this dynamic formulation is clumsy I prefer the static formulation but the former is more correct.

[83] Jevons 1863, p. 19

[84] See for instance p. 16.

to have *any* measure for that quantity. In order to have a convincing argument that monetary inflation is real, Jevons needs a reliable (or in my terminology, a vindicated) measure.

Jevons's examination proceeds in three stages. First, he asks, what does it mean to say that the value of gold has changed? "Value", Jevons characterises, "is a vague expression for potency in purchasing other commodities".[85] Now if prices in terms of a commodity *C on average* change, the value of *C* changes inversely proportionally. Second, Jevons asks, what does it mean to say that prices *on average* change? His answer is that the geometric mean of individual price changes is an adequate operationalisation of a price average (I will say much more about this in a moment). Third, Jevons asks, how do we ensure that an observed (measured) change of the price average is due to gold rather than other causes? His answer is essentially that if we eliminate the probable contribution of all other causes from the observed change, the remainder of the change must be due to gold. But let us look at this third answer in more detail.

Besides distinguishing temporary and permanent causes in the determination of the value of gold, Jevons further classifies "(1) [causes] which affect the supply, and (2) those which affect the demand for commodities".[86] Among the former Jevons lists for instance the *season* which temporarily and regularly affects mainly agricultural goods, and *wars* which temporarily and irregularly but sometimes very strongly affect many goods. The influences of the season are partly eliminated by including prices of goods whose variations are mutually independent. Jevons argues:[87]

> Alone, [the fluctuation of home-grown wheat] would afford no sure indication of such alteration in the value of gold, but we take it in company with hay, clover, and straw, with meat and butter, with cotton from several parts of the world, with sugar from the East and West Indies, with spices, dyewoods, and various other important products, each subject to its own independent natural fluctuations, but all subject to vary in price with the variations of value in gold of which we are in search.

The contribution of the Russian war, on the other hand, is eliminated by *ex*cluding the prices of various goods for certain years because these constitute, as we would say today, "outliers".

---

[85] *ibid.*, p. 20
[86] *ibid.*, p. 25

The greatest factor that operates on the demand side is what Jevons calls commercial fluctuation. Great booms in permanent investment are often accompanied by credit inflation which, in turn, can increase prices. These influences are eliminated by comparing the price average not on a quarterly or annual basis but in greater periods. This is because,[88]

> *Prices temporarily may rise of fall independently of the quantity of gold in the country; ultimately they must be governed by this quantity.* Credit gives a certain latitude without rendering prices ultimately independent of gold.

More precisely, Jevons compares *averages* of his price averages drawn from various parts of fluctuations with one another. Since not all commercial fluctuations resemble each other precisely and this factor is overshadowed by others such as wars we never know exactly at what point of the fluctuation we are. By using averages throughout the fluctuation, Jevons hopes to eliminate that influence.

From what was said here, I think it should be quite clear that Jevons aims at establishing the phenomenon of monetary inflation by methods comparable to Bogen and Woodward's *Control of Possible Confounding Factors*. All factors Jevons lists as "temporary" causes of price changes qualify as confounders in Bogen and Woodward's sense: they mimic the operation of the phenomenon of interest. Jevons tries to eliminate systematically the contribution of the confounders in various ways resembling the methods Bogen and Woodward describe. Excluding the prices of hemp and flax for the years 1853-55 is similar to *crafted isolation*. The confounder cannot operate because the phenomenon of interest has been shielded from it. In this case the shielding occurs not physically but by changing certain numbers.

Drawing an average falls under the various methods Bogen and Woodward describe that I have labelled *calculation*. In the summer some prices are lower than in winter. By drawing an average of prices of goods some of which originate in parts of the world where it is summer, and others which originate in parts of the world where it is winter controls for the factor *season*. A method similar to the one some CERN experimenters used to establish the reality of neutral currents (see Chapter 4) Jevons also employs. The CERN experimenters calculated an upper limit for the operation

---

[87] *ibid.*, p. 26
[88] *ibid.*, p. 32, original emphasis

of one possible confounder, the "neutron background". Jevons takes commercial fluctuations to be a possible confounder. He argues,[89]

> The lowest average range of prices since 1851 has indeed happened in the last year, 1862; but prices even then stood 13 per cent. above the average level of 1845-50; and it is most highly improbable that prices will long continue to fall; yet prices have continually stood above the high point that they reached in 1846!

We do not know the exact contribution of commercial fluctuations. Therefore, we cannot simply subtract its influence. But we have good (empirical) reason to believe that now we are near a commercial low. And still, prices are above the commercial high of the previous fluctuation. Thus, we can assume that some of the change of the price average (other factors having being eliminated previously) is due to gold. Consequently, and I think rightly, Jevons concludes,[90]

> *Examine the yearly average prices at any point of their fluctuations since* 1852, *and they stand above any point of their fluctuations before then within the scope of my tables!* There is but one way of accounting for such a fact, and that is by supposing a very considerable permanent depreciation of gold.

So much for the ontic problem. Let us then turn to the metric problem.

Vindicating a Measurement Procedure

Jevons's metric problem is to find a convincing argument that the change in the averaging formula he proposes really measures the change in the value of gold rather than either nothing or something else, and that it is a more or less accurate measure of it.

A measurement procedure has four ingredients: assumptions about specification and data, the data themselves and results. As pointed out above, the specification is the unweighted geometric average of price changes:

$$P^{01} = \sqrt[n]{p_1^1 / p_1^0 * p_2^1 / p_2^0 * \ldots * p_n^1 / p_n^0}, \tag{5.5}$$

[89] *ibid.*, pp. 48f.

where P^{01} denotes the change in the price average between periods 0 and 1 and p_i^j the price of good i in period j. The data come in two series. In Jevons's basic investigation he lists data series for thirty-nine goods, in what he calls the "extended proof" he adds another seventy-nine (he subsequently calls the first set the "chief goods" and the second set the "minor goods"). All data series are taken from the *Economist*. His results include the statement that,[91]

> If we take the average of the whole [*i.e.*, chief and minor goods], *the rise of prices is found to be in the ratio of* 100 *to* 110.25, *or by* 10¼ *per cent., corresponding to a depreciation of gold of* 100 *to* 90.70, *or by about* 9 1/3 *per cent.*

Does our causal and conceptual background knowledge imply that P^{01} actually measures the depreciation in the value of gold? The answer to that question can be naturally divided between issues regarding the *specification* and those regarding the *data*, *i.e.*, the choice of goods.

Jevons's most general argument that a price average of many goods should be able to measure the change in value of one good starts with the conceptual fact that a change in value of good C just *means* that the same amount of good C buys a different amount of other goods. Although Jevons does not draw this distinction explicitly, let us distinguish between good C's value relative to another good D, and its general value, or value *simpliciter*. The relative value of C to D changes whenever a different amount of C is required to buy some fixed amount of D or vice versa. The concept of the general value of a good, however, already involves the notion of an average. In a world with at least three goods, and where each may change in value between any two periods, the fact that C changes in general value may not show up in its relative value to D because D may have changed even more in the same direction. Saying that the general value of good C has changed, thus, means that C buys a different amount of other goods *on average*.

Jevons takes it to be a causal fact that the relative value of good C to D is determined by the relative conditions of supply and demand. If C is supplied more abundantly (other conditions being constant), for example, it will drop in relative value to D. In the same manner C's general value will drop when it is supplied more abundantly, given the conditions of supply and demand are constant for all other goods. The

90 *ibid.*, p. 49, original emphasis

phenomenon of monetary inflation is that of a depreciation in the (general) value of gold due to an increase in the stock of gold. Therefore, the change of an average of prices of all other goods in terms of gold measures monetary inflation if its change due to all other alterations in the various goods' supply and demand conditions has been controlled for.

However, there are many different kinds of average formulæ and the *concept* of value does not determine which one is to be taken. Jevons has various arguments to the effect that the geometric mean of price changes is the preferred formula. It does not make sense for him to use the arithmetic average of price changes. Imagine that cocoa doubles in price (rate of change = plus 100 per cent) and cloves halve in price (rate of change: minus 50 per cent). It would be erroneous to conclude the price average has increased by 25 per cent, the result which would be given by the arithmetic average. Rather, the average has not changed: this result is correctly given by the geometric average ($\sqrt{(1+1)(1-.5)} = 1$).[92]

This reasoning being assertive rather than argumentative, Jevons provides the arguments in favour of the geometric average in a contribution a few years later. He observes that if one good doubles in price and another remains unchanged, the arithmetic mean gives a rate of change of 50 per cent, the geometric of 41 per cent and the harmonic of 33 per cent and argues:[93]

> It is probable that each of these is right for its own purposes when these are more clearly understood in theory. But for the present approximate results I adopt the geometric mean, because (1) it lies between the other two; (2) it presents facilities for the calculation and correction of results by the continual use of logarithms, without which the inquiry could hardly be undertaken; (3) it seems likely to give in the most accurate manner such general change in prices as is due to a change on the part of gold. For any change in gold will affect all prices in an equal ratio; and if other disturbing causes may be considered proportional to the ratio of change of price they produce in one or more commodities, then all the individual variations of prices will be correctly balanced off against each other in the geometric mean, and the true variation of the value of gold will be detected.

---

91 *ibid.*, p. 53, original emphasis

92 Personally, I do not share Jevons's intuitions in this respect. After having studied the various approaches to index number theory for quite some time, I doubt that there is any *a priori* way to say this number or that should be the correct one.

93 Jevons 1884/1865, p. 121

Whereas the first two arguments are pragmatic (in the terminology of Chapter 4), the third one seems to try establish that under a plausible additional assumption the geometric mean is indeed the correct one to use. Although probably not known to Jevons, it was pointed out by Francis Ysidro Edgeworth[94] that the geometric mean is indeed the correct one to use if the prices are distributed according to the following form:

$$y = \frac{1}{x\sqrt{\pi}} e^{-h^2(\log x)^2}, \tag{5.6}$$

which is the log-normal distribution (Edgeworth called it the "law of the geometric mean"). Moreover, Edgeworth provided a number of arguments why it is reasonable that prices should be distributed according to (5.5).

The second consideration concerns the assumptions about the data. Jevons followed a general principle and a great number of particular thoughts. The general principle can be summarised by the following words:[95]

> *The only mode of eliminating [the temporary] fluctuations, is to render our inquiry, not more exclusive, but more inclusive.*

The greater the number of goods, the more likely it is that the temporary causes that act on individual goods cancel out. Thus, the principle is to take a *wide* average of goods.

Some more specific considerations regarding the treatment of the data I have described above. If our background knowledge entails for example that a temporary cause has influenced some price to a great extent, such as the case with hemp and flax during the Russian war, it should be excluded in order not to bias our measure. Similar considerations could lead to give greater weight to some prices, or to consider other prices *en bloc*.

---

94 Edgeworth 1925
95 Jevons 1884/1863, p. 26, original emphasis

Thus Jevons seems to have good reason to believe that his geometric average indeed measures the depreciation of gold rather than anything else. Does he also attempt to *vindicate* his measurement procedure?

He does indeed try to vindicate the procedure in various ways. The "proof" that gold has depreciated in the envisaged period is pointed out after calculating the average for the thirty-nine chief goods only. Jevons then calculates the average for the whole group (chief and minor goods) to find the claim of the first investigation supported. He also considers various weighing schemes:[96]

> It may seem to many persons absurd to take a mass of 118 commodities, and treat them as equally good measures of the value of gold, some being so greatly more important and more free from fluctuations than others. I have considered and tried many ways of deducing an average which should obviate this objection more or less perfectly. I proposed to give to each commodity a greater weight, as the range between the highest and lowest prices of the interval 1845-62 was less; but on applying this notion to the thirty-nine chief articles, I found that, always excepting silver, the highest price of nearly every one was just about double the lowest price, so that the method could give no result appreciably different from the simple average. [...]
>
> Another method which I have tried was to exclude all commodities which have undergone exceptionally great changes. [...]

A few years after these remarks, Jevons adopts the method of ordinary least squares in order to calculate a probability that the average has indeed risen.[97] An important point to note is that Jevons takes a very careful stance towards the accuracy of his measurement procedures. He knows that no amount of conceptual and causal background knowledge will yield a unique averaging formula, and that the different formulæ, all of which may be compatible with the background knowledge, give (slightly) different results. But then, Jevons seems to think, we can accept the conclusion only with the proviso that it is inexact to some degree. In the first chapter of the *Fall in the Value of Gold Ascertained*, he says,[98]

> *It must be confessed, however, that the exact mode in which preponderance of rising or falling prices ought to be determined is involved in doubt.* Ought we to take all commodities on an equal footing in the determination? Ought we to give most weight to those which are least intrinsically variable in value? Ought we to give additional weight to articles according to their

[96] *ibid.*, p. 57
[97] Jevons 1884/1869
[98] Jevons 1884/1863, p. 21, original emphasis

importance, and the total quantities bought and sold? [...] Fortunately, the conclusions I shall have to adopt may, I believe, be sustained under any and all modes of estimation which are likely to be proposed. I regard the fall of value as conclusively proved, although the exact nature of the problem is left amid the obscurities of economic science in general.

All I can pretend to *prove* in this inquiry is that, subject to the vagueness just referred to, the prices of commodities have risen, or that the rise of prices of those which have risen preponderates over the fall of those which have fallen. This *is and constitutes* the alteration of value of gold asserted to exist.

4.2 The Value of Gold and Epistemic Virtues

Jevons, it seems then, has devoted much attention to ascertaining that the depreciation of the value of gold due to the California and Australia discoveries is a real phenomenon. His model which predicts that (temporary fluctuations being held constant) the increase in the stock of gold would lead to a depreciation is phenomenally adequate in so far. But the depreciation is also *explained*. The model is a causal model because it claims that the conditions of supply and demand for each good are exclusively *causally* responsible for determination of the goods' value. By Jevons's various methods of eliminating temporary fluctuations due *other* supply and demand conditions, gold remains as the sole cause. It explains the depreciation because there is nothing left which could have brought it about. The analysis of economic explanation of Chapter 2 does not conflict with the claim just made that a causal model may have explanatory power. In that Chapter, the ability to unify and systematise thinking about phenomena has been identified only as a sufficient (or rather, INUS) condition for explanatory power. Thus it is well possible that there are other ways for a model to have explanatory power.

But we can observe a paving the way for unification and systematisation of our thinking about economic phenomena too. If we take the value of gold, or more generally, the value of money, as our alleged natural economic quantity, it appears (at least implicitly) in thousands of models that use "real" quantities (here the contrast is obviously between real and nominal rather than real and crank). All of these models, if they are to be tested empirically, must use a price deflator of the kind Jevons constructed. Thus, the value of gold/money helps us in building models that explain a great number of different phenomena.

Exactness I think is the most interesting epistemic virtue in the light of Jevons's investigations. He claims himself that there is a great deal of vagueness to his inquiry. But that depends on what we are looking at. Exactness, I have said above, is a derivative or secondary epistemic virtue. It is parasitic upon our interests, be they of a descriptive, explanatory or pragmatic (policy-making) nature.

The estimate Jevons calculates certainly *looks* very inexact. He writes in the Preface to the *Fall in the Value of Gold Ascertained*:[99]

> The lowest estimate of the fall that I arrive at is 9 PER CENT., and I shall be satisfied if my readers accept this. At the same time, in my own opinion the fall is nearer 15 PER CENT. It may even be more than this.

The difference arises mainly due to the fact that Jevons believes that the last year of his investigation the commercial fluctuation was at its low, whereas the base years are an average of a whole cycle. Hence prices are likely to rise more once the commercial fluctuation gains force.

However, given his own interest, which is to ascertain that that the value of gold has plummeted due to the gold discoveries the degree of exactness is sufficient. If no estimate that is plausible in the light of the causal and conceptual background knowledge gives a value below zero, his point can be made. And since nine per cent is the lowest estimate he can find (even taking into consideration work by other authors), he is safe to draw his conclusion:[100]

> While I must assert the fact of a depreciation of gold with the utmost confidence, I assign the numerical amount of it with equal diffidence.

Jevons hedges his bets; he does not allege more knowledge than he can reasonably assume to have. I think this is the main difference between him and others, whose models and methods I have examined in Chapters 3-5.

---

[99] *ibid.*, p. 17
[100] *ibid.*, p. 16f.

5 Conclusions

Why, we may ask, can Jevons in the 1860s avoid trade-offs between epistemic virtues that I have argued economists face more than 100 years later? Are those trade-offs an artefact of the interpretation of the various methods of economic concept formation given here? Or does he do so much better, and then why? Or do I give Jevons allowances not granted to contemporary economists?

In a sense I think that the trade-offs I have analysed in the previous chapters are very real, and Jevons cannot avoid them altogether. If anything it seems that Jevons's reasoning can establish that there is a real phenomenon of monetary inflation. But that model does not explain much; and if we want to use the value of gold as a quantity in other models, we will probably need to be able to measure it with more accuracy. On the other hand, we could take for example Jevons's estimation of the devaluation of gold of nine per cent as the "correct" one and thus would have a precise concept. But that by itself would not convince us that monetary inflation is real (*cf.* the analogous arguments in Chapter 4).

Still I believe there is a difference between Jevons's effort and much of what I have found in the contemporary literature. Jevons tried to find out about the existence of a real tendency law with empirical methods. This places him in the tradition of Bacon and Schmoller as interpreted in Chapter 1. There is comparatively little endeavour of this kind in the contemporary methods described in Chapters 3-5.

Recall the concepts "asymmetric information" and "trade costs". Hal Varian defines situations where the quantity of the former kind is present as "situations where one economic agent knows something that another economic agent doesn't"[101]; Obstfeld and Rogoff include among the latter "(among other things) tariffs, nontariff barriers [*e.g.* currency conversion costs and exchange rate uncertainty] and transportation costs"[102]. Now, I do not know of any situation of economic interest in which these two factors, thus broadly defined, are not present. It is not surprising, then, that they play a part in thought experiments which allegedly explain a great number of different phenomena—and hence are thought to "unify and systematise thinking about economic phenomena". But apart from displaying the possibility that these

101 Varian 1992, p. 440

102 Obstfeld and Rogoff 2000

factors may really contribute to the outcome of economic situations I don't think these thought experiments are very informative about real situations.

This is different with respect to Jevons's model. The claim it tries to establish, it tries to establish by a systematic empirical investigation. Whatever gold does, it is real gold that does it. And if Jevons is successful with his model, it has some explanatory power as well. And lastly, it is sufficiently exact in the light of Jevons's epistemic interest.

Paralipomena

Jevons's work on index numbers as well as his other statistical work (*e.g.* on trade cycles was an important element in a long chain of ideas that eventually led to the development of modern econometrics (see Morgan 1990). One of the aims of this Thesis was to give a Baconian interpretation to the kind of work Jevons achieved. Bacon urged the manipulation of nature in order to get her to reveal her real causal structure to us. In the case that Jevons analysed, he could not "manipulate nature". There are many reasons for this, one of them being that most events relevant to his study had already occurred. But what he could do was to manipulate data in a way that could achieve an (almost) equivalent result to he one he would have obtained could he have manipulated nature directly. In so doing, he came as close as possible at his time to establishing the naturalness of the quantity *monetary inflation*.

This interpretation can easily be extended to much of modern econometric analysis. Successful econometrics, in my interpretation, can determine which economic quantities are natural. But determining what a natural quantity is does not come cheap. As I hope to have shown in Chapters 2 and 3, it does not suffice to isolate causal factors *in thought* in order to establish the causal relations that are characteristic of natural quantities. Neither does it suffice to enter data into a computer and run regressions—as sophisticated as they may be—as I tried to argue in Chapter 4.

"Statistics has in many ways substituted the experiment that is lacking [in economics]; only it has created a sense of exactness and precision in this area of knowledge", Schmoller once said. But in order to use statistics and econometrics in a way that mimics a Baconian experiment, we need to apply methods that are equivalent to the methods that are constitutive of experimentalism. Some of them have been discussed above as exemplars, and I attempted to show that Jevons did exactly that. It is a long way from arguing, however, that Baconianism in economics is desirable (because it helps in achieving epistemic virtues that are demanded of that science) and possible (because Jevons demonstrated it) to showing how to use Baconian ideas in actual, contemporary economic investigations. This is the topic for another project.

Bibliography

(An entry with two dates (*e.g.* 1997/1992) refers to a text that has been republished or reprinted. The first date is the edition used and the second one that of the original edition.)

AKERLOF, GEORGE 1970: 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism', *Quarterly Journal of Economics* **84**(3): 488-500.

ALBERT, HANS 1998: *Marktsoziologie und Entscheidungslogik.* Tübingen: Mohr Siebeck.

ALDRICH, JOHN 1992: 'Probability and Depreciation: A History of the Stochastic Approach to Index Numbers', *History of Political Economy* **24**(3): 657-87.

ARMSTRONG, DAVID 1978: *Universals and Scientific Realism.* Cambridge: Cambridge University Press, 2 vols.

— 1997/1992: 'Properties', in: Hugh Mellor and Alex Oliver, *Properties.* Oxford: Oxford University Press, 160-172.

AUMANN, ROBERT 1985: 'What is Game Theory Trying to Accomplish?', in: Kenneth Arrow and Seppo Honkapohja, *Frontiers of Economics.* Oxford: Blackwell.

BACKHAUS, JÜRGEN AND REGINALD HANSEN 2000: 'Methodenstreit in der Nationalökonomie', *Journal for General Philosophy of Science* **31**: 307-336.

BACKHOUSE, ROGER 1998: 'If Mathematics Is Informal, Then Perhaps We Should Accept That Economics Must Be Informal Too', *Economic Journal* **108**(451): 1848-58.

BAILY, MARTIN NEIL AND JAMES TOBIN 1978: Inflation-Unemployment Consequences of Job Creation Policies, in: John Palmer, Creating Jobs: Public Employment Programs and Wage Subsidies. Washington (DC): Brookings Institution, 43-85.

BAKER, DEAN 1998: *Getting prices right: The Debate Over the Consumer Price Index.* Armonk (NY): Sharpe

BARNES, ERIC 1992: 'Explanatory Unification and Scientific Understanding', in: PSA 1992, vol. 1: 3-12.

BARTELS, ANDREAS 1990: 'Weshalb implizite Definitionen nicht genug sind', *Erkenntnis* **32** (2), 269-81.

BEENSTOCK, MICHAEL 1982: 'The Robustness of the Black-Scholes Option Princing Model', *Investment Analyst*: 30-40.

BENACERRAF, PAUL 1973: 'Mathematical Truth', *Journal of Philosophy* **70**: 661-80.

BLACK, FISCHER AND MYRON SCHOLES 1973: 'The Pricing of Options and Corporate Liabilities', *Journal of Political Economy* **81**(3): 637-54.

BLALOCK, HUBERT 1979: *Social Statistics* (2nd ed.). New York *etc.*: McGraw-Hill.

BLANCHARD, OLIVIER AND LAWRENCE KATZ 1997: 'What We Know and Do Not Know About the Natural Rate of Unemployment', *Journal of Economic Perspectives* **11**(1): 51-72.

BOGEN, JAMES AND JAMES WOODWARD 1988: 'Saving the Phenomena', *Philosophical Review* **97**: 302-352.

BOSKIN, MICHAEL *ET AL.* 1998: 'Consumer Prices, the Consumer Price Index, and the Cost of Living', *Journal of Economic Perspectives* **12**(1) 3-26

BOUMANS, MARCEL AND MARY MORGAN 2001: 'Ceteris Paribus Conditions: Materiality and the Application of Economic Theories', *Journal of Economic Methodology* 8(1): 11-26

BROWN, JAMES ROBERT 1991: *The Laboratory of the Mind.* London: Routledge

— 1994: *Smoke and Mirrors: How Science Reflects Reality.* London: Routledge

— 1999: *Philosophy of Mathematics.* London: Routledge

CAMPBELL, NORMAN ROBERT 1957/1922: *Foundations of Science* (formerly: *Physics: The Elements*). New York: Dover.

CAMPBELL, DONALD 1988/1969: 'Definitional versus Multiple Operationism', in: Campbell, Donald, *Methodology and Epistemology for Social Science*. Chicago (IL): University of Chicago Press, 31-36.

CAMPBELL, COLIN 1999: 'Large Electorates and Decisive Minorities', *Journal of Political Economy* **107**(6): 1199-1217.

CARNAP, RUDOLF 1966: *An Introduction to the Philosophy of Science*. Ed. by Martin Gardner. New York (NY): Basic Books.

CARTWRIGHT, NANCY 1983: *How the Laws of Physics Lie*. Oxford: Oxford University Press.

— 1989: *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press

— 1999: *The Dappled World*. Cambridge: Cambridge University Press.

— 2000: 'Measuring Causes: Invariance, Modularity and the Causal Markov Condition'. *Measurement in Physics and Economics Discussion Paper Series* DP MEAS 10/00, London: CPNSS, LSE.

CHANG, HASOK 2001: 'Spirit, Air and Quicksilver: The Search for the "Real" Scale of Temperature'. *Measurement in Physics and Economics Discussion Paper Series* DP MEAS 17/01, London: CPNSS, LSE

CHICK, VICTORIA 1998: 'On Knowing One's Place: The Role of Formalism in Economics', *Economic Journal* **108**(451): 1859-69.

COFFA, J. ALBERTO 1991: *The Semantic Tradition from Kant to Carnap*. Cambridge: Cambridge University Press.

COULTON, BRIAN AND ROY CROMB 1994: *The UK NAIRU*. London: HM Treasury Working Paper.

COURANT, RICHARD, HERBERT ROBBINS AND IAN STEWART 1996: *What Is Mathematics?*. Oxford: Oxford University Press.

CROMB, ROY 1993: 'A Survey of Recent Econometric Work on the NAIRU', *Journal of Economic Studies* **20**(1/2): 27-51.

CULP, SYLVIA 1994: 'Defending Robustness: the Bacterial Mesosome as a Test Case', in: David Hull, Micky Forbes and Richard Burian, *Philosophy of Science Association 1994*, vol. 1: 46-57.

DE MEZA, DAVID AND BEN LOCKWOOD 1998: 'Does Asset Ownership Always Motivate Managers? Outside Options and the Property Rights Theory of the Firm', *Quarterly Journal of Economics* **113**(2): 361-86.

DEATON, ANGUS 1998: Getting Prices Right: What Should Be Done?, *Journal of Economic Perspectives* **12**(1): 37-46.

DESCARTES, RENE 1997: *Principles of Philosophy*. Ed. by Enrique Chárvez-Arvizo, *Descartes: Key Philosophical Writings*. Ware: Wordsworth, 261-337.

DORNBUSCH, RUDIGER AND STANLEY FISCHer 1990: *Macroeconomics* (5th edition). New York *etc.*: McGraw-Hill.

DRETSKE, FRED 1977: 'Laws of Nature', Philosophy of Science **44**: 248-268

DUHEM, PIERRE 1954/1908: *The Aim and Structure of Physical Theory* (transl. by Philip Wiener). Princeton (NJ): Princeton University Press

DUPRÉ, JOHN 1996: *The Disorder of Things*. Cambridge (MA): Harvard University Press.

ECONOMIDES, NICHOLAS 1986: 'Minimal and Maximal Product Differentiation in Hotelling Duopoly', *Economic Letters* **21**: 67-71

EDGEWORTH, FRANCIS YSIDRO 1925: *Papers Relating to Political Economy* (vol. 1). London: Macmillan.

EISNER, ROBERT 1997: 'A New View of the NAIRU', in: Jan Kregel and Paul Davidson: *Improving the Global Economy: Keynesianism and the Growth in Output and Employment*. Cheltenham: Edward Elgar, 196-230.

ENGEL, CHARLES 2000: 'Comments on Obstfeld and Rogoff's "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?"', *NBER Working Paper* 7818 (July 2000).

BANK OF ENGLAND 1999: *Economic Models at the Bank of England*. London: Bank of

England.

ESPINOSA-VEGA, MARCO AND STEVEN RUSSELL 1997: 'History and Theory of the NAIRU: A Critical Review', *Federal Reserve Bank of Atlanta Economic Review* 2: 4-25.

FISHER, IRVING 1926: 'A Statistical Relationship between Unemployment and Price Changes', International Labor Review **13**(June): 758-92.

— 1976: 'I Discovered the Phillips Curve: A Statistical Relation between Unemployment and Price Changes', *Journal of Political Economy* **81**(2): 496-502.

FRIEDMAN, MILTON 1968: 'The Role of Monetary Policy', *American Economic Review* **68** (March): 1-17.

FRIEDMAN, MICHAEL 1974: 'Explanation and Scientific Understanding', *Journal of Philosophy* **71**(1): 5-19.

GALBRAITH, JAMES 1997: 'Time to Ditch the NAIRU', *Journal of Economic Perspectives* **11**(1): 93-108.

GOODMAN, NELSON *1983: Fact, Fiction and Forecast* (4th edition). Cambridge (MA): Harvard University Press.

GORDON, ROBERT 1977: 'The Theory of Domestic Inflation', American Economic Review 67 (Papers and Proceedings): 128-34.

— 1997: 'The Time-Varying NAIRU and Its Implications for Economic Policy', *Journal of Economic Perspectives* **11**(1): 11-32.

GRAYLING, ANTHONY C. 1995: 'Modern Philosophy II: The Empiricists', in: Grayling, Anthony C., *Philosophy: A Guide Through the Subject.* Oxford: Oxford University Press, 484-544.

GUYER, PHILLIP 1994: *Microeconomics*. University of St. Gallen lecture manuscript

HACKING, IAN 1975: *Why Does Language Matter to Philosophy?*. Cambridge: Cambridge University Press.

— 1983: *Representing and Intervening*. Cambridge: Cambridge University Press

HALE, ROBERT 1987: *Abstract Objects*. Oxford: Blackwell.

HARTMANN, STEPHAN 1997: 'Modelling and the Aims of Science', in: Paul Weingartner, Gerhard Schurz and Georg Dorn, *The Role of Pragmatics in Contemporary Philosophy* (Schriftenreihe der Wittgenstein-Gesellschaft). Vienna: Hölder-Pichler-Tempsky, 389-95.

HAUSMAN, DANIEL 1992: *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.

HÄGGQVIST, SÖREN 1996: *Thought Experiments in Philosophy*. Stockholm: Almqvist & Wiksell.

HEMPEL, CARL 1966: *The Philosophy of Natural Science*. Englewood Cliffs (NJ): Prentice-Hall.

HESSE, MARY 1964: Francis Bacon, in: O'Connor, Daniel, *A Critical History of Western Philosophy*, 141-152.

HILBERT, DAVID 1902/1900: 'Mathematische Probleme'. Bulletin of the American Mathematical Society **8**(1902): 437/79.

— 1959/1899: *Foundations of Geometry* (transl. by E. Townsend). La Salle (IL): Open Court.

HOOVER, KEVIN 1995: 'Is Macroeconomics For Real?', *The Monist* **78**(3): 235-57.

— 2001: *Causality in Macroeconomics*. Cambridge: Cambridge University Press.

HOROWITZ, TAMARA AND GERALD MASSEY 1991: *Thought Experiments in Science and Philosophy*. Savage (MD): Rowman & Littlefield

HORTON, MARY 1973: 'In Defence of Francis Bacon: A Criticism of the Critics of the Inductive Method', *Studies in the History and Philosophy of Science* **4**: 241-78

HOTELLING, HAROLD 1929/1990: 'Stability in Competition', in: Darnell, Adrian, *The Collected Economics Articles of Harold Hotelling*. New York *etc.*: Springer, 50-63.

HUDSON, ROBERT 1999: 'Mesosomes: A Study in the Nature of Experimental Reasoning', *Philosophy of Science* **66**: 289-309.

IMF 2001: *World Economic Outlook: Fiscal Policy and Macroeconomic Stability.* Washington (DC): IMF World Economic and Financial Surveys.

JEVONS, WILLIAM STANLEY 1871: *Theory of Political Economy.* London: Macmillan.

— 1874: *The Principles of Science: A Treatise on Logic and Scientific Method.* London: Macmillan.

— 1875: *Money and the Mechanism of Exchange.* London: Paul Kegan, Trench, Trübner & Co.

— 1884/1863: 'A Fall in the Value of Gold Ascertained, and Its Social Effects Set Forth', in William Stanley Jevons, *Investigations in Currency and Finance.* London: Macmillan, 13-118.

— 1884/1869: 'The Depreciation of Gold', in: William Stanley Jevons, *Investigations in Currency and Finance.* London: Macmillan, 151-159.

— 1965/1870: *Elementary Lessons in Logic.* London: Macmillan

— 1991/1890: *Pure Logic and Other Minor Works.* Bristol: Thoemmes.

KANT, IMMANUEL 1929/1887: *The Critique of Pure Reason.* (Transl. by Norman Kemp Smith). London: Macmillan.

KEYNES, JOHN MAYNARD 1936: *The General Theory of Employment, Interest and Money.* London: Macmillan

KINDLEBERGER, CHARLES 1989: *Economic Laws and Economic History.* Cambridge: Cambridge University Press

KING, ROBERT, JAMES STOCK AND MARK WATSON 1995: 'Temporal Instability of the Unemployment-Inflation Relationship', *Federal Reserve Bank of Chicago Economic Perspectives* **19**(3): 2-12

KING, MERVYN 1998, *Speech at the Employment Policy Institute's Fourth Annual Lecture.*

KITCHER, PHILLIP 1981: 'Explanatory Unification', *Philosophy of Science* **48**: 507-31.

KOOPMANS, TJALLING 1947: 'Measurement Without Theory', *Review of Economic Statistics* **29**: 161-7.

KREPS, DAVID 1990a: *A Course in Microeconomics*. London: Harvester Wheatsheaf.

— 1990b: *Game Theory and Economic Modelling*. Oxford: Clarendon Press.

KUHN, THOMAS 1981/1964: 'A Function for Thought Experiments', in: Ian Hacking, *Scientific Revolutions*. Oxford: Oxford University Press, 6-27.

KUSUKAWA, SACHIKO 1996: 'Bacon's Classification of Knowledge', in: Markku Peltonen, *The Cambridge Companion to Bacon*. Cambridge: Cambridge University Press, 47-74.

LAWSON, TONY 1997: *Economics and Reality*. London: Routledge.

LAYARD, RICHARD, STEPHEN NICKELL AND RICHARD JACKMAN 1991: *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford: Oxford University Press

LEAR, JONATHAN 1982: 'Aristotle's Philosophy of Mathematics', *Philosophical Review* **91**: 161-92.

LOCKE, JOHN 1689: *An Essay Concerning Human Understanding*. London: Ward, Lock & Co.

LOSEE, JOHN 1993: *A Historical Introduction to the Philosophy of Science* (3rd edition). Oxford: Oxford University Press (OPUS).

MAAS, HARRO 1999a: 'Mechanical Rationality: Jevons and the Making of Economic Man', *Measurement in Physics and Economics Discussion Paper Series* DP MEAS 6/99. London: CPNSS, LSE

— 1999b: 'Of Clouds and Statistics: Inferring Causal Structures from the Data', *Measurement in Physics and Economics Discussion Paper Series* DP MEAS

7/99, London: CPNSS, LSE.

MACH, ERNST 1920: *Erkenntnis und Irrtum: Skizzen zur Psychologie der Forschung* (4th edition). Leipzig: Barth

MACHAMER, PETER, LINDLEY DARDEN AND CARL CRAVER 2000: 'Thinking About Mechanisms', *Philosophy of Science* **67**(1): 1-25.

MACHOVER, MOSHÉ 1996: *Set Theory, Logic and Their Limitations.* Cambridge: Cambridge University Press.

MALHERBE, MICHEL 1996: 'Bacon's Method of Science', in: Markku Peltonen, *The Cambridge Companion to Bacon.* Cambridge: Cambridge University Press, 75-98.

MANKIW, N. GREGORY 1994: *Macroeconomics* (2nd edition). New York: Worth.

MARGENAU, HARRIET 1950: *The Nature of Physical Reality.* New York: McGraw-Hill.

MAS-COLELL, ANDREU, MICHAEL WHINSTON AND JERRY GREEN 1995: *Microeconomic Theory.* Oxford: Oxford University Press.

MÄKI, USKALI 1997: 'Universals and the "Methodenstreit"', *Studies in the History and Philosophy of Science* **28**(3): 475-495.

MCADAM, PETER AND K. MC MORROW 1999: *The NAIRU Concept—Measurement Uncertainties, Hysteresis and Economic Policy Role.* Bruxelles and Luxembourg Institution: European Commission Economic Papers 136.

MELLIS, CHRIS AND A. E. WEBB 1997: *The United Kingdom NAIRU: Concepts, Measurement and Policy Implications.* Paris: OECD Economics Department Working Papers 182.

MELLOR, HUGH 1995: *The Facts of Causation.* London: Routledge.

— 1997/1991: 'Properties and Predicates', in: Hugh Mellor and Alex Oliver, *Properties.* Oxford: Oxford University Press, 255-267.

MENGER, CARL 1976/1871: *Problems of Economics and Sociology* (transl. by: James Dingwall and Bert F. Hoselitz). Urbana: University of Illinois Press.

MILGROM, PAUL AND JOHN ROBERTS 1992: *Economics, Organization & Management*. Englewood Cliffs (NJ): Prentice-Hall.

MILL, JOHN STUART 1874: *A System of Logic* (8th edition). New York: Harper and Brothers.

— 1948/1830: *Essays on Some Unsettled Questions of Political Economy*. London: Parker.

MILTON, JOHN 2000: 'Francis Bacon', entry in *Routledge Encyclopedia of Philosophy*, online edition.

MODIGLIANI, FRANCO AND LUCAS PAPADEMOS 1975: 'Targets for Monetary Policy in the Coming Year', *Brooking Papers on Economic Activity* **1**: 141-63.

MORGAN, MARY 1990: *The History of Econometric Ideas*, Cambridge: Cambridge University Press.

MORGAN, MARY 2001: 'Experiments Without Material Intervention: Model Experiments, Virtual Experiments and Virtually Experiments', *Measurement in Physics and Economics Discussion Paper Series* DP MEAS 11/01. London: CPNSS, LSE.

— and Margaret Morrison 1999: *Models as Mediators*. Cambridge: Cambridge University Press

Morrison, Margaret 2000: *Unifying Scientific Theories*. Cambridge: Cambridge University Press.

MOSSELMANS, BERT 1998: 'William Stanley Jevons and the Extent of Meaning in Logic and Economics', *History and Philosophy of Logic* **19**: 83-99.

NAGEL, ERNST 1960: *The Structure of Science*. Indianapolis: Hackett.

NORDHAUS, WILLIAM 1998: 'Quality Change in Price Indexes', *Journal of Economic Perspectives* **12**(1): 59-68.

NOWAK, LESZEK 1980: *The Structure of Idealization: Towards a Systematic Interpretation of the Marxian Idea of Science*. Dordrecht: Reidel

OBSTFELD, MAURICE AND KENNETH ROGOFF 2000: 'The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?', *NBER working paper* 7777.

OKUN, ARTHUR 1978: 'Efficient Disinflationary Policies', *American Economic Review* **68**: 348-52.

PELTONEN, MARKKU 1996a: 'Introduction', in: Markku Peltonen, *The Cambridge Companion to Bacon*. Cambridge: Cambridge University Press, 1-24.

— 1996b: *The Cambridge Companion to Bacon*. Cambridge: Cambridge University Press.

PEMBERTON, JOHN forthcoming: 'Why Idealized Models in Economics Have Limited Use', in: Martin Jones and Nancy Cartwright, *Correcting the Model: Idealization and Abstraction in the Sciences*. Amsterdam: Rodopi.

PEREBOOM, DERK 1991: 'Mathematical Expressibility, Perceptual Relativity, and Secondary Qualities', *Studies in History and Philosophy of Science* **22**(1): 69-88.

PERRIN, JEAN 1916: *Atoms* (transl. by: D. Ll. Hammick). New York (NY): Van Nostrand

PÉRES-RAMOS, ANTONIO 1996: 'Bacon's Forms and the Maker's Knowledge Tradition', in: Markku Peltonen, *The Cambridge Companion to Bacon*. Cambridge: Cambridge University Press, 99-120.

— 1988: *Francis Bacon's Idea of Science and the Maker's Knowledge Tradition*. Oxford: Clarendon Press

PHELPS, EDMUND 1967: 'Phillips Curves, Expectations of Inflation and Optimal Unemployment over Time', *Economica* **34** (August): 154-81.

PHILLIPS, ALBAN 1958: 'The Relationship between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957', *Economica* **25** (November): 283-99.

PICHELMANN, KARL AND ANDREAS-ULRICH SCHUH 1997: *The NAIRU-Concept: A Few Remarks*. Paris: OECD Economics Department Working Papers 178.

POLLAK, ROBERT 1998: 'The Consumer Price Index: A Research Agenda and Three Proposals', *Journal of Economic Perspectives* **12**(1): 69-78.

POPPER, KARL 1972: Objective Knowledge: An Evolutionary Approach. Oxford: Clarendon Press.

PUTNAM, HILARY 1975: 'On Properties', in: Hilary Putnam, *Mathematics, Matter and Method: Philosophical Papers* (vol. 1). Cambridge: Cambridge University Press, 305-22.

QUINE, WILLARD VAN ORMAN 1964: 'Implicit Definition Sustained', *Journal of Philosophy* **61**: 71-73.

QUINTON, ANTHONY 1980: *Francis Bacon*. Oxford: Oxford University Press.

RAPPAPORT, STEVEN 1998: *Models and Reality in Economics*. Cheltenham: Edward Elgar

— 2001: 'Economic Models as Mini-Theories', *Journal of Economic Methodology* **8**(2): 275-86.

RASMUSSEN, ERIC 1994: *Games and Information* (2nd edition). Cambridge: Blackwell.

RASMUSSEN, NICOLAS 1993: 'Facts, Artifacts, and Mesosomes: Practicing Epistemology with the Electron Microscope', *Studies in History and Philosophy of Science* **24**: 227-65.

REISS, JULIAN 2001: 'Natural Economic Quantities and Their Measurement', in *Journal of Economic Methodology* **8**(2): 287-312.

RICHARDSON, PETE, LAURENCE BOONE, CLAUDE GIORNO, MARA MEACCI, DAVE RAE AND DAVE TURNER 2000: *The Concept, Policy Use and Measurement of Structural Unemployment: Estimating a Time Varying NAIRU Across 21 OECD Countries*. Paris: OECD Economics Department Working Papers 250.

ROGERSON, ROBERT 1997: 'Theory Ahead of Language in the Economics of Unemployment', *Journal of Economic Perspectives* **11**(1): 73-92.

SABATÉS, MARCELO 1994: 'Problems for Kitcher's Account of Explanation', *Philosophical Issues* **5**: 273-282.

SAINSBURY, MARK 1988: *Paradoxes*. Cambridge: Cambridge University Press.

SALMON, WESLEY 1992: 'Scientific Explanation', in: Salmon *et al.*, *Introduction to the Philosophy of Science: A Text by the Members of the Department of the History and Philosophy of Science of the University of Pittsburgh*. Englewood Cliffs (NJ): Prentice Hall, 7-41.

SAMUELSON, PAUL AND WILLIAM NORDHAUS 1992: *Economics* (14th edition). New York *etc.*: McGraw-Hill.

SCHLICK, MORITZ 1925: *Allgemeine Erkenntnislehre*. Tübingen: Mohr Siebeck.

SCHMOLLER, GUSTAV 1871: 'Ueber die Resultate der Bevölkerungs- und Moral-Statistik', *Sammlung gemeinverständlicher wissenschaftlicher Vorträge* (herausgegeben von Rud. Birchow und Fr. v. Holtzendorff), VI. Serie, Heft 123.

— 1900: *Grundriß der Allgemeinen Volkswirtschaftslehre*. Leipzig: Duncker & Humblot.

— 1998/1872: 'Eröffnungsrede auf der ersten Tagung des Vereins für Socialpolitik 1982 in Eisenach', in: Heino Heinrich Nau, *Gustav Schmoller: Historisch-ethische Nationalökonomie als Kulturwissenschaft*. Marburg: Metropolis, 67-74.

— 1998/1881: 'Ueber Zweck und Ziele des Jahrbuchs [für Gesetzgebung, Verwaltung und Volkswirthschaft im Deutschen Reich]', in: Heino Heinrich Nau, *Gustav Schmoller: Historisch-ethische Nationalökonomie als Kulturwissenschaft*. Marburg: Metropolis, 97-114.

— 1998/1911: 'Volkswirtschaft, Volkswirtschaftslehre und –methode', in: Heino Heinrich Nau, *Gustav Schmoller: Historisch-ethische Nationalökonomie als Kulturwissenschaft*. Marburg: Metropolis, 215-368.

SCHURZ, GERHARD 1988: *Erklären und Verstehen in der Wissenschaft*. München: Oldenbourg

— AND KAREL LAMBERT 1994: 'Outline of a Theory of Scientific Understanding', *Synthèse* **101**: 65-120.

SHAPIRO, STEWART 2000: *Thinking about Mathematics*. Oxford: Oxford University Press.

SHOEMAKER, SIDNEY 1997/1984: 'Causality and Properties', in: Alex Oliver and Hugh Mellor, *Properties*. Oxford: Oxford University Press, 228-254.

SNYDER, LAURA 1999: 'Renovating the "Novum Organum": Bacon, Whewell and Induction', *Studies in History and Philosophy of Science* **30**(4): 531-557.

SORENSEN, ROY 1992: *Thought Experiments*. Oxford: Oxford University Press.

STAIGER, DOUGLAS, JAMES STOCK AND MARK WATSON 1996: 'How Precise are Estimates of the Natural Rate of Unemployment?', *NBER Working Paper 5477.*

— 1997: 'The NAIRU, Unemployment and Monetary Policy', Journal of Economic Perspectives 11(1): 33-49.

STIGLER, STEPHEN 1987: 'Jevons as Statistician: The Role of Probability', *The Manchester School* **55**(3): 233-56.

STIGLITZ, JOSEPH 1997: 'Reflections on the Natural Rate Hypothesis', *Journal of Economic Perspectives* **11**(1): 3-10.

STIGUM, BERNT 1990: *Toward a Formal Science of Economics : The Axiomatic Method in Economics and Econometrics*. Cambridge (MA): MIT Press.

STURGEON, SCOTT 1995: 'Knowledge', in Anthony Grayling, *Philosophy: A Guide Through the Subject*. Oxford: Oxford University Press: 10-26.

SUPPES, PATRICK 1968: 'The Desirability of Formalization in Science', *Journal of Philosophy* **65**: 651-64

THIEBOUT, C. 1972: 'A Pure Theory of Local Expenditure', in: Jerome Rothenberg and Matthew Edel, Readings in Urban Economics. New York: Macmillan.

THIRLWALL, ANTHONY 1983: 'What are Estimates of the Natural Rate of Unemployment Measuring?', *Oxford Bulletin of Economics and Statistics* **45**:

173-9.

THOMAS, JIM 1999: 'Quantifying the Black Economy: 'Measurement without Theory' Yet Again?', *Economic Journal* **109**(456): F381-89.

TIROLE, JEAN 1992: *The Theory of Industrial Organization*. Cambridge (MA): MIT Press.

TOBIN, JAMES 1980: 'Stabilization Policy Ten Years After', Brooking Papers on Economic Activity 1: 19-71.

— 1997: *Supply Contraints on Employment and Output: Nairu Versus Natural Rate*. New Haven (CT): Cowles Foundation Discussion Paper 1150.

TOOLEY, MICHAEL 1977: 'The Nature of Laws', *Canadian Journal of Philosophy* **7**: 667-698

URBACH, PETER 1987: *Francis Bacon's Philosophy of Science*. LaSalle: Open Court

— AND JOHN GIBSON 1994: *Francis Bacon: Novum Organum*, Chicago and La Salle (IL): Open Court

VAN FRAASSEN, BAS 1980: *The Scientific Image*. Oxford: Oxford University Press.

VARIAN, HAL 1992: *Microeconomic Analysis* (3rd edition). London: Norton

WIMSATT, WILLIAM 1981: 'Robustness, Reliability and Overdetermination', in: Marilynn Brewer and Barry Collins, *Scientific Inquiry and the Social Sciences*. San Francisco *etc.*: Jossey-Bass, 124-63.

WOODWARD, JAMES 1989: 'Data and Phenomena'. *Synthèse* **79**: 393-472.

— 1997: 'Explanation, Intervention, and Invariance', in: Lindley Darden, *Philosophy of Science Association 1996* (vol. 2): S26-S41.

WOOLHOUSE, ROGER 1998: *The Empiricists*. Oxford: Oxford University Press.

YOUNG, S. MARK 1985: 'Participative Budgeting: The Effects of Risk Aversion and Asymmetric Information on Budgetary Slack', *Journal of Accounting Research* **23**(2): 829-42.