The Methodology of Economic Model Building

Methodology after Samuelson

Lawrence A. Boland

ROUTLEDGE
London and New York

First published 1989 by Routledge
11 New Fetter Lane, London EC4P 4EE

Reprinted 1990, Paperback 1991

The orignal was printed in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire and simultaneously published in the USA and Canada by Routledge, then a division of Routledge, Chapman and Hall, Inc., 29 West 35th Street, New York, NY 10001

This version has been retypeset at Burnaby, B.C., December 2000, and then produced as an Adobe PDF file by Lawrence A. Boland.

British Library Cataloguing in Publication Data
Boland, Lawrence A.
The methodology of economic model building: methodology after Samuelson.
1. Economic models. Design & applications
I. Title
330(?).0724

ISBN 0—415—00014—9
ISBN 0—415—06462—7

Library of Congress Cataloging in Publication Data
Boland, Lawrence, A.
The methodology of economic model building.

Bibliography: p.
Includes indexes.
1. Economics —Methodology. 2. Samuelson, Paul Anthony, 1915— . I. Title.

HB131.B65 1988 330(?).0l(?)8 88—23919
ISBN 0—415—00014—9
ISBN 0—415—06462—7

To JOSKE, my unofficial senior supervisor

Contents

Acknowledgements

I wish to thank many people for taking the time to read the manuscript of this book. Those deserving particular praise are Irene Gordon, Paul Harrald, Soren Lemche and Donna Wilson for their significant advice and criticism. My friends Peter Kennedy, Kevin Hoover, Nancy Wulwick and Shayam Kamath are to be commended for their gallant efforts towards setting me straight with regard to Chapter 8 and Robert Jones for his careful examination of Chapters 2 and 3. Since much of this book was developed over a period of twenty-five years in my many methodology seminars at Simon Fraser University, I would like to acknowledge my students and friends, Geoffrey Newman, Wayne Pack and Kathryn Goddard who provided the support and criticism that I needed to understand the methodologically important questions.

In addition, I wish to thank the editors of the three journals who were willing to publish the original versions of some of the chapters included here. A shorter version of Chapter 6 was published by the **Atlantic Economic Journal**. Chapters 4 and 5 are revised versions of articles that appeared in **Philosophy of Science**. Earlier versions of Chapters 1 and 7 (as well as Section 3, Chapter 3) were published by the **South African Journal of Economics**. Finally, I thank the current editors of these journals for their permission to use copyright material.

Preface

Does testability of a model really matter in economics? How would one know whether testability can ever matter? Why should testability matter to a model builder? These are the key questions addressed in this book. They involve the lofty concerns of methodology as well as the more mundane matters of model building itself.

In the past twenty-five years I have read or listened to hundreds of papers on economic theory or applied economics. I do not recall an author ever saying something like, 'I chose not to use a Cobb-Douglas production function because it would make my model less testable'. If testability really matters to economic model builders then there ought to be evidence which shows that the extent to which a chosen assumption affected the testability of a model was a determining factor in its use. Frankly, I doubt whether any model builder ever takes into account what a given modelling assumption does to the testability of his or her model despite all the lip-service given to high-sounding methodological pronouncements regarding the importance of testability.

Testability, of course, does not represent all methodological issues in economics nor does model building represent all activities in economics. In this book I have focused exclusively on these two topics as objects of a research programme in applied methodology. In effect, this book reports the results of that programme which has occupied my writing and research for more than twenty-five years. As a methodologist I try to be self-conscious. Whenever I discuss methodology, I find that I almost always need to talk also about how to study methodology. For example, since I am reporting specific results of a research programme in applied methodology, I feel compelled to report what I learned about research programmes in general, as well as about meta-methodology. This can be confusing both for me and the audience, so I have tried to separate applications from the general lessons.

The lessons I have learned about methodology as a research topic are discussed separately in the Prologue and the Epilogue of this book. The primary lesson was that methodologists are often misled by philosophers and can learn more by trying to understand what methodologically aware economists have to say. The Prologue and Epilogue are then merely slices of bread surrounding the meat. The 'meat', which comprises Chapters 1 to 8, presents all that I have learned about the role and application of testability in economic model building. The main finding is that if testability were as important as most model builders claim, very little in economic theory and model building would see the light of day. While none of the chapters are reprints, it should be noted that some of the 'meat' presented here is reprocessed. Specifically, earlier versions of Chapters 1 and 4 through 7 have been published. The remaining chapters are new or have not been published: Chapters 2 and 3 are adaptations of my unpublished PhD thesis and Chapter 8 is almost entirely new with only a few sentences reappearing from an earlier reply to some of my critics. Everything previously published has been rewritten so as to form a coherent whole.

As a beginning academic in the 1960s working on applied methodology, there were very few journals in the world which would publish my work. In fact, I had 55 rejections before my first article was published in 1968. I learned a lot from those rejections. While I would have preferred to publish in the AER or the JPE, at first only the SAJE was willing to put my work into print. By now things have changed and it is much easier to find someone to publish methodology although the JPE and the AER still do not find much room for it.

I have written this book for two audiences. The primary audience is intended to be anyone who is interested in the specific question of how testability matters in economic model building. The secondary audience is envisaged as that growing band of young would-be methodologists who I think *ought* to be more concerned with how grand notions about methodology actually affect the practice of economics than with whether philosophers will turn their heads to listen. For this secondary audience the book should be seen as a case study. For the primary audience I think I have provided ample evidence that methodological concerns do matter in the everyday affairs of economic model builders and moreover, that model builders ought not to take methodology for granted.

L.A.B.
Burnaby, British Columbia
9 March 1988

PROLOGUE

Methodology vs Applied Methodology

> What methodology can do is to provide criteria for the acceptance and rejections of research programs, setting standards that will help us to discriminate between wheat and chaff. These standards … are hierarchical, relative, dynamic, and by no means unambiguous in terms of the practical advice they offer to working economists.
>
> Mark Blaug [1980, p. 264]

> Economists do not follow the laws of enquiry their methodologies lay down. A good thing, too.
>
> Donald N. McCloskey [1983, p. 482]

This book is a methodological examination of model building in modern economics. The act of building a model in economic theory always involves methodological decisions on the part of the model builder. Such decisions are sometimes about large questions, such as whether to build an econometric model to deal with existing empirical data or an abstract mathematical model which ignores empirical data. At other times decisions about more mundane questions must be made such as whether to use ordinary least-squares or two-stage least-squares or to use simple linear models rather than more complex models using non-linear functions. As these decisions always depend on the intended purpose for the model, there are very few salient methodological principles for model building in general. Nevertheless, over the last thirty years a few common concerns of model builders have evolved. A central issue has been the concern for the testability of economic models. While some philosophers have made sweeping pronouncements on the question of testability, hardly anybody has examined the basis for making methodological decisions in modern economics. How a practicing model builder in modern economics deals with the question of testability and similar methodological questions is critically examined in this book.

1. Model building in modern economics

Almost all recent textbooks in economics include frequent references to ‘models’. In modern social sciences the methodology of ‘model building’ has virtually excluded the older, more literary approaches to analysis. Although model building in the more advanced textbooks involves considerable use of mathematics, the concept of a model is more elementary and also more common. For example, there are fashion models, model homes, model students, model airplanes, design models, prototype models, and so on. Let us consider what these have in common.

Models are somehow neither ‘realistic’ nor ordinary – usually they are, in some way, artificial. Models are the outcomes of conscious efforts on the part of model builders. All models are models of some*thing*. In academic studies that something is usually a theory or explanation of some phenomena.[1] The methodology of model building focuses on the *adequacy* of any given model. However, ‘adequacy’ is not an absolute or universal criterion. The adequacy of a model can be judged only in terms of the model’s intended purpose or the aims of the model builder. The purposes for building models seem to fall into two general categories:

(1) *Pure* or *abstract models* which are representations of the underlying logic of the theory being modelled,

(2) *Applied models* which are explicit, simplified representations of more general theories and which are designed to apply to specific real-world problems or situations.

In addition, there are two different types of applied models: models of explanation and models for deriving policy recommendations. To a significant extent there is an abstract model underlying every applied model and the interaction between them will be the focus of our attention throughout this book.

[1] It should be noted here that in economics literature there are differences in how models are distinguished from theories. The difference used in this book corresponds to the usage in engineering literature – models are more specific or particular than the theories being modelled. Mathematical logicians give the opposite meaning. In their literature, a model represents the meaningless logic underlying a theory such that the theory amounts to a specific interpretation of the logic of the model [e.g. see Papandreou 1958 and Bronfenbrenner 1966].

1.1. An example of an abstract model used in economics

The most common abstract model used in economics is that of ‘maximizing behaviour’. Producers are alleged to be ‘profit maximizers’ and consumers are alleged to be ‘utility maximizers’. Although they have the simple idea of maximization in common, most textbooks have a chapter on the behaviour of the consumer and another chapter on the behaviour of the producer (the firm). Both chapters go through all the logic of maximization but each uses different words. Alternatively, some textbooks discuss the logic of maximization in general and then deal with the consumer’s situation and the producer’s situation, each as a special case. This latter approach, which has been used many times in recent years, involves the building of an abstract model, usually the one based on the logic of maximization [e.g. Samuelson 1947/65].

Before considering an abstract model of the generic maximizer, I would like to develop one of the specific ‘special cases’ so we can be sure that we know what the abstract model is supposed to represent. Let us look at elementary consumer theory.

Economists have a view that every consumer considers all the quantities of any particular good, say tomatoes, that he or she could buy, given his or her tomato budget and the price of tomatoes, and then buys the one quantity which maximizes his or her ‘utility’ (i.e. the total measure of satisfaction the consumer will experience from eating the tomatoes purchased). To make the choice easier, let us say the consumer can have as many tomatoes as he or she wants (i.e. tomatoes are free). However, let us also say that the consumer must eat all the tomatoes chosen within a specified amount of time, say four hours. The economist says that the consumer will thus choose the one quantity of tomatoes which maximizes his or her utility, neither more nor less. Generally, the consumer will not eat all the available tomatoes.

For the most part, what an economist has to say about the behaviour of any consumer is merely a logical consequence of the assertion that the consumer is a ‘utility maximizer’. This view, for example, says that whenever the consumer *is maximizing* his or her utility while facing an unlimited budget for tomatoes (or when tomatoes are free), it must be the case that the consumer has chosen that quantity of tomatoes, say 10 pounds, such that if the consumer were to eat one *more* tomato, his or her total satisfaction (measured on some scale which we call 'utility') would decrease. Economists express this by saying the marginal utility of an extra tomato will be negative. Economists also say that for each pound up to the chosen 10 pounds, the consumer's total satisfaction increases with each additional pound. That is, the consumer's marginal utility is positive for each additional pound up to the tenth pound. Since (as they say) marginal utility is positive for each additional pound below ten and negative for each additional pound over ten, usually we can also conclude that the marginal utility must be zero at 10 pounds.

There are three distinct ideas used in this elementary *theory* of the tomato consumer facing unlimited choice.

The assumption that:

(a) utility is being maximized by the choice of the quantity of tomatoes.

The conclusions that:

(b) marginal utility is zero at the chosen quantity of tomatoes,

(c) marginal utility falls as the quantity of tomatoes rises.

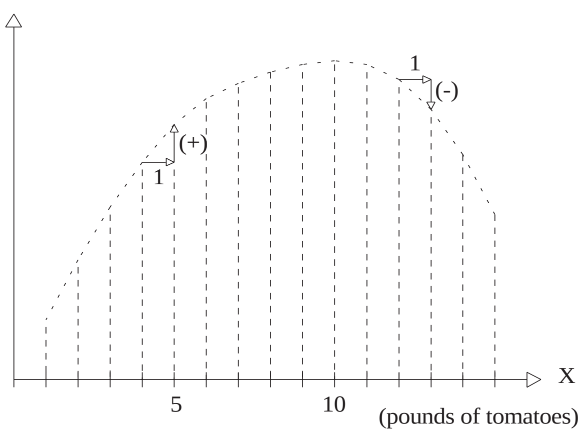

Figure 1 *Maximization Model*

It is said that (b) and (c) are logical consequences of (a). That is, if the condition of utility and chosen quantity are as indicated by (a) then *necessarily* both ideas (b) and (c) are true.

There are many different *models* of such a theory. They are different only because we choose to represent the basic ideas of the theory in different ways. Most often economists use diagrams to represent their ideas. In this elementary consumer theory case, the level of utility is represented with a diagram on which we record the levels of satisfaction or 'utility' that would be obtained by the consumer at each given amount of a good (see Figure 1). The curved dotted line connecting all the recorded points is supposedly determined by the consumer's psychologically given 'utility function'.

Whenever we say the consumer chooses 10 pounds of tomatoes because at 10 pounds his or her utility is maximized, we can represent that choice on the diagram by using the numbers along the horizontal axis to represent 'pounds of tomatoes' (i.e. **X** equals the 'quantity of tomatoes') and positioning the utility function such that the maximum occurs at **X** = 10, representing 10 pounds. This implies that the consumer is free to choose any quantity he or she wants. It is easy to see that ideas (b) and (c) will follow straight from this diagram. To the left of 10 pounds, as each pound is added to the quantity of tomatoes (e.g. the fifth pound), the total utility gets higher, i.e. the level of satisfaction received changes by a positive amount. This means, according to our diagram, that the marginal utility (MU) to the left of 10 pounds is positive. We can also note that for each additional pound, the amount by which the total utility goes up is less and less as we get closer to 10 pounds. From this we can conclude that 'marginal utility falls with rising quantity of tomatoes', i.e. idea (c) above. And since not only is the marginal utility to the right of 10 pounds negative, for a very small change in the neighbourhood of 10 pounds, it would be very difficult to detect any change in utility and thus we can say that marginal utility is zero at the chosen quantity of tomatoes, i.e. idea (b) above.

The above conclusions follow from the geometrical properties of our diagram and they would hold for *maximization* in general. That is, calling the horizontal scale 'pounds of tomatoes' is a specific model of the utility maximizing consumer. Thus we see that one possible *abstract* or *pure* model of utility maximizing consumers is merely the logic of the geometrical properties of the diagram without any labels on the axes. The logic of the diagram is more general than the specific case of choosing *tomatoes* in order to maximize *utility*. The logic holds for *any* case of maximizing any measure with respect to any product. Whether or not utility maximization is a true theory for the choice of tomatoes is a separate matter. Abstract models are intended to be true for all cases for which they apply but specific models may be true only for the case represented (e.g. for the choice of tomatoes but maybe not for wine).

1.2. Models of explanation and policy recommendations

In economics, we say that every explanation is of some (observable or observed) events. Each of these events is described as one of the many possible values or states for each of a specified set of variables (called *endogenous* variables since they are supposedly determined logically *within* the model explaining their values). The explanation of the particular observed events (or values) requires *both* another set of variables, each of which has many possible values or states (these are called the *exogenous* variables since their values are determined *outside* or independently of the model), *and* one or more universal principles (e.g. behavioural assumptions) relating to the two sets of variables in some specific way. The two sets of variables are distinguished only as follows: Endogenous variables are alleged to be *influenced* by exogenous variables as indicated by the universal principles, *but* exogenous variables are alleged *not* to be influenced by endogenous variables in any way. In this sense the exogenous set is called the ‘givens’ or sometimes the ‘initial conditions’. Now an explanation of one or more events (each represented by a list where there is one value for each endogenous variable) is accomplished by showing that, by means of the universal principles, *only those* events are logically compatible with a specified list of values, i.e. a list with one value for each exogenous ‘given’. The explanations are considered successful when the actual (observed) values of the ‘givens’ correspond to the logically necessary set of values. A different set of (observed) events would be compatible *only* with a different set of ‘givens’.

In the case of explaining a change in any endogenous variable (i.e. a change to a different list of values), we are limited, by the logic of this theory of explanation, to explaining a change as being the result of some change in the values of one or more of the exogenous ‘givens’.

Exogeneity is an asserted attribute of variables – it is not usually a property of the logical structure of a model. By itself, an equation of a model cannot represent exogeneity. Very often students find that it is possible to manipulate one or more equations of a simple macro-economic model and produce an equation which would *appear* to indicate that the observed values of the exogenous variables are determined by the observed values of the endogenous variables. Such an *interpretation* of an equation would contradict the intended meaning of exogeneity attributed to those variables.[2]

Recognition of the distinction between exogenous and endogenous variables is crucial in the building of models of the economy whenever those models are to be used as the basis for policy recommendations. One good reason why some variables are exogenous is that they are totally controlled by some autonomous institution or individual. Some examples of exogenous variables of this type might be tax rates, the level of advertising, the government’s budget, private investment or loans, and so on. These are examples of potential policy implements. Policy questions are about what would be the benefit or cost of changing one or more of these variables. For example, what would happen to the rate of inflation if the government changed the tax rate? If the tax rate were not an exogenous variable, this would be a meaningless question. We can talk only about directly changing something over which we have control.

In summary, when building models for either explanation or for policy recommendations, some variables must be exogenous. We cannot build models with only endogenous variables. Saying that there *must* be some exogenous variables means that we can never expect to be able to

[2] To illustrate, let us say that over the past ten years we observe that the average price of wheat each year has been inversely proportional to the amount of rainfall per year. We can put that in the form of a simple equation which would tell how to calculate the average price whenever you know what the expected rainfall is:

$$\mathbf{P_W = A/R}$$

In this equation, $\mathbf{P_W}$ represents the average price of wheat, **R** represents the annual rainfall measured in centimetres and **A** is a fixed proportionality parameter which translates centimetres into dollar prices. As a matter of algebraic (or logical) manipulations alone, we could reverse or ‘solve’ this equation for **R** and get:

$$\mathbf{R = A/P_W}$$

Now, even though we can solve or reverse the equation for $\mathbf{P_W}$ to make it an equation for **R**, we cannot interpret either to mean ‘the price of wheat determines the level of rainfall’. Stated another way, rainfall is an exogenous variable. It might make sense to say the level of rainfall influences the price of wheat, but it usually would not make sense to say that the price of wheat influences the level of rainfall.

explain everything. And going further, exogeneity by itself does not imply control. If we want to make policy recommendations with the explanatory models we build, some of the needed exogenous variables must represent variables over which someone has control. Deciding which variables are to be included in a model, and deciding which of them are to be exogenous, are the most fundamental methodological decisions that a model builder must make.

2. Methodology as a study of model-building methods

Since the early 1960s, graduate education in economics has been increasingly devoted to teaching the modelling methods that are widely accepted in the economics profession. It is all too easy for graduate students to reach the conclusion that the best graduate programmes are those that teach the most up-to-date modelling methods. And given such proclivities, when the graduate students become teachers, it is too inviting for them to try to convince their undergraduate students that there must be only one approved method. Some even turn to the study of methodology in hopes that philosophers can add support to their current choice of an approved modelling method.

When I was an undergraduate student of economics I read a book by Henry Briefs, **Three Views of Method in Economics** [1960], which I found liberating. For the first time I was encouraged to dismiss the common philosophical notion that there must be only one correct and authoritative method for economics and was free to consider many views of economic methodology. The feeling of liberation was due more to the rejection of the authoritarianism of most methodologists than to the specific views of method that Briefs discussed in his book.

While, morally speaking, one ought to allow for a plurality of reasonable views, such 'pluralism' can be debilitating. What if economists had to simultaneously allow that individual decision-makers can be assumed to be maximizers, non-maximizers and satisficers? Surely, economists wishing to apply economic theory would reach an insurmountable impasse. While some economic theorists might be able to survive by allowing for all possible behavioural assumptions, every applied economist learns quickly that to succeed he or she must adopt one position (e.g. one behavioural assumption) and give it a whirl.

Despite the efforts of many methodologists and philosophers, working economists eschew most disputes over the realism of their assumptions. The proof of the assumed pudding, they say, is always in the eating – or more specifically, in the results of the modelling based on an assumption in question. If you think you have a more realistic assumption, they will tell you to build a model using it and show that it yields superior results. By analogy, I think one can also say that if you think you have a superior methodology then show us. Apply that methodology to some important model building process and show that it yields superior models. Even if you are not convinced that you have found the world's best methodology, to practice economics you must adopt a single methodology. This book is about the application of one particular methodological approach to model building in economics. It is about the one methodological rule which has dominated economics since the early 1960s. The rule at issue is the methodological requirement that all economic models or theories, if they are going to be given serious consideration by practicing economists, must be shown to be testable where a successful test of a theory is defined as a falsification of that theory. A testable theory is a falsifiable theory.

2.1. Testability as falsifiability in economics

The methodological requirement of falsifiability is so commonplace today that it is difficult for most of us to think of economists ever putting forth theories or models that are not falsifiable. Falsifiability (or equivalently, refutability) was certainly in the air even during the 'years of high theory' (i.e. the 1930s). Most historians of economic thought credit Terrence Hutchison's 1938 book, **The Significance and Basic Postulates of Economic Theory**, with the explicit introduction of the methodological requirement of falsifiability to economics. Hutchison refers often to the philosopher Karl Popper in explaining the nature and significance of falsifiability. According to Popper, and in opposition to the commonplace view of his time, falsifiability rather than verifiability was to be stressed as the primary attribute that makes theories scientific. At about the same time (1937), Paul Samuelson introduced to economics a different methodological requirement – namely the requirement that economic models and theories must be 'operationally meaningful'. As Samuelson defines the issue,

> By a meaningful theorem I mean simply a hypothesis about empirical data which could conceivably be refuted, if only under ideal conditions. A meaningful theorem may be false. [Samuelson 1947/65, p. 4]

Unlike Hutchison, Samuelson makes no reference to Popper or any other philosopher to explain why falsifiability is methodologically important.

Hutchison's pronouncements did not seem to have much effect on the economics profession and Samuelson's contribution to economic methodology did not appear in print until 1947. But even as late as the mid-1950s, verifiability still held the attention of many economic methodologists [see Machlup 1955; Hutchison 1956]. However, a specific effort to make empirical testing central to economics and economic methodology was made by a group of young economists at the London School of Economics (LSE). The venue was the 'Staff Seminar on Methodology, Measurement, and Testing in Economics' [see de Marchi 1985/88]. For some reason, the idea of empirical testing was subsumed under the flag of 'positive economics'. While one of the members of the seminar, Richard Lipsey, made empirical testability a central concern of his 1963 textbook **An Introduction to Positive Economics**, another member, Chris Archibald, was pleading for a more palatable alternative to the severe requirement of empirical refutability [Archibald 1966].

Since the first edition (1961) of his famous textbook on the history of thought, **Economic Theory in Retrospect**, Mark Blaug has continued to promote falsifiability as the operative methodological rule in economics, culminating in his 1980 book, **The Methodology of Economics**, which expanded on the earlier textbook's treatment of methodology. Blaug makes an important distinction. He notes that 'economists frequently preach falsificationism … but they rarely practice it' [1980, p. 128]. Note carefully that for Blaug, any practice of what he calls falsificationism amounts not only to devising models which are in principle refutable but also to actively attempting to refute such models. With the possible exception of a brief moment in the LSE seminar, hardly any mainstream economists have advocated such a strict employment of the Popper-Samuelson methodological requirement of falsifiability. The view that anyone who advocates the employment of a requirement of falsifiability must also be advocating the pursuit of falsifying tests is a mistake, which incorrectly attributes concern for falsifiability exclusively to popular (mis)interpretations of Popper's philosophy of science. For practicing model builders, the issue of required falsifiability has more to do with avoiding vacuous tautologies than with philosophical concerns over the true purposes of science. Avoiding tautologies is quite explicit in Hutchison's 1938 book and the above quotation from Samuelson's 1947 book clearly shows that he thought by requiring refutability he was thereby assuring that theorems can be false. To say a statement can be false is just another way of saying that it is necessarily not a tautology.

So, the fact that mainstream economists advocate the methodological requirement of falsifiability yet do not spend all their time attempting to refute mainstream economic theories or models does not constitute an integrity failure as Blaug seems to suggest. It merely means that economists are more interested in what Samuelson had to say about methodology than what philosophers such as Karl Popper might think the purpose of economic science should be.

2.2. Against methodology as the study of Big Questions

The widespread adoption of the methodological requirement of falsifiability, whether it be implicit or explicit, makes an interesting topic for a research programme in economic methodology. But given the depth of concern for the philosophical aspects of the Popper-Samuelson requirement of falsifiability, it is important to keep the philosophical questions in proper perspective. Modesty and humility are essential for any methodological study of practicing economists.

Despite what a few well-intentioned methodologists have been arguing recently [e.g. Caldwell 1982], there is very little that an ordinary philosopher can do to help the typical model builder in economics. While Plato, Descartes or John Stuart Mill may have had profound ideas about the nature and purpose of human thought, such profound ideas will not likely help the practicing economist who is attempting to measure the level of unemployment in the forest industry or who is attempting to determine if individual utility maximization is consistent with the market determination of prices.

Unfortunately, one of the reasons why most people study methodology is that they are interested in the Big Questions about the nature and purposes of human activity. They are interested in what Don McCloskey [1983] would call 'Methodology with a capital M'. Very little has been published about lower-case methodology – i.e. the practical methodology of the practicing economist. Of course, it is somewhat risky to write about practical methodology since there may not be an audience. The ordinary philosopher or upper-case Methodologist will not find practical (lower-case) methodology of any interest since it never seems to answer Big Questions, and practicing economists will turn away because they think all methodology is restricted to the theory of Big Questions.

The schism between practice and theory has haunted academic economists for many decades. Alfred Marshall attempted to bridge the gap in 1890 with his famous **Principles of Economics**. John Maynard Keynes attempted the same thing in 1936 with his famous **General Theory of Employment, Interest and Money**. With their 1958 book, **Linear Programming and Economic Analysis**, Robert Dorfman, Paul Samuelson and Robert Solow tried to explain existence proofs and fixed-point theorems to the average business graduate student. While these books have been successful in the wider sense, none of them seems to have built a strong enough bridge and the schism remains.

One sees this schism clearest when it comes to using economic theory as a policy guide for governmental agencies. Consider Alain Enthoven's view of economic analysis in the US Department of Defense:

> the tools of analysis that we [government economists] … use are the simplest, most fundamental concepts of economic theory, combined with the simplest quantitative methods. The requirements for success in this line of work are a thorough understanding of and, if you like, belief in the relevance of such concepts as marginal products and costs in complex situations, combined with a good quantitative sense. The economic theory we are using is the theory most of us learned as sophomores. [Enthoven 1963, p. 422]

After conducting a survey of governmental economists, William R. Allen observed that:

> In performing their chores … government economists are subject to various constraints … [T]he world of the government economist can be, and typically is, very different in important respects from the world of the academic economist.
>
> No one seriously disputed the essence of [Enthoven's] characterization of governmental economics work done in the early 1960's. And few believed that the level of rigor and technical sophistication in economics … had increased strikingly during the following decade. [Allen 1977, pp. 56 and 73]

Virtually everyone Allen interviewed noted that there was seldom sufficient time to employ the sophisticated modelling techniques learned in graduate school and still meet the demands placed on practicing governmental economists.

The schism between upper-case Methodologists with their love of the Big Questions and the lower-case methodologists with their desire to be helpful is, I think, analogous to the schism that Allen observes. To be helpful to practicing model builders in economics, the grandiose methodological schemes that one might learn from reading the leading philosophers of science offer very little that can be applied to the practical methodological decisions that a model builder must make. Nevertheless, this schism too must be bridged and it is hoped that in what follows some progress will be made.

3. Applied methodology as a research programme

The chapters which follow represent a research programme that I began in the early 1960s. The programme was directed at bringing some commonplace notions to bear on some everyday methodological decisions of practicing economists. When engaging in applied methodology it is important to keep things simple. While philosophers were still bickering over the intellectual merits of falsifiability of scientific theories, I set about seeing whether testability or falsifiability matters to the practicing economist who is actively building models. In the following chapters, I present the results of my research programme so far. After the initial chapter which argues that testability matters to both applied and pure theorists in economics, I present the foundation of my research programme which is an uncritical application of both the methodological requirement promoted by Paul Samuelson and the philosophical demarcation developed by Karl Popper. My foundation is an operationalization of Popper's notion of degrees of testability. Chapters 2 and 3 show that rather ordinary modelling assumptions can lead to very different degrees of testability. In addition, I argue that if testability matters as much as most economic model builders claim, it must be recognized that some modelling decisions increase testability while others make testability a *practical* impossibility. One might be tempted to ask whether the impossibility is due to broader issues than those of practicalities, but this question is postponed until Chapter 7. Instead, in Chapters 4 and 5, I ask whether testability is being used to conceal more important concerns such as the truth status of economic models.

As one must not throw the baby out with the bath water, I return to a consideration of how testability can matter to those of us who are still concerned with the truth status of economic models. In Chapter 6, I show how testability can be made an essential part of any informative equilibrium model. If one's purpose for building models of a behavioural theory is to test that theory then, in Chapter 7, I show that there are some substantial obstacles precluding successful tests. Does this mean that the methodological foundation of model building in modern economics is an impossible dream? Is it really impossible to test the truth status of economic theories by testing models of those theories? In Chapter 7 things may look bleak, but in Chapter 8 a way is found to make falsifiability a realistic methodological requirement for model building in modern economics. The book closes with some suggestions for those economists who wish to engage in their own research programme based on the study of economic methodology.

PART I

Applications of the Popper-Samuelson Demarcation

1

Economic Understanding and Understanding Economics

> Most sciences make a legitimate distinction between their *applied* and *pure* aspects, even though the borders are fuzzy.
>
> Paul A. Samuelson [1962/66, p. 1668]

> Mainstream neoclassical economists … preach the importance of submitting theories to empirical test, but they rarely live up to their declared methodological canons. Analytical elegance, economy of theoretical means, and the widest possible scope obtained by ever more heroic simplification have been too often prized above predictability and significance for policy questions.
>
> Mark Blaug [1980, p. 259]

> The task of science is partly theoretical – *explanation* – and partly practical – *prediction and technical application*. I shall try to show that these two aims are, in a way, two different aspects of one and the same activity.
>
> Karl R. Popper [1972, p. 349]

Today it is safe to say that most economists are concerned with practical problems and that they view the purpose of all economic theories as helping to solve these problems. It is not as safe to say (but it is equally true) that today most, if not all, economists are instrumentalists. That is, most are less concerned with the truth status of economic theories and more concerned with whether their theories produce useful results or predictions. This currently dominant methodological and philosophical bias in favour of practical problems and results can be most disconcerting for the large and ever-growing band of economic theorists who are often called ‘pure theorists’. They are called ‘pure theorists’ because they (unfortunately) exclude the impurities of real-world complexities. Today we simply call them our 'theorists'.

Pure theorists are generally a hardy lot, hence we need not try to protect or defend them. There has been, however, a lot of needless controversy over certain methodological decisions which could have been avoided if we had been able to see that methodological controversies result from differences in purpose or that the objectives for many methodological decisions may be different. Given the prevailing bias regarding the purpose of theorizing and model building, it is quite difficult for the ordinary economist to see that the pure theorist is doing something quite different, and that his or her criterion for success may indeed be quite different.

The differences seem to be basically the following. Pure theorists are seeking to improve their *understanding* by creating new theories and models.[1] Having to assume something which leads to (one or more) false statements is totally unacceptable no matter how many true statements may validly follow. When they assume something which leads to statements which contradict known facts, they know that their bases for understanding (i.e. their assumptions) are, at best, inadequate. They would know that they do not really understand. Ordinary economic theorists – so-called 'applied theorists' – are seeking to solve practical problems. They attempt to create or modify theories and models to arrive at solutions to social problems or at specific policy directives. When they assume something, they would prefer to assume something which is known to be true, or 'approximately true', so as to guarantee that their conclusions are true or 'almost true' (assuming mistakes in logic are not made). Nevertheless, their operative criterion for acceptability of alternative assumptions is whether or not the assumptions lead to a (desirable) solution – the truth status of their assumptions is a secondary matter.

In this chapter I attempt to explain some of the methodological implications of this distinction in purposes, and particularly with respect to the question: Why do we assume what we assume? To assume one

1 If I take my TV set to an applied theorist (a TV repairman in this case), I do not question his understanding of electromagnetics – only his success at fixing my set. For all I know he may believe that there are little men in the transistors and thus he replaces transistors until he finds the 'culprit'. If he succeeds in finding and replacing the defective transistor, I can no longer distinguish his understanding of TV reception from that of any other repairman or even a modern electronics engineer.

thing and not another is a methodological decision that all model builders and all theorists (pure or applied) have to make. And it is a decision which should always be open to criticism if we desire to learn anything at all about our understanding of economic phenomena and economic problems.

Throughout this book I am particularly concerned with the role of models in both the study of economic phenomena and the understanding of given economic theories. The former role centres on practical problems and the latter on what is usually called 'axiomatics'. The distinction in roles is made because the methodological decisions a model builder will make (with respect to particular assumptions) depend primarily on the purposes for which the model will be used. Given different purposes, our models may differ considerably with respect to such characteristics as generality and completeness. And a requirement (such as testability, simplicity and universality) of one model builder need not apply to another. Realizing the inapplicability of some requirements can help us understand many methodological controversies in economic theory and practice. However, identifying differences in purposes does not mean that the pure theorist can at last feel safe in his or her 'ivory tower'. At the end of this chapter I present an argument for how pure theorists, who may be concerned more with modelling theoretical ideas than with solving practical problems, can still be in touch with the real world of the applied theorists.

1. The role of models in economics

A theory is a *system* of ideas – or more specifically, a collection of systematically interrelated ideas. Theories, or systems, are necessary because individual ideas will not do the job, whether it be an intellectual job of understanding (or explanation) or a practical job of recommending an economic policy. Of course, an important intellectual job to be done is explanation, namely the explication of a particular answer to a particular question (or a solution to a problem). A typical practical job might be providing a solution to a practical problem and thus may require an adequate description of the circumstances of the problem.

The usual methodological controversy over whether theories are descriptive, explanatory, or predictive, and so on, stems from disagreements over the purposes, that is, the jobs to be done. So long as one is not looking for *the* purpose, there should be no controversy here.

Theories created to do an intellectual job *may* be able to do a practical job. When applied economists are faced with a practical problem their task is to find a solution, or if they think they have a solution, they may wish to convince a policy-maker to employ their solution. One important source of solutions is the 'implications' of the various pure theories developed so far, thus the applied economist's task may become one of choosing between those theories.

Parenthetically, I should warn readers that economists too often use the terms 'implication' and 'logical consequence' interchangeably, even though this irritates some philosophers. A similar confusion is raised by the economists' use of the term 'tautology' when they really mean either a statement which is true by virtue of the definitions of its terms or any statement that is impossible to refute once certain conditions are accepted (e.g. the *ceteris paribus* condition) [see Boland 1981b]. While one might feel compelled to set economists straight on these distinctions, little is ever accomplished by trying to do so since these distinctions do not seem to matter relative to what economists intend to mean by their use. Nevertheless, whenever it is important to do so, I distinguish between how economists use these terms and their proper definitions.

What is the applied economists' criterion for choosing between theories? Is it the truth status of their theories? It need not be truth status unless the solution requires holistic, or very large-scale, changes [see Popper 1945/62]. The criterion need only be the success of a policy or solution implied by a theory. But, if there is no opportunity to try the policy or solution, an applied economist may rely on a comparison of the degrees to which the alternative theories have been 'confirmed' – i.e. on how much supporting evidence they have. Unfortunately, this criterion is not very reliable since it is not at all clear what is meant by 'confirmed' [see Agassi 1959, 1961; Popper 1965] and since the relative amount of supporting evidence can depend more on historical accident than the truth status of competing theories. Above all, whatever is meant by 'confirmed' it should never be used to defend against any criticism of a chosen policy since 'confirmed' never means verified (i.e. shown to be true). Obviously, if one of the alternative (pure) theories is true, it would be the theory to use, but of course we can rarely if ever prove a theory to be true because all theories involve universal statements (e.g. '*all* consumers are maximizers') which are by themselves not verifiable (no matter how much 'supporting' evidence we find) [see Popper 1959/61].

Most theories, and especially the individual ideas which constitute their ingredients, are non-specific with respect to formal relationships between the concepts involved in the ideas. For example, a theory may be based on the general idea of a downward sloping demand curve. We could assume an arbitrary curve, or be explicit and assume a particular shaped curve such as $\mathbf{P} = a + b\mathbf{Q}$ (where a and b are assumed to be fixed parameters) or even be more specific and assume, for example, $\mathbf{P} = 14.2 + 3.7\mathbf{Q}$. If our theories are in the form of explicit equations, then the policy implications of the theory can usually be made quite specific [cf. Brems 1959, Kuenne 1963]. Whenever one's theory is not explicitly stated, one way to determine if a theory is useful in a given practical situation is to build a 'model' of the theory much in the spirit of design engineering. For example, design engineers will often build a small model[2] of a new airplane design to test its aerodynamics in a wind tunnel. The design engineers – as applied theorists – will commit themselves to a specific model. Of course, many different models may be constructed (all based on the same new idea) by varying certain proportions, ingredients, and so on. Such opportunities for testing with scaled-down models seldom arise in economics. Thus when so-called 'applied economists' use a model, they should allow for error in a manner which will indicate the risk involved in being in error. One technique for such allowance is to use 'stochastic' models in which the possible error is explicitly represented in the model.

The word *stochastic* is based on the idea of a target and in particular on the pattern of hits around a target. The greater the distance a given unit of target area is from the centre of the target, the less frequent or dense will be the hits on that area. We might look at a model as a shot at the 'real-world' target. There are many reasons why we might miss the target, but they fall into two rough categories: (1) ours was a 'bad' shot, i.e. our model was false, and (2) the target unexpectedly moved, i.e. there is random *unexplained* variation in the objects we are attempting to explain (or use in our explanation). A stochastic model is one which systematically allows for the movements of the target. Stochastic models follow from a methodological decision *not* to attempt to explain anything *completely*. Non-stochastic models, by contrast, may attempt to give

2 Here, I speak of a *small* model to distinguish it from a prototype model which is a *full-scale* working model.

complete explanations. They do not allow for bad shots or moving targets.

To choose one approach over another involves a methodological decision. This decision is not open to theoretical criticism, but only methodological criticism. If our present theories are quite inadequate for practical purposes, then it might be wise to build stochastic models when we are interested only in practical usefulness. However, if we are interested in a *complete* understanding of phenomena, especially by improving on our present theories, building stochastic models may not be an appropriate strategy.

By constructing stochastic models, applied economists can try out various theories and compare them on the basis of the (potential) error (with respect to known data). The objective of the applied economist would be to use the policy implications of the model which minimizes the error (given the data at hand).

The applied economist assumes the truth of his or her model in order to apply it (see Chapter 3). The most common approach to dealing with the truth status of the model is to estimate the value of the parameters econometrically. In this approach one avoids specifying (a-priori) values for the parameters and uses actual data *ad hoc* to deduce the values of the parameters. The logically deduced values of the parameters are those of the hypothetical parameters of the posited model and are obtained only by assuming the model to be true (i.e. by assuming that together all the statements in the model form a true compound statement).

It must be stressed here that neither the applied economists' criterion of practical success nor their procedures will satisfy the pure theorists who routinely accept Karl Popper's philosophy of science. A false theory or model must surely be rejected regardless of how 'useful' it may be. If we do follow Popper in this manner, then in order to empirically criticize a theory we are required only to find one fact which is not compatible with that theory. Indeed, the existence of just one such fact indicates that at least one of the theory's assumptions is false. In other words, the discovery of even one fact which contradicts one of a theory's implications constitutes a refutation of that theory.

It might seem that non-stochastic *models* could play an important role in the empirical testing of theories which are not formally specific. However, does refuting a model of a theory necessarily constitute a refutation of that theory? The answer is clearly negative because in constructing a model we must add formal restrictive assumptions which are not necessary for the theory's job of explaining or justifying a conjecture. A model of a theory simply is not an implication of that theory, nor are the implications of a model necessarily identical to the implications of the theory it represents. To refute a theory using models requires that we show that *all* the possible models of a theory will be false.[3] It would seem that we have returned to Popper's requirement that we need to show at least one of the theory's implications to be false.

In the case of a theory which is explicit as to formal relationships – i.e. a theory which looks like a model – we could refute it by showing one of its solution statements[4] to be false. For the pure theorist who wishes to construct model-like theories to explain or demonstrate his or her conjectures it would be best to offer as many different (model-like) theories as possible (by varying the formal assumptions) in order to see the implications of the assumptions of particular forms of relationships. This, of course, would seemingly add to the intellectual efficiency since all the variations of the (model-like) theory could be tested simultaneously with the same data. But as with using models to test (more general) theories, the number of possible models (or in this case, possible variations of model-like theories) would be so enormous, and the individual models (or model-like theories) so specific, that we could expect them all to be false. If this is indeed the case – as we can be sure that it is – why do economists, including even some pure theorists, build model-like theories? For some it is a confusion as to methodological objectives – i.e. seeking solutions to theoretical problems vs seeking solutions to practical problems – but these economists can be excused for now. There are others who do understand the objectives of pure theory and still build model-like theories. And the reason clearly is that explicit models offer the opportunity to use a wealth of mathematical theorems. That is, by using mathematical theorems it is easier to show that a theory logically supports a particular conjecture. One of the methods of showing that a (model-like) theory does the job is to 'axiomatize' it and then demonstrate that it is consistent and complete –

3 Note that this requirement is similar to the verification of a universal statement (e.g. 'all swans are white'); it is a logical impossibility (we would have to show this for *all* swans that currently exist, ever did exist, or ever will exist anywhere in the universe).

4 When we say a solution for a system of equations, we mean a set of values, one for each endogenous variable. I speak here of a statement of the value for one of these variables given the values of the necessary parameters and exogenous variables – that is the 'solution statement' for that endogenous variable.

i.e. show that there exists a solution to the model and that the solution is unique. Being able to do a thorough and rigorous job in developing a theory is one of the advantages of building model-like theories.

2. On the foundations of economic ignorance: axiomatics

We assume when we don't know. The behavioural hypotheses upon which we build our economic theories are, so to speak, representations of our ignorance. Since the behavioural hypotheses are the foundation of our economic theories it seems reasonable to make them an object of study. One of the methods of studying the foundation of economic theory is called *axiomatics*. Before considering other uses for models, I will attempt to outline the nature of an axiomatic study and then discuss how model building may be a useful part of such a study.

Unfortunately, the only axiomatic studies completed so far have been by mathematicians and the more mathematically inclined economists – most notably Gerard Debreu, Kenneth Arrow, Abraham Wald and Tjalling Koopmans. Axiomatics as a distinct formal method has been around little more than 100 years. Originally it was concerned with the axioms and postulates of Euclid's geometry, and it is usually discussed in those terms. I shall attempt to refrain from that type of presentation here.

There has been little written expressly on axiomatics although little bits and pieces are to be found in advanced books on logic. It seems to be taken for granted that '[axiomatics] is a science in which we never know what we are talking about nor whether what we are saying is true' or that '[axiomatics] is the art of giving names to different things'. In all honesty one can say that its reputation comes from its being concerned only with form rather than substance and its deliberate attempt to create *systematic ambiguity*.

For the most part axiomatics is concerned with what are called 'primitive terms', 'primitive propositions' (postulates or axioms) which relate these primitive terms, and 'systems of postulates' (axiomatic systems). Axiomatics, then, is the study of the logical properties of such a system.

A system of postulates and primitive terms is analogous to a system of equations and variables. The problems associated with the properties of equation systems are similar to those which occur in axiomatics. When we solve a system of equations for **X**, we do so to answer the question: What determines the value of **X**? In axiomatics we might ask: On what basis can one logically derive the conclusion (or theorem) A? We might answer: (1) we can, using such-and-such axioms, or perhaps (2) given certain axioms, we cannot because the given system of axioms is either insufficient for such a derivation or involves a contradiction. Let us now look at some of the properties of axiomatic systems, namely the logical properties of consistency and completeness.

Consistency requires that the set of assumptions which make up a theory does not entail inconsistencies. As with systems of equations, a theory expressed as a system of axioms needs to be consistent, but with one important difference. Although a system of equations can be shown to be consistent, a system of axioms seldom can. However, a system of axioms can be shown to be inconsistent (as can a system of equations). This is an important consideration for the philosophy of science, since it is related to the difficulty of verifying a theory. To prove that a system is consistent it is necessary to show that *there does not exist* a statement for which both it and its denial can be deduced from the system.[5] To verify an explanatory theory (one which includes assumptions such as 'all consumers are maximizers') we must show that *there does not exist* a refutation, i.e. a false statement which is deducible from the theory. In each case, we may wish the positive but we can only prove the negative.

Consistency is obviously important since we cannot tolerate contradictions in explanations. For example, a theory which purports to explain resource allocations cannot imply that an economy at time t is both on and not on its production possibilities curve and be consistent. Consistency, however, does not rule out the possibility of a theory allowing for competing or contrary situations such as multiple equilibria. For example, all points on the production possibilities curve are potential equilibria differing only with regard to equilibrium price ratios. Any model which allows for flat spots on the production possibilities curve implies the possibility of more than one point being consistent with the same equilibrium price ratio.

Usually the question of consistency can be dealt with in a rather direct way by attempting to solve the system of equations constituting the

[5] One proof of an *in*consistency would be the case where a contradiction is possible, that is, where both a given statement and its denial are logically allowed by the theory.

model of the theory. If an acceptable solution cannot always be obtained, it may be possible to specify additional conditions to guarantee a solution. Or, if non-acceptable solutions (e.g. negative prices or outputs) are logically possible, it may be possible to eliminate them by further specification, as is done when specifying that the marginal propensity to consume in the elementary Keynesian model must be between 0 and 1. Eliminating non-acceptable solutions is a low-order 'completeness criterion', i.e. the model must be complete enough to exclude them but it may not be complete enough to allow only one acceptable solution.

Completeness is the requirement that an explanation does not allow for the possibility of competing or contrary situations. As such it rules out the possibility of a false explanation accidentally appearing to be true. That is, if a theory or model is complete and happens to be false, we shall be able to show it to be false directly. For example, by assuming or demonstrating that the production possibilities curve has no flat spot and is concave, our explanation of the economy's output mix and equilibrium prices is complete since each point on the implied production possibilities curve is compatible with only one price ratio, and each price ratio is compatible with only one point on the curve. In other words, in a complete theory any possible equilibrium point is unique given any price ratio. Should any other equilibrium point be observed for the same price ratio, our theory would be refuted. Note that when we explain equilibrium prices we do not usually require the model to be complete with respect to absolute prices, but only with respect to relative prices. Completeness then is always relative to what we wish to explain. The conditions which assure consistency are usually much less restricting than those which assure completeness. For this reason, the question of completeness can be an important source of fundamental criticism. This is explored more fully in Chapter 6.

While consistency and completeness are the most important logical attributes of any axiomatic system, there are some second-order considerations: independence of axioms within a system, economy of thought, and the so-called 'weakness' of the individual assumptions. The secondary considerations are sometimes more interesting because they are associated with intellectual adventure, or are claimed to be matters of aesthetics.

Independence and Economy of Thought. Here again there is a similarity between systems of equations and systems of axioms. We can have a consistent system of (linear) equations where the number of unknowns is one less than the number of equations, so long as at least one of the equations is a linear combination of one or more of the others. Or there may be a subset of equations which does not involve all the variables of the system, thus indicating an independence between some of the variables. In an axiomatic system there may be dependence between some of the axioms such that one or more of the 'axioms' can be deduced from some of the others, or some of the axioms may use the primitive terms of another subset of axioms within the system. In the 'art' of axiomatics it is considered desirable to have axioms be independent so that none can be deduced from others. Independence of axioms is considered to be evidence of economy of thought.

Another aspect of economy of thought is the number of primitive terms and axioms. The elimination of axioms in order to minimize the number of axioms usually comes at a cost, namely increased complexity of the individual axioms. Obviously there is a limit to this.

Weakness of Axioms. Weakness has a lot to do with generality and universality, i.e. the list of things to be explained and the limits of applicability for the items on the list. It is considered desirable that the axioms be as 'weak' as possible. It is not always clear what is meant by this. It would seem that it has to do with how limited an assumption is in terms of the logical constraints placed on its primitive terms. For example, an assumption of a variable or quantity being non-negative would be said to be weaker than an assumption of its being positive since the latter would exclude more possibilities, in this case the possibility of its value being zero.

A successful axiomatic study of a theory should produce an effect similar to that of putting on a pair of new glasses having suffered myopia or astigmatism for a long time. The clarity resulting from understanding the logical structure of a theory offers opportunities to investigate the theory's truth status – ideally this is a concern of the pure theorist. An axiomatic study offers an opportunity to 'see' the basis of our understanding and thus is very useful in a systematic criticism of economic theory [cf. Morgenstern 1963]. By requiring us to present all the *necessary* assumptions (i.e. necessary for completeness) an axiomatic study enables us to reject *any* theory which requires as an assumption a statement which is *known* to be false. And if an axiomatic study shows a theory to be inconsistent or incomplete, then clearly this would be an important criticism of that theory.

Before a theory (or axiomatic system) can be completed it is usually necessary to show that it is incomplete. Most of the theoretical analyses in traditional textbooks can be interpreted as results (i.e. failures) of indirect attempts to show the traditional theory to be incomplete. Abraham Wald's famous study of the incompleteness of Walrasian general equilibrium is an example of an axiomatic study [Wald 1936/51]. An *incomplete* theory or axiomatic system (as a whole) may still be testable if it entails at least one testable statement. For example Samuelson offers an example of a testable statement from traditional consumer theory: the sum of compensated demand elasticities of each consumer for each good is zero.[6]

Wald's study offered to complete an axiomatic structure of his Walrasian model by adding extra assumptions. He added an *ad hoc* assumption that demand prices are always positive (his condition 5). Although the inclusion of Wald's additional restrictive assumptions does the job of completing *an* explanation of prices and outputs, it does not follow that they are *necessary* for the original theory. As was later shown, the existence and uniqueness of solutions of an entire Walrasian system can be achieved using linear programming or activity analysis models which do not require such restrictive assumptions. Thus it would seem that demonstrating that any one of Wald's conditions is not satisfied (in the 'real world') does not necessarily refute the original incomplete theory. I shall return to Wald's study in Chapter 6.

From the methodological position entailed in either Popper's philosophy of science or Samuelson's methodology, this state of affairs is rather perplexing. We may wish to complete an axiomatic version of traditional consumer theory and then criticize it. But if our criticism only deals with those assumptions or clauses which we add (or complete), then we are not really criticizing traditional consumer theory. It would seem that this can be overcome by attempting to deduce testable statements from the incomplete theory and submit these to tests. And if we show any one of them to be false, the theory *as a whole* will be shown to be false, no matter how it is eventually completed. This is a very difficult task and not much has been accomplished so far. (We shall return to this in Chapters 7 and 8.)

From this methodological viewpoint it is important to realize that Samuelson's testable statement is deducible from a *set* of assumptions each of which is independently untestable, as will be discussed below. This being the case, we would conclude that it is certainly not *necessary* that the individual assumptions of our theories or models be testable in order to test the theory or model as a whole. Is the testability of each and every assumption even desirable? The answer is 'it depends'. For practical use of theories and models the answer may clearly be 'yes'. However, for purposes of intellectual adventure, economy of thought, aesthetics, etc., the answer may be 'no', since we may be more concerned with theories and models as *systems* of ideas [see Einstein and Infeld 1938/61, p. 30]. This question is considered again in the next section.

Having discussed economic models and economic theories, something needs to be said about what is called 'analysis'. What economists mean by analysis is not always obvious but it is possible to interpret their intentions. By recognizing explicit variables, a model builder is in effect analyzing the economic reality in question. An obvious example is the typical macroeconomic equation, $\mathbf{Y} = \mathbf{C} + \mathbf{I} + \mathbf{G}$. Here the GNP is being analyzed into its components, **C**, **I** and **G**. Similarly, by expressing theoretical ideas in the form of a system of simultaneous equations, the economic system in question is being analyzed into separate ideas or behavioural assumptions – demand behaviour is represented by one equation and supply behaviour by another. In the context of axiomatics we can perhaps see more clearly the traditional role of analysis in the development of economic theory. In particular what is usually called 'theoretical analysis' in economics is the process of attempting to derive certain propositions, such as downward-sloping demand curves, upward-sloping supply curves, conditions of efficiency, of equilibrium, of maximization, etc., from a certain set of primitive assumptions. For example, given neoclassical consumer theory we might ask: Does our explanation of consumer behaviour, based on the idea of utility maximization facing a given income and given prices, entail *only* downward-sloping demand curves? An attempt to derive *upward*-sloping demand curves is, in effect, a test of the completeness and consistency of our given consumer theory (if it is to be used in a neoclassical price theory). In the case of indifference analysis, such an attempt succeeds if

6 In Chapter 5 of his **Foundations of Economic Analysis**, Samuelson derives certain testable statements (such as the sum of a person's compensated demand elasticities for each good is zero) from assumptions which are not independently testable (e.g. *for every* person *there* exists a representable ordering on all goods) because the assumptions are of a logical form which does not permit falsification.

we do not rule out extremely inferior goods because then the possibility of upward-sloping demand curves arises. Specifically, the Hicksian assumption of a diminishing marginal rate of substitution along an indifference curve is insufficient for the avoidance of upward-sloping demand curves, although it is sufficient for consumer equilibrium! Unfortunately, exclusively downward-sloping demand curves may be *necessary* for the neoclassical theory of market prices [Boland 1977b, 1977d, 1986].[7]

One can see a positivistic bias in our traditional textbooks. Textbooks always seem to be telling us only about the propositions for which the traditional theory is sufficient. They seldom tell us anything about other relevant propositions for which the theory may be insufficient. If the traditional theoretical analysis is approached in a critical manner, it becomes a special case of axiomatic analysis. And if axiomatic analysis is approached in a critical way (by attempting to find important propositions for which the theory is incomplete), it can go a long way in helping to develop our economic theories, that is to say, our understanding of economic ideas and phenomena.

3. Beyond axiomatics

In the empirical social sciences, particularly economics, I think it is important that an axiomatic study also be concerned with the testability of the assumptions. This would at first seem to be an attempt to marry the two distinct approaches to economic theory discussed at the beginning of this chapter. However, I will suggest two different ways of considering the testability of assumptions. *Applied* theorists (or applied model builders) must be concerned with the testability of their assumptions. Preferably, their assumptions should be *directly testable*. On the other hand, the pure theorists or model builders who are interested in a model or theory as a *system* of ideas need only be concerned with the *indirect* testability of the assumptions – namely, in terms of the testability of the conjoint entailments.

The only time pure theorists require independent testability is when they find it necessary to augment their set of assumptions to complete a model or theory. Any additional (*ad hoc*) assumption must be criticizable on its own because the additional assumption may make the model or theory true in a quasi-tautological manner. A theory can always be made true by assuming *ad hoc* that any conceivable counter-examples are exceptional cases not to be considered [see Hahn 1965b]. The question for the economist interested in learning about (or criticizing) his or her understanding of economic phenomena is always whether or not the unaugmented system of ideas (i.e. the model or theory) contradicts facts. Untestable *ad hoc* assumptions would only insulate the model or theory from the real world [see Boland 1974]. If the concern of pure theorists is their understanding (there does not seem to be any other that is exclusively their concern) which is manifested in the system of ideas (i.e. in their theory or model), then the individual ideas need not be their first concern.

For the sake of discussion, let us just say that pure theorists (as opposed to applied economists) are interested in models and theories as systems of ideas. In the process of building their models, theorists select assumptions. Should they select only assumptions which can be independently tested? In this section, I argue that from the perspective of pure theory, the objective in constructing models (or theories) will be to choose assumptions that are *independently untestable* and still do the job. It should be noted that this view is slightly weaker than the commonly accepted requirement (attributed to Hobbes [see Watkins 1965]) that we should choose only assumptions that are not known to be false.

Given the widespread adoption of the Popper-Samuelson view that testability is everything, many readers may find my argument in favour of independently untestable assumptions to be rather surprising. There certainly would seem to be many alternative objectives which would be more acceptable, such as postulating assumptions which: (1) are (necessarily) testable and, best of all, verifiable *if true*, (2) are tautological hence are always true, (3) are approximations of the real world, i.e. assumptions with a ‘high probability of being true’, or (4) may or may not be ‘realistic’ so long as they make the implications of a theory ‘probably true’.

Each of these popular alternatives will be discussed in turn. I will try to show that each alternative methodological requirement (for model builders) has been based on a confusion of purposes or on a methodological error.

First I must explain what I mean by the phrase ‘and still do the job’ because it will play a prominent role in my arguments. I consider the job

7 That is, considering the arbitrariness of choosing between Marshallian or Walrasian stability conditions, only downward-sloping demand curves (and upward-sloping supply curves) will do the logical job of explaining market prices in *both cases*.

of an entire theory or model to be something more than just an exercise in (deductive) logic and by no means is it an exercise in 'inductive logic'. Theories are put forth in hopes of diminishing some of our ignorance.

Many economists believe that any theorizing is justifiable only to the extent to which its results are potentially useful [see Bronfenbrenner 1966]. They may be correct, but asking whether a theory will be potentially useful is superseded by the requirement of testability which in turn necessitates the requirement of falsifiability.

In economics we can go further in this logical chain. A falsifiable theory is not empirically testable unless it includes exogenous variables. This follows either from the avoidance of the identification problem or the desire for causal ordering [see Simon 1953]. However, the existence of exogenous variables in a theory immediately implies the potentiality of the theory being useful [p. 65]. But rather than just being useful in a practical sense (which may enhance our interest), the job of an entire theory is to help us to better comprehend phenomena and concepts (such as equilibrium [see Hahn 1973 on equilibrium theory]), i.e. to overcome the failures of our primitive comprehension of phenomena and concepts.

All theories can be characterized as attempts to 'justify' (i.e. to give reasons for) specific answers to specific questions. We assume what we assume in order to obtain our particular justification. The individual assumptions 'do the job' when they are logically sufficient for our justification! One of the purposes of *axiomatics* is to study their sufficiency and necessity for a particular justification. The *conjunction* of all the assumptions forms a specific representation of the ideas constituting the theory in question.

Taking a set of assumptions in conjunction permits us to deduce the implications of that theory. Now Popper's well-known falsifiability criterion requires only that at least one of these *implications* be (independently) falsifiable, but in no way does this requirement imply the necessity that any of the individual assumptions be independently falsifiable, let alone testable. Of course, following Hobbes, we realize that none of the assumptions should be known to be false! The task of deducing (testable) implications is merely an exercise in logic. Surely the job of the theorist is more than this. First, as indicated above, the theorist puts forth assumptions which do the job of establishing his or her desired justification (i.e. they are logically sufficient). There has been considerable controversy surrounding the question of whether all of

the assumptions should be necessarily testable [see Boland 1979a]. If they were, and we could show any one of them to be false, any logical analysis of their implications (deduced from their conjunction) would be beside the point (and perhaps uninteresting) even if they are logically sufficient for their intended job! Should we be able to show all of them to be true, the job of the theorist degenerates to the somewhat more interesting job of being a practitioner of logic alone. Furthermore, if we were to require our assumptions to be tautologies (i.e. statements which are always true) then the theorist's job will clearly be only that of a logician. Thus, we can dismiss objectives (1) and (2), listed above, because they both reduce the job of the pure theorist to that of only a logician.

Now let us consider the alternative objectives (3) and (4) for the selection of assumptions (p. 31). Both of these objectives presume that we want our theories (i.e. their entailments) to be 'probably true'.[8] (Note well: this presumption implies an acceptance of a theory being 'probably false'!) Regarding the assumptions of our theories, advocates of these objectives can adopt a 'strong' view or a 'weak' view. *Excluding* followers of Friedman's instrumentalism [see Friedman 1953, as well as Boland 1979a], most economists believe in the strong view that as a theory's entailments should 'approximate reality' (i.e. be 'probably true') so should a theory's assumptions be 'probably true'. The concept of 'probably true' is represented by the probability of a statement being true where the means of determination is in accordance with one of several widely accepted conventional criteria such as minimum R^2. It is the conventionality of the criteria of truth status that gives this methodological view the title of 'conventionalism' (this is dealt with further in Chapters 4 and 5). The weak view, which may be ascribed to Friedman's followers, is that it does not matter if the assumptions are 'unrealistic' (i.e. 'probably false') so long as the entailments (i.e. predictions) are 'sufficiently accurate'. At times this weak view may

[8] There is the obvious difficulty for those of us who accept the axioms of ordinary Aristotelian logic: (1) a thing is itself (identity), (2) a statement cannot be *both* true and false (contradictions excluded), and (3) a statement can only be true or false (excluded middle). Together axioms (2) and (3) tells us that a statement either is true or it is false but not both. This also means there is nothing else it can be (such as 'probably true'). See further Swamy, Conway and von zur Muehlen [1985].

suggest to some that there is a claimed virtue in the assumptions being false [Samuelson 1963]. It is this weak view to which Samuelson gives the name 'the F-twist' [see also Wong 1973].

Let us first look at the conventionalist (strong) view. Should we require that our assumptions be *approximations* of reality? That is, do we want our theories to be 'probably' true (and hence 'probably' false)? Here we reach a sensitive point with modern economists if we try to criticize the possibility of, or necessity of, the stochastic (or probabilistic) nature of economics. The difficulty in deciding this issue is due to the lack of a distinction between the objectives of the pure theorist and the objectives of the practitioner, namely the applied economist. The applied economist is primarily concerned with *success*, i.e. success in the practical job at hand (advising a government agency, a business, etc.). Surprisingly, the pure theorist is *not* concerned with success (at least not practical success). The pure theorist is more interested in what might be called 'intellectual adventure' [Agassi 1966b].[9] Unlike the applied economist, the pure theorist finds that he or she learns more from being wrong than from being right!

One important aspect of an intellectual adventure is the unexpectedness of the results, that is, on the basis of what we already accept (perhaps other theories) we should not expect a given new theory to be true. Now the traditional view is that if the assumptions were 'probably true', then the statements deduced from them may have a high probability of being true and therefore more useful [e.g. Bronfenbrenner 1966, p. 14], but this need not always be the case [see Boland 1982, Ch. 7]. Here again we find the job of the theorist reduced to that of a practitioner not of just logic, but of the logic of probability. There is a more fundamental problem with this view of the objectives of a theory and its assumptions. Do we (as theorists) want our theories to have an *ex ante* high probability of being true? If we are practitioners of economic theory, the answer is probably affirmative since success will be our criterion for evaluating a theory. From the viewpoint of the theorist interested in intellectual adventure it is more desirable for a theory to have unexpected results. That is, on the basis of what we already think we know (i.e. unrefuted previous theories), we would not expect the new theory to be true *and* if it were true then it would cast serious doubt upon

9 It might be called success in one sense – i.e. the success of creating a 'good' theory – but it is certainly not practical success.

the truth of our previous theories. A theorist could be more interested in a theory with a *low* rather than *high ex ante* probability of being true [see Popper 1959/61]. Thus 'pure' theorists would certainly not *require* their assumptions to have a high probability of being true.

Now let us look at the weak view. By saying that theorists do not *require* their assumptions to be true, or even have a high probability of being true, it is not intended to suggest that assumptions should be false! Obviously such a view would be ridiculous, as Samuelson has pointed out. This, I think, may be an unfair reading of Friedman's essay [see Boland 1979a]. Nevertheless we must be careful to avoid the possibility of this weak view being adopted. Whenever we require our assumptions to be only *probably* true (i.e. deemed to be true in accordance with some probability-based 'test conventions'), we must always leave open the possibility that the assumptions are false. If we were to have false assumptions then our understanding of the phenomena, or solution to the problem that the theory is supposed to justify, would surely be perverted [see Einstein 1950].

If a theory's assumptions are testable, then the possibility exists that we can (and should attempt to) show them to be false – or probably false in the case of stochastic theories (where 'probably false' means that the probability of being true is below a conventionally accepted minimum). If we are not interested in immediate practical usefulness, we can avoid this possibility by avoiding testability of our individual assumptions and let the burden of testability fall on the entailments of the theory as a whole. For example, the basic behavioural assumptions of traditional consumer theory may be axiomatically represented as follows:

(1) for *every* consumer *there exists* some non-economic criterion *such that* he or she is able to compare combinations of quantities of goods;

(2) for *every* consumer *there exists* an economic criterion *such that* he or she is able to compare combinations of quantities of goods;

(3) every consumer when confronted by two combinations which he or she can afford, defined in assumption (2), will buy the one which is 'better', defined in assumption (1).

Assumptions (1) and (2) are not testable. They are both incomplete; the criteria are unspecified and thus none can be ruled out.[10] Furthermore, assumption (3) is untestable because it depends on the other assumptions. The question here is whether the concept of 'better' is sufficient to be used to derive a testable statement from this theory. Samuelson specifically argued that it was sufficient to recognize that statement (3) presumes a choice is made and that the consumer is consistent in the application of the two criteria [see Samuelson 1938, 1947/65, 1950a].

We see then that the avoidance of individual (independent) testability amounts to saying that we make it impossible to know whether a *particular* assumption is true or false and leave it to be combined with other assumptions to form a theory. It is the theory as a *system* of ideas (rather than individual ideas or assumptions) which we are testing when we test a statement deduced from the conjunction of assumptions.

We can see another virtue of avoiding individual testability of assumptions and relying on the testability of systems of assumptions by considering another important aspect of an intellectual adventure, namely, the desire for economy of thought. If we accept Popper's classic means of demarcating scientific theories from non-scientific ones – i.e. the requirement of falsifiability [Popper 1959/61] – by requiring only that *at least one* statement derivable from a theory be falsifiable, then if we emphasize economy of thought in our development of a theory, we desire to *just* meet this requirement.[11] Clearly, if any one of the derivable statements is already testable, then, in terms of economy of thought, we will go beyond the requirements if we also require that one or more of the theory's assumptions is *also* falsifiable on its own.

It might be well and good to say that, from the standpoint of 'pure' theory (i.e. of systems of ideas) and of intellectual adventure, it is desirable to avoid individual testability of assumptions. It is quite another matter to show that it is possible to have individual assumptions untestable and still have at least one of the statements derivable from

[10] Statements of this form are sometimes called 'all-and-some statements' and as such are incomplete. One can complete them. The completed statement may be falsifiable but it is not a necessary outcome [see Watkins 1957].

[11] However, emphasizing economy of thought in this sense would seem to run counter to Popper's criterion for deciding which of two compatible theories is 'better', namely that the theory which is 'better' is more testable. But the testability of a theory is dependent more on the testability of any of its implications than on the number of implications. The testability of any of its implications is inversely related to the quantity of information, i.e. the 'dimension of a theory', required to test an implication. This topic is examined more fully in Chapters 2 and 3.

their conjunction (i.e. the theory formed by them) testable. But as noted above in my axiomatic example of consumer theory, Samuelson seems to have done this already (see the example above as well as his [1947/65, Ch. 5]). This means that it is indeed possible to satisfy Popper's criterion of refutability (a criterion which assures us of the possibility of empirically criticizing a theory) and still be concerned only with the theory or model as a system of ideas rather than empirical truths (i.e. 'realistic' assumptions).

As an axiomatic study, my version of traditional consumer theory is incomplete for two reasons: (1) it lists only the behavioural hypotheses, and (2) each individual assumption, as a member of a conjunctive set, is not complete with respect to the specification of necessary presuppositions and limitations entailed in the assertion that the assumptions *together* do the job of justifying the given answer to the given question. In particular, since all the assumptions for the behaviour of consumers are of the form of quantificationally incomplete statements, the metatheoretical assertion that they do the job requires that to complete the theory, we specify the necessary and sufficient conditions of maximization [see Boland 1981b]. What we can (or must) specify for a particular assumption depends on what the other assumptions say – i.e. it is a 'simultaneous argument' similar to a system of 'simultaneous equations'. In this manner traditional consumer analysis can be thought of as attempts to complete these clauses.

4. Testability for all

The title of this chapter indicates a distinction between two approaches to the study of economics. This distinction has implications for the methodological decisions involved in developing a theory or model. In particular, this distinction implies differences in the criteria of success of a model or theory, in the purposes for model building and in the epistemological requirements for the assumptions used in a model or theory.

Ideally there would be no need to distinguish between understanding economic phenomena and understanding economics. However, until we know all there is to know about economic phenomena, the distinction remains important. The intention is not to try to build a philosophical wall around either applied economists or pure theorists, but by recognizing the existence and the incompatibility of these two alternative approaches, to make it possible for more fertile criticism between them.

The models we build represent our understanding of specified economic phenomena. The individual assumptions of our models or theories form the foundation of our understanding. The implications of our models can be important for two different reasons. The implications can be useful (policy recommendations, etc.) and they can be a means of testing our understanding. Clearly, our understanding can be open to criticism. But if it is beyond direct criticism, either because the applied economist has chosen assumptions which appear to be realistic or because the pure theorist may have designed assumptions that are individually untestable so as to emphasize the model as a system, then we must rely on indirect criticism by testing the implications. So long as the implications are testable and none of the assumptions are known to be false, we need not be afraid of the pure theorist's apparent disinterest in the realism of his or her assumptions. But these are important conditions of acceptance. Axiomatic studies must take testability into consideration. And if we really are interested in understanding our economic understanding, we must reject methodological statements that suggest our assumptions can be allowed to be false.

So-called 'pure theorists' cannot insulate themselves by hiding behind a wall of aesthetics and scholasticism. They must put their systems of ideas to test and the only convincing test is against the 'real world'. Of course, here the applied economists can help a great deal, but they must attempt to appreciate what the pure theorist is attempting to do if the theorists are ever going to listen to their criticisms.

Similarly, applied economists should not be deluded by limited practical success or 'positive' results. Pure theorists may be able to offer useful criticism but they should not ignore the importance of practical success. If pure theorists wish to criticize an apparently successful model, they will never convince anyone until they also explain why that model is (or appears to be) successful.

2

On the Methodology of Economic Model Building

> To be of interest a scientific theory must have consequences. Upon hard-boiled examination, the theory of consumer's behaviour turns out not to be completely without interest. By this I mean: consumption theory does definitely have some refutable empirical implications. The prosaic deductive task of the economic theorist is to discern and state the consequences of economic theories.
>
> Paul A. Samuelson [1953, p. 1]

> the *refutability* or *falsifiability* of a theoretical system should be taken as the criterion of its demarcation. According to this view, … a system is to be considered as scientific only if it makes assertions which may clash with observations; and a system is, in fact tested by attempts to produce such clashes, that is to say by attempts to refute it.… There are, moreover, *degrees of testability*, some theories expose themselves to possible refutations more boldly than others.
>
> Karl R. Popper [1965, p. 256]

It would seem that the most important statement deduced from a multi-equation economic model is its solution. The solution statement specifies a certain relationship between that which we wish to explain (the endogenous variables) and that which we know (or assume to know) or can be independently determined (the parameters or the exogenous variables). The main concern of this chapter is the significance of the epistemological problem which is associated with the truth status of the *form* of an economic model. Unless we *know* in fact the values of the parameters, the solution is nothing more than a conditional statement.

Nevertheless, we can compare models (and the theories they represent) by comparing the forms of their solution statements.

1. Economic theories and the aim of science

The purpose of this chapter is to present the quantitative criterion of testability by which the forms of the different models can be compared. This criterion was developed in my 1965 PhD thesis [Boland 1966]. To develop this criterion I chose to depart from what I thought at the time was the typical philosophical viewpoint about the purpose of science and economic theories. The alternative which I chose is a version of what I now call the 'Popper-Socrates view of science' [Boland 1982, Ch. 10]. I did not consider just *any* alternative view but one which seemed to offer an opportunity to ask what appeared to be some new or different questions about the truth status of economic models.

Let us begin by presenting this alleged alternative view of science. After developing some useful concepts which follow from this view, I will illustrate them with some examples of simple economic models. In this context, the most important concept I develop is Popper's 'dimension of a theory'. Simply stated that is the number of observations *necessary* to refute a theory or model [cf. Samuelson 1947-8, pp. 88-90]. With this concept in mind, a few models will be compared on the basis of their different dimensions.

1.1. Science as an unending process

In the remainder of this introductory section, I present the view which in 1965 I wished to attribute to economic model builders. My view then was that economists as scientists do not seek theories which are 'true statements' or even 'almost true statements'. Thus I presumed that anyone who envisages a system of absolute, certain, irrevocably true statements, or 'probably' true statements, as the end-purpose of science would certainly reject the alternate view of science I used in my PhD thesis. To appreciate the Popper-Socrates view of science, consider the purpose for which we advance theories in economics. My view in 1965 was that science is not a singular act, i.e. an act of advancing just one theory – if it were, we would probably wish that each singular theory be put forth as a 'true statement'. To the contrary, with the alternate view of science as a guide, I claimed that science is an endless succession of revision, i.e. the replacement of an inferior theory with a less inferior theory which in turn is replaced by a slightly less inferior theory, and so on. Methodologically speaking, such a process must begin somewhere. So following Popper, I said that it begins with a conjecture, an attempt at explanation of some phenomena. However, this conjecture would *not* be offered as a (known) true statement but as something on which we may begin to make improvements. Thus I claimed that we advance theories because we need something to improve. Moreover, if we see science as a process of making advances, not only will the statement or conjecture which we offer not be a (known) true statement, but to get things going we may even need to begin with a false statement.[1]

As this may seem rather perverse to some readers, let us consider a few reasons. Most students of epistemology or methodology seem to assume (implicitly or explicitly) that there is something which might be called True Knowledge and that it is a *finite* quantity, although the process of obtaining it may be unending [see Boland 1982, Ch. 4]. Of course, one can attempt to describe or explain large or small portions of this quantity. Because it is claimed to be a finite quantity, some epistemologists and methodologists behave as if the Archimedes Principle[2] applies – they seem to think that by advancing theories which are in a small part true, there exists the possibility of a *finite* number of theories which when combined will describe Reality, i.e. add up to that finite maximum of knowledge, True Knowledge. Today, as in my 1965 thesis, I reject this view of knowledge because the conception of knowledge as a finite quantity represents a methodological bias in favour of inductivism [see Boland 1982, Ch. 1]. In its place, I will presume that truth status is a quality, i.e. all non-paradoxical statements must possess the quality or property that *either* they are true *or* they are false [see Boland 1979a]. Knowledge is *not* like wealth, since if it were, we would

1 After more than twenty years this seems rather naive to me. In later work I realized that many would interpret 'known true' to mean 'tautologically true' and thus see that all that was being required was the potentiality of being false rather than the naive claim that advances imply false conjectures.

2 The Archimedes Principle says that if there is a finite positive quantity A and there is a positive quantity B such that $B < A$, then there exists a finite number n such that $nB > A$.

always try to obtain *more* of it. Knowledge is more like health, to the extent that we can always try to *improve* it.[3]

With regard to economics and for the purposes of this chapter, I will treat Reality or the True World at any point of time as a 'Walrasian-like' system of an unknown number of equations (or a very large number for which there is always the possibility of adding more) and an indefinite number of unknowns [cf. Georgescu-Roegen 1971]. Economists, however, operate with the expectation that over time there are some salient features of this system which do not change very much, if at all [see also Hicks 1979, Ch. 4]. But here, too, economists do not expect to be able, *a priori*, to describe these features exactly, although one can expect to be able to show when a feature changes.

1.2. The Popper-Socrates view of science

The Popper-Socrates view of science shall be described as follows: Science begins by attempting to overcome certain inabilities to explain some unexpected phenomena (e.g. certain features of the economic system such as the capital stock and the level of GNP, which apparently change together) or it can begin with a mythological statement (e.g. the lack of a balanced budget is bad for the country) or it can begin with a statement of theory about some empirically observable phenomena (perhaps phenomena not yet observed, such as a predicted recession next year). Theories are advanced in order to overcome some real or conceptual problems in explanation. The starting point will not matter – so long as each succeeding theory presents new problems which must be overcome, i.e. new problems for which new theories are offered to solve the succeeding problems, inabilities or difficulties. If the new theories do not present new problems, the process would cease. Thus if science is to move along, we as theorists must seek new problems. And to do this we advance theories which we may perversely hope will be false, even though we offer them as potentially true theories.[4] The idea that a new theory should be able to explain some new phenomena is secondary to the survival of any science. Showing a new theory to be in some additional regard false gives us an important prize, namely a new problem to be explained or overcome. The necessity of a new theory's ability to explain what a previous theory can explain is, so to speak, a necessary but *not* sufficient condition for it to be an interesting new theory.

When one implication or conclusion of a theory (or model) is shown to be false, then clearly the theory *as a whole* is false. Thus in these strict terms it would have to be concluded that most (if not all) economic theories are false. Nevertheless, we as economists do not discard all of our theories. This is mainly because we do not have a 'better' theory with which to replace the false theories, although there may be many candidates [see Archibald 1966]. Applied theorists are concerned with the applicability or workability, but not necessarily the truth status [see Boland 1979a], of any theory which they wish *to use*. Some theories are better than others, or more than one theory will do the job required. In the absence of straightforward practical tests, how do 'pure theorists' decide when one theory is 'better' than the other?

The criterion with which theories are compared follows from one's prior view of the aim or purpose of science. With the Popper-Socrates view of science as a guide, any science may be considered a social activity and as such it needs some guidelines or 'rules'. However, these rules may merely be conventions rather than rules such as those of logic. The Popper-Socrates view prescribes rules in order that everyone can participate. Certainly more is accomplished by a team than by a group of individuals operating each with his or her own private conventions. A set of rules can make the social activity easier and more interesting but cannot assure success or even 'progress' (whatever this means). Any rules which are proposed must not be arbitrary, rather they must follow *logically* from the aims which are attributed to any science.

The Popper-Socrates view says that we conjecture hypothetical solutions or theories in order to attempt to improve them, which means that we may wish theories to be false to some extent. That is to say, there must be something to improve. However, when we put forth a theory we do not *know* that it is false – i.e. that any of its conclusions will be false. We only learn that a theory is false after we have tested it and have shown that it has not passed the test. Thus, if the discovery of errors or failures depends on the ability of our theories to be shown false, we can

---

3 And even if one were to have 'perfect knowledge', one would have to strive to maintain it (like perfect health) by countering any attempts to put forth false statements. Nevertheless, we need not pursue this analogy any further.

4 Of course, hardly anyone *expects* to find absolutely correct theories – e.g. a theory which would perfectly explain the year 1988. But it must be realized such a theory which does not is, 'as a whole', a false theory even though certain subparts of it may be true [see Samuelson 1964].

logically accept the property of falsifiability as one of our criteria. Of two competing and unrefuted theories, the one which is more falsifiable is considered to be 'better' to the extent that it is worth being the first to try to refute. On this basis, one can argue for the absolute requirement that, to be of scientific interest from the Popper-Socrates perspective, any suggested solutions to new problems (or answers to new questions) be in principle falsifiable or refutable.

1.3. Theory comparison and the dimension of a theory

In Section 2, an explicit argument for today's commonplace absolute criterion of falsifiability will be based on two reasons: (1) there is no point in discussing what economists call tautological theories since nothing can be learned with them, and (2) since the objective is to produce theories in order to criticize them, the theories must be capable of being false. In other words, only theories which are falsifiable or refutable are useful within the Popper-Socrates view of science. I stress that falsifiability is necessary but *not* sufficient for the purpose of science which I have attributed to Popper. The requirement of falsifiability is thus not sufficient for the description of the Popper-Socrates view of science. To put things in a broader perspective, it might be noted that there have been some methodologists who viewed science as the pursuit of true theories (although some may qualify the pursuit to be towards approximately true theories). According to this so-called 'verificationist' view, we present theories in order to verify them, which requires that theories be verifiable [see Machlup 1955]. However, verificationists may still require falsifiability since it would be trivial to verify tautologies. Falsifiability is necessary for *both* the Popper-Socrates view of science and the old-fashioned 'verificationist' view which Popper criticizes. (I will have more to say about this irony in Chapter 7.)

An additional criterion will be discussed by which a selection can be made between competing theories, i.e. those theories which explain the same phenomena. This criterion will involve two related properties of theories: degrees of falsifiability and degrees of testability.[5] Whenever there are two theories which explain the same phenomena (i.e. the same list of endogenous variables) or produce the same solution to the same problem, the more interesting theory is the one which is relatively 'more falsifiable'. If the competing theories are apparently of equal or incomparable falsifiability, then the more interesting theory is the one which is relatively 'more testable'. The degree of testability will be shown to depend on what Popper calls the 'dimension' of a theory, such that the lower the dimension, the higher the testability.

Popper's choice of the word 'dimension' may seem unfortunate because I will use it in the comparison of infinite sets. Let me, then, briefly digress to explain how I will be using the word dimension in a rather formal sense. For my purposes, the dimension of a set is *the dimension of its boundaries plus one* – e.g. the dimension of a cube is three (although it may contain an unlimited number of 'points') since the boundaries of a cube are squares which in turn have the dimension two. Likewise, a square has the dimension of two because it is bounded by unidimensional lines which are in turn bounded by points which have dimension zero. Although one may claim there are an infinity of points within the boundaries of the set defining a cube, the set has a finite dimension. (For a more rigorous discussion of this notion of dimension, see [Hurewicz and Wallman 1948].)

1.4. Science as a community activity

To investigate the 'real world' of a Walrasian-like system in small but *definite* steps can be incredibly difficult. Thus to make any reasonable advance one might choose to be methodologically efficient. It can be suggested that one means of being efficient is to offer as many solutions as possible when attempting to overcome a failure of a previous theory. Only after a reasonably complete list of possible solutions is obtained does the use of the criteria of falsifiability or testability begin. Unfortunately only when theories which cannot be eliminated by these criteria remain will it be necessary to devise new criteria such as 'crucial' tests or pragmatic simplicity. The smaller the steps or the improvements, the easier the compilation of possible solutions, i.e. the more manageable will be the task of science.

Since the Popper-Socrates view stresses the conception of science as a community activity rather than a private activity, the need for succinct presentation would seem almost crucial. Community efforts will appear to be best served if one explicitly spells out: (1) the problems (or

5 For now we may define these as the relative quantity of conceivably falsifiable statements, and quantity of empirical statements, respectively, which can be deduced from a theory.

questions) which a theory claims to solve (or answer), (2) the solution (or answers), (3) the means (arguments, models, assumptions, etc.) by which a particular solution (or answer) is to be derived, and (4) how one might attempt to refute the remaining theory or decide between the remaining theories. However, none of these guidelines guarantee anything.

This then is the Popper-Socrates view of science as I understood it in the 1960s. In the remainder of this chapter I will present my attempt to apply this view to some of the elementary methodological tasks of economic model building. Of special interest are those aspects of the Popper-Socrates view which will help in the comparison of the forms of various economic theories. My 1965 analysis began with a discussion of the so-called 'problem of demarcation', i.e. the alleged problem of distinguishing between scientific and non-scientific statements. This discussion was essential for anyone wishing to use Popper's philosophy of science, even though it is easy to see that such considerations are not necessary [see Bartley 1968]. While I think it is necessary to discuss Popper's demarcation criterion in order to discuss his more useful idea of the dimension of a theory, I must stress here that I am not interested in developing methodological rules or criteria which would restrict the freedom of economic theorists to ask certain questions. Rather, I am interested only in comparing the quality or *forms* of the theories and their representative models against a background that specifies the aim an economist has for his or her science.

2. Popper's famous demarcation criterion

Since it is commonly thought that the keystone of the Popper-Socrates view of science is its *requirement* of falsifiability of all truly scientific statements, the first task would seem to be a consideration of the implications of a demarcation between scientific and non-scientific statements [see Hutchison 1960]. Falsifiability is definitely the hallmark of many popular discussions of the relevance of Popper's philosophy of science for economic methodology [e.g. Blaug 1980, Caldwell 1982]. For Popper, a falsifiable statement is important because the truth of a falsifiable statement implies a prohibition of the truth of other statements. Although in many, but not all, cases statements may be both falsifiable and verifiable, for the purposes of the Popper-Socrates view of science, falsifiability is preferred to verifiability. This preference is established to exclude certain statements or questions from consideration on the basis that nothing can be learned from a statement or an answer to a question if that statement or answer cannot be wrong. Statements or answers to questions must be such that they can, in principle, be falsified.

Despite the efforts of many of us to credit Popper with revolutionizing economic methodology by his emphasis on falsifiability, it is only fair to recognize that falsifiability has long been the basis of Paul Samuelson's methodology [e.g. Samuelson 1947/65]. However, Samuelson does not explain why he is interested in falsifiable statements (which he calls 'operationally meaningful'). Consequently, the following discussion of what I will call the 'Popper-Samuelson demarcation' will have to be based only on Popper's explanations of what I have called the Popper-Socrates view of science.

2.1. Falsifiability vs verifiability

Let us review some of Popper's fundamental distinctions. First, falsifiability does not always imply verifiability. A statement such as 'all swans are white' can, in principle, easily be falsified, but it can never be verified since we must look at all swans that ever did or ever will exist. Similarly, the statement 'there exists at least one white swan' can easily be verified, but it can never be falsified because we would have to show that *every* swan that ever did or ever will exist is not white. Popper calls a statement such as the first, which uses the unlimited concept 'all', a *strictly universal* statement; and he calls a statement such as the second, which uses some form of 'there exists', a *strictly existential* statement [Popper 1959/61, p. 68]. Clearly the negation of a strictly universal statement is equivalent to a strictly existential statement and vice versa. Thus it can be said that the falsifiability of a theory is its ability to be used to deduce strictly universal or *non-existential* statements [p. 69]. In other words, a theory is falsifiable if one can deduce from it statements which prohibit certain conceivable phenomena. For example, one may deduce from a theory that the interest rate is non-negative (i.e. $i \geq 0$) or one may deduce that the rate of change of GNP with respect to time is always positive (i.e. $\partial GNP/\partial t > 0$).

The establishment of falsifiability as an operative demarcation criterion directly excludes existential statements and tautologies.

Tautologies are excluded because they are always true and thus can *never* be falsified. Emphasis on falsifiability also assures that self-contradictory statements will be excluded since they can be shown to be false by *any* observation or test. An example of a self-contradictory statement would be 'the interest rate is non-negative *and* the interest rate is negative'. The observation of a positive or zero interest rate would indicate that the second half of the statement is false while a negative interest rate would indicate that the first part is false. Since the interest rate is a real number the statement is *always* false, and hence will always be rejected [p. 91].

If the operative demarcation criterion were verifiability rather than falsifiability, tautologies would not be excluded necessarily since they can be easily verified. Also if we *only* looked for verification when testing statements or theories, the self-contradictory statement runs the minor risk of not being quickly eliminated. Of course it is easy to see the contradiction in our example, but when the contradiction is much more subtle or complex we run such a risk. The point here is that the concept of falsifiability is more convenient since the same restrictions are possible with the concept of verifiability *only* if it is *also* specified that tautologies and self-contradictory statements must be excluded from consideration. The difference between a demarcation using falsifiability and a demarcation using verifiability will be diminished somewhat since the Popper-Socrates view of science also requires that theories under consideration must be consistent [p. 75]. The importance of the requirement of consistency will be appreciated if it is realized that a self-contradictory (prohibitive) statement is uninformative. A consistent (prohibitive) statement divides the set of all possible observations or statements of fact into two groups: those which it contradicts and those with which it is compatible [p. 92]. I postpone discussion of 'informativeness' until Chapter 6.

2.2. Falsifiability vs false theories

If it is kept in mind that a theory or model consists of many assumptions which are logically conjoined to form an argument in favour of one or more propositions, we must now consider what the Popper-Socrates view of science means by a 'false theory'. If statement P follows logically from theory T and if P is false then the theory T *as a whole* is false. But it cannot be inferred that any one assumption used in the theory is, or is not, specifically upset by the falsification. Only if statement P is independent of some part T′ (or group of assumptions) used by theory T can it be said that the part T′ is not involved in the falsification. The problem of ambiguity of falsifications with regard to assumptions will be discussed further in Chapters 7 and 8.

Sometimes it appears to be possible to alter the theory slightly and eliminate the offending false implication. For example, let us say we have deduced from a growth model the following relation:

$$\mathbf{Y}_n = \mathbf{g}(n,\mathbf{Y}_0) \qquad [2.1]$$

where $\mathbf{Y}_n$ is the GNP of the nth period following period 0, $\mathbf{g}(\,)$ is a function involving the parameters of the model and $\mathbf{Y}_0$ is the GNP for period 0 (i.e. for $n = 0$). Also, let us say we know what $\mathbf{g}$ and $\mathbf{Y}_0$ actually are and we have 'tested' the model by observing GNP for various periods and have found that the calculated $\mathbf{Y}_n$ did not agree with the observed $\mathbf{Y}_n$; specifically, the actual GNP was always greater than $\mathbf{Y}_n$ by an amount equal to a constant multiple of n (i.e. by kn). We might consider adding to our model an 'auxiliary hypothesis' that actual GNP = $\mathbf{Y}_n + kn$. Such attempts are not unusual in the natural sciences [Popper 1959/61, p. 88]. Consider another type of modification. Let us say that our observation indicated that the solution represented by equation [2.1] was true after period 10. We can 'salvage' our model in this case by adding the restriction that [2.1] holds for $n > 10$. Either of these modifications (a transformation or a restriction) can 'save' a model, but only at the expense of its simplicity or generality. The Popper-Socrates view of science rules out such modifications on the basis that they are in conflict with the aims of a purely theoretical science [Popper 1959/61, pp. 82-3]. This does not mean that all auxiliary hypotheses should be rejected [see Boland 1974]. For example, consider the following solution deduced from a simple Keynesian model:

$$\mathbf{Y} = \mathbf{K}/(1 - b) \qquad [2.2]$$

where $\mathbf{Y}$ is GNP, $\mathbf{K}$ is the level of investment, and b is the well-known 'marginal propensity to consume'. If the simple model under consideration only specified that b is not zero, very little would be prohibited by equation [2.2] – i.e. the equation [2.2] is compatible with many different $\mathbf{Y}/\mathbf{K}$ ratios. An auxiliary hypothesis which says that b is positive but less than one would, in this case, improve the model by making it more susceptible to falsification. In general, with the exception of theories which have been modified by an auxiliary hypothesis which makes them easier to falsify, modified theories will be considered to be inferior to unmodified theories.

2.3. Consequences of using the Popper-Samuelson demarcation

On the basis of the demarcation criterion prescribed by both Popper and Samuelson, we can draw some conclusions about what type of theories or enquiries are prohibited. The demarcation based on falsifiability and non-falsifiability divides possible theories into two groups: (1) those theories from which non-tautological, non-universal prohibitive statements can be deduced, and (2) those theories from which either strictly existential statements or tautological statements or both can be deduced. Not much more can be said about the first group until a concept of 'degrees of falsifiability' is specified (this is our task in the following section).

The second group includes two types of statements which most scientists have little difficulty excluding from science. Falsifiability, however, offers something more than an intuitive basis for excluding them. As indicated, the two important members of this excluded group are: (a) tautological statements, and (b) strictly existential statements. We exclude both direct tautological enquiries (e.g. concerning statements that are true by virtue of their logical form alone – such as 'I am here or I am not here') and statements deduced from systems consisting of only definitional statements. Statements deduced from systems of definitions are quite common in mathematics (in fact this is one way to characterize pure mathematics). Economic theories, however, are based on more than definitions and tautological statements. Economic theories must always include one or more behavioural assumptions or conjectures, i.e. they must run the risk of falsely interpreting the behaviour in the 'real world'.

The second type of excluded statement, the existential statement, seldom occurs in the literature of modern economics. When it does, it is most often in the form of a prophetic statement. For example, consider a prediction or conclusion that there will be a social revolution. We can never refute such a claim because when we note that the revolution did not occur, its proponents will always say that it is still in the future.

Before concluding the discussion of Popper's idea of demarcation, one other requirement placed on the theorist should be noted. The demarcation based on falsifiability also leads to the requirement that economists who attempt to 'measure' capital, output, rates of return, capital coefficients, etc., must always specify the theory in which the particular concept of capital, or output, etc., appears. Without knowing the theory, one cannot know whether or not the economist's efforts only involve tautologies, hence one cannot *know* what will be learned by such a measurement [see Briefs 1960].

In the remainder of this chapter it will be assumed that the Popper-Samuelson demarcation criterion of falsifiability has been applied so that any remaining questions will be about comparisons or evaluations of falsifiable theories or models. Some falsifiable theories can be shown to be 'better' than others if we can measure a theory's 'degree' of falsifiability or testability.

3. Popper's subclass relations and the comparison of theories

Having now eliminated the 'unscientific' statements and theories by using the demarcation based on falsifiability, economic model builders need a criterion by which they can compare statements or theories that remain. For this purpose, Popper offers a criterion which logically follows from the Popper-Samuelson demarcation – namely, the concept of 'degrees of falsifiability' [Popper 1959/61, Ch. 6]. With this concept one theory can be judged 'better' than another on the basis that one theory is '*more* falsifiable' than another.

To understand the concept of degrees of falsifiability, one must understand the properties of a falsifiable theory or statement. According to Popper's view of science, a theory is falsifiable if it rules out certain 'possible occurrences' [p. 88]. A theory is falsified if any of these possible occurrences actually occur. To avoid possible semantic difficulties with the word 'occurrence' Popper chose to speak of the truth or falsity of the statement that a particular event has occurred or can occur [pp. 88-91]. He called such a statement a 'basic statement'.

At that time, Popper was attempting to explain his view of science in terms that would appeal to the reigning analytical philosophers [see Bartley 1968]. Therefore, he offered to provide an analysis of the falsifiability of a given theory by *analyzing* all the possible basic statements that can be deduced from the theory. To classify the basic statements that can possibly be deduced from a theory, following Popper, let P_a, P_b, P_c, …, represent elements of a class (or set) of occurrences which differ *only* in respect to individuals involved (i.e. to spatio-temporal locations or subclasses when they refer to similar things, for instance consumer goods). Popper called this class 'the event (*P*)' [Popper 1959/61, pp. 88-91].[6] To illustrate, one may speak of 'the event (prices of consumer goods)' where members of this class may include the price of apples (P_a), the price of butter (P_b), the price of coffee (P_c), etc. In Popper's terminology, the statement 'P_c is one dollar per kilogram' is called a 'singular statement' which represents the class 'the event (price of consumer goods)'. The singular statement 'P_c is one dollar per kilogram' asserts the occurrence of 'the event (price of consumer goods)' for coffee.

All singular basic statements which belong to one set or class of events (e.g. the relative prices for all consumer goods) will be called 'homotypic'. The singular statements which are homotypic *but* differ only in regard to their syntax (hence are logically equivalent) and thus describe the same event, will be called 'equivalent'. Although singular statements represent particular occurrences (i.e. particular observation statements), universal statements *exclude* particular occurrences. For example, the universal statement '*all* prices of consumer goods are positive' excludes the occurrence of negative or zero prices of all consumer goods.

Now in terms of this semi-formal framework we can say that a theory is falsifiable if and only if it always rules out, or prohibits, *at least one event*. Consequently, the *class* (or set) of all prohibited basic statements (i.e. potential falsifiers of the theory) will always be non-empty.

To visualize this Popperian notion of basic statements, imagine a horizontal line segment which represents the set of all possible basic statements – i.e. something like the totality of all possible economic worlds of economic experience. Let us further imagine that each event, or class, is represented by a subsegment, a 'slice', of our horizontal line segment. Now let all the homotypic statements which belong to a particular class be 'stacked' vertically like tiles (in any order) on our horizontal subsegment which represents that class (e.g. the price of a consumer good). Finally and to complete the picture, let us place in a segment of the horizontal plane all the basic statements (tiles) which are (logically) equivalent to one particular homotypic statement and let them be distributed directly behind and in the same vertical slice as that homotypic statement, thereby forming a wall of tiles. Thus the totality of all possible statements about economic events can be described by a three-dimensional space. In terms of this picture, the idea of falsifiability can be illustrated by the requirement that for every 'scientific' theory, there must be *at least one* vertical slice of a positive thickness in our diagrammatic space which the theory *forbids*. Similarly, for theories which are tautological there does not exist such a slice, and for theories which are self-contradictory the slice is the whole space.

Now the 'space' envisaged here consists of only Popperian basic statements. Before comparing various slices of this space which are prohibited by different theories, I will define more specifically what is meant by a 'basic statement' [Popper 1959/61, p. 100]. The statements which are included in the 'space' must satisfy the following two conditions if they are to be relevant for the demarcation criterion of falsifiability: (1) no basic statement can be deduced from a universal statement *without* initial conditions or parametric values, but (2) a universal statement and a basic statement can contradict each other. Together conditions (1) and (2) imply that a Popperian basic statement must have a logical form such that its negation cannot be a basic statement [p. 101]. In Section 2 of this chapter I discussed Popper's well-known view of statements with different logical forms, namely universal and existential statements. It was noted that the negation of a (strictly) universal statement is always equivalent to a (strictly) existential statement and vice versa. To speak of the occurrence of a single event is to speak of the truth of a singular basic statement (or singular non-existential statement). This suggests the following rule regarding basic statements: Basic statements (i.e. tiles in our imagined picture) have the logical form of singular existential statements. Singular existential statements differ from strictly existential statements in that the former is more specific. For example, where a strictly existential statement would be 'there are positive profits', a singular existential statement would be 'there are positive profits in the coffee industry'. Consider one more requirement which is less formal: The event which is the subject of a

[6] Unfortunately the term 'event' may seem a little vague when applied to economic concepts but this is because we are seldom succinct in our identification of economic variables. If we were to follow the suggestions of Koopmans or Debreu, we would have to say 'price of coffee *at place* A and *at time* T'. Thus our economic variables would appear to be more compatible with Popper's discussion of the concept of events.

basic statement must be, in principle, an observable event. In short, basic statements are observation statements.

Comparison of the falsifiability of two theories amounts to the comparison of two sets of classes of falsifiers (statements which if true would contradict the theory). In terms of our imagined picture of the world of economic experience, this means the comparison of slices of the 'space' of possible occurrences. Ideally, one would like to have a concept of measure or a concept of cardinality (or power) of a class. Unfortunately, such concepts are not applicable because the set of all statements of a language is not necessarily a metric space.

A concept which will work in some situations is the subclass (or subset) relation [pp. 115-16]. Consider two theories: theory *A* which is falsified by statements in slice [*a*] and theory *B* which is falsified by statements in slice [*b*]. In this regard three situations can be identified:

(1) If and only if the slice [*a*] includes *all* of slice [*b*], i.e. slice [*b*] is a proper subclass (or subset) of slice [*a*], then theory *A* is said to be 'falsifiable in a higher degree' than theory *B*.

(2) If the slices are identical (i.e. the classes of potential falsifiers of the two theories are identical) then they have 'the same degree of falsifiability'.

(3) If neither of the classes of potential falsifiers of the two theories includes the other as a proper subclass (or subset), i.e. neither slice contains all of the other, then the two theories have 'non-comparable degrees of falsifi-ability'.

In situation (1) the subclass relation would be decisive, in situation (2) it is of little help, and in situation (3) it is inapplicable.

In the following section, we shall be primarily concerned with situations (2) and (3). Ideally, all methodological questions about 'theory choice' [e.g. Tarascio and Caldwell, 1979] would be reduced to cases of situation (1), since all other criteria are not as decisive as the subclass relation (see further Chapter 5). To close this section, let us consider a rather simple example of the comparison of two theories (specifically, two models representing the theories) that can be based on the concept of the subclass relation. Consider the following variables:

$\mathbf{Y} \equiv$ GNP, i.e. aggregate spending,
$\mathbf{C} \equiv$ the portion of output demanded for use as consumer goods,
$\mathbf{K} \equiv$ the portion of output demanded for use as new capital goods,
$\mathbf{R} \equiv$ the interest rate.

For the purposes of this simple example, the comparison is between competing theories which differ only in regard to the list of endogenous variables.

Model 1

$$\mathbf{Y} = \mathbf{C} + \mathbf{K} \quad [2.3]$$
$$\mathbf{C} = a + b\mathbf{Y} \quad [2.4]$$
$$\mathbf{K} = k \quad [2.5]$$

where *a*, *b* and *k* are parameters.

Model 2

$$\mathbf{Y} = \mathbf{C} + \mathbf{K} \quad [2.6]$$
$$\mathbf{C} = a + b\mathbf{Y} \quad [2.7]$$
$$\mathbf{K} = d/\mathbf{R} \quad [2.8]$$
$$\mathbf{R} = r \quad [2.9]$$

where *a*, *b*, *d* and *r* are parameters.

Statements about **Y**, **C**, and **K** can be deduced from the model of the first theory and about **Y**, **C**, **K** and **R** from the model of the second theory. Thus the comparison between these theories satisfies the ideal conditions and, using Popper's views, it can be concluded that the second theory (Model 2) is 'more falsifiable' than the first theory (Model 1). Note well, this does not mean that it is *easier* to falsify, i.e. that it is more testable! In any case, such ideal situations are not very interesting. In the following section I will undertake the task of describing the determinants of the 'size' of the slice for a theory, especially in situations which are not ideal, namely situations (2) and (3). A more general discussion of theory-choice criteria will be postponed until Chapter 5.

4. Testability and Popper's dimension of a theory

In this section, I discuss some important methodological objectives and concepts which are compatible with the Popper-Samuelson demarcation criterion of falsifiability. Of most interest is the determination of how each can help in the comparison of the 'slices' of different theories when they do not fit the ideal situation. Unfortunately, there has been very little study of the mechanics of comparing economic theories so I shall borrow language from higher mathematical analysis because the concepts are intuitively similar. The most important concept to be developed is Popper's 'dimension of a theory'.

Let us consider some common concepts about theoretical statements and how they are related to falsifiability. Since the elements of the 'space' described in Section 3 were statements about occurrences, or better, observation statements, it is safe to conclude that these are empirical statements derivable from the theory in question. The class (or set) of potential falsifiers (i.e. our 'slice') may be called the 'empirical content' of the theory in question, such that empirical content is less than or equal to the 'logical content' (the class of all non-tautological statements derivable from the theory) [Popper 1959/61, pp. 113 and 120]. Since logical content is directly related to falsifiability, Popper says that empirical content is directly related to 'testability' [p. 121]. A theory which is easier to falsify is 'better testable'. Since what determines a theory's *success* is its ability to *withstand* severe 'tests', we will be concerned with the empirical content. When the empirical content of a theory is increased (i.e. the 'size' of the slice is increased), Popper says that it is made more testable and hence more falsifiable [pp. 119-21]. Stated this way, Popper's view may imply a misleading connection between testability and falsifiability. The distinction will be easier to discuss after we have considered some explicit economic models. For the purposes of this chapter, it will not be necessary to distinguish between falsifiability and testability with respect to the Popper-Socrates view of science.

Consider what Popper seems to think determines the empirical content and hence falsifiability of theories and statements. Popper notes that there are two common methodological objectives which may be reduced to the demand for the highest possible empirical content, namely the demand for the highest attainable 'level of universality', and the demand for the highest attainable 'degree of precision' [pp. 121-3]. Let us examine the following conceivable 'economic laws':

P: the time-path of the output of *all* goods can be described by a segment of a circle (i.e. by an equation of the *form*: $t^2 + y^2 + Et + Fy + G = 0$).

Q: the time-path of the output of *all agricultural* goods can be described by a segment of a circle.

R: the time-path of the output of all goods can be described by a segment of a conic (i.e. by an equation of the *form*: $At^2 + By^2 + Dty + Et + Fy + G = 0$).

S: the time-path of *all agricultural* goods can be described by a segment of a conic.

According to Popper, the deducibility relations holding between these four statements are as follows: From P all others follow; from Q follows S, which also follows from R; thus S follows from all the others.

Moving from P to Q the 'degree of universality' decreases. And Q says less than P because the time-paths of agricultural goods form a proper subclass of the time-paths of all goods. Consequently, P is more easily falsified than Q: if Q is falsified, then P is falsified, but *not* vice versa. Moving from P to R the 'degree of precision' (of the predicate) decreases: circles are a proper subclass of conics. And if R is falsified, then P is falsified, but, again, *not* vice versa. Corresponding considerations apply to the other moves: If we move from P to S then both the level of universality and the degree of precision decrease; from R to S the level of universality decreases; and from Q to S the degree of precision decreases. A higher degree of precision *or* level of universality corresponds to a greater (logical or) empirical content, and thus a higher degree of falsifiability.

On the basis of the desirability of universality and precision, Popper establishes the following rule: 'If of two statements both their universality *and* their precision are comparable, then the less universal or the less precise is derivable from the more universal or more precise; unless, of course, the one is more universal and the other more precise' [p. 123]. According to Popper, this rule demands that we leave nothing unexplained – i.e. that we always try to deduce statements from others of higher universality. He says that this follows from the fact that the demand for universality and precision can be reduced to the demand, or rule, that preference should be given to those theories which can be most severely tested.

This then is the foundation for a Popperian view of theory choice: choose the theory with the maximum degree of falsifiability (more precision, more universality, or both). Let us now see how Popper developed a concept of dimension which is associated with the testability of a theory. Generally the subclass relation will not be directly applicable to the comparison of economic theories and models. A quantitative criterion is needed which is not possible with such class or set concepts. As stated in Section 3, the set of falsifiers (our 'slice') includes all 'basic statements' which, if true, would falsify (i.e. contradict) the theory in question. It was also noted that Popper's basic statement is an observation statement, so let us now ask: What is the composition of an observation statement? Popper says that they are composed of '*relatively* atomic statements' (such as observed magnitudes for the variables or parameters of the theory) [p. 128]. They are 'relative' because where or how one might make these observations in order to test the theory must be specified. A set or class of such relatively atomic statements can be defined by means of a 'generating matrix' (e.g. the data published by the federal Department of Agriculture for the last ten years). The set of all these statements together with all the logical conjunctions which can be formed from them may be called a 'field'. A conjunction of *N* different relatively atomic statements of a field may be called an '*N*-tuple of the field'. For this '*N*-tuple' Popper says that its 'degree of composition' is equal to the number *N* [p. 129].

Consider a theory *A*, where there exists a field *F* of singular (although not necessarily basic) observation statements and there is some number *D* such that the theory *A* can never be falsified by a *D*-tuple of the field *F*, but it can be falsified by some *D*+1-tuple, then we call *D* the 'characteristic number' of the theory relative to field *F*. According to Popper, the class of all singular statements in this field *F* whose degree of composition is less than or equal to *D*, is then compatible with (i.e. permitted by) the theory regardless of their content [p. 129].

Since any singular statement (e.g. the list of all current prices for agricultural products) may be included in the field, it is possible to encounter difficulties and inconsistencies if some of these statements are irrelevant for the theory in question. Therefore, Popper spoke only of the 'field of application' for a theory. A rough idea of the concept of a 'field of application' is the list of variables and parameters of a model.

Now I can describe a basis for comparison of testability of theories which is associated with the characteristic number of a theory and will define what I call the *Popper-dimension* of a theory *T* with respect to the field of application *F*. A theory *T* will be called '*D*-dimensional' with respect to *F* if and only if the following relation holds between *T* and *F*: there is a number *D*, the Popper-dimension of theory *T* with respect to field *F*, such that the theory does not clash with any *D*-tuple of the field [pp. 285-6]. The concept of a field was not restricted to basic statements, but by comparing the 'dimensions' of the singular statements included in the field of application, the degree of composition of the basic statements can be estimated. According to Popper, it can thus be 'assumed that to a theory of higher [Popper] dimension, there corresponds a class of basic statements of higher dimension, such that all statements of this class are permitted by the theory, irrespective of what they assert' [pp. 129-30]. The class of permitted statements is the 'complementary set' with respect to the slice which was described in Section 3 – i.e. they are the complementary set with respect to the field of application.

Now, I shall illustrate the concept of the Popper-dimension with the statements Q and S (discussed at the beginning of this section) using some notions borrowed from elementary analytical geometry. The hypothesis Q – that the time-path of the output of *all* agricultural products is of the form: $t^2 + y^2 + \mathrm{E}t + \mathrm{F}y + \mathrm{G} = 0$ – is three-dimensional. This means that, if this statement is false, its *falsification* requires at least four singular statements of the field, corresponding to four points of its graphic representation. Likewise, hypothesis S – that the time-path of the output of *all* agricultural goods is of the form: $\mathrm{A}t^2 + \mathrm{B}y^2 + \mathrm{D}ty + \mathrm{E}t + \mathrm{F}y + \mathrm{G} = 0$ – is five-dimensional, since at least six singular statements are necessary for falsification, also corresponding to six points of its graphic representation.

At the beginning of this section it was possible to conclude that Q is more falsifiable than S because circles are only special cases of an ellipse or a conic (viz. an ellipse with eccentricity zero). That is, circles are a proper subclass of the class of all ellipses and the class of all conics. Note, however, a circle is *not* a special case of a parabola (for which the eccentricity is always equal to 1). Thus the subclass relation cannot be used to conclude that the hypothesis Q is more falsifiable than an alternative hypothesis S′ which differs from S by asserting that time-paths are parabolic. However using the concept of a Popper-dimension as the operative criterion, it can be correctly concluded that Q is more falsifiable than S′. Since a parabola is four-dimensional at least five singular statements are needed for the falsification of S′. Note also that the dimension of a set of curves depends on the number of *coefficients* which are free to be chosen. For example, a general second degree curve is algebraically expressed as: $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$. Setting the coefficients $A = B$ and $C = 0$, and eliminating two coefficients, yields the general equation for a circle. Setting *only* $4AB = C^2$ eliminates another coefficient and thereby yields the general equation for a parabola. In the case of a circle, only three coefficients are left to be freely chosen. For a parabola, only four coefficients are free. Thus Popper says that the number of freely determinable coefficients of a set of curves (or time-paths as in our example) by which a theory is represented is characteristic for the degree of testability of that theory [p. 131].

The concept of dimensions of curves can be extended to surfaces. This is convenient since the solution (specifically the parametric equations which describe the solutions) of a model can be generally described as a surface. Consider the following simple model.

Model 3

$$\mathbf{Y} = \mathbf{C} + k \qquad [2.10]$$

$$\mathbf{C} = b\mathbf{Y} \qquad [2.11]$$

where **Y** and **C** are endogenous variables and *k* and *b* are considered to be exogenous parametric variables.

Since we do not *know* that *b* is the same for all observations, we must allow for all possible values. Hence we must treat *b* as if it were a variable. The solution of this model is:

$$\mathbf{Y} = k/(1 - b) \qquad [2.12]$$

$$\mathbf{C} = bk/(1 - b) \qquad [2.13]$$

Let us illustrate the concept of the Popper-dimension by considering equation [2.12]. This is a second-degree equation or more properly, it is *at least* a second-degree equation. Since we do not *know* that it is higher, for convenience sake, I will choose the lowest possible case. Expressing a relationship between all the possible values of **Y**, *k* and *b* (the unknowns) yields a general equation of the following form:

$$A\mathbf{Y}^2 + Bk^2 + Db^2 + Ekb + F\mathbf{Y}b + G\mathbf{Y}k + H\mathbf{Y} + Jk + Lb + M = 0 \qquad [2.14]$$

Rearranging equation [2.12] yields:

$$k - \mathbf{Y} + \mathbf{Y}b = 0 \qquad [2.12a]$$

If in equation [2.14] we let $J = 1$, $H = -1$, and $F = 1$, using [2.12a] we can reduce equation [2.14] to:[7]

$$A\mathbf{Y}^2 + Bk^2 + Db^2 + Ekb + G\mathbf{Y}k + Lb + M = 0 \qquad [2.15]$$

For equation [2.15], six coefficients are free to be chosen. Thus the Popper-dimension of this model is six since it would take at least seven observations (non-coplanar points in the graphical space) in order to falsify this model. Equation [2.12] can be interpreted to say that the solution of the model asserts that the relationship between the endogenous variables and what are considered to be the exogenous variables forms a special case of a second-degree surface, namely one which also satisfies [2.12a]. I am concerned here only with the formal aspects of a model's solution, thus whenever [2.12a] holds [2.15] must hold as well. From a methodological point of view it would be helpful to consider all parameters, such as *b* in [2.11], as unlimited exogenous *variables* whenever their values (or the limits of their values) are not actually *known*.

Consider Model 2 (discussed in Section 3) with respect to the solution for **Y** (i.e. GNP) – its Popper-dimension is 51 (as I will show in Chapter 3). This may seem to be a rather large dimension for such a small model. Clearly it might seem to some that the higher the Popper-dimension, the greater the risk of the theory becoming virtually tautological [see Popper 1959/61, p. 127]. Perhaps we should consider ways to reduce the dimension of a theory. The most common method is to specify initial conditions.

In effect, the specification of initial conditions is the requirement that the 'solution surface' passes through particular points of the graphical space. In this sense, every set of initial conditions reduces the Popper-dimension by *at least* one. An example of an initial condition would be to require that in our statement Q above, the time-path pass through the origin of the graph.

The Popper-dimension of a theory or statement can be reduced in another way. For example, in statement Q, circles in general may be changed to circles with a given curvature. The change reduces the dimensions of statement Q by one since its falsification would now only

7 This method is only an approximation. A more involved but more accurate method is to consider the relationships between the coefficients under all possible translation or rotation transformations. This would help to eliminate some of the free coefficients and thus reduce the estimate of the Popper-dimension.

require three observation statements. Popper calls this method of reduction (i.e. of changing the form of the curve) a 'formal reduction' of the dimension of Q. He calls the other method (i.e. of specifying initial conditions) a 'material reduction' of the dimension of Q [p. 133]. Both methods have a similar effect on the algebraic representation of the statement – namely both determine one of the coefficients thus leaving us with one less coefficient which we may freely choose.

Note that two statements may be of equal dimensions. For example, if our statement R must hold for two sets of observations (i.e. certain outputs at two points in time), the dimension of R is reduced to three, which is also the dimension of statement P. However, statements P and R are not equivalent since statement R still allows for conical paths and statement P only circular paths. Thus it can be said that R is still more 'general' than P, in that a circle is a special case of a conic.

In economics we are usually interested in unique solutions of models – that is, given a set of values for all the parameters and exogenous variables, there is only one set of values for the endogenous variables. If there were more than one set of values for the endogenous variables which would not contradict a model, it might be said that the model in question is 'more general' than a model with a unique solution. If the values of all the parameters and exogenous variables were known then generality might be desirable. Since in this chapter such knowledge is being excluded, a criterion such as 'generality' will be disregarded. In other words, I am noting here that generality and what might be called 'specificity' are not necessarily opposites. A specific hyperbola hypothesis is no more 'general' than a specific straight-line hypothesis if both can be refuted with one observation.

The concept of a Popper-dimension will be illustrated in Chapter 3. The idea of reducing the dimension of a theory can play an important role in economic model building. I will, however, have to be more careful than Popper was about applying these concepts. The solution of an economic model is a set of parametric equations – i.e. a set of equations each of which expresses one endogenous variable as a function of parameters or exogenous variables. In the illustration using Model 3, only the dimension of one of these parametric equations or statements was discussed. If Popper's criterion is extended, it must be concluded that the Popper-dimension of an entire theory (assuming there are no separable independent parts) is the minimum dimension of all the possible parametric solution statements. In Model 3, it so happens that the Popper-dimension of both parametric equations, namely equations [2.12] and [2.13], were the same. In general this will not be the case. For example, in Model 2, the minimum dimension is 7. This extension is possible because, as Popper says, on the basis of *modus tollens*, if one statement P is shown to be false, then the theory *as a whole*, which was required for the deduction of P, is falsified [see Popper 1959/61, p. 76].