THEORIES OF EXCHANGE RATES AND THE METHODOLOGY OF ECONOMICS

THESIS

Submitted in fulfilment of the
requirements for the Degree of
DOCTOR OF PHILOSOPHY
of Rhodes University

by

DUNCAN HODGE

December 1996

ABSTRACT

This thesis is an exercise in applied methodology. Ideas in the history and philosophy of science which have proved to be influential in the methodology of economics, and in shaping economists' self-image in this regard, are selected for closer analysis and criticism. The main ideas that are addressed are those of empiricism, with emphasis on the methodological falsificationism of Karl Popper and Imre Lakatos, and Laudan's problem solving model of scientific progress. The thesis examines the relationship between empirical evidence, in the form of both econometric test results and stylized facts, and the development of theories about exchange rates and the open economy. This analysis begins with Cassel's formulation of purchasing power parity theory in 1916, through the elasticities, absorption, and Mundell-Fleming models of exchange rates and the balance of payments, up to the present day monetary and asset market models. This is done with regard to the broad methodological issues examined earlier in the thesis. Some of the main empirical and methodological difficulties in testing such theories are addressed, with particular reference to the role played by the Duhem-Quine thesis and the *ceteris paribus* assumption. Although some of these difficulties may be regarded as a matter of degree compared to similar problems in the natural sciences, it is argued that this difference is significant for the workability of falsification in economics. Moreover, the presence of hypotheses about expectations in many economic theories appears to be a substantive difference such that the difficulties posed by the Duhem-Quine thesis apply with much greater force in a social science like economics. The main conclusions are that neither the Popperian nor Lakatosian versions of falsification are seriously practiced in the area of exchange rate economics and that, unlike the position taken by advocates of falsification such as Mark Blaug, it would be inappropriate and

misguided to do so. A tentative case is made, with some reference to the theories surveyed in this thesis, for the possibly greater relevance of Laudan's more pragmatic problem solving model for the methodology of economics, particularly as regards his analysis and emphasis on conceptual problem solving in the progress of knowledge.

CONTENTS

Page

Page

Page

Page

LIST OF TABLES AND FIGURES

Page

Tables

Figures

PREFACE

One cannot but be struck by the evident unpredictability of macroeconomic variables for which it is the unenviable task of economists, particularly non-academic economists, to produce knowledgeable and accurate forecasts. This is the case even with forecasts of relatively stable variables such as economic growth rates but especially for those variables which are determined in financial markets like the stock, bond, and foreign exchange markets. Recognition of the growing significance of the latter category of macroeconomic variables has been accompanied by a large increase of research output in this area and exchange rate economics is no exception. Due to the exchange rate being regarded as perhaps the single most important price in open economies and also for the lessons that research in this area may hold for macroeconomic theories generally, exchange rate economics has experienced greatly intensified research efforts over the past twenty-five years.

The apparent 'unpredictability' of macroeconomic variables like the exchange rate and the empirical difficulties encountered in this area of research led me to question the idea that economists are able to 'test' economic theories in the same way as scientists can test theories in the natural sciences, as is implied by the unity of science thesis. Of course, I was aware that adherents to this latter view freely acknowledged practical differences in the way that the two types of scientists went about testing their theories - such as the ability of physical scientists to conduct controlled experiments while social scientists must rely mostly on non-experimental observations - but that such differences were regarded as a "matter of degree" rather than qualitative or substantive differences which seriously undermined the unity of science as regards its basic research strategy and methodology. However, my understanding of research practice in exchange rate economics made me wonder if this was an entirely satisfactory

response to the manifest difficulties economists face in testing such theories - and whether it could properly explain the reliance on *a priori* deductive methods in economics and the attendant emphasis on conceptual-analytical puzzle solving.

A central pillar of the modern unity of science thesis is methodological falsificationism: the idea that scientists should, and mostly do, strive to falsify their theories in the sense that if theoretical predictions or implications collide with the 'facts', then it is the theory that must give way to the facts; and that *ad hoc*, face-saving modifications to a theory designed to prevent it from being falsified are unacceptable. This raises the question, which many economists have disputed, whether such prescriptions with their emphases on excess empirical content, prediction of novel facts, and independent testability are seriously upheld in economics and, if not, why not. At least as regards theories of exchange rates and the open economy, the first part of this question was relatively easy to answer. The second part requires an explanation of what exactly *is* different about the subject matter that makes falsification an inappropriate methodology for economics. Thus the main thrust of this thesis is to understand the nature of the problems economics faces in this regard, with specific reference to developments in the general area of exchange rate economics since Cassel's formulation of purchasing power parity theory in 1916.

Since this work straddles two very different branches of economics I am indebted to many colleagues who helped me get to grips with the fundamental issues in each area. In particular, I would like to thank Professor Arthur Webb as my supervisor at Rhodes University for his encouragement and support, as well as the other staff members and friends in the Department of Economics who made me feel so at home during my stay at Rhodes in 1995 while I was on sabbatical leave. However Greg Farrell, now at Birmingham University in the UK, deserves a special mention and word of thanks for reading through a preliminary draft of this thesis and providing many helpful comments and

suggestions, particularly with regard to the econometric techniques that have been used to test theories of exchange rates. Similarly, I would like to thank Francis Williamson, who lectures in the Philosophy Department at Rhodes University, for valuable comments on a first draft of chapter two concerning ideas in the history and philosophy of science that have been applied in the methodology of economics. A much shorter version of this chapter has been published as a journal article, "The meaning and limitations of falsification in the methodology of economics" in *Acta Academica*, April 1995.

Finally I would like to extend my gratitude to the University of North-West (UNW) for granting me sabbatical leave to conduct the necessary research and to the Foundation for Research Development (FRD) for granting financial assistance towards the cost of this research. However, the opinions expressed and the conclusions arrived at are not necessarily those of either UNW or the FRD.

CHAPTER ONE

INTRODUCTION

The attempt to discover rational criteria which can be used to demarcate science from non-scientific research endeavour has preoccupied some of the most influential thinkers in the history and philosophy of science this century. Of central concern in this thinking, reaching even further back to Bacon, Hume, Locke, and the other British empiricists, has been the nature of the relationship between theory and empirical evidence as regards scientific method. Is a scientific theory or explanation derived from a careful observation of all the available facts relevant to a particular field of inquiry, as recommended by Francis Bacon? Or, is science characterized by the hypothetico-deductive method whereby preconceived theories have testable implications which are either confirmed or disconfirmed by the cumulative weight of empirical test results? Or, to draw on the seminal thinking of Sir Karl Popper, does science proceed by an interminable succession of conjectures and refutations: a kind of survival of the fittest in which only those theories remain which have thus far resisted strenuous attempts to find any crucial contrary evidence that would disprove or falsify them? Or, yet again, taking our cue from Imre Lakatos, does the significance of empirical test results lie in the corroboration of novel facts predicted by a new theory?

These questions have mostly been formulated with the natural sciences in mind, especially physics and mathematics but also to some extent chemistry and biology. As regards the social sciences it is fair to say that economics is the most methodologically self-conscious discipline in the field and has positioned itself as the "hardest" of the "soft" social sciences. Partly this has been due to the greater degree of quantification and mathematical expression of theory that is possible in economics; and partly to the attempt to bolster the

credibility and scientific status of economics in the public arena. As might be expected, economics has borrowed heavily, and some would argue too readily and uncritically, from ideas in the history and philosophy of the natural sciences. Particularly since the advent of logical positivism in the 1930's there has been an accelerating rate of research in the methodology of economics. The seminal contributions to the history and philosophy of science by Kuhn, Popper, and Lakatos gave an added impetus to this trend.

A fundamental issue in the methodology of economics derived from these ideas has been the role played by empiricism in the progression of knowledge in economics. Caldwell (1982) highlights this issue well by contrasting the positions of Robbins, Hutchison, and Machlup in a series of head-to-head debates: Robbins defending economic theories as derived from self-evident intuitive truths of which he regarded empirical tests as a spurious exercise as far as the validity of the theory was concerned; Hutchison introducing positivist ideas to the economics profession for the first time; and Machlup accusing Hutchison of a debilitating "ultra-empiricism". More recently, with the contributions by Popper and Lakatos, the debate has shifted emphasis to the meaning and application of falsification in the methodology of economics.

The aim of this thesis is to carry this debate forward with reference to theories of exchange rates and the open macroeconomy, duly emphasizing the cluster of theories comprising the monetary and asset market approach. The novelty of this study is that it conducts a close analysis of the empirical test results of selected theories of exchange rate determination. This is done to focus on and clarify the relationship between empirical evidence and the theories concerned, and to explain the research strategies that have been adopted by economists in the field. In brief, this thesis is intended to advance understanding of how economic theories evolve and are modified or replaced.

Although methodologists recognize the importance of the link between empirical results and theories in economics, very few

have actually closely examined the actual test results for particular theories and the research responses to them with the above objective in mind. Many papers in the methodology of economics discuss in general terms pertinent issues such as falsification or the application of Lakatos' MSRP without paying close attention to the actual empirical tests and findings that have been produced by practitioners in the field (Blaug's popular book *The Methodology of Economics, or How Economists Explain* which critically examines nine research programmes in cameo fashion is a good example). This approach may be deficient in that it tends to gloss over significant problems and issues for the methodology of economics which arise when specific theories are tested empirically. By focusing on particular groups of theories in economics and the related empirical research, this thesis may help to understand why economists adopt certain research practices.

Why theories of exchange rates? Firstly, the theories have evolved around analytically coherent models such as the relatively simple purchasing power parity relationship, the Mundell-Fleming model, and the monetary model. Whether these broadly stylized theories may be regarded as examples of a Lakatosian research programme in economics is not examined closely in this thesis. However, it is instructive from a methodological perspective to analyse the conceptual breaks between these models and the way in which empirical considerations have influenced (or have notably failed to influence) the evolution of such theories. Secondly, given the nature of the subject matter, theories of exchange rates have inevitably had to grapple with the manifest problems of modelling expectations. The pervasive influence and significance of expectations in economics has been widely acknowledged. What has perhaps not been fully recognized, however, are the implications of expectations augmented theories for the testability of economic theories and the methodology of economics. The various theories of exchange rate determination examined in this thesis thus serve as a useful vehicle for exploring these implications more closely.

Exchange rate theories also bring into sharp relief the methodological problems of econometric tests. Such problems arise whenever an abstract theory is cast into a testable form as an empirical model. Few economists have been prepared to accept the evidence of econometric tests as either proving or falsifying the underlying theory. The reliance on historical time series data in such tests and the distinction between in-sample and out-of-sample tests also helps to illustrate some noteworthy issues as regards the nature of prediction in economics. Further insight as to the nature of prediction in economics is gained by an analysis of the foreign exchange market as an efficient asset market in which the behaviour of exchange rates in many studies resembles that of a random walk.

It should be clear from the above that this thesis analyzes selected methodological issues in economics within the dominant Keynesian and neoclassical synthesis. There are, of course, alternative schools of thought that critically address questions about method in economics as elucidated by, for example, the Austrian school, the institutionalists, Simon's behavioural approach, and Marxian or radical perspectives. The latter deny that what passes for science is characterized by a rules driven methodology at all - witness the so-called *Strong Programme in the Sociology of Knowledge* at the University of Edinburgh which looks to class interests and actual personalities as the decisive influences on the direction of scientific research. These alternatives may be construed as outsiders' critiques of neoclassical economics. The present study, by way of contrast, is an insiders' critique of some received views on methodology applied (or which have failed to be applied) within the mainstream. Even if we take mainstream economics on its own terms there are significant questions regarding methodological issues which remain unanswered, or to which the response has been unsatisfactory, which this thesis seeks to clarify and perhaps redress.

Chapter two introduces some relevant ideas in the history and philosophy of science that have been applied to the methodology of economics. This chapter provides the epistemological backdrop against which the theories of exchange rates in

subsequent chapters are appraised. Due to the burgeoning research output in the general area of methodology chapter two is of necessity a highly selective but, it is hoped, not a superficial survey of the most important ideas and issues. Attention is focused on the application of methodological falsificationism in economics. The seminal work of both Popper and Lakatos is examined in this regard and the way in which their ideas have been received by economists. The basic questions posed in the chapter are: have economists genuinely practiced falsification in their research efforts? And more importantly, if not, why not? And further, is it desirable or appropriate that such methodological prescriptions be taken more seriously?

Mark Blaug is chosen as a stalwart protagonist of falsification in economics. Acknowledging that economists do not practice what they preach and roundly castigating the profession for relying on "innocuous falsification" he asserts that it is both feasible and desirable to uphold the principle of falsification in economics more strictly. He concurs with Popper that the bottom line should be that honest scientific research requires specifying in advance exactly what empirical test results or other evidence would lead one to disbandon a theory. In contrast to Blaug a central theme running through this thesis is that falsification, in either the Popperian or Lakatosian versions, is inappropriate given the subject matter peculiar to economics. This contention is illustrated with reference to theories of exchange rates and the foreign exchange market.

Of course, many economists have had qualms about the slavish attempts to model the methodology of economics on the natural sciences and the attendant emphases on mechanical determinism, causation, and prediction - a condition sometimes referred to as "scientism". A basic schism regarding the appropriate method for the social sciences was highlighted by both the *Methodenstreit* and the *Verstehen* doctrines put forward by German economic historians and philosophers in the nineteenth century. A deep seated unease with the role played by the empirical testing of theories in the methodology of economics has been expressed more recently by John Hicks:

"Once it is recognized that economic theories (those which are not mere tautologies) can offer no more than weak explanations - that they are always subject to a *ceteris paribus* clause - it becomes clear that they cannot be verified (or 'falsified') by confrontation with fact. We are told that 'when theory and fact come into conflict, it is theory, not fact, that must give way'. It is very doubtful how far that dictum applies to economics. Our theories, as has been shown, are not that sort of theory; but it is also true that our facts are not that sort of fact" (Hicks 1983: 371-2).

Unfortunately for economists like Hicks, it has proved difficult to satisfactorily articulate the precise reasons why economic theories are "not that sort of theory" and why facts in economics are "not that sort of fact". Blaug, for example, was easily able to demolish Hicks' arguments and concluded that "it is impossible to extract any coherent methodology of economics from the writings of Hicks" (Blaug 1988: 194). The arguments advanced against methodological monism have generally failed to convince most mainstream economists. Such economists have a somewhat quixotic attitude to research practice, by holding up falsification as an ideal which cannot in reality be met, but that we shouldn't stop trying. It is hoped that this thesis, by concentrating on the specific ways in which theories of exchange rates and their relationship to empirical evidence have developed, will articulate the intuitive misgivings felt by economists such as Hicks more clearly.

Chapter two briefly outlines the main points of Popper's falsificationism and Lakatos' sophisticated version thereof within the context of his methodology of scientific research programmes (MSRP). Both versions of falsification are critically evaluated as applied to economics generally. The problems that the *ceteris paribus* assumption and the Duhem-Quine thesis pose for a social science like economics are also discussed. In this regard a key idea explored in chapter two is that because many economic theories are combined with hypotheses about expectations they cannot be conclusively refuted. Any test of such a theory is necessarily a test of a joint hypothesis. If the test results are negative it usually cannot be decided whether this amounts to a disproof of the economic theory, based on the underlying fundamental variables,

or of the way in which the relevant expectations have been modelled. Thus the economic theory being tested may escape wholesale rejection at least partly due to the ambiguities in the test results implied by the operation of expectations. Only if a satisfactory model of expectations were available could we be more confident of attributing the contrary empirical evidence to deficiences in the economic theory alone. A discussion of the main approaches to modelling expectations, including the rational expectations hypothesis, suggests that no generally acceptable model currently exists and that such a model is likely to prove unobtainable given the nature of the subject matter.

The problems posed by expectations as regards the interpretation of test results and other empirical evidence is particularly relevant in financial markets such as the stock, bond, and foreign exchange markets. But it must therefore also be problematic for macroeconomic theories generally as any such theory has to explain the linkages between the real and financial sectors of the economy. A significant variable in open economy macroeconomic models is the exchange rate (and/or the balance of payments). Thus by looking more closely at the way in which theories of exchange rates have evolved, including the role played by expectations, some significant lessons regarding the methodology of economics may become apparent. Such an examination of exchange rate theories may reveal additional insights as to how such theories have evolved historically in relation to empirical evidence.

Mention was made above of the Duhem-Quine thesis. The thesis asserts that it is generally not possible to refute an isolated theory because it is invariably combined with auxiliary hypotheses concerning both the operation of exogenous variables (held constant by the catch-all *ceteris paribus* clause) and the empirical measurement of the endogenous variables. Both sets of hypotheses are necessary to make the theory testable. A negative test result points to the entire complex of hypotheses (the explanans) and thus cannot conclusively falsify the particular theory or hypothesis in question. Besides the general problem in economics that, to a far greater and more

significant extent than in the natural sciences 'everything depends on everything else', the fact that many economic theories are combined with a hypothesis about expectations means that the Duhem-Quine thesis applies with particular force in economics. These and other related differences considered in chapter two support the view that economics may be better construed as an explanatory discipline, as Hicks believed, rather than a deterministic predictive science.

Chapter two also considers the problem solving model of scientific progress proposed by the philosopher Larry Laudan. The main differences between Laudan's model and the methodology of falsification are explained. Laudan's ideas are also applied to the theories examined in subsequent chapters, particularly his emphasis on conceptual problem solving as opposed to the empiricist emphases of falsificationists such as Popper and Lakatos. It is suggested that Laudan's model may be more appropriate in economics than methodological falsificationism.

Chapter three explains the main versions of purchasing power parity theory (PPP) which have been tested empirically. The historical origins and doctrinal aspects of the theory are examined, and its connection with the monetary approach to the balance of payments and exchange rates. The chapter explains the conceptual confusion that has surrounded the meaning of PPP. A range of test results from various empirical analyses of the theories are reported and the significance of deviations from the predictions of the theory are discussed. This section includes the most recent work using unit root tests and cointegration techniques whereby PPP is interpreted as a long-run equlibrium condition. The methodological implications of the above are then examined.

Since the initial formulation of the theory by Gustav Cassel in 1916, the basic idea that there is a relationship between exchange rates and relative price levels has persisted to this day and has been the subject of ongoing, intensive empirical investigations. It is instructive from a methodological perspective to try to understand why this concept has endured so tenaciously when it is considered that most empirical tests

of this relationship, in its various forms, have proved negative. Corroboration of the testable implications of the theory have been the exception rather than the rule, including those tests that interpret PPP as a long-run equilibrium relationship rather than a deterministic law. Even if we disregard naive versions of falsificationism which emphasize knockout blows to a theory from one or two crucial experiments or counterexamples, the cumulative weight of empirical evidence reported in this chapter still appears to tilt the scales firmly downwards. Evidently, since the theory remains very much with us, economists have not taken falsification very seriously in this case. The question then is why not, and should they have? It is suggested that to do so would have been inappropriate and counterproductive.

Two underlying interpretations of PPP are distinguished. The commodity arbitrage version, when applied at the level of individual internationally traded goods, implies the law of one price: that the exchange rate adjusted prices of identical goods should be everywhere the same. The analysis of empirical research into the law of one price reveals some noteworthy methodological issues: what is the significance of observed deviations from what the theory predicts; should the law of one price be interpreted as an intuitive, self-evident truth - an economic axiom of the kind Robbins felt that it was unnecessary to test the validity of empirically; or, with the DuhemQuine thesis in mind, do economists simply end up testing for the extraneous influence of one or more of the many auxiliary hypotheses which are invariably combined with the law of one price itself - for example, the hypotheses that transport costs are zero; perfect information; zero arbitrage risks; absence of product differentiation, *etcetera*?

When the law of one price is applied to collections of goods, additional problems arise concerning the appropriate price or cost indices used to measure national price levels. Moreover, once we move from narrow traded goods indices to broader indices including non-traded goods and services the very meaning of the theory itself changes - from a commodity arbitrage to a monetary interpretation. The commodity arbitrage

version of PPP hypothesizes a link between relative price parity levels and the exchange rate based on traded goods arbitrage, whereas the monetary interpretation hypothesizes a link between excess national money supplies and exchange rates. The latter, it may be argued, was the original meaning Cassel gave to his purchasing power parity theory in 1916, as suggested by his choice of words in naming the theory. The two interpretations may be reconciled by interposing the quantity theory of money in the determination of national price levels although this, in Frenkel's view, was counterproductive:

"In retrospect it seems that the translation of the theory from a relationship between moneys into a relationship between prices - via the quantity theory of money - was counterproductive and led to a lack of emphasis on the fundamental determinants of the exchange rate and to an unnecessary amount of ambiguity and confusion"
(Frenkel 1978b: 5).

Conceptual difficulties like these point to the operation of the *ceteris paribus* clause and the distinction between the long-run and the short-run in economics. Because of the catch-all nature of the *ceteris paribus* assumption attached to theories like PPP, deviations from what the theory predicts are not usually construed as falsifying evidence. In economic theories such as PPP any given set of explanatory variables can only provide a partial or weakly causal explanation of variation in a dependent variable like the exchange rate. Under these circumstances it is argued that empirical research and "tests" may play an important exploratory role in helping to clarify the limitations and range of application of preconceived deductive theories, for particular historical sample periods, but that serious attempts at falsification would be misguided.

A further problem in trying to apply falsification to economic theories is the inevitable presence of a time dimension attached to such theories. This imparts a qualitatively different dimension to the study of economic phenomena and processes compared to a pure natural science like physics. Causation in most physical theories is of short or even instantaneous duration (although there are exceptions as, for

example, in theories of evolution). Cause and effect relationships in economic theories, however, generally take place over a much longer time period. Moreover, the effects of any particular cause may differ depending on the time period - short-run effects may be different to long-run effects. These considerations are clearly apparent in the analysis of theories of PPP outlined in chapter three. In particular, most economists have regarded PPP as a long-run relationship of some kind and this is one reason why negative empirical evidence has been treated with a large degree of circumspection. The notion has persisted that exchange rate deviations from the levels predicted by PPP are due to temporary short-run disturbances which, given sufficient time for the fundamentals to reassert themselves, will disappear in the long-run. However, the long-run is rarely if ever defined operationally. Thus, depending on one's point of view, the use of this concept in economics may either be interpreted critically as a conventionalist strategem to protect a theory from falsifying evidence - or simply accepted as a necessary adjunct to many economic theories because of the nature of the subject matter.

Recent empirical tests of both PPP and the monetary model have adopted more powerful unit root and cointegration techniques, compared to the standard regression techniques used in earlier studies. Economists such as Ronald MacDonald and Mark Taylor argue that due to the parameter constraints imposed on standard regression equations early tests may have failed to detect evidence of mean reversion which, if statistically significant, implies that exchange rates do not stray from their theoretical PPP levels indefinitely. Some of the more recent cointegration tests using the Johansen estimation technique have indeed found empirical evidence of mean reversion in exchange rates, with the conclusion that PPP (and the monetary model) is valid when interpreted as a long-run equilibrium condition.

A pattern that emerges from the battery of 'tests' that have been run on PPP is that test results, whether negative of positive, are provisional because they are subject to ongoing innovations in econometric techniques. For example, early tests by Frenkel (1978a) could not reject the null hypothesis of unit

elasticity of exchange rates with respect to the relevant price ratios, which were thus interpreted as highly favourable to PPP. However later unit root and Dickey-Fuller tests, which avoid the problem of spurious regression when there are common time trends, could not distinguish the real exchange rate from a random walk and could not find evidence of mean reversion to PPP even in the long-run. Recent cointegration tests using the Johansen maximum-likelihood estimation procedure, on the other hand, do find evidence of a long-run equilibrium relationship. The mostly negative results of the unit root tests (and of initial Engle and Granger cointegration tests) are explained as by MacDonald and Taylor above. It seems clear that if prior empirical results are subject to serious reservations on the basis of the latest innovation in econometric techniques, then genuine attempts at falsification of the underlying theory are inappropriate.

Chapter four examines analytical and conceptual developments in theories of the balance of payments and exchange rates since the 1920s. In roughly chronological order the main theories examined are the elasticities model, the income-expenditure (absorption) approach, the Meade-Mundell-Fleming model, and the monetary and asset market models. Most of the attention is focussed on the latter, particularly the transition from the Mundell-Fleming model to the monetary model. The historical context and doctrinaire aspects of these developments are discussed with some reference to the Keynesian synthesis presented in James Meade's *The Balance of Payments* and the pioneering work by Robert Mundell, Jacob Frenkel and Harry Johnson. It is not assumed that the reader is necessarily familiar with all the relevant theories so their main features, differences, and implications are described and explained. Some preliminary methodological issues are addressed in the course of this analysis and at the end of the chapter. The empirical research results and methodological implications are analysed further in chapter five.

The main fault line between these theories occurred towards the end of the 1960s with a clear conceptual shift favouring the monetary approach *vis-a-vis* the Mundell-Fleming model, which

represented the most sophisticated open version of Keynesian macroeconomics. This shift may be understood as an extension of the broader monetarist research programme in macroeconomics that had been initiated earlier in the decade and the resurgence of classical thinking generally in the form of the New Classical Economics. Thus the monetary approach may be regarded as a subsidiary research programme which extended monetarism from closed to open macroeconomic analysis and which constituted a theoretically consistent rival to the established Keynesian open economy model. The monetary and asset market approaches also developed some distinctive conceptual features of their own which went well beyond the conventional closed economy arguments concerning the slopes of the IS and LM curves, stability of the demand for money (velocity), and exogeneity of the money supply.

The most distinctive conceptual feature of the monetary approach was that it made international money market equilibrium the centrepiece in its analysis of open economies. Replacing the Keynesian emphasis on income, expenditure, and trade flows was an analysis based on flows of money across the balance of payments and the attainment of stock equilibrium in the money and foreign exchange markets. The careful theoretical distinction between stock *versus* flow equilibrium processes clearly demarcated the monetary approach from earlier Keynesian models. In the Mundell-Fleming model, for example, changes in international interest rate differentials are regarded as bringing about a more or less permanent change in net capital flows which can offset adverse trade flows indefinitely. By contrast, under the monetary approach changes in interest rates trigger off temporary changes in capital flows which cease once desired adjustments in asset portfolios are realized. A related conceptual difficulty of the Mundell-Fleming model, but which becomes an integral feature of the monetary and asset market models, is the treatment of expectations and speculative capital movements. In the monetary approach speculative capital flows - actual flows if the exchange rate is fixed and incipient flows under a flexible exchange rate - reflect asset market disequilibrium. Such capital flows and/or exchange rate adjustments are temporary (virtually instantaneous in the case

of freely floating exchange rates) phenomena which re-establish asset market equilibrium in the sense that at the then prevailing interest rates and exchange rates, asset stocks are willingly held.

An important methodological issue in this regard is that a flow model like the Mundell-Fleming model can, in principle, be modified to account for the effects of changes in expectations and speculative capital movements: by allowing sufficiently large exogenous shifts in the relevant balance of payments (FF) schedule or by admitting perfectly elastic capital flows. However, the point is that the model is then no longer a genuine flow model in which income, expenditure, and trade flows are the main explanatory variables. The dependent variables of the model become completely determined by the tacking on of an exogenously determined variable, that of expectations induced speculative capital flows. Thus the adaptation of the model becomes *ad hoc* in the Lakatosian sense that it does not conform with the spirit or positive heuristic of the research programme of which it is part (Lakatos 1970: 174-7 and 182-7).

Of particular interest is the way in which the monetary and Mundell-Fleming models may be regarded as rivals and in what sense, if any, the former can claim superiority over the latter. As outlined in chapter two, both Popper and Lakatos emphasize empirical criteria as the arbiter of good theories and scientific progress. However, the reasons for the transition from the Mundell-Fleming model to the monetary approach suggests that very little of the compelling story told by Popper and Lakatos was relevant.

From its position as the most sophisticated and complete explanation of the balance of payments and exchange rates within the Keynesian theoretical framework, the Mundell-Fleming model was unceremoniously dumped as a serious contender in this regard less than a decade after it was developed. With hindsight the reasons for this rejection appear to be conceptual and analytical rather than empirical. Very little "serious" empirical testing of the Mundell-Fleming model or its

implications was attempted during the 1960s. As regards the monetary model, as made clear in chapter five, the bulk of econometric test results have proved negative and yet the model in its various forms has persisted (with ever more "tests" being run on it) for over twenty-five years. The discussion in chapter four suggests that the reasons for preferring the monetary model appear to be based more on conceptual grounds than specific empirical considerations. *A priori*, on the basis of thought experiments, the conceptual shift from thinking in terms of flows to the full implications of monetary stock-flow equilibrium seems to have proved decisive - with the significance of empirical tests relegated to the lesser order of smalls. Put another way, the predictive capability of the model desirable from a falsificationist position played second fiddle to its analytical power as a conceptual model.

Chapter five examines some of the empirical evidence regarding the monetary model of exchange rate determination. In so doing a distinction is made between the results of econometric tests and the evidence as regards stylized economic facts, loosely defined as broad empirical regularities that are regarded as requiring a consistent theoretical explanation. It is argued that empirical models may fail specific econometric tests and yet the underlying theory may be able to explain certain stylized facts better than a rival theory. The chapter also analyses the efficiency hypothesis in the context of the foreign exchange markets. The implications of this hypothesis for the use of structural models to predict exchange rates are also addressed. The last section examines some pertinent methodological issues arising from the above.

Three broad phases of empirical research into the monetary and other asset market models may be distinguished. The initial phase (1968 - 1978) had some success in constructing econometric models that fitted the historical time series data reasonably well. However the second phase (1979 - 1985), which concentrated on testing monetary and asset market models for the recent floating exchange rate period, was much less successful in this regard. A third phase of research (1985 to date) appears at this stage to have revived the fortunes of the

basic monetary model in that some cointegration tests have reported evidence of mean reversion in exchange rates towards long-run monetary (and PPP) equilibrium.

A notable feature of the second phase of tests was the general inability to reject the null hypothesis of a unit root in time series data on real exchange rates, and the great difficulty in finding a structural model that could outperform naive random walk models in out-of-sample predictions of the exchange rate. For example, Frankel (1983) regarded the results of his tests of the mark/dollar exchange rate on post-1978 sample data as a "disaster" for the monetary model; and, as noted in chapter three, Frenkel (1981b) concluded that purchasing power parities had "collapsed" during the 1970s. These results would presumably be more than sufficient to falsify and reject the monetary model yet testing persevered with a third phase of empirical research. Interestingly, the initial Engle and Granger cointegration tests of this phase also proved mostly negative - it is only with the later adoption of the Johansen cointegration technique, which allows for unconstrained simultaneous multivariate estimation of the model parameters, that more positive empirical support for the monetary model (and PPP) has emerged. The nature of these tests also raises some pertinent epistemological questions regarding the distinction between the economic meaning and statistical significance of the empirically estimated relationships.

Despite the generally negative econometric test results of the second phase it may nevertheless be argued, with some justification, that the monetary approach can better explain certain stylized facts about exchange rates and the balance of payments than other rival theories. For example, genuine flow models of the foreign exchange market cannot easily explain the essentially random walk behaviour of floating exchange rates in a theoretically consistent way, whereas the monetary model can readily accommodate this fact (more accurately, the later asset market versions thereof). It may also be argued that the monetary model could account for some noticeable policy failures, such policies based on the orthodox Keynesian theory, that occurred during the Bretton Woods era - the evident

failure of devaluation policies to significantly improve the balance of payments when needed (the British devaluation of 1967 being a particularly acute example thereof). Significantly, the monetary model could do so in a theoretically consistent manner. In a nutshell, the revealed preference for the monetary/asset market approach seems to have been based on a combination of *a priori* conceptual considerations and the ability to explain certain stylized facts as outlined above.

Of the various reasons given for the empirical failure of the monetary model, the lack of a satisfactory explanation of how expectations are determined seems paramount. Without such a model it cannot readily be ascertained whether negative test results are due to incorrect hypotheses about expectations or to misspecification of the fundamentals of the structural model itself. Rational expectations may help by endogenizing expectations in a theoretically consistent way but this approach leads to a *cul-de-sac* because it assumes that speculators base their expectations about the exchange rate on the parameters of the chosen monetary model. If speculators act on some other set of fundamentals or if their expectations diverge on the basis of competing views of how the economy works then the rational expectations approach is of little use when it comes to interpreting empirical test results. Depending on how expectations and speculative capital flows are actually determined, the observed changes in exchange rates could be consistent with any structural model. Moreover, hypotheses about such expectations may very well be historically circumscribed since speculators may not consistently react to any given matrix of fundamentals at all times.

The way in which theories of the balance of payments and exchange rates have developed suggests that the empirical requirements demanded by the versions of falsification put forward by Popper and Lakatos - with their emphases on excess content, independent testability, and the prediction and corroboration of novel facts - is misplaced. The relevant facts to be explained were hardly novel or unexpected. Rather than predicting anything new, such theories were judged by their

ability to postdict commonplace observations (such as the policy failures of the 1960's and the volatility of floating exchange rates) in a theoretically consistent fashion. Theoretical developments were thus *ad hoc* in the sense that they did not include any testable predictions beyond the potential falsifiers of their predecessors. However, they were not *ad hoc* in the sense that they conformed to the positive heuristic of their respective Keynesian and monetarist research programmes as applied to the open economy.

A final point made in chapter five concerns the implications of the efficiency hypothesis as regards prediction on the basis of macroeconomic models. If, for example, a model or econometric technique were found which for a time consistently outperformed a random walk in out-of-sample predictions of the exchange rate (as claimed for the Johansen variant of cointegration techniques), the hypothesis implies that it would not be ignored by competitive financial markets. Thus, like any other publicly available information, its use would soon erode any excess returns to speculation in the relevant markets and bring exchange rates back to a state of near unpredictability. This implies that any mechanical formula, including cointegration techniques, would have only fleeting success in predicting asset prices like the exchange rate.

Chapter six draws some tentative conclusions and implications for the methodology of economics from the above. The basic conclusion, with reference to the theories and empirical tests analysed in this thesis, is that not only has falsification not been practiced seriously but that it would be inappropriate and misguided to do so. The reasons for this conclusion are briefly examined. It is further argued that Laudan's more pragmatic problem solving model of scientific progress, despite some drawbacks, generally fits the intellectual history of economic theories far better than alternative falsificationist accounts thereof.

CHAPTER TWO

HISTORY AND PHILOSOPHY OF SCIENCE AND THE METHODOLOGY OF ECONOMICS*

This chapter introduces some selected ideas in the history and philosophy of science and how they have been applied in the methodology of economics. Methodology here is taken to mean the approach by which a discipline attempts to explain its subject matter; and how explanatory theories are appraised. It is thus a part of epistemology, that branch of philosophy dealing with theories of knowledge. These ideas are applied to the analysis of theories of exchange rates and the open economy in subsequent chapters.

There is a vast literature on these topics in the history and philosophy of science which thus necessitates a selective precis of the most relevant issues. This chapter focuses on empiricism and the testability of theories. Section 2.1 addresses Karl Popper's methodological falsificationism and this is followed up in section 2.2 with an examination of the meaning and limitations of Popperian falsification in the methodology of economics. Section 2.3 extends some of these issues further with reference to the role played by expectations in economic theories. Section 2.4 distinguishes Imre Lakatos' version of falsification within the context of his methodology of scientific research programmes and section 2.5 explains how MSRP has been received in mainstream economics. A similar pattern is followed in sections 2.6 and 2.7 where Larry Laudan's problem solving model of scientific progress is explained and its possible relevance for economics investigated respectively.

---

* A much shorter version of this chapter has been published by the author in the South African journal *Acta Academica* (April 1995) as "The meaning and limitations of falsification in the methodology of economics".

2.1 Karl Popper's methodological falsificationism

Paul Johnson opens his wide-ranging book *A History of the Modern World* with the provocative statement that:

"The modern world began on 29 May 1919 when photographs of a solar eclipse, taken on the island of Principe off West Africa and at Sobral in Brazil, confirmed the truth of a new theory of the universe" (Johnson 1983: 1).

He relates how Einstein refused to accept the validity of his relativity theory until it had withstood these tests, and describes Eddington's highly publicized expedition to Principe for this purpose. There was a dramatic quality to the experiment, from the thunderstorm which almost prevented the required photographs being taken, to the announcement of the results at a meeting of the Royal Society in London later that year where the august gathering acknowledged the first major revision, after more than two hundred years, to Newton's theory of gravitation, time and space - the application of which laws had helped usher in the Industrial Revolution and indelibly shaped our conception of the world.

It was this type of bold, high risk theory, vulnerable to refutation from a single crucial experiment, that so excited the imagination of Karl Popper in his formative years as a philosopher of science. For in sharp contrast to the all embracing theories of Marxism and psychoanalysis which were fashionable at that time, Einstein's theory of relativity specifically forbad certain observable events - its great power lay, not in being able to explain anything and everything that occurred, but in providing explicit testable implications and subject to refutation if it failed such tests. The idea that theories should be open to falsification, in the sense of specifying in advance precisely which test results would disprove them, became Popper's methodological demarcation criterion between science and non-science:

"Criteria of refutation have to be laid down beforehand: it must be agreed which observable situations, if actually observed, mean that the theory is refuted" (Popper 1963: 38).

It is important to distinguish between Popper's view of empiricism and that of both logical positivism and logical empiricism, which held sway from about 1925 to the middle 1950s (see Caldwell 1982: 11-67). The logical positivists attempted to brand all non-analytic statements and terms as literally meaningless if they could not be independently verified, at least in principle, by empirical data. This radical philosophical view, exemplified in physics by the operationism of Percy Bridgman, only allowed the use of metaphysical concepts and theoretical constructs in scientific theories if they were defined instrumentally in purely observational terms. This unduly restrictive view of science gave way to a more accommodating version of empiricism from about the mid 1930s, that of confirmation. The logical empiricists focused their attention on the hypotheses that were deducible from an internally consistent theoretical framework. If these were repeatedly verified by observation and experiment then the theory and, indirectly, the metaphysical concepts and constructs embedded in it, were confirmed. The greater the number of positive test results, the greater the degree of confirmation and faith in the theory, and correspondingly the greater the probability of its laws.

Popper's approach, although at first glance superficially similar to these views, on closer examination is a very different, and in some respects opposing, methodological prescription. Firstly, his demarcation principle is not intended to discern meaningful from meaningless statements but to separate science from non-scientific endeavours. Metaphysical and even mythical explanations are regarded as meaningful, perhaps even true, and many scientific theories have originated from them. However, only when expressed as falsifiable theories do they qualify as scientific explanations.

Secondly, Popper's methodology attempts to demonstrate not a logic of proof, as attempted by logical empiricism, but a logic of disproof. As argued so cogently by David Hume (1777: 25-55), no number of individual observations can establish the logical certainty of any future observations or of any generalizations

inferred from them. Still more devastating, for a logical empiricist, is the conclusion that not even the probability of such statements is increased by repeated confirmations. We feel sure, from custom and habit, that the sun will rise tomorrow because it has done so for all our yesterdays; but since there are yet an infinite number of tomorrows the logical probability of the sun rising again is not increased by our past confirmatory experiences, no matter how numerous they may be. As Popper made clear, applying a relative frequency concept of probability to a hypothesis is wholly unsatisfactory:

"For example, one is led to the definition which attributes a probability 1/2 - instead of 0 - to a hypothesis which has been falsified a thousand times; for this attribution would have to be made if the hypothesis is falsified by every second result of its tests" (Popper 1959: 316).

Accordingly, science should not be seen as an established body of knowledge that is either true or even probable, but as a system of inspired guesses or conjectures that are accepted only for as long as they survive incessant and exhaustive tests.

Popper resolves, or rather circumvents, the problem of induction by asserting the fallibility of all theories and our knowledge about the world based on them. Although no number of confirming instances can increase the certainty of any scientific hypothesis it can be logically disproven, that is falsified, by contrary empirical evidence. No observed number of white swans can logically prove the hypothesis that "All swans are white". However, a single black swan is sufficient, by the logic of *modus tollens*, to disprove the statement. Falsification thus circumvents the problem of induction at two levels: the old Baconian logic of discovery, in which the collection of facts supposedly precedes theory and from which the hidden regularities of nature or generalizations can be rationally inferred; and the logic of justification whereby confirming instances of a prediction are thought to increase the probability of a hypothesis.

Popper also demands of science (his "second requirement") that a refuted hypothesis may only be superceded by a rival

hypothesis if it predicts novel facts, that is, new facts that could not be predicted, or are even forbidden, by the refuted hypothesis. The rival theory must, to some extent, be independently testable. It is not enough for it merely to explain the failures of the refuted theory, as the same result could be achieved by the *ad hoc* addition of auxiliary hypotheses, or changes in the definitions of terms, or by some other convenience.

A "third requirement", moreover, is that at least some of the novel facts predicted by the theory are corroborated. However, only theories with a high empirical content (in the sense of excluding many possible observations and therefore with a high risk of refutation) which generate unambiguous predictions that have repeatedly withstood critical tests are worthy of consideration. Low risk statements like "It will rain" contain little specific empirical content and may thus easily fail to be falsified but would be regarded as trivial and uninformative. Scientists are more interested in bold predictions like "It will rain in London tomorrow afternoon" that are both high risk and yet survive searching tests (Magee 1973: 35). The corresponding theory, however, is never proved; our knowledge of the world based on such theories remains provisional and open to revision or refutation in the light of future experience. The notion of corroboration is an important part of Popper's methodology for without it the weight of purely negative test results would crush the life out of continued scientific endeavour:

> "...I contend that further progress in science would become impossible if we did not reasonably often manage to meet the third requirement; thus if the progress of science is to continue, and its rationality not to decline, we need not only successful refutations, but also positive successes" (Popper 1963: 243).

In *Logick der Forschung*, Popper maintained that his methodological falsificationism could avoid the use of metaphysical concepts such as "truth" and "falsity". Theories are either provisionally "accepted" or "rejected" according to the rules. By playing the game of science according to his methodological rules, the best we can hope for is the detection

and elimination of error. The positive successes of corroboration, important though they are, can function only as a goad to scientists to devise yet more searching tests to refute their theories. The question of what is the aim of this never ending scientific game is left hanging. As Lakatos complains:

"Scientific 'progress' is increased awareness of ignorance rather than growth of knowledge. It is *'learning'* without ever *knowing*" (Lakatos 1978: 155).

Only about twenty five years later[1] did Popper address these concerns directly, by introducing the simple concept of verisimilitude or "truthlikeness", derived from the correspondence theory of truth which had been rehabilitated by Alfred Tarski during the intervening years. In short, as successive theories are refuted and replaced by new hypotheses with greater corroborated empirical content so is a greater approximation to the truth achieved, that is, the greater their verisimilitude. The corroboration of theories may thus be interpreted as signposts by which positive progress towards truth may be recognized. Although ultimate truth may be unobtainable, more accurate correspondence with reality and successively better approximations to the truth are possible.

Criticism of Popper's falsificationism has emerged from various quarters and alternative methodological approaches have developed since the early 1960s. The most significant of these are the growth of knowledge theorists of which Thomas Kuhn, Paul Feyerabend and Imre Lakatos are the main exponents. It is noteworthy that in each case, as with Popper, their ideas are primarily derived from a study of the natural sciences, particularly physics. This chapter largely ignores the alternative methodological positions of Kuhn and Feyerabend as their ideas are not applied significantly in the remainder of this thesis. However, Lakatos' version of falsification is examined in some depth, particularly as regards the issues of theory preference and choice. Laudan's problem solving model is also investigated as to its possible relevance in economics in this regard.

2.2 The meaning and limitations of Popperian falsification in the methodology of economics

Mainstream economics is peculiar in that, although it is a social science and supposedly shares the same characteristics and unique problems inherent in the study of human behaviour as, for example, sociology and psychology, it often appears more akin to a natural science in its degree of mathematical abstraction and quantification. A layman reading a micro-economics textbook may be forgiven in thinking that he has mistakenly picked up a book on physics or geometry; and this impression would be reinforced by textbooks on welfare economics, international trade, and financial economics. Moreover, since the publication of the journal *Econometrica* in the 1930s a whole branch of economics, that of econometrics, has developed with the express aim of formalizing economic theories and testing them with the aid of increasingly sophisticated statistical techniques. This is generally taken as a sign of the greater maturity and progress that has occurred in the subject and for this reason economics is often regarded as the "hardest" of the "soft" social sciences.

It is thus tempting to conclude that the criterion of falsification is strongly upheld in economics. However, although it is true that most mainstream economists uphold the principle of critically testing their theories (see, for example, Lipsey *et al* 1993: 29-31) few theories in economics have ever been discarded simply because they appear to be refuted by negative test results. Mark Blaug, a prominent advocate of falsification in economics, deplores neoclassical economists who "preach the importance of submitting theories to empirical tests, but (they) rarely live up to their declared methodological canons" (Blaug 1980: 259) and he characterizes the usual research practice of economics as "innocuous falsificationism" in the sense that apparently damaging empirical evidence is rarely taken seriously:

> "My own contention, by way of contrast, is that the central weakness of modern economics is, indeed, the reluctance to produce theories that yield unambiguously refutable implications, followed by a general unwillingness to confront these implications with the facts" (Blaug 1980: 254).

A major theme of this thesis, using the history of theories of exchange rates and the open economy as an illustrative case study, is to explain why falsification is an inappropriate methodology for a social science like economics and thus why few economists have ever actually practiced it seriously[2].

A general difficulty in applying Popper's falsificationist methodology to economics is the problem raised by Duhem (1914) and Quine (1980), originally as regards theories in the natural sciences. The Duhem-Quine thesis is that negative test results cannot refute an individual hypothesis because it is always tested, explicitly or implicitly, in conjunction with auxiliary hypotheses. By the logic of *modus tollens*, a negative test result points to the entire *explanans* and thus cannot conclusively disprove the particular hypothesis concerned. For example, Newton's Second Law of Motion would be refuted in terrestrial applications unless account were taken of the effects of air resistance and other sources of friction. Such *prima facie* refutations can be denied by asserting either that the auxiliary hypothesis concerning the effects of friction is false, or that the instruments used to measure friction were faulty, or even that there was a mistake in recording or reporting the test results[3]. Hence, test results are always subject to interpretation and appraisal against these considerations; and the decision to accept or reject them as genuine refutations of a theory is largely a matter of convention.

Popper was aware of this conventionalist argument and devised additional methodological rules to limit the "immunizing strategems" that scientists might use to prevent refutation of their theories. These are meant to bar purely *ad hoc* changes to a theory, whether of concepts or the auxiliary hypotheses, which have the sole intention of saving the theory from falsification. Changes in, or the introduction of, new auxiliary hypotheses are only allowed if they "do not diminish the degree of falsifiability or testability of the system in question, but, on the contrary, increase it" (Popper 1963: 83). The question to be asked is whether these rules are feasible given the peculiar subject matter of economics.

Although the Duhem-Quine problem was posed with the natural sciences in mind, the difficulties it poses for a social science like economics are greatly magnified. In the natural sciences it is possible to conduct controlled experiments whereby exogenous variables such as friction, temperature and pressure are specified and can be held constant. Alternatively, if it is not possible to hold the exogenous variables constant, or if they are unknown, it can usually be reliably assumed that they do not systematically affect the endogenous variables from one test situation to another. In economics it is generally not possible to control exogenous variables under experimental conditions. We cannot, for example, ask consumers to obligingly maintain their given tastes and preferences, or disallow changes in their incomes, while checking the law of demand. In testing for the slope of the demand curve, these and other exogenous variables are held constant (or, more accurately, not to change systematically over time) by the *ceteris paribus* assumption. However, in economics there can be far less faith in this assumption being met as the exogenous variables, and the hypotheses regarding their effects on the theory in question, usually cannot be as clearly isolated or specified as those in the natural sciences. Moreover, the variables concerned may change systematically over time. In short, the universal laws and precise, checkable initial conditions required for conclusive test results to disprove a theory are largely absent in economics.

Popper seemed to be aware of such problems:

"In physics...the parameters of our equation can...be reduced to a small number of natural constants...This is not so in economics; here the parameters are themselves in the most important cases quickly changing variables. This clearly reduces the significance, interpretability and testability of our measurements" (Popper 1957: 143).

Unfortunately, as commented upon by Hutchison (1976: 188), Popper failed to elaborate the implications of this difference for the methodology of economics and held that the "laws" of economics were comparable to those of the natural sciences. There are very few economic laws, however, that are comparable to laws in the natural sciences as regards their reliability,

accuracy and generality, partly due to the presence of numerous (and sometimes unspecified) exogenous variables and partly due to the indefinite time period allowed for the operation of the law to take effect. An increase in the relative price of oil will cause, according to the "law" of demand, a reduction in the quantity of oil demanded; but it is an unsatisfactory law that cannot say reliably when or by how much this will occur; and that it will depend on the behaviour of numerous exogenous variables some of which are known yet others are unspecified; and further that any contrary empirical observation means only that the long run has not yet been reached - that we need to wait yet a while longer for the operation and effect of the law to become apparent.

Hutchison (1976: 189) suggests that if "there are no reliable laws but only historical trends or patterns on which to base predictions, then the question arises as to the relevance and suitability, for economics, of the strongly anti-inductive emphasis in Popper's methodology". In other words, the subject matter of economics usually only permits the weaker inductivist methodology of confirmation:

> "Extrapolation of trends by a kind of induction is a method which has obvious weaknesses. But beggars can't be choosers, and if, in some important branches of economic prediction, inductive extrapolation is an inevitable or demonstrably superior method, *because of the nature of the material*, then it must be recognised, and the best must be made of it" (Hutchison 1977: 23)[4].

A further difficulty presented by the subject matter of economics is that the empirical data may not adequately represent the corresponding theoretical constructs. Machlup (1978: 159-88), for example, asks how the price of steel must be measured to test the law of demand in the steel market. This seems easy enough, but steel is not a homogenous product sold at a uniform price. There are many different grades of steel and of varying quality. Steel may be sold in various shapes and sizes, such as sheet metal or ingots. There may be different terms of delivery and payment depending on the supplier or market in which it is sold. Thus the empirical counterpart is only an approximation to the theoretically clean concept of

"the steel price". Such problems are particularly acute in macroeconomics where it is necessary to give empirical proxies to far broader aggregates. Many of the empirical and conceptual problems concerning tests of purchasing power parity, for example, are the result of a mismatch between the actual indices used to measure national price levels and their theoretical counterparts (see chapter three).

These problems are reflected in the evident failure of predictions based on econometric models. Although equations may be found which provide a good explanatory fit to historical data very few, if any, of these models have succeeded in providing *repeatably* accurate predictions in the sense of physical laws in the natural sciences. As noted by Thurow:

"In the end, econometric testing did not prove up to the assigned task. Equations and coefficients were not stable-robust. In other words, equations that were good at tracking historical experience proved to be poor predictors of the future; equations did not stand up over time; and changes in parameter values were frequent and dramatic" (Thurow 1983: 106)

If the predictive power of an econometric model depends upon which sample is used it is specious to maintain that test results can be used to refute economic theories. Also, more than one economic theory may be able to account for the historical data yet none of them may prove superior as regards their predictive record. Much econometric research can, in fact, be construed as a procedure for discovering and estimating functional economic relationships rather than as genuine 'tests' of an economic theory, which has led to criticisms of such research as 'crude empiricism' and 'cookbook econometrics'[5].

Boland (1988) and Morgan (1984; 1988) have shown that the terms "model" and "theory" in economics are not identical from a methodological viewpoint. Economic theories are derived from a set of behavioural assumptions about economic agents or institutions. Such theories are not directly testable. Only in conjunction with a further set of simplifying assumptions (for example, whether the relationship between the endogenous variables is linear or non-linear) can a testable model with

specific parameters be constructed and the relevant empirical data gathered. Numerous representative models may be constructed from the same basic theory so that each and every conceivable model must be tested and falsified before the underlying theory can be refuted. To illustrate the implications thereof, Boland (1988: 166-189) performed such an exercise for eleven simple aggregate demand models and calculated well over 100 000 possible testable versions for some of them. The more realistic the model the greater the number of endogenous variables there are, the more complex the relationships between them, and the less testable the model becomes.

The statistical nature of most relationships in economics compared to the deterministic laws of physics is, of course, a matter of degree as the development of quantum theory, the Schrödinger equation and Heisenberg's uncertainty principle all show. But even here there is an important qualitative difference in the way such probabilistic laws are expressed and the confidence we may have in them. With physical laws, even where expressed in stochastic form, it is usually possible to define an unambiguous chain of causation between the relevant variables whereas econometric relationships, even when statistically significant, can generally uncover no more than the presence of a *correlation* between the "dependent" and "independent" variables. For example, in testing the relationship between the exchange rate and relative national price indices it cannot be decisively ascertained whether changes in the price levels determine changes in the exchange rate, or vice versa, or whether they are jointly determined[6]. The test results could be consistent with any of these hypotheses. Moreover, the decision to accept or reject the relationship as statistically significant in the first place is again a matter of convention. The null hypothesis is rejected (and the alternate hypothesis accepted) if the likelihood of the observed correlation occurring by chance is less than some arbitrary probability (5% and 1% being the most commonly accepted borderline). Thus the relative bias towards committing a Type I error (rejecting a true null hypothesis) rather than a

Type II error (accepting a false null hypothesis) is also a matter of convention[7].

These concerns are not meant to imply that the empirical estimation and testing of econometric relationships should be eschewed. However, it does imply that genuine attempts at falsification may be inappropriate. Any equation in economics, even the iron "laws" of demand and supply, if pushed hard enough will yield falsifying test results. As suggested by Hutchison above, empirical economics cannot escape the use of inductive methods and is thus subject to all the problems of inductivism concerning the stability and reliability of results generalised from finite sample data. Although these problems are, of course, present in the natural sciences the difficulties are greatly magnified in a social science like economics. Blaug (1980: 256) observes that the bulk of empirical research in economics is much like "playing tennis with the net down". The answer to this complaint, as suggested above, is that we aren't playing tennis, that in many important respects economics is just a different ball game. This argument gains further support when we consider the role played by expectations in economic theories.

2.3 Expectations in economics

Expectations play a crucial theoretical role in modern economics. For example, in macroeconomics the significance of expectations is highlighted in the determination of unemployment, inflation, interest rates and exchange rates. Keynes (1936) recognized the significance of expectations in determining the level of desired investment spending in the economy and thereby influencing the level of aggregate demand, output and employment. However, he did not provide a theory of the determination of expectations and treated them as exogenous in his *General Theory of Employment, Interest and Money*.

There has been relatively little comment on the implications of expectations augmented theories for the methodology of economics (that is, methodology with a big 'M' as opposed to

econometric methods). Expectations are usually treated as an exogenous variable confined to the set of 'initial conditions' and background knowledge necessary for valid inferences to be drawn from empirical tests of a theory. However, the presence of a hypothesis about expectations in economic theories (especially in macroeconomics) constitutes a substantive difference to theories in the natural sciences. The expectations of economic agents, however they are modelled, have a significant independent effect on the relevant dependent variables in an economic theory. Where the dependent variables are determined in asset markets like the stock, bond, or foreign exchange markets such expectations are the dominant (perhaps even exclusive) explanatory variable.

The main problem in modelling expectations is that they are unobservable and their presence can only be inferred from *ex post* changes in economic variables. There are two main approaches to explaining the determination of expectations in modern economics. The adaptive expectations approach suggests that economic agents anticipate the future purely on the basis of past experience. For example, with adaptive expectations the expected inflation rate is determined by the past behaviour of the inflation rate, perhaps with recent experiences weighted more heavily. Thus economic agents adjust their expectations of the next period inflation rate according to the gap between the inflation rate expected in the previous period and the actual inflation rate in that period. If the actual inflation rate had been lower than expected, current expectations of inflation will be revised upwards and vice versa[8].

This approach has been criticised as an inadequate explanation of how economic agents learn from, and act upon, all the relevant information available to them. If such agents learn of, say, a link between budget deficits, money supply growth and the inflation rate, then such a link would presumably be used in forming expectations about inflation. Indeed, the mere announcement of changes in fiscal and monetary policy may be sufficient to trigger changes in the expected inflation rate. Adaptive expectations ignores this possibility whereas the efficient use of all relevant information is a central concern

of the rational expectations approach[9]. Under this approach, the collective wisdom of markets does not for long ignore observable systematic regularities and relationships in the economy. This does not necessarily mean that efficient markets have perfect foresight and that expectations are never mistaken; only that such markets do not err systematically in their forecasts.

In modelling the economy, rational expectations theorists take the efficient markets insight to its logical conclusion by deriving expectations from the endogenous variables and parameters of the model itself. Thus, economic agents are assumed to know the relevant model and base their expectations on it. This has the advantage of endogenizing expectations within an internally consistent model and thereby avoiding *ad hoc* assumptions about how expectations are formed. However, it implies that expectations differ depending upon which model or theory is chosen (for example, Keynesian or Monetarist versions of the economy). To the extent that there is no true model of the economy, neither can there be a true model of expectations.

Whichever hypothesis or model of expectations is used, a difficulty related to the Duhem-Quine thesis may arise in testing economic theories. For example, we may want to know what determines changes to the exchange rate. The chosen model may include fundamental variables such as net trade flows, interest rate and inflation rate differentials and would specify the hypothesized relationships between them. The model predicts the equilibrium exchange rate to change by x for given changes in the fundamental variables. However, today's (spot) exchange rate also depends on the expected exchange rate, since changes in the latter will rapidly produce a speculative adjustment in the spot exchange rate. Hence, the model must also include a hypothesis about how such expectations are determined. Therefore, tests of the model are necessarily tests of a joint hypothesis: If the observed exchange rate does not change as predicted this can be attributed *either* to faulty hypotheses about the fundamental variables, *or* to an incorrect model of expectations. Clearly, this is a special case of the Duhem-Quine thesis which is especially pertinent in economics

and possibly other social sciences as well, which renders even a diluted methodology of falsification impractical and inappropriate. Until an adequate theory of expectations is developed, which at this stage appears to be unobtainable, many significant theories in economics cannot be refuted. As noted above, the rational expectations approach does not resolve the problem. Although it advances our knowledge of how expectations are formed, by recognizing that economic agents do not rely solely on past experience, rational expectations theorists make the heroic assumption that such agents know the chosen, or "true", model and base their expectations on it. This approach fails if economic agents uniformly entertain a different model of how the economy actually functions, or if some groups entertain different models such that their expectations diverge (see, for example, Runde and Torr 1985).

2.4 Imre Lakatos and the methodology of scientific research programmes: soft focus falsification?

Lakatos, a colleague of Popper in the history and philosophy of science at the London School of Economics since the early 1960s, sought to rework Popper's methodology of falsification within a more liberal and tolerant methodological framework. He saw (Lakatos 1968; 1970; 1976; 1978) that the history of scientific progress in mathematics and physics did not conform to a strict methodology of falsification; that what with hindsight was accepted as good science and theory (for example, the development of quantum theory) had often survived amidst a sea of anomalies and experimental contradictions. To have insisted on severe testing and refutation of these theories at an early stage of their development would have halted the progress of science.

Lakatos' methodology of scientific research programmes (MSRP) redefines falsification and the appraisal of scientific theories. In his interpretation of the Duhem-Quine thesis, Lakatos suggested that test results are not used to appraise a particular hypothesis or even a single theory but an entire, interconnected and loosely consistent series of theories; a theoretical framework of ideas called a scientific research

programme. Each SRP may be identified by its "hard core" and a "protective belt". The former is the set of assumptions and metaphysical presuppositions that researchers in the field have agreed, by convention, not to question. The hard core is thus shielded by a negative heuristic. The protective belt contains the set of auxiliary assumptions and hypotheses which must bear the brunt of empirical test results and which, by a positive heuristic, may be adjusted, revised or even discarded accordingly.

How is falsification supposed to work in the context of SRPs? Lakatos proposes a version of *sophisticated* falsification whereby an existing theory is only discarded, no matter how damaging the empirical evidence, if it can be superceded by a superior theory:

"For the naive falsificationist a theory is falsified by a (fortified) 'observational' statement which conflicts with it (or which he decides to interpret as conflicting with it). For the sophisticated falsificationist a scientific theory T is falsified if and only if another theory T' has been proposed with the following characteristics: (1) T' has excess empirical content over T: that is, it predicts novel facts, that is facts improbable in the light of, or even forbidden, by T; (2) T' explains the previous success of T, that is, all the unrefuted content of T is included (within the limits of observational error) in the content of T'; and (3) some of the excess content of T' is corroborated" (Lakatos 1970: 116).

The positive heuristic guides the modification of theories in the protective belt by the addition of auxiliary hypotheses and conceptual changes. This series of theories is 'theoretically progressive' if it is characterized by (1) and (2). If the excess content of each successive theory is corroborated then the series is also 'empirically progressive'. New theories in the sequence must at least be theoretically progressive to qualify as scientific rather than *ad hoc* and pseudo-scientific. Prediction of novel facts, in contrast to Popper's openness to falsification, is Lakatos' demarcation criterion between scientific and pseudo-scientific theories.

An SRP which advances our knowledge by meeting all three criteria is called "progressive" while a "degenerating" SRP routinely accounts for serious anomalies in an *ad hoc* manner,

without producing any excess content or novel facts. Thus, the demarcation between science and pseudo-science becomes a matter of adopting a progressive rival SRP over a degenerating one. Lakatos maintained further that the history of physics can be rationally reconstructed in these terms whereby scientists do, in fact, eventually adopt progressive over degenerating SRPs; and that other methodologies such as justificationism, conventionalism and Popper's version of falsificationism are refuted by the history of science (Lakatos 1978: 102-138).

For Lakatos the process by which a degenerating SRP is eventually superceded by a progressive rival SRP is a protracted affair which may be drawn out over "decades" and which may be characterised by many inconclusive battles before the rival emerges victorious - Lakatos was critical of naive falsificationism with its methodology of "instant rationalism" and instant learning. He noted that a discredited SRP which has been overtaken by a progressive rival may even stage a successful comeback at a later date. Thus, in contrast to naive falsification, crucial experiments are rarely seen as such at the time. Only with long hindsight are particular experiments regarded as "crucial" from the perspective of a rival SRP which has superceded a degenerating research programme.

In this brief sketch of MSRP, both similarities and sharp contrasts with Popper's methodological falsificationism are apparent. Both methodologies embrace conventionalism by recognizing that the refutation of theories is a matter of deciding how to interpret test results. Assume, as a naive falsificationist, that Popper's first requirement has been met, so that empirical criteria of refutation have been set out, and the results of an experiment recorded. Before the relevant theory can be either falsified or corroborated there are four main types of decisions that must be taken:

i) Decide which 'observations' or 'basic statements' are acceptable in the light of fallible experimental theories. All test results depend upon existing observational theories. For example, Galileo's discovery of mountains on the moon depended upon the optical theory of the telescope - which was hotly

disputed by his critics (Lakatos 1970: 98). The decision to accept such statements allows them to become part of the "unproblematic background knowledge".

ii) Decide on the truth of such statements given the possibility of human error, faulty instrumentation and measurement, freak events and the like (see endnote 3). In the natural sciences, this can normally be achieved by conducting a control experiment.

iii) As regards probabilistic theories, decide on appropriate rejection rules which may be used to make the statistically interpreted evidence 'inconsistent' with the probabilistic theory.

iv) In testing the *conjunction* of a theory and its *ceteris paribus* clause, decide whether to accept a refutation also as a refutation of the specific theory alone. This also means deciding whether to relegate the *ceteris paribus* clause to the realm of 'unproblematic background knowledge'.

Lakatos' sophisticated version of falsification softens the implications of (iv), as explained below. For a naive falsificationist the conventionalist implications are more problematic, especially in a subject like economics (see section 2.2). Only if the exogenous variables which are believed to influence the endogenous variables of the model are specified, tested, and refuted can the *ceteris paribus* assumption be corroborated; and only then can the specific theory itself be refuted by a negative test result. The normal practice in empirical economic research, however, rarely meets, and is unable to meet, this requirement. Thus, most econometric practice ends up estimating the parameters of a model and its goodness-of-fit to the data along the lines of decision three above. In testing the model out-of-sample, corroboration of the *ceteris paribus* assumption is the exception rather than the rule, not because economists perversely ignore this crucial exercise but because it is usually not possible to do so given the subject matter of economics, particularly as regards changes in tastes and in expectations (see section 2.3). If the

ceteris paribus assumption cannot be corroborated, then negative test results need not be (and usually are not) interpreted as a refutation of the theory.

The methodological problem raised by the presence of *ceteris paribus* clauses applies as much to theories in physics and the other natural sciences as it does to theories in a social science like economics: the Duhem-Quine thesis attests to as much. However, theories in physics are subject to far fewer exogenous variables which *can* usually be specified and their possible influence on the model or theory tested. Indeed, Popper believed that although we may not be certain exactly which specific part of a theoretical system was challenged by an empirical counter-example, that it was possible, in such cases, to guess successfully which part was in fact responsible. He also gave a partial counter-argument to the applicability of the Duhem-Quine thesis in a rigorously axiomatized science like physics:

"It is possible in quite a few cases to find...which part, or group of hypotheses, was necessary for the derivation of the refuted prediction. The fact that logical dependencies may be discovered is established by the practice of *independence proofs* of axiomatized systems; proofs which show that certain axioms of an axiomatic system cannot be derived from the rest" (Popper 1963: 239).

Lakatos also borrows Popper's ideas concerning excess content, novel facts and corroboration and, in fact, makes them the cornerstone of his methodology, for it is only through the corroboration of a *new theory* that an established theory can be falsified. For Lakatos, in sharp contrast to Popper and naive versions of falsification, there can be no falsification before the emergence of a better theory - one which explains the unrefuted content of its predecessor, offers new predictions not possible (or, even better, forbidden) by the old theory, some of which are successfully confirmed:

"...it is the - rather rare - corroborating instances of the *excess* information which are the crucial ones; these receive all the attention. We are no longer interested in the thousands of trivial verifying instances nor in the hundreds of readily available anomalies: the few crucial *excess-verifying instances* are decisive" (Lakatos 1970: 121).

It is important to see that sophisticated falsification makes the fourth methodological decision required by naive versions of falsification redundant. It is no longer necessary to decide, either by informed guesswork or Popper's independence proofs, which part of a theoretical system is refuted by empirical counter-evidence. If a new, rival theory offers corroborated novel facts (and explains the unrefuted content of its predecessor) the whole complex of hypotheses constituting the older theory may be replaced, including the metaphysical presuppositions of its hard core if necessary. Thus the difficulties placed in the way of naive falsification by the troublesome *ceteris paribus* assumption are greatly reduced under the methodology of sophisticated falsification. Theories are no longer appraised in isolation but as a related cluster of theories, that is an SRP, which, if degenerating, may be superceded by a rival, progressive SRP.

Mention was made above of undesirable *ad hoc* changes to a theory. Lakatos (1970: 125, 142-3, 175) distinguishes three types of *ad hoc* modification: *ad hoc*$_1$ and *ad hoc*$_2$ are defined in a similar way to Popper's demand for new theories to have excess empirical content and to be independently testable. In Lakatos' terms a theory is *ad hoc*$_1$ if it does not lead to the prediction of any new facts. This would be the case if the theory were modified only to account for the empirical failures of its predecessor. In Lakatosian terms a problem shift is 'theoretically progressive' if it is not *ad hoc*$_1$. A problem shift that predicts new facts but none of which are corroborated is *ad hoc*$_2$. Only theories which are not *ad hoc*$_{1-2}$ also count as 'empirically progressive' theories.

Lakatos also distinguishes a third type of *ad hoc* adjustment which is not considered by Popper. A modified theory is *ad hoc*$_3$ if it does not accord with the spirit of the positive heuristic of the research programme of which it is part. The reason for disallowing *ad hoc*$_3$ modifications is to prevent the tacking on of auxiliary hypotheses to a theory which detract from the organic unity and continuity of scientific research programmes. Such modifications may be both theoretically and empirically progressive but are undesirable from this point of view and

cannot be excluded by Popperian notions of ad hocness alone. Lakatos' definition of *ad hoc*$_3$ thus constitutes a third way in which a modified theory can be progressive: if the modification is in accordance with hard core presuppositions or mechanisms the theory is heuristically progressive or non-*ad hoc*$_3$. As explained in section 2.5 *ad hoc*$_3$ concerns are at least as significant in economics as *ad hoc*$_{1-2}$ considerations.

Some of the concepts developed by Lakatos are similar to those of Thomas Kuhn, another philosopher in the growth of knowledge tradition. Archibald (1979: 304), for example, suggests that MSRP may be thought of as a combination of the ideas of Popper and Kuhn. Although a close examination or critique of Kuhn's work is not attempted here, it is worthwhile highlighting the main methodological differences between them, if only because Kuhn's ideas have also been applied to economics. Just as we talk easily about 'research programmes' in economics (even if we haven't read Lakatos) so has economic discourse picked up Kuhnian terms such as 'paradigms', 'scientific revolutions' and 'counter-revolutions'. It is helpful to understand the basic Lakatosian critique of Kuhn in the context of rationality and scientific progress.

In his reconstruction of the history of science (again mainly physics and chemistry) Kuhn (1970) discerned long periods dominated by a particular theoretical framework or 'paradigm', a concept which resembles the Lakatosian hard core of an SRP. Workaday researchers accept the authority of this paradigm unquestioningly and devote themselves to empirical and theoretical puzzle-solving, that is, applying the paradigm to resolve the many unexplained anomalies and inconsistencies that inevitably arise. Kuhn calls this activity 'normal science'. If these anomalies and inconsistencies resist explanation and proliferate, normal science breaks down and an intellectual crisis develops, heralding the onset of extraordinary or 'revolutionary science' during which a new paradigm, incommensurable with its predecessor, establishes its authority and a new phase of normal science. 'Normal science', 'crisis' and 'revolutionary science' correspond, although not perfectly, with Lakatos' explanations of 'positive heuristic',

'degenerating research programme' and 'problemshifts' respectively.

Kuhn stressed the importance of normal science for the growth of knowledge and, like Lakatos, criticised naive falsification with its implication of instant rationality. Both Kuhn and Lakatos sought to explain the apparent continuity of science and the tenacity of scientific theories. However, Kuhn eschewed prescriptive methodological rules and gave a purely descriptive account of what scientists do, claiming from his reconstruction of the history of science that falsification in any form was a chimera. For Kuhn there is no logic of discovery but only a social psychology of discovery. There are no rational grounds for converting from one paradigm to the next, nor any rational explanation for slavish submission to the authority of a paradigm during periods of normal science. This contrasts sharply with both Popper and Lakatos who may be regarded as critical rationalists. As noted above, Lakatos suggested that scientists follow (unknowingly perhaps - more of this in section 2.5) a sophisticated version of falsification whereby historically they have eventually adopted progressive over degenerating SRPs.

2.5 MSRP and the methodology of economics

Lakatos' softened version of falsification provides an attractive methodological framework for appraising theories in economics. As argued by de Marchi (1991), economists' self-image corresponds closely to that of the "good scientist" following the strictures of MSRP. Particularly attractive to economists is the idea that science may progress by the proliferation of rival theories without first refuting individual theories in the series. Any new theory if it predicts novel facts, some of which are corroborated, and which accounts for the successes of a rival theory may be permitted by MSRP. This contrasts with the linear approach of naive falsification where theories must first be refuted by experimental evidence; and helps explain the tendency in economics, given the difficulties in testing theories, to rely

heavily on abstract model building and innovative mathematical techniques.

Various researchers have applied MSRP as a conceptual framework for evaluating research programmes and methodological issues in economics. Lakatos' ideas were brought to the attention of economists in the collection of papers edited by Latsis (1976). In this volume, for example, Latsis identified the cluster of neoclassical theories of the firm as a research programme in economics, characterized by the analytical technique which he called situational determinism, in which the behaviour of firms is constrained by 'single exit' or 'straight jacket' solutions; and de Marchi gives a masterly analysis of the Leontief paradox in international trade. In a generally dismissive review of the Latsis volume, where he seriously questioned the applicability of MSRP to economics, Archibald singles out de Marchi as an exception:

"Not only does de Marchi make sense, but I doubt if the same sense could be made within another methodological programme...the story lends itself admirably to Lakatosian reconstruction" (Archibald 1979: 306).

Despite the criticisms of Archibald (1979) and Agassi (1979) amongst others, MSRP has continued to be applied to various areas of economics. Wong (1978) applied MSRP to a critical analysis of Samuelson's revealed preference theory; and Blaug (1980) discusses nine cameo SRPs, including the Heckscher-Ohlin theory of international trade, with an eye to assessing how well they have matched up to Popper's falsificationist criteria. Further examples include Cross (1982), and Backhouse (1991) and Weintraub (1985a, 1985b) as applied to monetarism and general equilibrium analysis respectively. In surveying the use of MSRP to evaluate research programmes in economics Fulton (1984: 204) concluded that "The MSRP employed in the proper context and in the proper way will provide a more than useful method of historical appraisal". At a later stage de Marchi (1991: 1-2) commented that "economists' self-image is extraordinarily close to the ideal portrayed in MSRP" and that "we have been shown, with the help of Lakatosian questions, that there are problems in reconciling economists' self-image

with the reality of economic practice, but also that, even within the set of Lakatosian concerns, there is much about economic practice that has not begun to be explored".

The following criticisms of MSRP are made in the context of its application to economics. Although there is considerable overlap with the natural sciences, some of these criticisms are especially acute, or even unique to economics as a social science. As suggested above, MSRP fits the self-image of economists and the language of MSRP has become part of everyday economic discourse. However, it is in the details and technicalities of such applications that some significant difficulties arise. Lakatos has been criticised for not giving sufficiently clear definitions of many of his key concepts and this has proved to be problematic in economics. For example, the distinction between "hard core" and "protective belt" has been difficult to apply in economics. Economists disagree on what the fundamental assumptions are supposed to be, even within a particular school of thought. Thus, Cross (1982: 330) points out that a proposition like "unemployment is independent of aggregate demand in the long-run" sounds like a hard core assumption which fundamentally defines a monetarist SRP. However, there is sufficient leeway in the term "long-run" and perhaps in the way "unemployment" is defined and measured to make it a dispensable auxiliary hypothesis within the protective belt. The distinction is important because it is only propositions within the protective belt that are susceptible to falsification. Cross suggests that hard core assumptions are only identified *ex post* such that there is no accepted negative heuristic protecting them *ex ante*. He recommends that the hard core/protective belt distinction in economics should be dropped altogether and that research programmes such as monetarism can be identified and appraised purely by the positive heuristic which guides the programme.

Without a clear cut definition of the hard core it has also proved difficult to demarcate RPs in economics. Researchers such as Archibald (1979) have regarded the whole of neoclassical economics as a research programme while others have attempted to demarcate narrower RPs such as human capital

theory or consumption theory, but still falling within the neoclassical tradition. This issue has important implications for the notion of falsification in the MSRP. The broader the RP the more entrenched it may become, and the more innocuous falsification becomes, since a thicker protective belt is available to shield it from damaging empirical evidence. The longevity of the RP is limited only by the ingenuity of its adherents in adjusting the auxiliary assumptions and hypotheses. For this reason the idea of a refutable research programme in economics may only be workable for a more narrowly defined area.

The problem of choice between theories is complicated by the degree of incommensurability that often exists between them (Kuhn 1970; Feyerabend 1962,1970). Critics of Lakatos feel that it is unduly restrictive to insist that a new theory should explain *all* the unrefuted successes of its established rivals. It is unlikely that the conceptual terms of reference of a new theory will exactly match those of its competitors. Thus there is an inescapable element of judgement in deciding whether the novel facts predicted and corroborated by the newcomer outweigh its failure to explain all the successes of its rivals. To some extent, therefore, the notion that theories or RPs are direct competitors may be inappropriate in the same sense that a race between a sprinter and a middle distance runner would not satisfactorily decide who is the better athlete. It would all depend upon the distance selected for the race.

The point here is that Lakatos' soft focus version of falsification may not prevent indefinitely self-perpetuating research programmes, at least in a social science like economics. Classical/neoclassical economics has survived as the orthodox master RP in economics since the time of Adam Smith. On the one hand it may be argued that the most serious threat to neoclassical economics this century, Keynesian economics, has in retrospect been more or less subsumed by the former albeit with some major additions and adjustments in the protective belt as regards sticky prices, imperfect information, uncertainty and expectations; or it may be argued that Keynesian economics has simply evolved as a new research

programme, incommensurable with the neoclassical programme, with these diametrically opposing assumptions at its core. Marxian and radical economics has not, at least in Lakatosian terms, produced any corroborated novel facts which could make it a contender for the mainstream orthodoxy. Indeed, most Marxist economists have been too busy explaining away the crippling anomalies and counter-evidence to their historical, class defined theories which currently bears all the hallmarks of a degenerating RP.

Nonetheless it can be argued that Marxian economics, despite clear signs of degeneracy, has not withered and died - it appears instead to be reinventing itself and settling for a period of coexistence with the mainstream neoclassical orthodoxy rather than an outright victory. Other RPs, as elucidated by the school of institutional economics, the Austrian school and Simon's behavioural approach similarly seem destined to coexist without ever seriously threatening its rivals or being threatened by them. Lakatos certainly wanted his more liberal methodology to allow time for budding research programmes to develop without being cut down prematurely by naive forms of falsification. However, it is unlikely that he would want to countenance the indefinite time allowed for most RPs in economics to bud and proliferate without severe pruning and the removal of some choking weeds. These observations are not intended as a complaint against the plurality of non-competing research programmes in economics: given its subject matter such pluralism is to be expected. It does mean, however, that MSRP does not fit the intellectual history of economics as well, or in the same way, that it does the natural sciences.

The prediction of novel facts, as we have seen, is a central requirement of good theories by MSRP. However, the historical time series nature of data and empirical evidence in economics means that new theories are often valued for their ability to *post*dict well known but unexplained (or inadequately explained) phenomena or what are sometimes called stylized facts. For example, the efficient markets hypothesis and asset market theory helped to explain the observation that floating exchange rates have tended to follow a random walk, an observation that

could not easily be explained by the older flow theories of exchange rates. Keynes' *General Theory* also did not predict any novel facts - rather its attraction lay in explaining, in a theoretically consistent way, the stylized fact of involuntary unemployment which the Classical theory was unable to do (see also section 2.7 below). Thus in economics, old theories are not always falsified by new theories that predict novel facts. Instead, they may be replaced by new theories that are more powerful in explaining the common knowledge of the past. It is often sagely commented in economics that hindsight is an exact science. However, the above account of theoretical explanation in economics suggests that economists do not, unfortunately, even enjoy this small luxury!

These difficulties all seem to point to a central concern: what guidance or advice can MSRP give practicing economists, saddled as they are most of the time with negative test results and *prima facie* empirically refuted theories? This concern arises from a marked ambivalence in Lakatos' work: his belief that MSRP is corroborated in its rational reconstructions of the history of science, yet his reluctance to derive any prescriptive rules or guidance to those working within a particular research programme. Budding research programmes are to be treated leniently (Lakatos 1970: 157 and 179). How leniently? Editors of scientific journals should refuse to publish papers of researchers in a degenerating research programme. But Lakatos concedes that such programmes cannot be written off and may even stage a triumphant counter-attack. The best he can come up with is:

"When a scientific school degenerates into pseudoscience, it may be worthwhile to force a methodological debate in the hope that working scientists will learn more from it than philosophers" (Lakatos 1978: 137).

Well, the Queensbury Rules are unlikely to be observed when a research programme is engaged in a street fight for its very survival. It appears that the critical rationalism of Lakatos is also, curiously, blind rationalism. Historically things have worked out: progressive research programmes are seen with hindsight to supersede degenerating research programmes (in

physics at least; given the problems referred to above this may not strictly apply in economics). However, researchers in the field do not and, it seems, need not, recognise the wisdom distilled by Lakatos from the history of science. Science is what science does; and scientists do what scientists must for the survival of their programme.

2.6 Laudan's problem solving approach to scientific progress

A central theme running through both the Popperian and Lakatosian versions of falsification is that progress in science only takes place after certain empirical criteria have been met. The strong empirical demands of Popper's three requirements (see section 2.1) appear to be softened somewhat by Lakatos' distinction between empirically and theoretically progressive problem shifts - but as pointed out by Shearmur (1991), since this theoretical progress consists of making novel empirical predictions (while at the same time preserving all the empirical successes of the predecessor theory) the bedrock empirical criteria in appraising theoretical developments remains intact. At the same time the metaphysical suppositions of the hard core are held immune from criticism by the negative heuristic of the research programme. Thus at a fundamental level both Popper and Lakatos are firmly rooted in the empiricist tradition in their view of science and scientific progress.

As suggested in the brief critique of their views earlier in this chapter it is this strong empirical bias which has proved especially problematic in the application of their ideas to the methodology of economics (and, presumably, the other social sciences as well). Despite these problems, however, the predominant methodological precepts in economics and self-image of economists remains very much in the Popper-Lakatos mould (see, for instance, the volumes edited by de Marchi 1988 and de Marchi and Blaug 1991). With the above problems in mind it is perhaps surprising that economists have not much heeded the significant contributions by Larry Laudan to epistemology and specifically methodological issues. Laudan's ideas remove a

large part of the empirical burden placed on the scientific acceptability of theories and emphasize the significance of purely conceptual advances in science: they are thus of particular relevance to economics as these are exactly the basic weaknesses of the received views on the methodology of economics.

Laudan's approach is set out in his book *Progress and Its Problems* (Laudan 1977) and is elaborated further in a more recent volume *Beyond Positivism and Relativism* (Laudan 1996). Like Popper and Lakatos, Laudan's ideas are derived from the history and philosophy of the natural sciences. However, as suggested above, his model of scientific progress appears to be more generally suitable for application to other disciplines, including the social sciences. Indeed Laudan is sceptical of attempts to demarcate science from non-scientific endeavour, previous attempts at which he calls "an unqualified failure" (Laudan 1996: 85-6). In the following precis of Laudan's model only the most pertinent features for the methodology of economics are addressed - a full evaluation of his ideas and their implications is beyond the scope of this thesis. Subsequent chapters, where appropriate, give some indication of the possible relevance of Laudan's approach to economics with reference to theories of exchange rates and the open economy.

Laudan's basic premise is that the aim of science is problem solving and that science progresses when new theories solve more problems than their predecessors. Although this seems to be a somewhat obvious anodyne, the emphasis on problem solving has important consequences for methodology. Firstly, it avoids conflating progress with a transcendent goal like "truth" (or successive approximations to the truth). It may well be that theories with greater problem solving effectiveness are nearer to the truth than those with lesser such capabilities but this is not a necessary condition for the more pragmatic aim of scientific progress. Secondly, the problem solving approach allows for a broader range of responses to theories other than the conventional ones of "acceptance" or "rejection" thereof. Laudan makes an important distinction between the rationality of acceptance and the rationality of pursuit - in its early

stages a theory may be unacceptable on various grounds but depending on its rate of progress in solving significant problems may warrant further pursuit[10]. This distinction helps remove the dichotomy in alternative methodologies between scientific progress, which is temporal, and criteria of rationality, which are atemporal: Laudan reverses the reasoning that progress depends on successive rational theory choices by asserting instead that rationality depends on progress in choosing those theories with greater problem solving effectiveness. And, maintains Laudan, objective criteria exist for determining when such progress has occurred.

Laudan distinguishes two types of problems: empirical and conceptual. Empirical problems are regarded as first order problems in that, although they usually arise within a theoretical context of inquiry (and are thus nonveridical), they are primary questions about empirical phenomena. Conceptual problems, in contrast, are second order problems that arise within or between theories concerning their intuitive and logical "well-foundedness". It is clear from Laudan's work that he regards the resolution of conceptual problems to be at least as important as the solving of empirical problems in the progress of science; and that traditional empiricist accounts of the history and philosophy of science have either ignored or misconceived the significance of such problems.

Solving an empirical problem is not, according to Laudan, the same as explaining a fact. Many facts about the world are unknown and thus do not constitute an empirical problem, that is, facts have to be discovered before they can conceivably generate concern. Moreover to count as an empirical problem, beyond an idle curiosity, there must be felt a premium on solving it - since the number of facts is infinite, empirical problems are a subset of nontrivial facts. Some genuine problems may even turn out to be counterfactual (for example, the attempt by early medical theorists to explain the "fact" that bloodletting cured some illnesses).

Laudan classifies three types of empirical problems according to the role they play in theory evaluation. Unsolved problems are those which no existing theory has been able to solve whereas anomalous problems are those which a particular theory has not solved but which one or more competing theories has done so. Solved problems are those empirical problems which a theory has been able to solve successfully. Thus one of the signs of scientific progress is the transformation of anomalous and unsolved problems into solved ones. Unsolved problems do not represent as much of a cognitive threat to a theory as anomalous problems - only those unsolved problems which become the solved problems of a rival theory in an alternative research tradition pose a serious threat.

Laudan's interpretation of anomaly contrasts with the conventional methodologies of both confirmation and falsification. He argues that anomalies may raise doubts about a particular theory but need not compel theoreticians to abandon it. His justification for this is uncontroversial, relying as it does on the insights of the Duhem-Quine thesis explained earlier in this chapter. However, Laudan further maintains that some facts may pose an anomalous problem for a theory even if they are consistent with that theory[11]:

"Such situations arise when a theory in some field or domain fails to say anything about a kind of problem which other theories in the same domain have already solved"
(Laudan 1977: 28).

For example, Laudan points out that Newton's theory of planetary motion could not explain why all the planets move in the same direction around the sun whereas those of Kepler and the Cartesian astronomers could do so. The point is not that Newton's theory made a false prediction about the direction of orbit (any direction would be consistent with it) but that it remained altogether silent on the problem.

As regards conceptual problems Laudan shows that many of the key debates in the history and philosophy of science have concerned non-empirical issues and not only during the early stages in the development of new theories. Empiricist

methodologies, including those of Popper and Lakatos, are criticised for their failure to explain the conceptual nature of theory choices in the many historical examples where the empirical problem solving abilities of competing theories have been almost the same. Laudan (1977: 49) suggests that conceptual problems can arise in two ways:

"1. When T exhibits certain internal inconsistencies, or when its basic categories of analysis are vague and unclear; these are internal conceptual problems.

2. When T is in conflict with another theory or doctrine, T', which proponents of T believe to be rationally well founded; these are external conceptual problems."

Laudan (1996: 79) adds two further possibilities:

"3. When T violates principles of the research tradition of which it is part...

4. When T fails to utilize concepts from other, more general theories to which it should be logically subordinate".

In his later work Laudan broadens the second clause to allow for the possibility of conceptual conflict when T "makes claims about the world which cannot be warranted by prevailing epistemic and methodological doctrines" (Laudan 1996: 79).

Although the removal of internal ambiguities, circular reasoning, and logical inconsistencies through a process of conceptual clarification is important for the progress of science, Laudan places greater emphasis on external conceptual problems. Conceptual conflict or "tension" between theories can arise in three ways: logical inconsistency; joint implausibility; and lack of theoretical reinforcement. Laudan uses examples from the natural sciences to illustrate these difficulties. Taking some cursory examples from economics instead, the initial conflict between utility and cost of production theories of value seems to be a good illustration of an apparent logical inconsistency between two rival theories - which was only satisfactorily resolved by Marshall's marginal supply and demand analysis and synthesis. Marxian and Keynesian theories of unemployment, on the other hand, may be thought of as jointly implausible in the sense that although Marxian

theory does not deny that deficient aggregate demand may lead to unemployment (and is thus on this score compatible with Keynesian economics) its explanation of the reasons for deficient demand are quite different to those of Keynesians. Thus the acceptance of either theory lessens the plausibility of the other. The reinforcement thesis rests on the perceived organic unity of science. It would be problematic, for example, if a neurologist did not use electrochemical theory to explain brain functions even if this were possible and did not actually conflict with the latter. Similarly a conceptual problem would arise in different branches of neo-classical economics if, say, a theory of the labour market ignored the accepted elementary principles of price theory.

The tensions between two or more theories result in different degrees of cognitive threat. Laudan (1977: 54) suggests a taxonomy of cognitive relationships between theories as follows:

"1 *Entailment* - one theory, T, entails another theory, T_1.
2 *Reinforcement* - T provides a "rationale" for (a part of) T_1.
3 *Compatibility* - T entails nothing about T_1.
4 *Implausibility* - T entails that (a part of) T1 is unlikely.
5 *Inconsistency* - T entails the negation of (a part of) T1"

As is self-evident from this taxonomy, the degree of cognitive threat increases from 2 to 5.

Not all pairings of different theories qualify as a source of external conceptual problems - otherwise it would be possible to create an artificial conflict between a scientific theory and any arbitrary belief. Laudan considers three classes of difficulties which may lead to external conceptual problems. Intra-scientific difficulties occur when a new theory includes presuppositions which are incompatible with an independently well-established and accepted scientific theory. The decision to reject it must usually entail an obligation to provide an adequate supporting theoretical framework for the victor. For this reason the outright rejection of an established theory (as opposed to a heightened pursuit of the rival theory) may be a protracted affair which, like the emergence of a victorious

Lakatosian research programme, only becomes evident with hindsight.

Laudan notes that intra-scientific conflict raises presumptive doubts about both theories. This may ultimately lead to the rejection of one of them but there are circumstances where both theories may be retained. For example, the nineteenth century conflict between uniformitarian geology and evolutionary biology on the one hand (both theories supported by a considerable accumulation of evidence), and the laws of thermodynamics on the other (which was more recent but had solved a number of important physical problems), was resolved by the discovery of radioactivity which allowed all three theories to be retained. It would appear that this type of situation is quite common in economics where many theories begin by providing what is in retrospect only a partial explanation for a given phenomenon only to be subsumed (as opposed to being totally rejected) by a later more comprehensive theory. This would apply, for example, to the early elasticities and purchasing power parity theories of exchange rates and to the prototype monetary models of the early 1970s. All these partial explanations have at different stages been subsumed by more comprehensive macroeconomic theories - but may yet be regarded as valid over their relevant range of application (see chapters three and four).

Empiricist methodologies like those of Popper and Lakatos are often used as normative metatheoretical criteria to judge the rationality and progress of science. Laudan makes the important point that such philosophers have ironically failed to recognize that conflict between a theory and accepted methodological norms (which have not been permanent historically) itself plays a part in the appraisal of such theories. When a promising new set of theories or research tradition (see below) conflicts with prevailing methodological norms this poses acute conceptual problems for both the theories and the norms. The eventual outcome of this conflict may well be that it is the prevailing methodology that is rejected and replaced, rather than the theory, or perhaps a modification of both. It cannot be presumed that it is always

the theory that must give way to an unalterable methodological canon.

In economics this situation brings to mind the *Methodenstreit* in the late nineteenth century between the inductivist German Historical school (represented by Gustav Schmoller) and the deductive method of the emerging marginal utility theory of value (defended by Carl Menger). A similar conflict between deductivism and positivism took place during the 1930s in the classic debates between Robbins and Hutchison and thereafter between Hutchison and Machlup. The highly self-conscious nature of this intense debate, and the consequences for both theory and method in economics, clearly support Laudan's thesis.

Laudan further suggests a third source of external conceptual problems, conflict with widely accepted non-scientific beliefs or worldview including metaphysics, ethics, and theology. A good contemporary example of this is the conflict between scientific theories which hypothesize possible racial differences in intelligence and prevailing egalitarian social and political norms. Again the problems are symmetrical - Laudan does not claim that scientific theories should necessarily be rejected whenever they conflict with the worldview or vice versa: he merely notes that such tensions have and do exist and that the seriousness of the problem for the theory depends on how entrenched these non-scientific beliefs are and the problem solving potential that would be lost by abandoning the theory.

From the above discussion of problem solving it is evident that, in contrast to empiricist methodologies, Laudan sees scientific progress as significantly more than a succession of theories with ever increasing empirical content. Laudan suggests that scientific progress and growth depends upon maximizing the number and importance of solved empirical problems while minimizing the number and importance of anomalies and conceptual problems. His "problem effectiveness" demand of scientific theories and view of scientific progress is thus much richer and broader than the narrow empiricist criteria put forward by Popper and Lakatos:

"If it counts in favor of a theory when it can accumulate solved empirical problems (as the standard view allows), then it should also count *against* a theory if it generates anomalous and conceptual problems. *Indeed, the problem-solving effectiveness of a theory depends on the balance it strikes between its solved problems and its unresolved problems*" (Laudan 1977: 67).

Thus scientific progress can occur in the conventional way of solving more empirical problems. But progress can also take place without any increase in the number of solved empirical problems and, indeed, even when the domain of such problems contracts if this is accompanied by fewer anomalies or conceptual problems. Hence the empirical requirement of "saving the phenomena" is not a necessary condition for progress. Moreover some theory changes may be regressive even when the scope of solved empirical problems is increased if the changes create more acute anomalies or conceptual problems for the new theory compared to its predecessor.

Like Lakatos' research programmes and Kuhn's paradigms, Laudan also identifies theory complexes as distinct from specific theoretical hypotheses. Laudan's concept of a research tradition is similar in some respects to a Lakatosian research programme but there are also important differences. The brief treatment of these differences here is justified as the methodological analysis in subsequent chapters does not rely much on the peculiarities of Laudan's research tradition approach *vis-a-vis* that of Lakatos or Kuhn.

Laudan's definition of a research tradition is "*a set of general assumptions about the entities and processes in a domain of study, and about the appropriate methods to be used for investigating the problems and constructing the theories in that domain*" (Laudan 1977: 81). Like an RP a research tradition has a hard core set of metaphysical and methodological commitments which distinguish it from others, and contains specific archetypal theories which exemplify it. Laudan also sees research traditions as enduring for long historical eras, unlike its individual theories which are usually transient. However, unlike Lakatos, Laudan emphasizes that research traditions do not have an immutable hard core: its

presuppositions undergo different stages of development and reformulation. For example, the hard core essence of the early Newtonian research tradition was quite different to the later stages thereof. Similarly the essence of the Marxist research tradition in the late nineteenth century was substantially different to what are now regarded as the essential basic assumptions of Marxism.

Research traditions are regarded by Laudan as a set of guidelines for the development of specific theories which "reduce" all empirical problems in the relevant domain to the entities and metaphysical dictates of the research tradition. However, these guidelines do not logically entail the specific theories or hypotheses of which it is comprised. In contrast to Lakatos, Laudan's approach allows for mutually inconsistent theories to claim allegiance to the same research tradition - and different research traditions can give metaphysical support to a given theory. This approach seems to fit economics better than a Lakatosian RP which permits two theories to be in the same RP only if one of the theories entails the other. As noted in section 2.5 there is a great difficulty in clearly identifying and demarcating RPs in economics. For example, are Keynesian and monetarist theories really different RPs or do they both fall under the broader mantle of neoclassical economics? To some degree they are mutually inconsistent theories but at a more fundamental level they share neoclassical abstractions about basic economic processes like equilibrium, supply and demand, the unit of analysis and level of aggregation, maximization or minimization of utilities and costs and so on. Laudan's approach would naturally allow both Keynesian and monetarist theories as part of the same neoclassical research tradition. In contrast, Marxist-radical schools of economic thought and the institutionalist school would comprise alternative research traditions because of their distinctively different presuppositional base.

It is clear from the above account of research traditions that, unlike individual theories, they do not imply specific empirical predictions and are not directly testable. Thus appraising the success of a research programme, as opposed to

its truth or falsity, depends crucially on the range of empirical and conceptual problems its constituent theories have been able to solve. The more such problems that are solved the more progressive the research tradition.

2.7 Appraisal of Laudan's model and its relevance to economics

Despite some significant drawbacks it is argued here that Laudan's problem solving approach to the questions of theory choice, rationality, and the progress of science, has a competitive advantage when applied to economics. The main reason for this is the recognition and explanation of the role played by conceptual problem solving in the progress of scientific theories. This contrasts sharply with the preoccupation of conventional methodologies such as those of Popper and Lakatos with more or less strict empirical criteria in this regard. By these standards the response to a theory which fails empirical tests is either outright rejection (Popper) or, in the case of a degenerating research programme, an ambiguous skirting of the issue (Lakatos). The attraction of Laudan's approach for economics is that progress may be achieved by solving significant conceptual problems as well as, or even independently of, any progress made in increasing the scope of its solved empirical problems. Also, theories may be held unacceptable because they fail empirical tests and face significant unsolved conceptual problems but may yet be worthy of rational pursuit. This would be the case if, for example, a new theory shows promise by a high rate of problem solving progress in its early stages (even though in its current state it cannot compete with a more established theory in the domain for general problem solving effectiveness).

Moreover, Laudan's model does not require new theories to entail all the empirically corroborated (or unrefuted) content of their predecessors. It thus fits the history of science (and of economic theories) more naturally, given the observation that different theories (or research programmes or paradigms) are often incommensurable (see section 2.5). However, Laudan's model also complicates the appraisal of theories by having to

count and weight the problem gains and losses from one theory to another. He does not give a complete example of this calculus for the complex of theories comprising a whole research tradition. In practice it may not be a simple matter to apply the mini-max procedure suggested by Laudan particularly when, in comparing different theories, it is necessary not only to simply count problem gains and losses but to weight their significance as well. To argue that a conceptual problem solving gain is more valuable or significant than a corresponding conceptual (or empirical) loss is unavoidably a matter of judgement and perception. Laudan (1977: 64-66) does give some clues as to how to weight conceptual problems but these are not entirely satisfactory. Still, if Laudan's reconstruction of the history of science and the role played by conceptual problem solving is even roughly accurate then this difficulty cannot be avoided - any alternative epistemology would also need to resolve it.

Because Laudan's model is concerned with problem solving rather than the assignment of truth or falsity to a theory and its auxiliary hypotheses, it largely avoids the ambiguity of empirical tests implied by the Duhem-Quine thesis which, as argued earlier, has proved especially problematic in economics. Empirical anomalies, by Laudan's account, do not falsify a theory but do raise doubts about its problem solving effectiveness (and of all the auxiliary hypotheses in the theory complex). By symmetrical reasoning empirical confirmation of a theory cannot, as Popper observed, logically tell us anything about the truth of a theory, but if a theory complex solves an empirical problem this counts as a solved problem for all its nonanalytic component parts. Instead of trying to isolate the specific hypothesis(es) that falsify (or confirm) a theory, Laudan suggests the opposite by generalizing blame or credit to all the member hypotheses of that theory. Thus, contrary to the position taken by some philosophers, Laudan maintains that it is possible to appraise a theory as long as this concerns its problem solving effectiveness and not questions about its truth or falsity status.

Laudan is also much softer than either Popper or Lakatos on those *ad hoc* modifications to a theory which solve only the unsolved empirical problems or refuting instances of its predecessor. Such *ad* hoc_{1-2} changes are condemned by falsificationists but are welcomed by Laudan as "perfectly consistent with the general aim of increasing our problem solving capacities. *Ad hoc* modifications, by their very definition, are empirically progressive" (Laudan 1977: 115). He notes that it would certainly be a bonus if such theory changes could also solve new problems as demanded by Popper and Lakatos, but to insist that a theory which fails these criteria is unacceptable is counterproductive - if taken seriously it would foil most scientific progress. The history of science is replete with examples of theories which were *ad hoc* in this sense but were subsequently highly successful problem solvers.

However, Laudan is harder on theory modifications which are *ad hoc* in the sense of Lakatos' *ad* hoc_3: modifications that do not accord with the spirit of the research programme and violate the negative heuristic of the hard core. Laudan rephrases this by noting that a modified theory may create more conceptual problems than it solves existing empirical problems and thus reduce the overall problem solving effectiveness of the theory. Unlike the Lakatos approach, the problem solving model distinguishes between spurious and legitimate interpretations of *ad hoc*ness and provides a way (albeit flawed) of assessing the seriousness of the cognitive threat to a theory in each case.

The *ad hoc* debate is important in economics. Blaug's concern with innocuous falsification in economics is partly attributable to the common practice of *ad hoc* changes to a theory when it meets *prima facie* refuting instances or fails empirical tests. Few theories in economics, however, have passed the falsificationists acid test of independent testability by predicting novel facts, some of which at least must subsequently be corroborated (Popper's second and third requirements respectively). On this score many successful theories in economics could be dismissed, in the Lakatosian sense, as *ad* hoc_{1-2} - theories which could satisfactorily

resolve empirical difficulties or explain certain stylized facts which were problematic for an established rival but which failed to predict any new or unexpected facts (see below). Laudan's approach thus rationalizes what is common practice in economics and avoids the rather bizarre conclusion for falsificationists that most of what economists are doing most of the time is unacceptable and even irrational. Economists may quite rationally prefer one theory to another on non-empirical conceptual grounds. In this case they may persist with the theory even if it appears to have nothing extra to offer empirically. Given the huge empirical difficulties in testing economic theories explained earlier in this chapter it would be quite surprising, even perverse, if economists took falsification seriously.

Like Lakatos, however, Laudan is strict on problem shifts that are *ad hoc*$_3$. The desire for theories that are not *ad hoc*$_3$ is consistent with Laudan's requirement that theories should minimize the number of conceptual problems they face in their domain. Hands (1988) shows that *ad hoc*$_3$ manoeuvres *are* taken much more seriously as a criticism of economic theories. The reason for this is that such modifications may lead to conceptual or analytical problems which threaten the theoretical integrity of the research programme. Hands gives some convincing examples of economists' aversion to such modifications from the New Classical Macroeconomics, general equilibrium theory, and Keynesian economics and concludes:

"Thus, for economic theorists, the sin of ad hocness seems to be infidelity to the metaphysical presuppositions of the neoclassical programme rather than face-saving adjustments in response to recalcitrant data. Of course, this is not to say that economic theorists do not actually adjust their theories in an ad hoc1-2 manner, only that recent theorists seem to consider ad hoc3ness to be a more damning criticism" (Hands 1988: 132).

Subsequent chapters flesh out some of these issues with regard to theories of exchange rates and the open economy. However, it is perhaps worthwhile illustrating the more general relevance of Laudan's approach to economics (and the unsuitability of falsification) by interpreting the success of the Keynesian revolution in problem solving terms. The following brief sketch

is highly selective and is only intended as a suggestive application thereof - a more thorough analysis and reconstruction is beyond the scope of this thesis.

Much, of course, has been written about what was really different about Keynes' theories compared to those of the classical economists. Many post-Keynesians have argued that what is routinely passed off as distinctively Keynesian theory is a travesty of what Keynes really said or tacitly had in mind in writing the *General Theory*, and have tried to pin down the true basis of his theoretical insights - as attempted, for example, in the book by Axel Leijonhufvud pithily titled *Keynesian Economics and the Economics of Keynes* and the reappraisal of the Keynesian "counter-revolution" by Robert Clower. Such research has focused on different areas of Keynes' work: the principle of effective demand (Clower's dual decision hypothesis); expectations and uncertainty (the point of departure taken by Joan Robinson and GLS Shackle from the opening section of Keynes' restatement of his General Theory of Employment in the 1937 edition of the *Quarterly Journal of Economics*); or on the Hicks-Hansen income-expenditure version of the *General Theory* (labelled by Coddington "hydraulic Keynesianism"). It is clear from a survey of this literature that its primary emphasis is on clarifying exactly what conceptual problems Keynes was really grappling with and whether in fact his theories actually solved them. For example, the Keynesian assumption of downwardly rigid wages and prices may be regarded as an *ad hoc*$_3$ empirical hypothesis rather than the basis of a serious cognitive threat to neoclassical economics. The assumption of rigid wages and prices, it is argued by some, simply makes the Keynesian model a special case of neoclassical theory.

It is also evident that Keynes himself placed great emphasis on what he believed were logical inconsistencies and conceptual flaws in the classical theory and the superiority of his own theory in this regard. In chapter two of the *General Theory* Keynes makes explicit what he thinks are the fundamental postulates of classical macroeconomics as the basis for his attack on the explanation of output and employment derived

therefrom. It may be argued that the central message of Keynesian economics (see Blaug 1985: 670-2) results from making explicit and solving a formal logical error in the classical theory. In Laudan's terms progress was achieved by the invention of a modified theory which solved a serious internal conceptual problem for the classical system *besides* the more obvious empirical problem of involuntary unemployment. Keynes states his basic point of departure from the classical economists in chapter fourteen of the *General Theory* regarding the relationships between changes in the interest rate, saving, investment, and the level of income:

"The classical theory of the rate of interest seems to suppose that, if the demand curve for capital shifts or if the curve relating the rate of interest to the amounts saved out of a given income shifts or if both these curves shift, the new rate of interest will be given by the point of intersection of the new positions of the two curves. But this is a nonsense theory. For the assumption that income is constant is inconsistent with the assumption that these two curves can shift independently of one another. If either of them shift, then in general, income will change; with the result that the whole schematism based on the assumption of a given income breaks down"
(Keynes 1936: 179).

The fundamental logical inconsistency of the classical schema is reiterated throughout the chapter:

"But this is the point at which definite error creeps into the classical theory" (*ibid* 178).

"But, in fact, the classical theory not merely neglects the influence of changes in the level of income, but involves formal error" (*ibid* 179).

"Thus the traditional analysis is faulty because it has failed to isolate correctly the independent variables of the system" (*ibid* 183).

"The reader will readily appreciate that the problem here under discussion is a matter of the most fundamental theoretical significance and of overwhelming practical importance"
(*ibid* 184).

It would be very difficult to construe the type of progress Keynesian economics made over the classical system in terms of a falsificationist approach because it does not count purely conceptual advances as progress. Laudan's problem solving

model, however, fits our intuitions about progress in economics far more readily because it not only accepts but emphasizes the importance of conceptual problem solving, without requiring any concomitant change in empirical content.

One escape route for the falsificationists is to argue that Keynes' theory explained an empirical fact - that of involuntary unemployment - which the classical theory had failed to do and which instead tried to pretend, by various conventionalist strategems, that the high levels of unemployment observed during the Great Depression were in fact voluntary. But by their own insistence on independent testability Keynes' theory fails this test because it does not "lead to the prediction of phenomena which have not so far been observed" (Popper 1965: 241). In other words, Keynes' theory was successful not because it predicted a novel or unexpected fact but because it explained a known and much debated stylized fact in a theoretically consistent way.

The serious falsificationist would have to admit that by all accounts Keynes' theory was an *ad hoc* theory *par excellence* as its very rationale and great success was to explain the stylized fact of persistent involuntary unemployment. A stylized fact cannot, by definition, also be a novel fact. Given the importance of stylized facts (see also chapter five) as a testing ground for economic theories, the unsuitability of falsificationist criteria in the appraisal of such theories should be self-evident. By comparison, Laudan's model is far more direct with regard to empirical problems. As regards Keynes' theory, for example, the fact that it was seriously *ad* hoc_{1-2} does not count against it. As a theoretically consistent explanation of involuntary unemployment it had solved a weighty empirical problem and thus constituted progress over its rival.

The inconclusive debate over the significance of the high levels of unemployment during the Great Depression is also consistent with Laudan's approach. The classical theory could afford to ignore or explain away the problem of involuntary unemployment largely because there was no alternative theory thereof[12]. However, once Keynes and his followers had provided

such a theory the problem became far more serious - in Laudan's terms it changed from an unsolved problem to a more weighty anomalous problem. The Keynesian model of the economy, although it could not logically compel the abandonment of the classical theory, did place it in sufficient doubt on this score to replace it as the new mainstream economics.

Notes

1 A chronological note is usual here. Popper first published his *Logik der Forschung* in German in 1934. It was translated into English as *The Logic of Scientific Dicovery* and republished in 1959. The latter book contains various amendments and introduces new material such as Tarski's correspondence theory of truth and Popper's own concept of "verisimilitude". It was only after the 1959 publication that Popper's ideas became more accessible to a wider audience and began to gain ascendancy over logical positivism and logical empiricism.

2 In a methodological postscript to his magnum opus *Economic Theory in Retrospect*, Blaug (1985: 702-705) appears to take a less forthright position but remains reluctant to abandon falsification altogether. He concludes somewhat ambivalently in a footnote: "At any rate, it would be fair to say that the status of the falsifyability criterion in economics is about halfway between its status in psychoanalysis and its status in nuclear physics". This paper suggests that even this watered down compromise is invalid.

3 This is not merely idle speculation. Vernon Smith (1994) relates how the foundations of relativity theory in Einstein's famous paper "On the Electrodynamics of Moving Bodies" were "refuted" by the physicist Kaufman later the same year and in the same journal. Einstein rationalized the refutation by suggesting "a not yet recognized source of errors" that might become apparent after a further "diverse body of evidence becomes available". Kaufman's apparatus was later found to have been faulty.

4 Hutchison's earlier work (for example, Hutchison 1935; 1956) introduced positivist ideas to economists in a series of methodological debates with Robbins and, later, Machlup. This should be seen in the context of the deductivism that strongly inluenced thinking in the methodology of economics at that time.

5 Mayer (1980) strongly criticises the practice of applied econometrics and suggests improvements which could help economics achieve the status of "hard science".

6 There are some well known econometric tests that can determine the apparent direction of causation between two or more endogenous variables as, for example, in the tests by Granger (1969); Sims (1972); and Geweke, Meese and Dent (1982). These tests define causality temporally and test whether a change in X regularly preceeds a change in Y or vice versa (in the sense of a statistically significant regularity). However, these tests do not get around the difficulty that there may be a third variable (or more complex set thereof) which independently causes the changes in both X and Y. This has been

a problem in tests of PPP (see chapter three) where changes in excess national money supplies may affect both the domestic price level and the exchange rate independently. A more basic difficulty is that the observed relationship between X and Y could merely be a chance correlation for the particular time-series sample data. Such difficulties reflect the fundamental philosophical problem of induction common to all the sciences - see, for example, the discussion of Mill's methods of inductive inference in Copi (1972: 369-421) of which Mill's 'Method of Concomitant Variation' is especially pertinent to economics. Given these difficulties, statistical methods thus cannot be a substitute for *a priori* theoretical considerations in attempting to establish the "true" cause-and-effect relationships. A full examination of the nature of causality is beyond the scope of this thesis. In this regard perhaps it will suffice here to note that proponents of the unity of science thesis regard the perceived differences between the natural sciences and a social science like economics as simply a "matter of degree". The main point, however, is that even if these differences are only a "matter of degree" this may nevertheless be significant as regards the applicability of falsificationist criteria in economics.

7 Blaug (1980: 23) notes that, astonishingly, Popper ignored, in his later philosophical work, the highly relevant Pearson-Neyman work on statistical inference published between 1928 and 1935 which is completely dependent on methodological falsificationism.

8 The seminal work on adaptive expectations is Phillip Cagan's study of the relationship between hyperinflation and excess money supplies - see Cagan (1956).

9 The interested reader may consult the books by Redman (1992) and Sheffrin (1996) for a thorough, modern, and relatively non-mathematical review of the origins and conceptual development of the rational expectations hypothesis in economics.

10 As noted in section 2.4 Lakatos was also critical of the search for 'instant rationality' and his MSRP also allows for the rational pursuit of theories that initially may not be empirically progressive (although they would have to involve a problem shift that is at least theoretically progressive).

11 This idea is similar to a late amendment of MSRP by Lakatos where he broadens the concept of a theoretically progressive problem shift to include the explanation of known facts that are merely fortuitous in terms of earlier theories in addition to his original emphasis on the prediction of novel facts (see Lakatos 1978: 179-86).

12 See, for example, the classic arguments used by Benjamin and Kochin (1979) to explain British unemployment during the inter-war years as 'voluntary'.

CHAPTER THREE

THE PURCHASING POWER PARITY PUZZLE

The title of this chapter is taken from a recent paper by Rogoff (1996). Rogoff's concern was to evaluate the evidence from wide-ranging and intensive empirical studies against the hypothesis that deviations from purchasing power parity revert back to their PPP values - either in the short, medium or long-run. The main concern of this chapter, however, is to understand from a methodological perspective why PPP theory has not been falsified despite the accumulated weight of generally negative test results from most of the post Second World War era. In so doing, it also provides some historical and doctrinal background to the monetary approach to the balance of payments and exchange rates (which is the focus of chapters four and five), of which a version of PPP is usually an integral part. There is an extensive and growing literature on research into PPP and this chapter does not attempt an exhaustive survey thereof. Only those issues which are most pertinent to the methodology of economics have been selected for closer analysis and comment.

As made clear below, there is no single PPP theory of exchange rates but many different versions and interpretations of the basic concept. The meaning of PPP varies depending on the context of the problem and the type of question being asked. Partly because of this conceptual confusion and partly due to empirical problems in testing PPP, most test results have proved inconclusive despite almost a century of research in the area. The basic question whether it is possible or appropriate to formally "accept" or "reject" PPP, itself does not have a clear answer. It is uncertain exactly what would have been achieved, in the sense of progress in or growth of knowledge in economics, even if this was done.

Despite the mostly negative empirical evidence against PPP as a short-run relationship, economists have been reluctant to abandon the concept altogether and have persisted with the idea that PPP is valid when interpreted as a long-run equilibrium condition. By this interpretation PPP theory can only offer a partial explanation of exchange rates in the short-run so that other dynamics must be specified to account for such deviations from PPP. To this end, PPP is included as a necessary long-run equilibrium constraint within a broader structural model of exchange rate determination such as the monetary or asset market approaches. The methodological issues raised by this type of development, which can be further illustrated in other areas of economics, are noteworthy. Empirical work in exchange rate economics does not appear to be primarily concerned with proving or disproving theories. The outcome of most empirical research suggests that its role is to explore the applications and limitations of economic theories in particular contexts; and to help clarify and refine our understanding of the theoretical concepts themselves. To discard an intuitive organizing concept like PPP because negative test results "falsify" an empirical application thereof is to make the mistake of throwing the baby out with the bathwater.

Section 3.1 outlines the three main versions of PPP theory, while section 3.2 discusses its historical origins and doctrinal aspects, particularly its connection with the monetary approach. Section 3.3 examines more closely the conceptual confusion surrounding the meaning of PPP. Section 3.4 looks at how different models of PPP have been tested and how the test results have been interpreted, with close attention to the significance of deviations from PPP. Section 3.5 regards the implications of these issues for the methodology of economics. In particular the question whether falsification is actually practiced in this area of economic research is addressed and, if not, whether it should be taken more seriously.

3.1 A brief statement of purchasing power parity theory

The basic idea behind all theories of purchasing power parity is that there is a relationship between prices and exchange rates. This relationship may be expressed in different ways. For individual, internationally traded goods it may be interpreted as the law of one price, which simply states that in perfectly competitive markets with no arbitrage costs or risks the price of an identical good will be the same anywhere in the world, when expressed in a common currency. If this concept is extended to collections of goods, including non-tradables such as services, then it suggests a simple equilibrium relationship between national price levels such that:

$$SP^* = P \qquad (3a)$$

where **S** is the spot exchange rate[1] between two currencies and **P** and P^* are the domestic and foreign price levels respectively. The direction of causation is not necessarily specified. A change in **S** could cause a change in the domestic price level; or a change in relative price levels could cause an adjustment in the exchange rate; or price levels and the exchange rate could be interdependent. Equation (3a) is mostly used in comparing national price levels and real incomes of different countries.

A basic difficulty with equation (3a) is static price differentials resulting from barriers to trade and arbitrage costs such as transportation. The theory can be amended to avoid this difficulty by expressing the relationship in relative form - between the rates of change rather than the absolute levels of the variables concerned. In this chapter and subsequently rates of change of variables are written in lower case letters and the absolute levels thereof in upper case, unless otherwise stated. Equation (3a) may thus be rewritten as:

$$s + p^* = p \qquad (3b)$$

Thus, for example, a 10 percent depreciation in **s** means that either the domestic price level must rise by 10 percent, or the foreign price level fall by 10 percent, or some combination of the two. To the extent that the depreciation is not fully offset by changes in national price levels there is a deviation from PPP and an exception to the theory.

The idea that changes in exchange rates are determined by changes in relative national price levels was popularized by Gustav Cassel after the First World War. This version of PPP is thus often called the inflation theory of exchange rates and may be written[2]:

$$s = p - p^* \quad \text{.....(3c)}$$

In this version the direction of causation *is* specified - from inflation differentials to changes in the exchange rate. In empirical applications and tests of the theory it must be decided whether the statistical relationship between the variables is significant and, if possible, whether there is indeed a causative rather than a merely correlative relationship at work.

More recent research into PPP has also studied variance in real exchange rates. The real exchange rate may be thought of as the nominal exchange rate adjusted (or "deflated") by relative national price levels. In absolute form:

$$R = SP^*/P \quad \text{.....(3d)}$$

where **R** is the real exchange rate. If PPP holds, then **R** should be constant. Hence, research showing significant changes in **R** over time is also evidence against PPP theory.

3.2 Origins of the purchasing power parity doctrine

The onset of the First World War in 1914 decisively ended a long era of *laissez faire* capitalism. The combatant countries imposed controls which regulated the previously unhindered

flows of international trade, mobility of capital, and payments of gold as deemed necessary to meet the economic and financial demands of conducting the war. In particular the gold standard, which had greatly facilitated the expansion of international trade and foreign investment for most of the 19th century, was abandoned[3]. The ensuing disruption to international monetary and trade relations continued well after the war had ended and the relatively smooth operation of the gold standard[4] was not regained, contrary to expectations, even after Britain returned to gold in 1925 with sterling at its old pre-war parity.

A significant development in the theory of international trade resulting from the war time experience was the purchasing power parity doctrine advanced by the Swedish economist Gustav Cassel. His theories were introduced in contributions to the 1916 edition of *The Economic Journal* and subsequently elaborated and refined in Cassel (1918; 1921; 1922; and 1928). The ideas behind Cassel's purchasing power parity doctrine were not entirely new. For example, the English bullionists recognized a distinction between real and nominal differences in exchange rates, as is clear from the 1810 *Bullion Report* in England. In the 19th century Wheatley, Ricardo, Mill, Goschen, and Marshall all appeared to hold some notion of PPP. Earlier members of the Classical school held an inflation theory of exchange rates based on the quantity theory of money: depreciation of the exchange rate was thought to result from an increase in paper money via a simultaneous (or preceeding) rise in prices. Marshall, on the other hand, accepted the possibility in his *Official Papers* that speculation on the foreign exchanges could result in the external value of the currency depreciating *prior* to an increase in home goods prices and a fall in the currency's internal value. The 19th century history of economic thought on PPP is thoroughly reviewed in Viner (1937) and Officer (1984).

Cassel, however, extended these basic concepts well beyond the theoretical significance imparted to them by the Classical economists and welded them into a more sophisticated and rigorous theory amenable to empirical testing. By comparison, Viner (1937) argued that since the early Classical economists

had not developed the abstract concept of a "price level" they could not have had a genuine theory of PPP in mind. It was Cassel who first gave the purchasing power parity theory its name and provided a precise formula for its calculation:

"Thus the following rule: When two currencies have undergone inflation, the normal rate of exchange will be equal to the old rate multiplied by the quotient of the degree of inflation in the one country and in the other. There will naturally always be found deviations from this new normal rate...But the rate that has been calculated by the above method must be regarded as the new parity between the currencies, the point of balance towards which, in spite of all temporary fluctuations, the exchange rates will always tend. This parity I call *purchasing power parity*" (Cassel 1922: 140).

Despite the effective abandonment of the gold standard during the First World War, the belligerent countries did what they could to protect and maintain their gold reserves with the central banks. Great efforts were made to advertise the quantity of gold reserves in the banks' vaults, to preserve the public's confidence in the currency. The general perception was that the gold standard still operated in some mysterious fashion and that the value of the national currency remained backed by gold. Cassel wrote at some length to dispel these popular misconceptions. To the extent that paper money was no longer redeemable into gold, or that the free exportation and importation of gold was prevented, or the melting down of gold coin prohibited[5], there was no relation between the amount of gold in the vaults of the central banks and the value of their currencies:

"From the moment of the outbreak of war, the various currencies had in the main to be regarded as free paper currencies which were not limited to any metal, and therefore were not in any relation to each other" (Cassel 1922: 1).

Given the effective abandonment of the gold standard it was thus naive for the public to believe that the pre-war gold parity values of the various countries would naturally be restored after what was thought to be a purely temporary disruption resulting from the exigencies of the war. According to Cassel (1922: 19-32), the primary cause of the persistent depreciation in the belligerent countries' currencies was the sustained increase in credit and inflation of the money supply,

made necessary by the limited real means of payment for the war effort. Without the discipline imposed by the gold standard, the value of the floating paper currencies was determined by the extent to which one country created fresh artificial purchasing power relative to another. Hence the only way in which the pre-war parities could be regained and thereafter sustained at these levels was for the countries concerned to permanently reduce the rate of increase in their paper money.

It is instructive to digress a little further in this matter as it concerns the debate about what is meant by the "undervaluation" or "overvaluation" of the exchange rate and what the proper empirical measure or proxy for the price levels should be. The question of the appropriate values for the major currencies (primarily the pound sterling against the dollar, mark, and franc) assumed great practical significance after the war. In Germany, the hyperinflation of the early 1920s required the creation of an entirely new currency. This was achieved by means of the Dawes Plan and with the aid of the Dawes Loan, which in 1924 enabled the new currency to be set at the pre-war gold parity of the mark. Earlier that year, Sweden had been the first European country to restore its currency to the pre-war gold standard. Together with the stabilization of the German currency, this helped foster expectations that sterling would soon follow suit. These expectations were eventually met after lengthy and controversial public debate, by the decision of the Chancellor of the Exchequer, Winston Churchill, in consultation with the Governor of the Bank of England, to return to the gold standard at sterling's pre-war parity of $4.86, as announced in the Budget speech of April 1925.

From a strictly purchasing power parity viewpoint it could be argued that this parity level overvalued sterling. In 1925 Keynes, for example, argued in his *The Economic Consequences of Mr Churchill*, that the Chancellor's decision overvalued sterling by at least 10-12 percent. On these grounds Keynes suggested that the decision was an unfortunate mistake that would require a painful disinflation in Britain and would worsen unemployment. However, the degree of overvaluation of sterling depended on which price or cost indices were used to

calculate the hypothetical PPP exchange rates. Keynes used the rather narrow Massachusetts index, a retail price index based on only one state in the USA but quoted regularly in the *Federal Reserve Bulletin*. Other indices, such as the US Bureau of Labour index, suggested little if any overvaluation of sterling in 1925. Export price indices for the US and the UK, which excluded non-traded goods and services, suggested that sterling was significantly *undervalued* in 1924 and slightly so in 1925. As made clear below, the theoretical and empirical problems concerning the appropriate choice of price or cost indices have yet to be satisfactorily resolved.

Moggridge (1969) noted that the practical decision to return sterling to its pre-war gold parity largely ignored deviations from these theoretical PPP calculations. Policy makers felt that expectations of sterling's imminent return to gold at or near its pre-war parity had taken hold to such an extent that to have delayed the decision any longer would have precipitated a renewed crisis of confidence. A wider set of criteria were used to determine a desirable valuation for sterling. Of primary concern were the implications of broad changes in wages and productivity, international competitiveness, the balance of trade, capital flows, and domestic unemployment. Yet in many of these respects, Britain was in a relatively weaker international position than before 1914. The increase in American tariffs in 1921 also disadvantaged British goods in a significant and growing export market. Thus, despite inconclusive purchasing power parities, Moggridge (1969: 70) reached roughly the same conclusion as Keynes in 1925: that an exchange rate of perhaps 10 percent lower than $4.86 would have been more appropriate for sterling.

It was commonly believed that England's economic problems after 1925 were a direct consequence of this "overvaluation", as had been argued by Keynes in his earlier criticism of Churchill's decision to return sterling to the pre-war gold parity. From 1924 to 1931 the value of British exports stagnated and unemployment appeared to worsen, particularly in the coal mining and cotton textile industries. Employers responded to declining demand with wage cuts and retrenchments which in turn

provoked repeated strikes and mass marches by disaffected workers.

Cassel, on the other hand, denied that these problems were caused by a premature decision to return to gold and the overvaluation of sterling. To argue that there was a direct cause and effect relationship between the two events was to commit the *post hoc ergo propter hoc* logical fallacy. Cassel (1936) maintained that the roughly 10 percent domestic disinflation made necessary by the overvaluation could be managed without unduly harming the British economy. Sweden, for example, had adjusted to a far more severe deflationary process between 1920 and 1922 and had no trouble in maintaining the pre-war gold parity of the kroner after 1924.

Cassel (1936) argued that it was only events *after* 1925, that could not have been readily foreseen at the time, which precipitated the economic crisis and eventually made the restored gold standard unworkable. In particular, the stock market crash and the collapse in prices which subsequently took place in America forced similar deflationary pressures in Britain and elsewhere[6]. By September 1931, in the wake of bank runs throughout Europe, the drain of gold from London proved intolerable and the British government released the Bank of England from its obligation to redeem its notes for gold. Despite various attempts to remedy the situation other countries soon followed suit, effectively ending the restored gold standard.

3.3 What does purchasing power parity mean?

From the outset, conceptual confusions have prevented a single unambiguous theory of purchasing power parity:

"Much of the controversies concerning the usefulness of the PPP doctrine is due to the fact that the doctrine does not specify the precise mechanism by which exchange rates are linked to prices. Rather, the PPP doctrine may be viewed as a short-cut; it specifies a relationship between two variables without providing the details of the process which brings about such a relationship. As a result, the doctrine has been subjected to different interpretations" (Frenkel 1978a: 169).

Without such a theory the results of empirical tests have proved ambiguous and thus do not allow economists to decisively "accept" or "reject" PPP theory. The meaning of PPP largely depends on the precise question being asked or the problem one is trying to solve. For instance, is the theory intended to explain how exchange rates are determined, or is it being used to explain differences in price levels and real incomes between countries? Is the relationship between relative price levels and exchange rates causative and, if so, what is the direction of causation? Is PPP meant to be interpreted narrowly as pertaining to traded goods only, or should the widest possible range of goods and services be used in the index of price levels? Is PPP valid in the short-run or is it only to be regarded as a long-run relationship? If so, what exactly is meant by "the long-run"?

As should be clear from these questions there is no consensus as to "the" meaning of PPP. Many interpretations exist. Despite conceptual and empirical problems with the various interpretations, however, PPP remains an important part of the workaday economist's theoretical repertoire. This, in turn, raises a number of relevant issues in the methodology of economics. Of particular interest is the relationship between empirical tests and PPP theory. Do such tests decisively knock out (or confirm) theories of PPP? If not, what are the grounds for preferring one theory and interpretation of PPP to another? Are the rationalizations of deviations from PPP an example of undesirable *ad hoc* modifications to an economic theory that has been falsified by the empirical evidence? And what do the various innovations in econometric techniques at the different stages of testing PPP imply as regards the interpretation of such test results? The ensuing discussion is intended to clarify some answers to these and related questions in the methodology of economics.

A review of the literature reveals two fundamentally divergent views on the meaning of purchasing power parity. These may be labelled the 'commodity arbitrage' view and the 'monetary' view respectively. The commodity arbitrage view may be thought of as an extension of the law of one price to international trade:

the price of an identical good at different locales must be the same if there is perfect information, no barriers to trade and zero arbitrage costs. For example, if an identical television set fetched $600 in Japan but only $400 in America (after converting the yen price by the spot dollar/yen exchange rate) then profit maximizing arbitragers would purchase such TV sets in America to sell in Japan. This would increase demand and the price of TV sets in America while increasing supply and lowering their price in Japan. The arbitrage process continues until the prices of the TV sets in the two countries are equalized, when expressed in a common currency. This is the equilibrium international price. Residual discrepancies in prices can be explained by the various costs, risks and barriers to arbitrage such as transport costs, insurance, foreign exchange risk, tariffs, quotas and other non-tariff barriers. Alternatively, factors such as brand loyalty and subtle product differentiation may also result in price discrepancies. Empirical research which uncovers deviations from the law of one price is not usually interpreted as disproving the law. Instead it prompts further research to explain such deviations by discovering these (possibly hidden) costs, risks, trade barriers, and product differences. In the case of assets like gold or financial securities, which are virtually perfectly homogenous and do not have to be physically present for trade to occur, the law of one price closely approximates the facts (see section 3.4 for some empirical studies that have been conducted in this regard).

The commodity arbitrage view does not necessarily specify a unidirectional causal relationship between relative prices and exchange rates. The equality of exchange rate adjusted prices can be interpreted as an equilibrium condition which must be met if the law of one price is upheld. Changes in exchange rates may cause offsetting changes in the domestic currency prices of traded goods between two countries or vice versa. Thus, a depreciation of the rand/dollar exchange rate could be offset by increases in the rand price of gold and other internationally traded commodities. On the other hand, if traded goods prices in SA increased at a faster rate than

abroad the theory implies an offsetting depreciation of the exchange rate, which re-establishes the equality of prices.

Some economists have regarded PPP as an implication of the law of one price, reasoning that if commodity arbitrage equalizes the exchange rate adjusted prices of individual traded goods, it must logically do the same for price levels calculated from collections of goods:

"The proposition that general price levels in different currencies are connected through the prices of internationally traded goods is the foundation of the purchasing power parity doctrine"(Officer 1976a: 558).

An extreme commodity arbitrage view questions the relevance of using an aggregate price index at all:

"Foreign exchange rates have nothing to do with the wholesale commodity price level as such but only with individual prices" (Ohlin 1967: 290).

By way of criticism, Keynes (1930) and Samuelson (1974) noted that the traded goods, commodity arbitrage view is virtually a truism:

"Of course, under perfect competition, free trade without tariffs, quotas or exchange controls, relative prices of one good could not deviate regionally if transport costs were zero" (Samuelson 1974: 602).

Accordingly, some economists maintain that a genuine PPP theory should exclude parities calculated only from traded goods prices. Instead a broad price index including both traded and non-traded goods and services, such as a consumer price index, should be used. This was the view of, amongst others, Hawtrey (1919), Cassel (1928), and Haberler (1945). This line of argument can be taken to the other extreme - only non-traded goods should be included in the price indices. The logical conclusion of this approach is to calculate cost or factor price parities of least traded service commodities such as labour (see Officer 1974). However, there are also various problems with this interpretation, particularly the measurement of returns to the different factors of production and the need to account for differences in productivity (see section 3.4).

Using a broad price index including non-traded goods suggests a monetary interpretation of the law of one price, in which the exchange rate adjusts to maintain equilibrium between the internal and external value of the currency. Since the value of money is inversely proportional to the *general* price level, it is possible to redefine the law in terms of the relative purchasing power of different currencies. This was the essence of Cassel's approach:

"The whole theory of purchasing power parity refers to the internal value of the currencies concerned, and variations in this value can be measured only by general index figures representing as far as possible the whole mass of commodities marketed in the country" (Cassel 1922: 33).

By this interpretation, a unit of currency should purchase the same representative basket of goods and services anywhere in the world, once adjusted by the relevant exchange rate. In other words, the internal value and the external value of a currency should be the same in equilibrium. The direction of causation is an integral part of the theory: from changes in relative price levels, and thus the internal values of the currencies concerned, to changes in the exchange rate.

Some proponents of the monetary view, however, question the significance of the hypothesized causative link between relative price levels and the exchange rate:

"Since in general both prices and exchange rates are endogenous variables that are determined simultaneously, discussions of the link between them provide little insights into the analysis of the determinants of the exchange rate" (Frenkel 1978b: 4).

"In retrospect it seems that the translation of the theory from a relationship between moneys into a relationship between prices - via the quantity theory of money - was counter productive and led to a lack of emphasis on the fundamental determinants of the exchange rate and to an unnecessary amount of ambiguity and confusion"
(Frenkel 1978b: 5).

Thus in the modern monetary interpretation of PPP, changes in the relative quantities of and demands for money are the fundamental independent variables which determine the exchange rate directly, rather than indirectly via their effect on the

price level. Frenkel asserts that classical economists such as Wheatley (1803) and Ricardo (1821) intuitively understood this and that they emphasized the importance of monetary influences on the value of the currency. The modern monetary view regarding the issue of causation was anticipated more explicitly by Wicksell:

"Does the difference in the level of prices cause the rate of exchange to deviate from par; or is the converse true? This question can only be answered in regard to each individual case. Either process is conceivable; and both of them are - as we shall shortly explain - quite capable of being combined with a third, which is the ultimate or underlying cause" (Wicksell 1919: 233).

"The ultimate cause must again in both cases be sought in the mismanagement of the currency in the country in which it has deteriorated" (Wicksell 1919: 234).

Cassel, in his original statement of PPP in 1916, emphasized the relative quantities of money in determining the exchange rate. He used price indices as a proxy for inflation of the money supply only because data concerning the latter was, at that time, generally either unavailable or inaccurate (see also the quote from Cassel 1922: 140 in section 3.2):

"Nevertheless, it has been possible to prove that the advance of the general level of prices in the United Kingdom...is, broadly speaking, proportional to the increase of the circulating medium of the country, and thus that the enhancement of prices is essentially to be regarded as an expression of the inflation of money" (Cassel 1916: 63).

However, this was later translated into a relationship between relative price levels and exchange rates via an application of the quantity theory of money. Such economists recognized the significance of monetary disturbances but regarded changes in relative price levels as the prime independent cause of changes in exchange rates:

"But as the present European currencies have only a direct purchasing power inside the issuing country, it is perhaps easier to put this truth in the form that the rates of exchange are determined by the relative quantities of money issued by different countries, though as the price level is influenced by other things as well as the quantity of money, this latter way of speaking is the less accurate" (Gregory 1930: 83).

and

"In other words, the theory asserts that...the relative price levels, ultimately determine the rates of exchange" (Gregory 1930: 84).

Later empirical work tended to focus on the relationship between relative price levels and exchange rates rather than the relationship between money supplies and demands and exchange rates. This neglect was only rectified in the late 1960s and early 1970s with the revival of monetarism and the formulation of the monetary approach to the balance of payments and to flexible exchange rates respectively.

If relative price levels are not the transmission mechanism through which changes in money supply and demand determine the exchange rate, it is pertinent to ask how exactly monetary disturbances are supposed to influence exchange rates. This is explained in chapter four where the role of the exchange rate in maintaining stock equilibrium in domestic asset markets is considered more fully, including the crucial role played by expectations and speculation in the foreign exchange markets.

3.4 Empirical tests of deviations from the law of one price and purchasing power parity

This section examines some of the empirical work that has been carried out on the relationship between prices and exchange rates, and the implications of the results thereof for theories of PPP. Given the large amount of research done in the area, this examination is necessarily selective and an exhaustive survey is not attempted here. However, the research selected is indicative of the general empirical findings regarding PPP and is sufficient to draw some preliminary conclusions in the methodology of economics. The methodological implications are considered in section 3.7.

Dornbusch (1988) suggested two basic categories of deviations from PPP: structural and transitory. Structural disparities arise through systematic changes in real economic factors which cause permanent changes in equilibrium relative prices. Lasting changes in productivity and the effect of differences in economic growth rates are the two most common examples.

Transitory departures from PPP may arise if goods and assets markets adjust at different speeds to disturbances such as capital account shocks. This may occur whenever wages and prices are not perfectly flexible. Cassel (1922: 147-162) recognized three possible sources of deviations from PPP: changes in expected inflation rates, new barriers to trade, and autonomous shifts in international capital flows. However, Cassel regarded these deviations as limited and/or temporary. Since he emphasized PPP as a benchmark against which to assess the value of a currency, he neglected any further investigation of such disparities.

There are good reasons for believing *a priori* that many exceptions to the law of one price exist, even at the level of individual traded goods. As noted in section 3.3 the law of one price is only expected to hold on the basis of restrictive, unrealistic assumptions such as zero transport costs, no barriers to international trade and perfect information. To the extent that these assumptions are not met in the real world we should expect price discrepancies between what appear to be 'identical' goods. Where, however, these assumptions are met more closely the more positive are the empirical test results. For example, Genberg (1975) compared quarterly changes in the common currency wholesale prices of cacao, jute, rubber, and tin at eight different locations worldwide and found the price changes to be very nearly the same. Similar findings are apparent in highly traded standardized commodities such as base metals traded on the London Metal Exchange. Commodities with a financial asset aspect to them, such as gold and silver, exhibit even tighter price ranges between locales.

For manufactured goods and with aggregate index price data, the positive results of Genberg's (1975) study tend to be the exception rather than the rule. There is a large body of empirical research into deviations from the law of one price and PPP: the general conclusion to be drawn from this research is that such deviations are significant (whatever version of PPP is tested) and persist beyond most definitions of the short-run. For example, Isard (1977) compared the exchange rate adjusted prices of disaggregated SITC (Standard International

Trade Classification) categories of internationally traded goods and concluded that:

"In reality the law of one price is flagrantly and systematically violated by empirical data...exchange rate changes substantially alter the relative dollar-equivalent prices of the most narrowly defined domestic and foreign manufactured goods for which prices can readily be matched" (Isard 1977: 942).

In a careful econometric study of price and exchange rate relationships between Canada and the US (1965-1974), Richardson (1978) also concluded that the law of one price failed, in all of the 22 SIC (Standard Industrial Classification) commodity categories he studied. For a given category, changes in US prices and the bilateral exchange rate were found to have only a weak influence on Canadian prices: the hypothesis that the regression coefficients on US prices and the exchange rate were unity had to be rejected at the 95% level of significance. Interestingly, however, the hypothesis that the coefficients were equal could only be rejected in 3 cases. For the other 19 categories, Canadian prices responded much the same way to changes in US prices as to changes in the exchange rate[7].

Kravis and Lipsey (1978), in a thorough empirical investigation of international price behaviour, reached conclusions supporting Isard's and Richardson's results. Using cross-section data, they found that the common currency prices of comparable export goods from different countries varied substantially and that these departures from the law of one price persisted for most of the 1953-64 sample period of their study. For example, in 1963 Japanese iron and steel prices were on average found to be 30 percent less than for the US, while German and UK prices were 24 percent and 22 percent less respectively (Kravis and Lipsey 1978: 230). At a narrower SITC classification level, they reported iron and steel wire prices as much as 43 percent lower for Japan compared to the US; and 40 percent lower than the US for German exports of bars and rods, and tube and pipe fittings. These price differences persisted for roughly the entire 1953-64 period. Japanese and German market shares in these goods increased during the same period, although that of the UK fell.

The authors also show that relative export prices of comparable goods from different countries sometimes differ substantially over time. If the law of one price always held then the relative prices of such tradables should not change significantly and should be unaffected by changes in exchange rates. In comparing the German/USA export (dollar) price ratio, however, they found that it was 45,5 percent higher in 1975 than in 1969. This trend they attribute to the strong appreciation of the mark beginning in 1969, which rose by 59,7 percent over the six year period. Kravis and Lipsey (1978: 231) concluded that "there is very little similarity between the changes in German and US export prices when both are expressed in dollars".

Kravis and Lipsey (1978: 206-219) also tested the PPP relationship by comparing changes in exchange rate adjusted price levels (with implicit GDP deflators used as an index for price levels) for the major industrial countries. For the twenty year period 1950-70 they found that the PPP relationship was a reasonably good approximation for France, Belgium and the UK when paired against the US (relative dollar price levels remained close to the base value of unity at the end of the period) but that for the other countries, deviations ranged from a minimum of 10 percent for Germany to 21 percent for the Netherlands. For pairings within Europe the disparities were mostly even higher, with a maximum of 34 percent for the Netherlands/Italy.

The authors also constructed a revealing frequency distribution of the deviations from the PPP relationship for 11 countries *vis-a-vis* the US in which all possible 2,3,4,7, and 10 year periods between 1950-70 were examined, as set out in Table 1. From the table it is evident that the expected PPP relationship of unity, where relative inflation is exactly offset by exchange rate changes, generally holds for short durations (although the authors do point to individual cases with deviations greater than 10 percent). But for longer durations the relationship diverges significantly from unity. For the ten year periods, a fifth of cases showed disparities of 20 percent or more while only 14 percent of the sample showed disparities

of 5 percent or less. When wholesale or consumer price indices were used instead of implicit GDP deflators, the PPP relationship tended to be closer to 1. However, even with the wholesale price index (which has a greater weighting of traded goods) large deviations were still present for some pairs of countries, particularly when they did not include the US.

Table 1: Frequency table of absolute deviations from PPP using implicit GDP deflators for 11 countries against the US, 1950-70

Duration (years)	Number of Periods	Number of Observations	Percent of cases in which absolute deviation was: ≥20%	10-19,9%	5-9,9%	≤5%
2	19	202	1	5	19	74
3	18	191	1	15	21	63
4	17	180	1	24	22	53
7	14	147	10	32	32	27
10	11	114	20	42	24	14

Source: Kravis and Lipsey (1978: 213)

More recent work on the law of one price has tended to confirm the earlier findings reported above. For example, Knetter (1989; 1993) compared variations in the cost of 7 digit SIC export goods from a single source between multiple destinations and found that they differed significantly. Such cost comparisons also varied considerably over time. Giovannini (1988) used an even more disaggregated data set in comparing goods prices between the USA and Japan and found sharp price deviations not only in quality manufactured goods but also in more basic commodity manufactures such as screws, nuts, and bolts.

The empirical evidence against the law of one price appears to be quite damning. In a recent survey of the main empirical studies done in this area Rogoff summed up as follows:

"Overall, it is hard to read the empirical evidence without concluding that outside a fairly small range of various homogenous goods, short-run international arbitrage has only a limited effect on equating international goods market prices" (Rogoff 1996: 654).

3.5 Explaining deviations from purchasing power parity

From the above it is clear that deviations from the predictions of the law of one price and PPP theory are the rule rather than the exception. In other words, there appears to be more than sufficient evidence available to falsify the relevant theories. However, economists have generally not responded to this evidence by abandoning the concept of PPP. Instead the response has been to reinterpret the results in various ways and to "save" the concept from an untimely death. Alternatively, some researchers have investigated exogenous variables that might help explain changes in real exchange rates. It is worthwhile to look at some examples of how this has been done before examining the methodological issues in section 3.7 more closely.

How large must deviations from PPP be to be regarded as significant? Most regression analyses of PPP, in common with most other statistical tests, conventionally assign probablistic "levels of significance" to the results, usually of 1 or 5 percent. This simply means that the observed goodness-of-fit of the data to the hypothesized relationship would have only the assigned probability of occurring by chance alone. However, not all economists have accepted this definition of "significance". An alternative approach to assessing the significance of deviations from PPP has been taken by Genberg (1975; 1976). He tested the PPP relationship by comparing the price indices of internationally traded goods with those traded *intra*nationally. He found no significant difference in price correlations between OECD countries and those between regions within the USA. Therefore, he concludes, if we regard the USA as a unified integrated market in goods, so should we view the OECD. His approach is noteworthy because it sets an alternative benchmark against which to compare deviations from the law of one price. However, his approach seems to beg the question by assuming that a nation represents an integrated market in which the law of one price is a good approximation to the truth. It could be argued that the law of one price is not necessarily upheld intranationally either, particularly in a geographically large and diverse country such

as the USA where the various states have different taxes, trade restrictions and other regulations governing commerce.

Engel and Rogers (1995), on the other hand, provide convincing evidence against Genberg's findings. They found that price differences for similar goods categories were significantly larger and more variable across the borders of the countries they studied than within them. Engel and Rogers examined data on 14 categories of disaggregated consumer price indices for 23 cities in the USA and Canada and found large differences in relative price volatility between cities on opposite sides of the border compared to cities on the same side of the border, even after allowing for distance. Rogers and Jenkins (1995) also show that such price differences across borders are more persistent than within the borders.

Other economists have queried the relevance of deviations from PPP. For example, Magee (1978) suggests that observed serial correlation of deviations from PPP may be spurious due to multiperiod contracting. Although exchange rate adjusted prices may be identical when goods contracts are signed, tests using these contract prices rather than ruling spot prices are likely to show evidence of serial correlation. However, this merely demonstrates the use of inappropriate price data and not necessarily inefficient goods arbitrage. Substituting spot commodity prices for contract prices should, according to Magee, resolve this issue.

Roll (1979) argues that there must be many goods whose exchange rate adjusted prices do differ between locales, to provide an incentive for trade. The price differential must be sufficient to compensate for the various costs and risks of arbitrage and to provide the required return on capital. Otherwise, none of the observed trade patterns, where goods persistently flow in one direction, would exist. The law of one price should fail for these goods. Roll points out that only the average real return on all goods traded between two countries should be zero with perfect foresight. Relaxing the simplifying assumption of perfect foresight allows PPP to be interpreted within a stochastic efficient goods markets framework. Given a time lag

between the purchase and sale of a commodity, arbitragers cannot know with certainty the exchange rate adjusted price that it will fetch. Thus *actual* average real returns from such arbitrage are not necessarily zero. Instead, only the *expected* returns from arbitrage should be zero. This means that the anticipated exchange rate adjusted prices of traded goods should be equal and that deviations from parity values be serially uncorrelated. Thus relative form PPP may now be expressed as:

$$E[\ S_t\{1 + p_t\}/\{1 + p^*_t\}\ ] = S_{t-1} + u_t \quad \ldots\ldots\ldots\ldots (3e)$$

where p_t and p^*_t denote the rates of change in domestic and foreign prices of traded goods respectively, **E** denotes the expected value of the bracketed term and u_t is the random error term. The equation states that the best forecast of the next period's inflation adjusted (real) exchange rate is the current spot rate. As in other efficient markets, all information believed to be relevant in predicting the real exchange rate should already be reflected in the current spot exchange rate. Roll (1979) tested a regression version of equation (3e) and could not reject the null hypothesis of a unit root on the real exchange rate. His results suggested an absence of mean reversion and thus essentially random walk behaviour in the real exchange rate over the short to medium run (see section 3.6 for more on unit root tests and section 5.2 of chapter five for tests of the efficiency aspects of the foreign exchange market).

Other studies have searched for exogenous variables which might explain variations in real exchange rates (or exchange rate adjusted price levels in comparisons of international incomes). These include various structural factors such as differences in real per capita incomes, productivity, fiscal deficits, and current account balances. Some of these studies and their findings are examined below.

Kravis, Heston and Summers (1978) conducted a cross-sectional study of 34 countries with a wide range of national incomes. Using real GDP per capita to classify countries, they found

that 1975 dollar price levels in high income countries were more than double those of low income countries. In a later paper, Kravis and Lipsey (1983) drew on the data from this research to show that long-run structural variables were important in explaining differences in exchange rate adjusted price levels. Using multiple regression analysis they discovered a statistically significant positive relationship between real GDP per capita; measures of openness of the economy; the nontradables share of GDP, and the price level (the dependent variable). For real GDP per capita alone they reported a highly significant t statistic for a beta coefficient close to unity and an R^2 of 0,801. They concluded that real income per capita was the major source of variation among countries in the price levels of both tradable and nontradable goods and in the total price level (Kravis and Lipsey 1983: 29). The other two independent variables were also found to be statistically significant: higher degrees of openness and a higher share of non-tradables in GDP were both associated with higher price levels.

Changes in common currency adjusted price levels imply that the real exchange rate does not remain constant. Hsieh (1982) developed a simple model which explained variation in the real exchange rate by differences in labour productivity growth rates between countries in the traded and nontraded goods sectors. His work drew on previous research into the PPP effects of such productivity differences by Balassa (1964), Samuelson (1964), and Officer (1976b). However, Balassa and Officer used cross-sectional empirical techniques while Hsieh uses a time series approach. Hsieh (1982: 360) points out that cross-sectional techniques do not account for country specific factors such as tastes and resource endowments. If there are large differences in these factors between countries then the results of cross-section regressions are unlikely to be as significant as those reported by Officer (1976b). Time series regressions avoid this difficulty if the country specific factors remain roughly constant over time. Hsieh's research is noteworthy: instead of merely pointing to evidence of deviations from PPP, he attempts to explain them using a testable structural model.

In his model, changes in the real exchange rate and deviations from PPP can occur if non-traded goods prices differ between countries; or if the weights of traded vs non-traded goods in the countries' price level indexes differ. This will be the case even if the law of one price holds for all individual traded goods. In Hsieh's competitive two country model, prices of traded and non-traded goods are reduced to unit labour costs. The nominal wage rate is the same for both sectors at home due to mobility of labour between sectors, but may differ between countries. Given these simplifying assumptions, changes in the real exchange rate may be expressed in regression equation form as:

$$r_t = c + d[a^T - a^N]_t - e[a^{T*} - a^{N*}]_t + f[w - s - w^* + a^{T*} - a^T]_t + e_t \quad \text{.....(3f)}$$

where a^T and a^N denote rates of change in the average (and marginal) product of labour in the traded and nontraded goods sectors respectively, and w the change in wage rates. Asterisks denote foreign variables and e_t is the sample error. The equation states that changes in the real exchange rate r_t are explained by the differences between labour productivity growth rates in the traded and nontraded goods sectors at home and abroad (the bracketed second and third terms respectively); and the difference in growth rates of unit labour costs (prices) of traded goods between countries (fourth).

For both Germany and Japan (1954-76) against their main trading partners, Hsieh found the intercept coefficient c not to be significantly different from zero while d and e were between zero and unity, as predicted. However, f was estimated to be significantly less than the expected value of unity in all the regressions tested. Hsieh suggests that this could be due to "measurement error" such as inaccurate classification of goods into the traded and nontraded goods sectors. However, it could also imply that differences in the prices of traded goods between countries were not fully offset by changes in the exchange rate, contrary to Hsieh's assumption that the law of one price holds for traded goods. Hsieh (1982: 361) interprets his overall results as "a more favourable confirmation of the

productivity differential model than the cross-section regressions of Officer" which, as noted above, did not account for possible differences in country specific factors. Hsieh's findings are in keeping with those of Kravis *et al* discussed above.

In general the results of empirical tests of the Balassa-Samuelson hypothesis seem to show that it works passably well in comparisons of rich versus poor countries but that it breaks down in comparisons of countries with similar levels of real per capita GDP. There are, however, exceptions to the latter as suggested by Hsieh's analysis of variations in the yen and mark exchange rates. Rogoff (1996) also finds evidence favouring the Balassa-Samuelson effect for the yen/dollar exchange rate.

3.6 Asset markets, random walks, and purchasing power parity as a long-run equilibrium relationship

With the advent of floating exchange rates during the 1970s it became increasingly evident that the behaviour of foreign exchange markets resembled that of other stock markets in which financial assets are traded, rather than the flow markets for goods and services. The exchange rate came to be seen as an asset price which adjusts to maintain stock equilibrium in asset markets. The role played by expectations, the efficiency of foreign exchange markets and related issues were also explored in greater depth (see chapters four and five). The asset market approach was helpful in explaining the stylized fact of marked volatility in flexible exchange rates during the 1970s (and subsequently) which had not been anticipated by economists at the time. It was thought by many economists that speculation would prove to have been more of a problem under a fixed exchange rate system and would be a stabilizing influence on freely floating exchange rates[8].

The stock market view of the exchange rate led to renewed questioning of PPP. Initial studies by Frenkel with regard to the monetary model of exchange rate determination, for example, provided evidence favouring PPP for the franc/dollar, pound/dollar, and franc/pound during the early 1920s (February

1921 - May 1926). For relative PPP, using three different price indices, Frenkel (1978a) could not reject the hypothesis of unit elasticity of the exchange rates with respect to the relevant price ratios, except for the dollar/pound exchange rate. The data were in most cases also consistent with absolute PPP. However, the results of a Sims (1972) test of causality between two time series showed that the direction of causation ran from exchange rates to prices rather than vice versa meaning that in the regression equations the price ratios should be specified as the dependent variable and exchange rates as the independent variable placed on the right hand side of the equation. Frenkel noted that this "pattern of causality" was consistent with the asset market view that both exchange rates and commodity prices are endogenous variables which adjust to changes in other common factors, but with adjustment in asset markets occurring more rapidly than in commodity markets[9].

Later research, much of it by the same author, was far less favourable for the PPP theory. Frenkel (1981a), for example, compared the time series properties of relative price levels and exchange rates between 1973 and 1979. He showed that the monthly changes in exchange rates were serially uncorrelated and appeared to follow a random walk, while relative price levels and their monthly changes were far less volatile. For Germany, France, and the UK, absolute mean monthly changes in exchange rates against the dollar were about 2 percent while the changes in the ratios of national price levels (relative to the US) ranged from an average of only 0,3 percent for France to 0,7 percent for the UK. Unexpected changes in interest rate differentials were found to be a key variable in explaining the high volatility of exchange rates.

Frenkel also found that exchange rate deviations from PPP appeared to follow a first-order autoregressive process. Such autocorrelation suggests that economic factors are at work to bring exchange rates back to their PPPs in the long-run. However, the first differences of the deviations were not found to be serially correlated, which implies that the deviations did not tend to diminish with time. Frenkel (1981a: 699)

concluded that "the data do not provide sufficient evidence to reject the alternative hypothesis of a random walk".

In a later paper that year, Frenkel (1981b) concluded more strongly that purchasing power parities had "collapsed" during the 1970s. In this paper Frenkel estimated regression equations of the simple PPP relationship of the (log) form:

$$s_t = a + b(p_t - p^*_t) + e_t \quad \text{.....(3g)}$$

to test the alternate hypothesis that $b = 1$. Contrary to the positive evidence he found for many hyperinflationary economies (see Frenkel 1978c), PPP performed badly for the more moderate inflation experience of industrialized countries during the 1970s. The b estimates typically varied far from one, with some countries even yielding negative coefficients. For a time Frenkel's conclusions became the consensus view of economists as regards the performance of PPP, at least for the floating rate period of the 1970s. This was reinforced by the results of early unit root tests on real exchange rates and initial cointegration tests of the PPP relationship (see below).

The asset market view suggests that short-run changes in the exchange rate resemble a random walk but leaves open the possibility that there may be a long-run equilibrium PPP relationship, despite Frenkel's (1981a; 1981b) unfavourable preliminary findings. If such a relationship exists then the results of univariate regression analysis should show that real exchange rates do not have unit roots and that there is evidence of mean reversion. In other words, exogenous shocks to the nominal exchange rate should be corrected in the long-run.

The generally negative results from the above tests of the simple PPP relationship, and concerns over possible nonstationarity of relative price levels and exchange rates (which invalidates standard regression tests of the null hypothesis), thus led to an alternative approach in the early 1980s which tested for unit roots in *real* exchange rates. In log form:

$$R_t = a + bR_{t-1} + e_t \quad \text{.....(3h)}$$

In these tests the null hypothesis $b = 1$ is imposed (rather than estimated as in Frenkel's tests of equation 3g above). If the null hypothesis cannot be rejected then there is evidence against mean reversion, thus suggesting that the real exchange rate follows a random walk. The alternative hypothesis is that mean reversion brings about PPP in the long-run. However, initial tests using this approach could not reject the null of a random walk either. For example, tests by Roll (1979), Darby (1983), and Adler and Lehman (1983) could not reject the random walk hypothesis of a unit root. Things could hardly be worse for defenders of the PPP faith. Not only did the evidence from previous tests suggest that PPP did not hold in the short-run but these latest tests suggested that it was not even valid as a long-run relationship!

Later research in the 1980s used a range of alternative techniques to distinguish the real exchange rate from a random walk. A popular modern technique is the Dickey-Fuller (DF) and augmented Dickey-Fuller (ADF) tests[10]. The standard DF test (see Dickey and Fuller 1979) uses the autoregressive equation:

$$R_t = b_0 + b_1 t + b_2 R_{t-1} + e_t \quad \text{.....(3i)}$$

where t represents a time trend. The null hypothesis of a random walk real exchange rate is that $b_2 = 1$. Under the alternative hypothesis that PPP holds in the long-run, $b_1 = 0$ and $b_2 < 1$. However, many of these tests also proved negative. Using standard DF tests Meese and Rogoff (1988), for example, could not reject the null of a unit root for monthly dollar exchange rates against the pound, yen, and mark over the recent floating rate period. Mark (1990) tested intra-European bilateral exchange rates for the period 1973-88 and came closer to rejecting a random walk (although only for the Belgian franc/mark real exchange rate could a unit root be rejected at the 5 percent level of significance). His better results were possibly due to the fact that many of the currencies he tested were maintained within the European Exchange Rate Mechanism (ERM) for much of this time. Following this train of thought

Chowdhury and Sdogati (1993) compared various bilateral ERM currencies against the mark and the US dollar for the period 1979-90. They were able to reject the random walk hypothesis for the intra-European exchange rates against the mark but not against the dollar, thus also suggesting an exchange rate regime effect.

A problem with the above unit root tests is their low power. Because the real exchange rate is so volatile, particularly for the post Bretton Woods era of floating exchange rates, it is difficult to distinguish between slow mean reversion and random walk behaviour in the real exchange rate. Two main approaches to resolving this difficulty evolved during the 1980s. One approach was to use much longer sample periods or broader cross-sectional exchange rate data. The other was the adoption of cointegration techniques which are more powerful than standard unit root tests alone in discerning whether a long-run PPP equilibrium relationship exists. Cointegration techniques have also been used to test for a long-run equilibrium relationship in the monetary model of exchange rate determination, discussed further in chapter five.

Cointegration techniques were first applied in econometric work by Engle and Granger (1987) and Engle and Yoo (1987). These techniques are designed to uncover long-run equilibrium relationships for which the short-run dynamics are unspecified. Tests for unit roots in real exchange rates are used in univariate autoregressive equations such as (3h) whereas cointegration is appropriate for a multivariate expression of PPP:

$$s_t = a + bp_t - b^*p^*_t + e_t \qquad (3j)$$

If two or more nonstationary variables can be combined to get a stationary variable then the nonstationary variables are said to be cointegrated[11]. Any linear combination is permissible. Cointegration tests of the exchange rate against domestic and foreign price levels thus relax the symmetry and proportionality restrictions that $b = b^* = 1$ imposed on equation (3i) by many of the earlier tests. These restrictions

may be significant because, for example, the measured price index data used in actual tests may not correspond to the "true" or theoretical price level[12]. This may explain why previous researchers were unable to reject the null hypothesis that the real exchange rate has a unit root. The greater power of cointegration tests over the early unit root tests thus derives from the relaxing of these restrictions.

Again, however, the results from early cointegration tests were mixed. Using the Engle and Granger three-step cointegration procedure[13] Enders (1988), Taylor (1988a), Mark (1990), and Patel (1990) could not reject the null hypothesis of no cointegration. On the other hand Taylor and MacMahon (1988), Kim (1990), and Pippenger (1993) were able to reject the null and thus concluded that there was evidence supporting PPP as a long-run equilibrium relationship.

A pattern emerging from the empirical research into PPP suggests that if you only hang on long enough a new econometric technique will come to the fore, holding out the promise of salvation. Thus, for example, MacDonald (1993; 1995) recommended the use of a yet more sophisticated cointegration technique proposed by Johansen (1988; 1991) and Johansen and Juselius (1992). MacDonald attributes the mixed results of early cointegration tests to deficiencies in the Engle and Granger three-step procedure which are avoided using this latest technique. Johansen's estimation technique is described as a one-step full-information maximum likelihood test which simultaneously estimates the b coefficients and tests for the presence of a unit root in the regression equations (see Froot and Rogoff 1995: 1662-1667). Like the Engle and Granger three-step method it avoids possible measurement error due to the proportionality restriction. Unlike the Engle and Granger method (which often retains the symmetry restriction that $b = -b^*$) the coefficients are estimated in totally unconstrained form and the maximum likelihood estimates do not depend on which variable is put on the left hand side of the equation.

Using the Johansen approach, MacDonald (1993) tested both strong-form and weak-form PPP. Strong-form PPP means that

exchange rates and relative price levels are cointegrated *and* that there is proportionality of the exchange rate with respect to relative price levels (therefore this is a bivariate test). Because of the measurement problem, MacDonald emphasizes weak-form PPP in which the latter condition is dropped (a trivariate test). MacDonald tested bilateral US dollar exchange rates against the Canadian dollar, franc, mark, yen, and the pound using monthly data (January 1974 - June 1990). A wholesale and a consumer price index series were used to measure relative price levels. Strong support for weak-form PPP was reported: all but two of the PPP relationships estimated had a cointegrating vector, with the UK appearing to have two such vectors. However, in all the cases tested MacDonald had to reject the homogeneity hypothesis that a one percent increase in relative price levels causes an equiproportionate increase in the US dollar bilateral exchange rates. Using implicit European cross-rates he found evidence suggesting that the rejection of strong-form PPP was due to transport costs and trade restrictions.

MacDonald was also able to estimate how rapidly long-run PPP is regained after an exogenous shock to the exchange rate. The estimated coefficients of the loading (or alpha) matrix (see Johansen 1988) can be interpreted as percentage "weights" by which, on average, PPP disequilibrium is corrected[14]. MacDonald found, for example, that France had the most rapid adjustment speed, with about 7 percent of any deviation from PPP being corrected in the current month, while the UK had the smallest coefficient of only 0,2 percent.

MacDonald's findings concur with the results of other researchers using the Johansen approach. Cheung and Lai (1993) and Kugler and Lenz (1993), for example, also reported significant evidence of a long-run PPP equilibrium relationship using multivariate cointegration techniques. However, not all the results of recent research in this area are favourable. Serletis (1995) had mixed results with the Johansen approach. Using quarterly data over the sample period January 1973 - January 1992, for most of the 17 OECD countries he investigated

Serletis rejected the null hypothesis that the PPP relation is stationary.

In their survey of empirical work done on PPP, Froot and Rogoff (1995: 1665) note that cointegration studies reveal some systematic empirical features. Firstly, the null hypothesis of no-cointegration is rejected more easily across pairs of currencies that are fixed than those that are floating. Under fixed rates, deviations from PPP can only be removed by adjustments in the domestic price level. To the extent that such adjustments occur, cointegration is more likely than with a more volatile floating exchange rate buffeted by "news". Secondly, tests which proxy national price levels using the CPI reject the null less frequently than tests using a WPI (or PPI). This is probably because the CPI is more heavily weighted by non-traded goods and services than the WPI. Lastly, unconstrained trivariate tests (using, for example, the Johansen procedure) tend to reject the null more often than bivariate tests (as in the Engle and Granger method) where the symmetry restriction is retained.

The latest findings of more positive results using the Johansen technique have been the subject of a fair amount of debate and controversy. The main reason for this is that the maximum likelihood estimates of the *b* coefficients vary widely from one study to another. Moreover, the magnitudes of the estimates are often far too high to be given a plausible economic explanation. The difficulty in giving a sensible economic interpretation to the estimated coefficients has also been evident in cointegration tests of the monetary model of exchange rates. This discussion is thus carried further in chapter five (section 5.4). Froot and Rogoff (1995: 1667) conclude that the results of cointegration tests of PPP, besides being possibly misleading, do not yield any new insights from those of earlier unit root tests using long-horizon data sets.

Cointegration techniques are one way of improving the low power of early unit root tests. A more direct route is to use long-horizon data sets or bigger cross-section samples of exchange

rates. Frankel (1986; 1990) used annual data on the dollar/pound exchange rate and price levels for the period 1869-1984 and was able to reject the null of a unit root in the real exchange rate, using standard Dickey-Fuller tests. He estimated that real exchange rate deviations damped out at an average rate of 14 percent per year, implying a half-life of deviations from PPP of 4,6 years. His results find support from similar studies by Abuaf and Jorion (1990), Diebold, Husted and Rush (1991), and Cheung and Lai (1994). There is a surprising degree of consistency in the reported rates of reversion to PPP amongst these studies, with an estimated half-life of about 5 years for the dollar/pound exchange rate and about 3 years for most other currencies. Shiller and Perron (1985) argue that the power of cointegration tests is also improved by increasing the length of the sample period, more than by increasing the number of observations by using, say, monthly instead of quarterly data.

However, Rogoff (1996: 657) notes that all these long-horizon studies combine fixed and floating exchange rate periods. Since real exchange rates have been far more volatile under the latter (see Mussa 1986) it is still not clear whether PPP outperforms a random walk under these conditions, particularly for the post 1973 period where capital account flows have increased greatly relative to trade flows. These doubts are reinforced by the work of Meese and Rogoff (1983; 1984) amongst others who find that popular models of exchange rate determination (including the monetary model in which PPP is an important component) cannot outperform naive random walk models in out-of-sample forecasts for this period (see chapter five).

PPP tests based on expanded cross-country data sets have also produced mixed results. Hakkio (1984) was the first to suggest this approach to improve the power of unit root tests. He jointly tested four industrialized countries' exchange rates against the dollar but failed to reject the random walk hypothesis. More recent tests, however, have had better success in rejecting a random walk and finding evidence of mean reversion in real exchange rates (see Lothian 1994; Frankel and Rose 1995; and Wei and Parsely 1995). Virtually all of these

studies find evidence of mean reversion with a half-life of deviations from PPP between 4 and 5 years. This lends support to similar findings based on the long-horizon studies.

3.7 Implications for the methodology of economics

A central criticism of the methodology of economics noted in chapter two is that empirical tests of economic theories amount to little more than "innocuous falsification" or "playing tennis with the net down" as Blaug so colourfully puts it. The examination of the law of one price and purchasing power parity theory in this chapter well illustrate this tendency in economics. However, rather than deplore this practice and exhort economists to change their research strategy to fit the dictates of falsificationism (or any other methodology) a preliminary attempt is made here to explain why this is the case. The theories examined above provide some clues as to why economists have little alternative, given the nature of their subject matter, to conducting their research in this way.

At first glance the law of one price and theories of PPP appear to be good candidates for Popper's demand for bold, high risk theories with a high degree of empirical content (in the sense of excluding many possible observations or states of the world). The theories can be stated simply as mathematical formulas with a limited number of key explanatory variables, similar in appearance to that of a physical law. However, it is here that the similarity ends. As shown above most empirical tests have produced results that show significant deviations or exceptions from what is predicted by the theory. It is at this point, where empirical evidence meets theoretical predictions, that a seemingly endless process of qualifications, interpretations and amendments is initiated. In most cases these would be regarded by a follower of Popper or Lakatos as undesirable *ad hoc* changes to the theory, conventionalist stratagems which are made solely to immunize it from falsification. The question to be asked of falsificationists is: why haven't economists long ago discarded PPP theory? Why

do economists persist with apparently discredited theories? There are probably a combination of reasons for this practice.

In the first place, economists appear to regard the law of one price as an economic axiom rather than a genuine empirical law, in the same sense that the proposition "The shortest distance between two points is a straight line" is held to be a self-evident, intuitively certain or *a priori* truth. The response to a claim that this proposition had been tested and found wanting would lead to a re-examination of the conditions under which the test measurements were made, the reliability of the instruments used, or even the credibility of the "experimenter" rather than to construe the evidence as falsifying the axiom itself. Even if an exhaustive investigation and repeated experiments appeared to support the claim, the intuitive certainty of the axiom would not be diminished thereby. In the same way, evidence which appears to falsify the law of one price is usually not construed as such and perhaps explains why economists like Keynes and Samuelson regarded the law as a "truism".

It may be argued that economists do not in fact regard the law of one price as a self-evident truth, given the numerous empirical tests thereof. Genuine tests intended to either prove or disprove the law of one price would be superfluous if it was indeed intuitively certain and self-evident. The answer to this is that tests of the law of one price are, in the Duhem-Quine sense, tests of a complex of hypotheses. The axiomatic truth of the law is hedged about by a host of auxiliary hypotheses and it is essentially these hypotheses which the economist ends up testing. The statement that "The common currency price of an identical good must be everywhere the same" is subject to if-then qualifiers such as "transport costs are zero" and "there are no tariff or other barriers to trade" and "there are no exchange rate or arbitrage risks" and "arbitragers have perfect information". In other words, negative test results mean that it is one or more of these hypotheses that has been falsified, not the axiomatic law of one price itself. Empirical economic research thus clarifies the range of application and limitations of the law of one price but cannot result in a

genuine falsification thereof. A further consideration that may be significant in the appraisal of PPP is the stylized fact that for homogenous or standardized, highly traded commodities the law of one price appears to be a very close approximation to the available data (see section 3.4). Although this positive evidence has been more than outweighed by potentially refuting instances as regards more heterogenous manufactured goods (and at higher levels of aggregation), the limited confirmatory evidence may be sufficient on *a priori* grounds to treat anomalies in other markets innocuously rather than as falsifying the law itself.

Rather than be denied a fruitful avenue of empirical research into the integration of international markets, economists have responded rationally by adopting an inductive *ad hoc* strategy whereby, faced with damaging *prima facie* evidence, they have creatively devised a variety of auxiliary hypotheses to explain such anomalies. These hypotheses have guided continuing empirical research and further tests, in the process uncovering new anomalies leading to alternative auxiliary hypotheses and so on in an unending cycle. Of course one could always choose, as a matter of convention, to take the *prima facie* evidence seriously at some point and declare the law or theory "falsified". But this would simply lead to a sterile *cul-de-sac*. Understandably economists have not taken this approach, as revealed by the ongoing intensive empirical research into PPP (if they had taken falsification seriously what would have been the point of persisting so doggedly with successive waves of new 'tests')?

Moving from the the law of one price for individual traded goods to theories of PPP involving collections of goods raises similar difficulties and others besides. Whichever version of PPP is used, it is clear from section 3.4 that large and persistent deviations from the predicted (or, more accurately, *post*dicted) parity values of the exchange rate are the norm. To a great extent this results from the catch-all nature of the *ceteris paribus* assumption which is implicitly part of all such theories. The nature of the *ceteris paribus* clause in economics suggests that PPP should not be regarded as a deterministic law

and that, by adding a random error term, it is preferable to restate it as a probablistic or stochastic relationship, as in some of the theories discussed above. In reduced form:

$$S = f(P, P^*, u) \quad \text{.....(3k)}$$

where **P** and P^* are but two independent variables hypothesized to systematically influence the exchange rate and **u** is the random error term expressing the resultant of the vector of all other economic forces which may influence the exchange rate. If the error terms are serially correlated and the deviations from PPP are large then it implies that variables other than national price levels exert an independent systematic influence on exchange rates. By this interpretation, PPP theory is only "falsified" if these other variables can be specified and shown to have greater weight than price parities in determining the exchange rate. This is the modern statistical equivalent of the view by some 19th century economists that, unlike the deterministic laws of Newtonian physics, many economic relationships are really statements about tendencies and thus should only be thought of as "tendency laws".

Faced with damaging anomalies, the economist does not simply abandon PPP theory. Instead, the deviations are interpreted or explained by auxiliary hypotheses about the variables assumed constant by the *ceteris paribus* clause. For example, Hsieh's (1982) paper (section 3.5) sought to explain movements in the real exchange rate (which implies exchange rate disparities) by differences in productivity trends in the traded and nontraded goods sectors between countries. Frenkel (1981a), on the other hand, explained deviations from PPP by unexpected changes in interest rate differentials which produce volatile asset market adjustments in the exchange rate. A large number and variety of auxiliary hypotheses can be invoked to explain such anomalies. The revealed purpose of empirical research in this area is thus to discover and clarify the effect of other variables by means of these hypotheses. Economists have little alternative but to interpret and explain anomalies in this *ad hoc* fashion rather than embrace falsification as such.

All this points to a significant difference between a social science like economics and the natural sciences even if, as some critics maintain, it is only a matter of degree. As noted in chapter two, it is easier to make a clean incision of the natural environment and to isolate the key explanatory variables in a particular theory. In economics this is generally not possible as to a far greater extent "everything depends on everything else". Moreover, the exogenous variables held constant by the *ceteris paribus* assumption are often not specified. Thus, confronted by negative test results, the economist has little choice but to hypothesize which exogenous variables, or combinations thereof, have produced the anomalous results rather than conclude that the theory has been falsified. The argument that, as regards method, the difference between the subject matter of economics and that of, say, physics is just a matter of degree appears to be specious. This is like saying that the difference in size between a mouse and an elephant is also merely a matter of degree. By analogy, a mousetrap, however effective it is in catching mice, is unlikely to be very successful in delivering an elephant to the gamekeeper.

A unique difficulty in many areas of economics, entirely absent in the natural sciences, is that many economic theories include a hypothesis about the nature and operation of expectations (or, as did Keynes, they are treated as an exogenous variable within the ambit of the *ceteris paribus* clause). Frenkel's (1981a) work, for example, suggests that the exchange rate roughly follows a random walk because it is influenced by expectations which respond to the unpredictable release of "news". The exchange rate is seen as a price which adjusts rapidly to changes in expectations, to maintain stock equilibrium in asset markets. The exchange rate thus becomes completely determined, in the short-run, by speculative capital movements triggered by changes in expectations. It is then an open question what exactly determines expectations. Fundamentals such as PPP considerations are just one of many other variables that speculators may recognize in forming their expectations about exchange rates. Without a satisfactory model of expectations (see chapter six) genuine predictions in

important areas of economic theory are unobtainable. Negative test results indicate that either the structural model is wrong or the hypothesis about expectations is invalid - but it usually cannot be clearly discerned which part of the theory has been 'falsified'. Thus the full implications of the Duhem-Quine thesis for the testability of economic theories may have been underestimated by falsificationists like Blaug.

The operation of expectations in asset markets like the foreign exchange market also implies that in trying to construct structural models of the exchange rate we are attempting to predict the unpredictable. Speculative foreign exchange flows induced by changes in expectations are reflected in the current exchange rate which thus largely anticipates the future. Any new information believed relevant in determining exchange rates is quickly reflected by the foreign exchange markets in the spot exchange rate. Therefore, prediction of exchange rates based on a model of fundamentals like PPP and the attempt to falsify such predictions becomes a somewhat pointless exercise - the best prediction of tomorrow's exchange rate is simply today's spot rate (or perhaps today's forward exchange rate). Prediction in such markets is thus akin to prophecy, in which the prophet will (as Keynes observed) probably have better success at guessing what other speculators think is important than by trying to establish the fundamental determinants of asset prices like the exchange rate.

If PPP theories are supposed to help predict the exchange rate then the question of cause-and-effect also has to be resolved. If, for example, the exchange rate were undervalued by, say, 15 percent relative to its estimated PPP value one may be led to predict that the exchange rate would duly appreciate by roughly this amount. However, if the cause-effect relationship is from exchange rates to price levels it could be predicted that either the domestic price level will rise by 15 percent, or that the foreign price level will fall by this amount, or some combination of the two. In short, either the exchange rate could rise to meet its parity value or the parity value could fall to meet the exchange rate. As mentioned in section 3.2, similar questions were debated prior to Britain's decision to

return to the gold standard in 1925 at sterling's pre-war parity value. The mistaken expectation was that the British price level would have to fall only marginally and that most of the adjustment would take place via inflation of foreign prices.

The general approach to testing an exchange rate theory like PPP is to estimate the parameters of an econometric model and to determine the goodness-of-fit of the model to the time series data. Most structural economic models are estimated and tested this way. The power of the model thus depends on how well it *post*dicts the empirical evidence rather than its success in predicting, for example, what the exchange rate will be for any given relative price levels. Thus economic models are essentially backward looking rather than forward looking. Out-of-sample tests of the predictions of PPP models were not examined here but, as shown with the monetary model in chapter five, such predictions are often no more accurate than those of a naive random walk model. This is because the structural parameters of the models change over time, unlike the natural constants of the physical sciences. At a still lower level of knowledge where there is insufficient or no information about the future values of the independent variables, the forecasts of econometric models are notoriously inaccurate and unreliable. As regards PPP, for example, even if the direction of causation runs consistently from relative price levels to exchange rates and the parameters of the model do not change, a prediction of the long-run exchange rate would require a similarly long-run forecast of domestic and foreign inflation rates. The less accurate that forecasts of macroeconomic variables like the inflation rate are, the greater will be the error in predicting exchange rates. The widely acknowledged failure of macroeconomic models in this regard is well documented. Thus the attempt to second-guess the markets is akin to prophesy or fortune telling as it is based on the prediction of one unknown from another.

Some economists agree that changes in the exchange rate are unpredictable in the short-run but maintain that they are predictable in the long-run and that PPP theories are useful

for this purpose. Deviations from PPP are thus regarded as temporary aberrations which will disappear with time. Some, but not all, of the very recent cointegration estimates of the PPP relationship discussed above lend some credibility to this view. However, the long-run is rarely defined operationally. If no benchmark time interval is specified against which the significance of persistent deviations can be measured, it is unclear what exactly is meant by "the long-run". The long-run becomes an empty analytical concept, the import of which is something like "The length of time necessary for the exchange rate to conform to what my model predicts". It is invoked when economists have no good alternative explanation for anomalous test results.

A feature of empirical research into PPP is the extent to which it has been conditional on the development of new techniques in econometrics - that is, methodology with a small 'm'. The basic PPP theory, although it may be expressed in various forms and has been interpreted in sometimes subtly different ways, has not changed fundamentally since Cassel's formulation in 1916. What has undergone significant successive changes are the econometric techniques used to test the theory. At least six distinct methods can be discerned (with examples of specific studies or tests in parentheses):

1) The basic OLS regression tests of disaggregated categories of internationally traded goods (Richardson 1978).

2) Estimation of the simple PPP regression equation (3g). If the alternate hypothesis b = 1 is accepted, the findings are favourable for PPP (Frenkel 1978a ; 1981b).

3) Early unit root tests on real exchange rates where the inability to reject the null hypothesis that $b = 1$ is unfavourable for PPP, that is, there is no evidence of long-run mean reversion and the real exchange rate cannot be distinguished from a random walk (Adler and Lehman 1983).

4) Dickey-Fuller unit root tests and other tests for distinguishing the real exchange rate from a random walk. If

the null cannot be rejected there is no evidence that PPP holds as a long-run equilibrium relationship (Meese and Rogoff 1988).

5) Engle and Granger cointegration tests of PPP in multivariate form (equation 3i). If the null hypothesis of nonstationarity in the residuals cannot be rejected then prices and exchange rates are not cointegrated thus also rejecting PPP as a long-run relationship (Taylor 1988a).

6) Johansen unconstrained maximum likelihood cointegration tests with the null as above (MacDonald 1993).

At each stage, the empirical evidence resulting from tests of PPP using these econometric innovations has been at best mixed. In most cases the test results have been negative (this has also been the case with the monetary model of exchange rates - see chapter five). Yet, significantly, they have never been interpreted as falsifying the basic PPP theory itself. All one could say is that for a *specific test* the null hypothesis had been rejected (or had failed to be rejected in the case of real exchange rates). In other words, the econometric tests have been used to clarify and explore the range of application of the theory under particular circumstances. Such 'test' results are thus always provisional and never amount to a genuine falsification of the theory.

Models of exchange rate determination, like most other economic models, are historically circumscribed: a model which explains exchange rates for a particular time and place may fail, and as evident from section 3.4 usually does fail, for another time and place. The discussion in this chapter thus suggests that there are grounds for regarding economics as an explanatory science (or discipline) rather than a genuine predictive science. Popper's deductivist resolution of the problem of induction by the logic of falsification outlined in chapter two may, therefore, be less appropriate in a subject like economics than in mathematics, physics, chemistry or biology.

Another way of putting all this is to invert the supposed methods characterizing the logic of justification and discovery

respectively. The falsificationist view is that the logic of discovery is based on the intuitive, creative acts of the imagination - the method *a priori* (Popper's bold, high risk theories); whereas the logic of justification is based on exclusively empirical considerations (which even Lakatos'softened version of falsification does not avoid). In economics the reverse logic appears to be just as plausible: an inductive logic of discovery (empirical research used as an exploratory tool) and an *a priori* or deductive logic of justification based (or perhaps a messy combination of the two methods in both the discovery and justification of economic theories).

The basic differences between falsificationist methodologies and Larry Laudan's problem solving model were outlined in chapter two. In concluding this chapter, some remarks concerning Laudan's model and his conception of scientific progress and rationality are in order (see also the methodology sections of chapters four and five). The empirical tests of PPP examined in this chapter would not, for the most part, meet the prescriptive criteria of falsificationism (nor even the purely descriptive characterization of science in Lakatos' version thereof). The unavoidably *ad hoc* inductive nature of much of this research (in the sense of Lakatos' *ad* hoc_{1-2}) is, however, quite consistent with Laudan's more pragmatic emphasis on problem solving. To the extent that face saving auxiliary hypotheses or innovative econometric techniques are able to solve existing empirical problems (in a theoretically consistent non-*ad* hoc_3 fashion), then this represents scientific progress and economists should be deemed to be using rational research strategies. Because Laudan is less concerned with transcendental objectives like "truth" than with the more pragmatic aim of problem solving, adoption of his model allows us to avoid the perverse conclusion of falsificationists that the best practice research strategies of economists are irrational.

Notes

1 Unless stated otherwise, the exchange rate is expressed as the domestic price of the foreign currency unit. Also, uppercase letters refer to levels of the variables and lowercase letters to the rates of change thereof.

2 For arithmetic convenience, econometric models often express regression equations in logarithmic form. Absolute PPP would thus be expressed as:

$$\mathbf{ln}S_t = a + b\mathbf{ln}P_t - b^*\mathbf{ln}P^*_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where **ln** represents natural logarithms. The relative version of PPP may be written as:

$$\Delta\,\mathbf{ln}S_t = b\,\Delta\,\mathbf{ln}P_t - b^*\Delta\,\mathbf{ln}P^*_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

where Δ is the first difference operation symbol. The constant term *a* in the relative version of PPP is hypothesized to be zero, as implied by equation (2). Strict PPP, where the proportional change in the exchange rate equals the proportional change in relative price levels, implies that $b = b^* = 1$.

3 Britain had effectively been on the gold standard since 1717, when the Bank of England's selling price for gold was set at 77s.10 1/2d per ounce, except for a short period during and after the Napoleonic Wars when it was restored in 1830 (see Moggridge 1969: 11). America had persevered with a bimetallic silver and gold standard during the 19th century. It was only from about 1870 that gold triumphed over silver to become fully acceptable as the international payments standard.

4 McCloskey and Zecher (1976) argue that the operation of the gold standard was not the sole or even central mechanism of adjustment in prices and interest rates amongst those countries on the gold standard. They maintain that the 1880-1913 era of *laissez faire* capitalism was characterized by integrated global markets, unified by extensive arbitrage in traded goods; and that intervention by central banks and the small flows of gold between these countries were not an indispensible part of the arbitrage process:

"The world's economy determined the prices and interest rates prevailing in each nation's economy and it was the flow of gold itself that re-established equilibrium in the money market by satisfying the demand for money that prompted the flow in the first place" (McCloskey and Zecher 1976: 363).

The assumption of competitive global markets in which the law of one price holds continuously is a feature of global monetarism, discussed further in chapter four.

5 Moggridge (1969: 12) notes that although the gold standard still existed in a legal sense, these and other restrictions meant that it had effectively ceased to exist. The refusal by the British government to include gold under the war risks insurance scheme was crucial in this regard as it prevented private arbitrage and the shipping of gold abroad.

6 A significant factor contributing to the downfall of the gold standard was the decision by France in 1928 to return the franc to a gold par value of only 20 percent of its pre-war value (Cassel 1936: 46-51). By virtually any measure this implied a large undervaluation of the franc which put other trading nations such as Britain at a competitive disadvantage. Other deflationary forces included the Wall Street stock market crash of 1929 (and the overly cautious monetary response to it); the ultraprotectionist Smoot-Hawley Tariff Act of 1930 in America; and the halting of further reparations payments by Germany.

7 This finding contradicts those of Curtis (1971) and Dunn (1970; 1972) which suggested that Canadian prices respond less to exchange rates than to US prices.

8 Nurkse (1944) assessed the brief inter-war period experience with floating exchange rates and opened the debate about stabilizing *versus* destabilizing foreign exchange speculation. He concluded that a floating exchange rate system was vulnerable to destabilizing speculation and that the exchange rate should not be determined solely by the unfettered market forces of supply and demand. During the Bretton Woods era, economists such as Friedman (1953: 157-203), Baumol (1957), and Sohmen (1969: 59-80) were of the view that profitable speculation would be stabilizing and recommended the replacement of fixed exchange rates with a flexible exchange rate system.

9 Tests of PPP usually do much better when high inflation economies are studied. See, for example, Frenkel (1978b; 1978c). Apart from hyperinflations, however, most standard OLS regression tests have rejected PPP.

10 The ADF test includes a term specifying lagged changes in R_{t-1}. If the coefficient on this term is estimated to be significantly different to one it implies serial correlation of the DF residuals, indicating the presence of non-random changes in the real exchange rate. The confidence intervals against which the significance of the estimates in equation (3h) are tested are wider than for standard OLS estimates. The appropriate intervals are given by Dickey and Fuller (1979). Besides the DF tests, other techniques used to judge whether the real exchange rate follows a random walk include variance test ratios (see Poterba and Summers 1986) and fractional integration tests (see Diebold, Husted and Rush 1991).

11 A stationary variable has a uniform variance over time. Most economic time series are nonstationary because they have a time

trend. However, a linear combination of two or more nonstationary variables may exist which produces a stationary time series. If so, the variables are said to be cointegrated and should not drift too far apart in the long-run.

12 For example, there may be a trend in the relative price of traded and nontraded goods or the index of nontraded goods prices may be subject to the familiar index problems of fixed weights and new goods bias. See Bryant and Cecchetti (1993) who show how the presence of these problems can lend an upward bias to measured inflation.

13 Step one tests the exchange rate and the two price level series for unit roots using the ADF technique. If the random walk hypothesis cannot be rejected, step two estimates the cointegrating regression equation (3i) using OLS. Step three tests the null hypothesis of stationarity in the OLS residuals obtained from step two. A DF test (with the time trend omitted) is used to test for the null that $b_2 = 1$. If the null cannot be rejected then the exchange rate and prices are not cointegrated whereas if $b_2 < 1$ they are cointegrated under the alternative hypothesis.

14 There is a close connection between cointegration models and earlier error correction models (which relate one variable to changes in another variable as well as to the past levels of the two variables, and which can estimate the mean rate of reversion in deviations from a long-run relationship).

CHAPTER FOUR

THE DEVELOPMENT OF MACROECONOMIC THEORIES OF THE BALANCE OF PAYMENTS AND EXCHANGE RATE DETERMINATION

This chapter examines some important milestones in the development of theories about the balance of payments and exchange rate determination since the Second World War, from a methodological perspective. The emphasis, however, is on the monetary and asset market approaches which have preoccuppied economists working in this area for most of the recent era of floating exchange rates. This is followed up in chapter five which considers selected empirical issues as regards the monetary and asset market models of exchange rate determination.

Section 4.1 places the monetary approach in historical context by comparing it with the elasticities, absorption, and Mundell-Fleming models. Section 4.2 looks more closely at the transition from the orthodox approach, in particular the Mundell-Fleming model, to the monetary approach and the analytical differences between them. Section 4.3 describes the basic monetary model and contrasts the main conclusions and implications thereof against those derived from the earlier flow models of exchange rates. The role and significance of expectations in the determination of exchange rates is also introduced. Section 4.4 briefly surveys four main variants or extensions of the basic monetary model: global monetarism; the overshooting exchange rate hypothesis; the real interest rate differential model; and the portfolio balance model. Section 4.5 draws out some preliminary methodological implications from the foregoing regarding the theoretical development of the different analytical models. The relationship of the monetary approach to the broader research programme of monetarism is also discussed. A more complete evaluation and methodological analysis is postponed until the empirical issues have been addressed in chapter five.

4.1 Historical context of the monetary approach

The modern monetary approach to the balance of payments emerged in the early 1970s as a serious rival to the then orthodox Keynesian oriented income-expenditure and relative price or "elasticities" approaches, which had dominated thinking about balance of payments adjustment during the Bretton Woods era of pegged exchange rates. With the decision by the Nixon administration to end official US dollar-gold convertibility in 1971, and the adoption of floating exchange rates by the major western industrialized countries in 1973, attention shifted to the determination of exchange rates. The conceptual framework and insights of the monetary approach were soon applied to flexible exchange rates.

Two books, *The Monetary Approach to the Balance of Payments* and *The Economics of Exchange Rates*, both edited by Jacob Frenkel and Harry Johnson (1976; 1978) proved enormously influential at both a theoretical and, indirectly, a public policy level. The two books collated the most important research work of the monetary approach, much of which had been published earlier as papers in various journals or more general books on international trade and monetary economics. These collections of papers explain the basic theory of the monetary approach, its roots in earlier research traditions, and other doctrinal aspects. As pointed out by Frenkel, the monetary approach was not to be thought of as a radically new or "revolutionary" approach in international monetary economics since many of its insights could be found in the works of the early classical economists and later up until the early 1930s:

"It is appropriate, therefore, to view the recent revival of the monetary approach as a natural evolution rather than a revolutionary change in views" (Frenkel 1978b: 6)

and:

"The continuity of its development, however, was reversed and the approach suppressed in international economic theory for upwards of a quarter of a century by the events of the 1930s" (Frenkel and Johnson 1976: 29).

These events included the exceptional international monetary and economic collapse of the Great Depression in which adjustments to quantities of output and employment cleared markets rather than changes in wages and prices. Thereafter, the increasing influence of Keynesian thinking in monetary economics saw it become established as the economic orthodoxy for most of the post Second World War era. For most of this time controls over international trade and capital movements obscured the underlying process of international monetary adjustment.

These developments meant that the full implications for the monetary view of PPP discussed in chapter three as regards theories of the balance of payments and exchange rates, were neglected. For most of the Bretton Woods era of pegged exchange rates, attention was focused on the effects of devaluation on the trade balance and latterly the indirect effects on domestic and foreign spending, output and employment. This analysis took place within a broadened Keynesian model of income-expenditure flows applied to the open economy, in which devaluation and changes in net exports altered the level of national income and employment via the foreign trade multiplier. The conceptual and analytical flaws inherent in the Keynesian approach were obscured by the insularity (see McKinnon 1981) of the national economies which emerged from the Second World War. Only with the later removal of trade and exchange controls in the 1960s and the increasing global integration of national economies did these flaws become more apparent and in need of a satisfactory theoretical explanation.

Within the Keynesian theoretical framework, three distinct stages in the evolution of thinking about balance of payments adjustment and exchange rate determination can be identified. In rough chronological order these may be referred to as the relative price or "elasticities" approach to devaluation; the income-expenditure or absorption approach; and the Mundell-Fleming model which integrated capital as well as trade flows and which represented the most sophisticated Keynesian theory of the open economy. These are discussed briefly to put the monetary approach in historical context and to compare and

contrast the different models from a methodological perspective. The elasticities and absorption approaches are concisely explained and evaluated in Michaely (1960), Johnson (1977), and Krueger (1983). The seminal papers of the Mundell-Fleming model, which is described in virtually all modern textbooks of international economics, are Mundell (1960; 1961; 1962; 1963) and Fleming (1962).

The elasticities model was originally developed by Alfred Marshall after the First World War as an application of his partial equilibrium supply and demand analysis to the foreign exchange market (Marshall 1923). Consider a country facing a large trade deficit. Under a fixed exchange rate and without compensating capital inflows it will experience an outflow of foreign exchange. If foreign exchange reserves are limited the country will thus face a balance of payments crisis. In this situation devaluation may be used in an attempt to improve the trade balance and to avert the payments crisis. The success of such a policy depends on the responsiveness or elasticity of imports and exports to changes in the currency's external value.

The following simplifying assumptions are usually made:

a) A two country world in which both countries are small, open economies. This assumption ensures that devaluation cannot change the foreign price of goods imported by either country - each country is a price taker for the foreign goods in which it trades.

b) Devaluation causes a decline in the terms of trade - the relative price of exports in terms of imports decreases. In models which assume competitive world markets unified by arbitrage, like those falling under the global monetary approach, the law of one price holds such that the terms of trade are exogenous and cannot be changed by devaluation. However, these models may allow some other set of relative prices to change, for example between traded and nontraded goods.

c) The capital account is ignored and the balance of payments is equated with the trade balance.

d) Income and expenditure effects and relationships are ignored.

In a simple partial equilibrium model where the terms of trade are endogenous, the volume of imports by the home country is necessarily reduced by the devaluation if they are normal goods. Thus derived foreign currency payments are also reduced. However, the effect of devaluation on foreign currency receipts from exports is indeterminate. The lower foreign prices paid for home exports reduces such receipts while greater foreign demand increases them. The net result depends upon the price elasticities of home demand for imports (e_m) and foreign demand for exports (e_x). The two elasticities may compensate for each other but the condition for an improvement in the trade balance is that the sum of the elasticities must be greater than one ($e_m + e_x > 1$). This is the familiar Marshall-Lerner condition. Put another way, the improvement in the net volume of trade must be sufficient to compensate for the fall in the value of trade caused by the devaluation. If assumption (a) is relaxed to allow for a less than perfectly elastic supply of exports a more complex condition for an improved trade balance results[1].

If the elasticities model is applied to floating exchange rates then the Marshall-Lerner criterion becomes the condition for stability in the foreign exchange market. The exchange rate is determined by the supply of and demand for the foreign currency, derived from the flows of exports and imports respectively. If the price elasticities sum to less than unity then any disturbance to an initially balanced trade account would become magnified - a growing trade deficit would coexist with a continually depreciating exchange rate (and *vice versa* for a trade surplus). Alternatively, even if the Marshall-Lerner condition were satisfied the elasticities could still be very low, requiring large exchange rate adjustments to equilibrate the market.

A noteworthy feature of the elasticities model is that it sees the balance of payments as a flow rather than a stock phenomenon, with relative prices being the major influence on flows. The flows of foreign exchange per unit time are derived from the ongoing flow demands for imports and exports. Thus an implication of the model is that devaluation should permanently improve the trade balance (if the Marshall-Lerner condition is met) because it permanently decreases the flow demand for imports while increasing that of exports.

As the Keynesian research programme gathered momentum after the Second World War, the relevance of the elasticities approach was increasingly questioned as an adequate explanation of balance of payments behaviour. Keynes' income-expenditure model of the economy, a macroeconomic model which made the interrelatedness and general equilibrium of markets its central feature, had major implications for external as well as internal equilibrium. In particular, it was realized that devaluation (or depreciation under flexible exchange rates) may have secondary consequences which could at least partially reverse any improvement in the trade balance resulting from a change in relative prices. Given the Keynesian assumptions of sticky wages and prices and underemployment, devaluation would increase demand and expand output in the export and import substitution industries. This initial expansion, in turn, generates successive increases in income and expenditure via the multiplier mechanism. As aggregate spending increases, imports rise according to the marginal propensity to import.

Unless wages and prices are perfectly inelastic, the increase in aggregate demand also tends to push up marginal costs and prices. Devaluation also raises import prices directly which further increases production costs. Thus, in the longer-run at least, the competitive edge gained by devaluation would be eroded. This further encourages imports while reducing the demand for exports. Together these reversal effects might, for a small open economy, completely offset the relative price effect. Moreover, if the opposite classical assumptions of full employment and perfectly flexible prices are made, output cannot rise to meet the increased demand for goods produced by

the export and import substitution industries. In this case, costs and prices will rise instead to completely offset the competitive advantage bestowed by devaluation or depreciation. Only if domestic spending is reduced commensurately can output in these industries expand without self-defeating rises in wages and prices, thereby allowing the trade deficit to improve. The elasticities approach ignores such income-expenditure effects and can thus only offer a partial explanation of balance of payments adjustment.

These strands of thought were pointedly brought together by Sidney Alexander (1952) using the concept of absorption. He noted that the difference between the value of output (or income) generated by the economy and total domestic expenditure (saving) must *ex post* equal the trade balance, by definition:

$$\mathbf{Y - A = X - M = B} \qquad \text{(4a)}$$

where **Y** is total output or income, **A** is aggregate domestic expenditure (absorption), **X** and **M** are the values of exports and imports respectively, and **B** is the trade balance. **A** includes the usual components of domestic aggregate demand: consumption, investment and government spending.

Independently of each other the Australian economists Salter (1959) and Swan (1960) developed the theory of the policy mix required for internal and external balance to be achieved simultaneously. Since capital flows were ignored this was essentially a theory of the current account with only two policy instuments: expenditure/absorption, which could be altered by discretionary fiscal policy, and devaluation. Two cases may be considered here. Where the economy faces both a trade deficit and unemployment, it presents an ideal situation for devaluation/depreciation to be effective. In this case, devaluation stimulates greater employment in the export and import substitution industries without raising wages and prices, and the trade balance improves (if the Marshall-Lerner condition is met). However, with fully employed resources devaluation alone cannot improve the trade balance. If a trade deficit coexists with full employment, devaluation will be

effective only if combined with restrictive fiscal and/or monetary policies which reduce aggregate domestic spending. Without such measures the devaluation simply results in a higher domestic price level and the terms of trade remain unchanged. It should be noted that there is no logical or conceptual contradiction between the elasticities and absorption approaches. Only the operation of intervening background variables need be made explicit for the two approaches to be consistent (see Michaely 1960 and Dornbusch 1975).

Although the absorption approach has the advantage of allowing a macroeconomic general equilibrium analysis of the external sector, balance of payments adjustment is still viewed as a flow per unit time rather than reflecting discrete adjustments of stocks of assets. It is assumed, for example, that devaluation (if it succeeds in altering the terms of trade) will permanently improve the trade balance in subsequent periods so long as domestic aggregate demand is restricted by some combination of monetary and fiscal policy. In the event of a trade surplus, the effect of the accumulation of foreign exchange reserves on domestic asset markets is ignored or assumed to be sterilized by the monetary authority. Trade deficits, on the other hand, are assumed to continue until the country runs out of foreign exchange reserves unless discretionary expenditure switching and/or expenditure reducing policies are used. The effect of the deficit on the domestic money market and the extent to which other countries are willing to accumulate stocks of the national currency is not explained. These features result from the absence of any explicit consideration of the role of money in the balance of payments adjustment process.

By ignoring the capital account of the balance of payments, the elasticities and absorption approaches imply that external equilibrium is reached only when the trade balance is zero. Economists such as Meade (1951) and Swan (1960) showed that with a fixed exchange rate and rigid prices, the two objectives of internal and external balance might be incompatible no matter what combination of monetary and fiscal policy was used.

Trade balance, for example, may require domestic unemployment. However, the liberalization of exchange controls and the greater international mobility of capital which occurred from the early 1960s helped provoke a modified theoretical approach to the problem of internal and external balance. As in Mundell (1961; 1962; 1963) and Fleming (1962), a natural escape route is to introduce the role of capital flows in compensating for trade imbalances and thus remove any inconsistency between employment and balance of payments objectives. Unless capital flows were assumed to be perfectly interest rate inelastic, Mundell showed that internal balance could be achieved using fiscal policy measures and external balance by an appropriate monetary policy. The Mundell-Fleming model also shows the relative efficacy of monetary and fiscal policy under fixed and flexible exchange rates.

The Mundell-Fleming model is essentially an extension of the Hicks-Hansen IS/LM analysis to an open economy, with capital flows included as an integral part of the theory. The model remains to this day as a standard textbook analysis of the open economy under the Keynesian assumptions of a rigid price level and less than full employment. It is thus not necessary to fully explain the mechanics of the model and adjustments to various policy measures. For the purposes of this discussion, however, the FF curve in figure 4.1 below does require further comment:

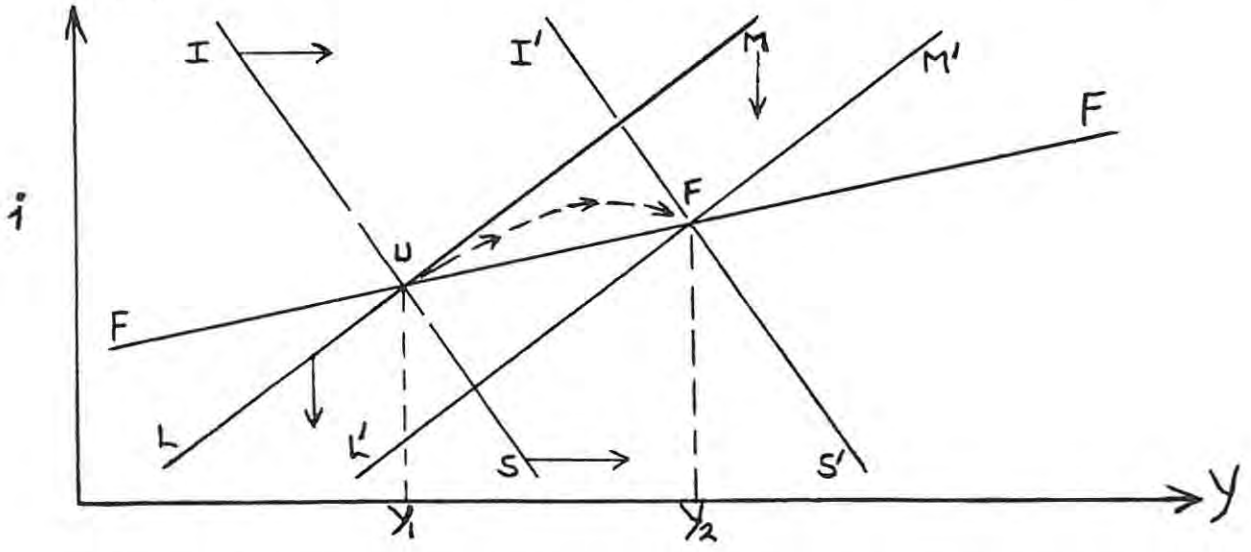

Figure 1: The Mundell-Fleming model of an open economy with inelastic capital flows

All points along the **FF** curve represent a balanced (zero) net foreign payments position arising from the summation of trade flows and capital account flows. If capital flows are less than perfectly elastic as in figure 4.1 then the **FF** curve slopes upwards to the right, signifying that a reduced trade balance resulting from a higher level of national income may be offset by capital inflows attracted by a higher interest rate. In this case it is possible to achieve external balance without sacrificing full employment (at Y_2) by applying an appropriate mix of fiscal and monetary policy which shifts the **IS** and **LM** curves as illustrated in the diagram. The conclusions of the model are that under a fixed exchange rate, fiscal policy is more effective than monetary policy in achieving a higher level of national income and employment, but vice versa under a flexible exchange rate system[2]. The less the interest rate elasticity of capital flows (the greater the slope of the **FF** curve) the less the disparity in the potency of monetary and fiscal policy.

4.2 The transition to the monetary approach

It is fair to say that the modern revival of the monetary approach centered on the research activities of economists associated with the University of Chicago, although significant contributions were also made by researchers at other institutions such as the International Monetary Fund, the London School of Economics, the Graduate Institute of International Studies of Geneva, and the Board of the Federal Reserve Bank. A central figure at the University of Chicago credited with "the core of the development of the monetary approach as commonly understood" (Frenkel and Johnson 1976: 31) was Robert Mundell. It is of some interest in the history of economic thought that the same person who helped develop the most sophisticated Keynesian model of the open economy, was also a key figure in the development of its most powerful rival[3]. The change in Mundell's thinking evolved gradually and was scattered in various journal articles, later collated in his two famous textbooks, *International Economics* and *Monetary Theory*:

"The theory of the fiscal-monetary mix, though important for policy theorising in the early 1960s, was essentially a by-product of Mundell's gradual realisation that with capital mobility the money supply ceased to be exogenously determined by the monetary authority, and monetary policy's prime role was to influence the international flow of reserves. This realisation was accompanied by a switch of analytical method from the Keynesian multiplier analysis to analysis in terms of the Walras' Law relation between excess demands and supplies in various markets, the need for equilibrium simultaneously in two out of three markets in a three market system, and the dynamics of reaction of a system out of equilibrium, a method of analysis popularised by Patinkin, but going back to Hicks' *Value and Capital*" (Frenkel and Johnson 1976: 32).

The integration of capital flows by Mundell and Fleming, albeit within a Keynesian theoretical framework, thus led to the question of what would happen if international capital flows were perfectly mobile. Under a fixed exchange rate, the implication of perfectly interest rate elastic capital flows is that the money supply becomes exogenous. The monetary authority cannot alter total reserves but only the division between domestic and foreign reserves. A flexible exchange rate, on the other hand, is completely determined by a perfectly elastic supply of capital, in which case the trade balance is a redundant explanatory variable. This is tantamount to accepting the essential irrelevance of trade flows in determining the exchange rate if international capital is perfectly mobile. Furthermore, the Mundell-Fleming model was unable to explain why desired capital movements do not take place instantly in response to changes in interest rate differentials, rather than as a continuing flow per unit time:

"Conceptually, there are difficulties with treating the flow of assets as a function of the interest rate or interest rate differentials. Presumably, the desired level of asset holdings is a stock, and one would have to explain why stock adjustment to the desired level was not instantaneous" (Krueger 1983: 82).

It was these and related conceptual difficulties which led to the abandonment of the Mundell-Fleming model, barely ten years after it had been developed. It is perhaps somewhat ironic, and instructive from a methodological view, that the crowning achievement of the Mundell-Fleming model - the explicit treatment of capital flows in a theoretically consistent model and its derived policy conclusions - contained the seeds of its

own demise. As global economic integration increased through the 1960s and 1970s the full implications of more elastic capital flows and the importance of stock adjustments became increasingly apparent and set the stage for the modern revival of the monetary approach.

It should be noted from a historical perspective that the conceptual difficulties of the Mundell-Fleming model were inherited from James Meade's (1951) earlier monetary synthesis of the relative price and Keynesian income-expenditure analyses of the balance of payments. At a purely theoretical level the conditions for the insights of the monetary approach to take root were arguably present much earlier than the development of the Mundell-Fleming model:

"The modern revival of the monetary approach may be said to have begun, in an important but indirect sense, with James Meade's *The Balance of Payments*" (Frenkel and Johnson 1976: 30).

Meade's model was regarded by Mundell as a valuable contribution to international monetary theory but which remained conceptually flawed at a fundamental level of analysis:

"That Meade's own model, buried in the mathematical supplement to *The Balance of Payments*, incomplete and unsatisfactory as it is, is still one of the best attempts at the integration of monetary and real analysis in the theory of international trade, is the most eloquent testimony to the importance of his contribution and at the same time the most devastating criticism of the current state of theoretical work in international monetary economics" (Mundell 1976: 67).

The problem with Meade's analysis was, specifically, that he confused the marginal propensity to save (related to income flows) with hoarding (a temporary stock adjustment of assets). Although he distinguished between stock and flow adjustments across the capital account in the text of his book[4], he did not reflect this in the model he set out in the mathematical supplement thereto (Meade 1951: 103). The Mundell-Fleming model is based on Meade's mathematical exposition and thus similarly fails to distinguish between stock and flow adjustments to changes in interest rate differentials.

Also, by defining monetary policy as the setting of the level of interest rates, Meade implicitly assumed that the monetary consequences of trade imbalances are absorbed or sterilized by the monetary authority. Maintaining a fixed interest rate implies, for example, that the inflow of money from a balance of payments surplus is siphoned off by open market sales of government securities or by some other means.

The systematic and schematic presentation of Meade's analysis, although unsatisfactory from an economic point of view, made it easier for later economists to detect its fundamental weaknesses or to extend the analysis where it was incomplete, as in the Mundell-Fleming model. In a sense, therefore, the Meade synthesis had already suggested the source of the basic conceptual problems with a Keynesian analysis of the balance of payments and, indirectly, set the stage for the revival of the monetary approach.

Besides the contribution of Mundell to monetary theory in an open economy, two papers by Johnson (1961; 1972) were especially influential in the transition to the monetary approach. *Towards a General Theory of the Balance of Payments* was an early classic which emphasized the monetary aspects of balance of payments adjustment and the significance of the distinction between stock and flow equilibrium in this process, particularly as regards the policies used to correct a deficit. Johnson did this by generalizing the implications of the absorption approach more fully and by relating the deficit to the operation of the economy as a whole:

"The distinction between 'stock' and 'flow' balance - of - payments deficits is important for both theory and practical policy, though refined theoretical analysis has generally been concerned with 'flow' deficits, without making the distinction explicit" (Johnson 1976a: 52).

As discussed above, this was regarded as a central weakness of the Meade-Mundell-Fleming analysis. Johnson's (1972) paper, *The Monetary Approach to Balance-of-Payments Theory*, contrasted the emergent monetary approach with the dominant Keynesian model more sharply, particularly as regards the policy recommendations of the two models (see section 4.3):

"The monetary models suggest that it may be very misleading to rely on the Keynesian model as a guide to policy-making over a succession of short-run periods within each of which the Keynesian model may appear to be a reasonable approximation to reality" (Johnson 1976b: 166).

During the 1970s, the monetary approach was applied to the determination of floating exchange rates and has established itself (as the more general asset market view of exchange rate determination) as the dominant explanatory model. In most respects, the monetary model of exchange rates is symmetrical with the monetary approach to the balance of payments. In broad terms, flexible exchange rates are determined by the amounts or stocks of national monies relative to the demands or willingness to hold them, and the domestic money supply is exogenous; while under a fixed exchange rate, relative money supplies and demands determine the balance of payments and the money supply becomes endogenous. However, the role played by expectations and speculative capital movements in determining exchange rates has called forth an asset market view of the foreign exchange market which is not entirely symmetrical with the earlier monetary approach to the balance of payments. For this reason it is perhaps more accurate to regard the post 1970s monetary models as a branch of the more general asset market approach to exchange rate determination (see Dornbusch 1980 and Mussa 1984).

4.3 The simple monetary model and its comparison with flow models of the balance of payments and exchange rates

The monetary approach to the balance of payments and exchange rate determination includes a variety of models which differ according to the assumptions or auxiliary hypotheses made about variables such as: wage and price flexibility; aggregation and classification of goods such as traded versus nontraded goods; relative speeds of adjustment in goods and asset markets; and expectations. As with any exercise in product differentiation, however, a basic design or family resemblance exists. This section outlines the essential assumptions, characteristics, and implications of the monetary approach. Selected examples

are used in section 4.4 to show how the basic model has been extended. As noted above, monetary models are a subset of the more general asset market approach. The crucial simplifying assumption demarcating the monetary model is that domestic and foreign bonds are perfect substitutes such that money is the only asset that is explicitly considered. The subset of models which relax this assumption form the more recent portfolio approach and are discussed briefly in section 4.4.

i) *Fixed exchange rates.* Under a fixed exchange rate system such as a gold standard, the monetary approach becomes a theory of the balance of payments and of how, for example, a devaluation will effect it. The approach derives from David Hume's (1752) price-specie-flow mechanism, although it differs in some important respects as discussed below. Hume showed that the quantity of money in a country would be adjusted automatically to the demand for it, through surpluses or deficits in the balance of payments. For instance an excess national supply of money would, according to Hume, result in an increase in the domestic price level relative to abroad. The increased relative price of home goods would then lead, via arbitrage, to an increased inflow of goods imported from abroad (and possibly also a reduction in exports) and thus a worsening of the balance of trade. An equivalent outflow of money (gold specie) would occur, in payment for the excess of imports over exports. This process would necessarily continue until the excess supply of money was eliminated so that, in equilibrium, the equality in the demand for and supply of money would be restored. Hume's intention was to show that attempts along mercantilist lines to generate a balance of payments surplus and the accumulation of national 'treasure' could only be successful temporarily and must be self-defeating in the long-run. The inflow of money corresponding to such a surplus would, in due course, bring about a reverse outflow of money in the manner just described thereby eliminating the surplus.

The monetary approach differs in two key respects from Hume's price-specie-flow mechanism. First, the overall balance of payments including both trade and capital account items is assumed to adjust in response to money market disequilibrium,

not just the balance of trade alone. Second, the monetary approach downplays differences in relative price levels as the cause of changes in the balance of trade. This is especially true of global monetarism where the common currency prices of traded goods (and of assets) are assumed to obey the law of one price. Instead, the monetary approach emphasizes the direct effect of excess money supply (or demand) on the balance between the total sources of funds or receipts (from income and borrowing) and the total uses of funds or payments (from consumption and lending), and thus the balance of payments as a whole. In this respect it is similar to the absorption approach with which, as noted by Johnson, it can be reconciled under certain conditions. Despite these differences, however, the monetary approach retains Hume's idea of payments imbalances reflecting temporary stock disequilibrium in the money market and the automatic, self-correcting nature of such disequilibrium.

In common with domestic monetarism, the monetary approach regards the demand for money to be a stable function of a limited number of variables. As noted in chapter five, one of the main reasons given for the well-documented empirical failure of the monetary model has been the hypothesized instability of the demand for money. Since this is also a key concept in the monetary approach, a brief explanation of Milton Friedman's (1956) restatement of the quantity theory of money is necessary. Friedman's basic postulate is that money is but one form in which wealth may be held and that the analysis of the demand for money is similar to that of the demand for a normal good. Thus the demand for money depends on (i) The total amount of wealth, equivalent to the budget constraint in price theory (ii) The real return on money and alternative assets and (iii) The preferences of wealth holders. The only qualitative difference from the demand for goods is that prices of the assets in (ii) are expressed as the inverse of their real returns per unit time. There is a link between income flows and the stock of wealth in that $\mathbf{Y} = \mathbf{r}.\mathbf{W}$ where $\mathbf{Y}$ is income, $\mathbf{r}$ is the total rate of return on assets, and $\mathbf{W}$ is the stock of wealth.

Friedman distinguishes five assets between which wealth may be divided: money; bonds; equities; goods; and human capital. Each of these assets commands a rate of return or yield either as services in kind or money, ideally viewed as a stream of income as defined above and maintained over the lifetime of the owner. Although money balances may earn interest income, money is generally valued for the stream of services that it renders in the form of convenience, liquidity, and security[5].

The demand for money depends on the competing returns available on the other assets. The higher the returns on these assets the lower the demand for money and vice versa. Total income is defined as above and where wages are regarded as the returns to the stock and productivity of human capital. Unlike the returns on bonds and equities, the demand for money bears a negative relationship to the level of income. The price level also enters the function in that it affects the level of nominal income. Moreover, the real demand for money is assumed to be unaffected by any change in the nominal units used to measure the variables in the demand function. Thus any change in the unit of account by which prices and income are measured is assumed to result in a proportionate change in nominal money demand. The inverse of the function with respect to **Y** yields the income velocity of money used in the quantity equation, **MV = PY**.

From the viewpoint of both domestic monetarism and the monetary approach to the balance of payments and exchange rate determination, the stability of the money demand function is critical. If the variables determining the demand for money are relatively independent of those determining the money supply and if the demand function is stable, then it is theoretically possible to predict the effect of a change in the money supply on output, the price level, and/or the balance of payments:

"A stable demand function is useful precisely in order to trace out the effects of changes in supply, which means that it is useful only if supply is affected by at least some factors other than those regarded as affecting demand"
(Friedman 1956: 17).

It is also important to note that the quantity theory does not require velocity to be a numerical constant:

"For the stability...is in the functional relation between the quantity of money demanded and the variables that determine it, and the sharp rise in velocity of circulation of money during hyper-inflation is entirely consistent with a stable functional relationship" (Friedman 1956: 16).

Accordingly the monetary approach does not deny the importance of variables such as interest rates, income or growth in their effect on the balance of payments or exchange rates. However, unlike the elasticities and absorption approaches it insists that, at least in the long-run, such variables exert their influence through monetary channels, specifically through their influence on the demand for money. In this sense the demand for and supply of money are seen as the proximate determinants of the balance of payments and the exchange rate.

Under fixed exchange rates the stock of money is supported both by domestic and foreign reserves:

$$M = m(D + R) \quad \text{.....(4b)}$$

where **M** is the stock of money, **m** is the composite money multiplier, **D** is the domestic credit (reserves) created by the central bank and **R** is the domestic currency equivalent of the foreign exchange reserves. This is the money base identity. The monetary approach then assumes that the money market always tends to equilibrium in the sense of the existing stock of money being willingly held:

$$L = M = m(D + R) \quad \text{.....(4c)}$$

where **L** is the demand for money. Changes in **L** and **M** can occur rapidly and are interpreted as stock shifts. By rearranging terms:

$$mR = L - mD \quad \text{.....(4d)}$$

Stated in this form the basic adjustment mechanism of the monetary approach is apparent. Excess demand for or supply of money in the domestic money market is cleared by a discrete

change in foreign exchange reserves (ΔR). Thus an excess supply of money will be got rid of by an outflow of reserves ($-R$) corresponding to a balance of payments deficit while an excess demand for money will attract an inflow of reserves ($+R$) corresponding to a balance of payments surplus.

The composition of a payments imbalance between the trade and capital accounts is not made explicit in the simple monetary model. Wealth owners may, for example, eliminate an excess supply of money by purchasing more foreign goods and/or foreign assets. Thus the overall balance of payments is made synonymous with the change in the official settlements balance. The change in foreign exchange reserves, then, is the mechanism by which the domestic money market regains stock equilibrium. The monetary approach does not stipulate how long the adjustment in R takes. The balance of payments flows continue until stock equlibrium is regained and cease thereafter. The flows are thus temporary and self-correcting under a fixed exchange rate. It is, however, implicitly assumed that the monetary authority cannot or will not sterilize such flows indefinitely. For example, a drain of reserves corresponding to a balance of payments deficit may, at least in the short-run, be offset by open market purchases of securities by the central bank. Consequently the domestic money market can remain in equilibrium without any tendency to correct itself. The monetarist reply to this is that such sterilization is not possible in the long-run because the central bank must eventually run out of foreign exchange reserves or that foreigners will refuse to accumulate the deficit country's domestic currency indefinitely.

The policy implications of the simple monetary model are in sharp contrast to those of Keynesian flow models. This is an important methodological consideration since these differences may provide grounds for the searching empirical tests necessary, for a falsificationist, to choose between the two models, that is, to refute one of them. The historical record of the success or failure of the respective policy measures, where they have been applied, may be regarded as an example of a stylized fact in economics. Repeated failures of particular

policies should constitute sufficient empirical evidence to refute the underlying theory especially if (as in Lakatos and Laudan) an alternative theory in the same domain can satisfactorily explain such failures and where its policy implications have been corroborated. One of the reasons for the adoption of the monetary model over the Mundell-Fleming model, for example, was the perceived failure of devaluation and interest rate policies (notably in the UK) during the 1960s to remedy balance of payments problems. These and related issues are discussed further in section 4.5. The following explains the main policy differences of the monetary approach compared to the general conclusions of the Keynesian models.

Under a fixed exchange rate the main implication of the monetary approach is that the monetary authority cannot control the domestic money supply. This conclusion follows directly from the money-base identity (4b) and the equilibrium condition for the domestic money market. If, for example, the central bank attempts to expand the money supply in excess of the demand for it the increase in the domestic component of the money-base **(+D)** will be offset by an equivalent decrease in foreign exchange reserves **(-R)**. Thus domestic monetary policy can influence the composition but not the overall amount of reserve money. By contrast, Keynesian models assume that changes in the supply of money have only an indirect influence on the balance of payments via the liquidity effect on interest rates and, in turn, the effect of the latter on expenditure and income flows.

The monetary approach also implies that devaluation cannot permanently improve the balance of payments, even without the assumption of the law of one price made by the global version of the monetary model. Assume that devaluation is successful in lowering the terms of trade and that the elasticities condition is met. The resulting improvement in the trade balance generates an inflow (or reduced outflow) of foreign exchange reserves. However, by creating an excess supply of money this inflow of reserves disturbs stock equilibrium in the domestic money market, *ceteris paribus*. The excess money supply is eliminated by increased expenditure on foreign goods and/or

assets, thereby reversing the temporary improvement in the balance of payments. The inflows vanish altogether once stock equilibrium in the money market is reattained.

If the law of one price is taken for granted, devaluation is matched by a proportionate increase in the domestic price level which precludes any change in the terms of trade. If, however, the monetary authority does not accommodate the higher price level by expanding the domestic money supply this may create an excess demand for money and a temporary balance of payments surplus. But the inflow of reserves ceases once wealth holders achieve their desired quantity of money balances. Although the monetary approach generally assumes that the law of one price holds and that devaluation cannot alter the terms of trade, some models admit the possibility that devaluation increases the relative price of traded versus non-traded (home) goods. In this case resources switch from home goods industries to the export oriented and import substitution industries and this improves the trade balance. The improvement in the balance of payments remains temporary and the "switching" of production simply accelerates the adjustment to money market equilibrium.

These implications conflict sharply with those of elasticities models where it is assumed that a devaluation should permanently improve the trade balance. However, it is possible to reconcile such implications with those of the absorption approach. The main theoretical difference is that the monetary approach assumes spending to be a function of real money balances while the absorption approach assumes expenditure to be a function of income. The former thus suggests a stock equilibrium view of the adjustment process while the latter suggests a flow equilibrium view.

According to the monetary approach tariffs, import quotas, and exchange controls have similar effects to devaluation under the assumption of the law of one price. Each measure tends to increase the general price level and the nominal demand for money. This temporarily improves the balance of payments, *ceteris paribus*, but only until stock equilibrium in the domestic money market is restored. In summary, devaluation and

protective measures such as tariffs cannot permanently improve balance of payments flows unless they are accompanied by a restrictive monetary policy. Ultimately only a lower growth rate in domestic credit creation relative to that of the country's major trading partners will be effective in preventing recurrent balance of payments crises.

Keynesian models such as the Mundell-Fleming model assume that an increase in the domestic interest rate will improve the balance of payments by increasing the net inflow of capital and, by reducing domestic spending, lowering the level of imports. This also contrasts with the predictions of the monetary approach: an increase in the (real) interest rate reduces the demand for money which temporarily worsens the balance of payments as individuals and companies attempt to rid themselves of excess money balances. A similar contrast is apparent as regards differences in economic growth rates. Keynesian models predict that higher relative growth rates worsen the balance of payments according to the economy's marginal propensity to import while the monetary approach suggests that such growth increases the demand for money and thus improves the balance of payments.

ii) *Flexible exchange rates*. The monetary approach to the balance of payments can be readily extended to the determination of exchange rates and there is a degree of symmetry between the two cases. Under fixed exchange rates, changes in relative supplies of and demands for different national monies determine the balance of payments. With a flexible exchange rate, monetary disturbances give rise to *incipient* balance of payments deficits or surpluses which are rapidly eliminated by changes in the exchange rate. Thus the basic reduced form equation for the monetary model of exchange rate determination:

$$s_t = (m - m^*)_t - a(y - y^*)_t + b(i - i^*)_t \quad \text{.....(4e)}$$

where the lower case letters indicate rates of change and the latter two terms are derived from the demand for money function. An increase in relative real income $(y - y^*)$

increases the demand for money and thus leads, *ceteris paribus*, to an appreciation of the exchange rate s_t; while an increase in relative nominal interest rates decreases the demand for money and results in depreciation[6]. This is counter to the prediction of the Mundell-Fleming model that an increase in the interest rate should attract a greater net capital inflow and thereby appreciate the exchange rate[7].

Equation (4e) can be derived from the purchasing power parity condition for traded goods and many of the monetary models adopt this approach. In this case the negative empirical test results and other criticisms of PPP discussed in chapter three indirectly also cast doubt on the basic monetary model of exchange rate determination. However as made clear by Mussa, PPP is not a necessary condition of the monetary model:

"This equation can be derived without explicit reference to purchasing power parity; indeed, it can be derived from a model that allows explicitly for divergences from purchasing power parity. Moreover, some empirical studies...have argued that the conditions of money market equilibrium are more immediately relevant for determining the exchange rate (which is a freely adjusting asset price) than they are for determining national price levels" (Mussa 1984: 22-23).

Emphasizing the relationship between relative price levels and the exchange rate, besides being unnecessary, also appears to introduce a subtle but conceptually significant asymmetry between the monetary approach to fixed and flexible exchange rates. With fixed exchange rates an excess supply of money is assumed to be eliminated by purchases of either foreign goods (trade account) or foreign assets (capital account) or both. Deriving the monetary model of flexible exchange rates from the PPP condition, however, implies that an excess supply of money causes the exchange rate to depreciate in response to incipient trade flows alone, which thus restricts the generality of the monetary approach.

Under a fixed exchange rate, money market equilibrium is restored by more or less gradual flows of foreign exchange reserves. Flexible exchange rates, on the other hand, are assumed to adjust instantly such that existing stocks of monies are willingly held. This difference results from the different

way in which expectations and speculative capital flows function under the two regimes. Under a fixed exchange rate speculators take a view about the sustainability of the official rate that the monetary authority is willing to defend. If, for example, on the basis of fundamentals speculators regard the currency as overvalued at the official rate of exchange they will take a position against the currency[8]. The greater the perceived divergence between the official rate and the market evaluation, and the less the perceived ability or intent of the monetary authority to defend the official rate, the more intense the speculative pressure on the currency. Under these circumstances private speculation will ultimately prevail over opposing speculation by the central bank. An instructive recent example was the futile attempt by the Bank of England to prevent the depreciation of sterling beyond the limits imposed by the European exchange rate mechanism (ERM)[9]. Thus under fixed exchange rates, speculative pressure manifests itself as large changes in autonomous flows of short-term capital.

Flexible exchange rates, however, are assumed to adjust instantly to changes in expectations such that actual (as opposed to incipient) capital flows are largely forestalled thereby. If most speculators anticipate an appreciation in the exchange rate then the increase in demand for that currency will force an immediate appreciation of the spot exchange rate and vice versa. Thus the question of what determines the exchange rate reduces to the question of what determines expectations:

> "The problem is to identify the things which influence expectations concerning the returns to holding particular currencies and to explain why expectations have tended to fluctuate" (Mussa 1978: 54).

It was noted in section 4.2 that the Mundell-Fleming model is conceptually flawed because it regards changes in interest rate differentials as the cause of permanent changes in capital flows rather than a temporary (even instantaneous) stock adjustment to desired asset holdings. The Mundell-Fleming model encounters similar conceptual difficulties when it is amended

to deal with expectations. Then the capital flow component of the FF curve may be defined as:

$$K = k\{i - i^* - (S - S^e)/S\} \quad \text{.....(4f)}$$

where **K** is the value of capital flows and S^e is the expected spot exchange rate, with all values in the equation normalized to a standard unit of time. The coefficient *k* measures the joint elasticity of capital flows with respect to interest rates and exchange rate expectations. If $S^e = S$ (static expectations) then capital flows are simply a function of the interest rate differential $(i - i^*)$. The problem the model faces is to explain why capital movements do not occur instantaneously when $S^e \neq S$ rather than as a uniform flow per unit time. Clearly, no appeal can be made to a flow of saving explanation since capital movements based on expectations are inherently speculative. One way in which the model can be adapted to deal with this criticism is to remove the expectations component from equation (4f) and to allow the FF schedule to shift with changes in such expectations. The difficulty with this approach is that expectations become an *ad hoc* exogenous variable which can be invoked as a convenient immunizing strategem whenever the basic model fails an empirical test.

Mussa (1979) argued that any flow model of exchange rate determination which includes speculative capital movements faces similar conceptual difficulties. Consider the following diagram:

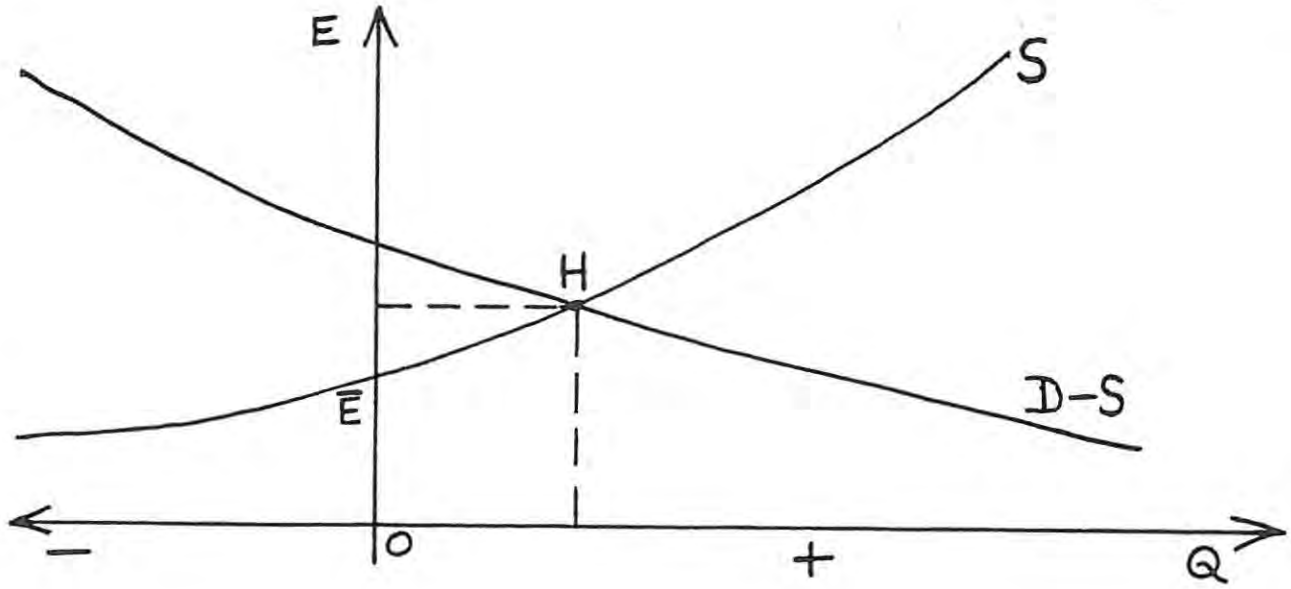

Figure 2: An extended flow model of the foreign exchange market including currency speculation
Source: Mussa (1979: 32)

The vertical axis measures the exchange rate as the domestic price of foreign currency while the horizontal axis measures the desired quantities of foreign currency demanded and supplied per unit time. The **D-S** curve shows the excess demand for foreign currency derived from trade flows, as a function of the exchange rate. The **S** schedule shows the supply of speculative foreign capital as a function of the gap between the ruling exchange rate **E** and the expected equilibrium exchange rate $\bar{\mathbf{E}}$. If $\mathbf{E} > \bar{\mathbf{E}}$ speculators anticipate capital gains and take a long position in the domestic currency, and vice versa. The greater the gap between **E** and $\bar{\mathbf{E}}$ the more profit speculators anticipate by taking a position in the domestic currency, reflected in the positively sloped **S** curve. Equilibrium in the foreign exchange market is achieved at point **H** where the excess flow demand for foreign currency by goods traders just equals the flow supply of foreign currency by speculators.

If the **S** curve is perfectly elastic with respect to expectations then the exchange rate is fully determined by speculative capital flows. The theory of exchange rate determination then becomes purely a matter of explaining movements in the speculative supply of foregn capital and, therefore, how exchange rate expectations are formed. Anticipated trade flows may play an indirect role as information used in forming rational exchange rate expectations. Similarly, anticipated changes in interest rates may influence expectations. However, apart from the difference in the assumed speed of adjustment this is essentially the asset market view of exchange rates and not a genuine flow model thereof. Expectations in such flow models are not an integral part of the theory - they are tagged on as an *ad hoc* auxiliary hypothesis in the sense of Lakatos' *ad hoc*$_3$.

The main implication of the flexible exchange rate monetary model is that persistent currency depreciation is the result of growth in the domestic money supply exceeding that of the demand for money relative to major trading partners. Consequently the only effective way to halt such depreciation in the long-run is to permanently reduce growth in the domestic money supply. In a rational expectations augmented model (see below) the mere announcement of a tighter monetary policy may immediately appreciate the exchange rate. This depends on the credibility of the monetary authority and may be reversed if the announced targets are not adhered to.

In the monetary model, real factors may also influence the exchange rate but primarily through their effects on the demand for money. In countries with rapid real income growth the demand for money increases and this results in an appreciation of the exchange rate. In developing countries, changes in the prices of primary exports may also cause fluctuations in the exchange rate and/or the balance of payments, but as noted by Mussa (1978: 55) this reflects a situation where domestic consumption does not fluctuate by the same amount as export earnings, that is, saving has increased (this may occur via a reduced budget deficit from the passive increase in tax revenues accruing to the state).

Flexible exchange rates allow the central bank to control the money supply and hence the long-run domestic inflation rate. The money supply becomes an exogenous policy variable, unlike with fixed exchange rates where it is endogenous. However, floating exchange rates may not fully insulate the economy from foreign disturbances. For example, reduced aggregate demand in the US will also lower demand for German exports resulting in a depreciation of the mark. But such depreciation will fail to restore US demand for German goods if the price level in Germany rises proportionately such that the dollar prices of traded goods remain unchanged. Thus at least part of the reduction in US demand may be transmitted to Germany.

Under a fixed exchange rate a policy conflict is likely to arise whenever one country wishes to expand aggregate demand while another wants to contract. The expanding economy will tend to lose reserves to the contractionary economy. With flexible exchange rates this problem can be overcome by allowing the exchange rate to depreciate. However, policy conflict may still occur if both countries wish to increase aggregate demand or reduce inflation simultaneously. To avoid self-defeating rounds of competitive depreciation in the former, policy coordination is required. Primarily this involves the coordination of monetary policies but fiscal policies may also need coordination to the extent that the financing of budget deficits or investment of surpluses have monetary consequences.

4.4 Extensions and refinements to the basic monetary model

The basic monetary model described above can be extended and refined almost without limit by combining it with various auxiliary hypotheses. As noted in section 4.3, the models may be classified under the general asset market approach by their different assumptions and auxiliary hypotheses. This section does not attempt an exhaustive survey of the research in this area but describes the main theoretical developments of the monetary approach since the prototype models of the early 1970s. Four models are selected for analysis: global monetarism

or the 'Chicago' model; the Dornbusch overshooting hypothesis; Frankel's real interest rate differential model; and the basic portfolio balance model. An important point made in this section is that the implications of the different models vary depending upon which version of expectations is used. The Dornbusch and Frankel models rely on regressive expectations: the expected exchange rate is a weighted function of the current exchange rate and an estimate of the long-run equilibrium exchange rate - in other words the exchange rate is expected to converge towards a long-run equilibrium determined by PPP or some other set of fundamentals. The rational expectations hypothesis, which can be applied to any fundamental model of exchange rate determination, is discussed with reference to the Chicago model and the portfolio balance model. In the latter a distinction is made between the conditional probability or expected value version of the rational expectations hypothesis and the certainty equivalence or perfect foresight version thereof.

Global monetarism[10] includes two main auxiliary hypotheses. Firstly, perfect wage and price flexibility is assumed such that full employment prevails. Secondly, the law of one price is assumed to hold in both goods and capital markets which thus integrates diverse national markets into one world market. These assumptions rule out the possibility of any independent changes in output, relative prices, or interest rates. Changes in relative money supplies or demands have no effect on any real variables such as income, even in the short-run. Thus the real exchange rate remains constant unless it is disturbed by changes in other real variables. An expansion in the domestic money supply must result in an equiproportionate increase in the price level and exchange rate depreciation. Laffer and Miles (1982) contrast the more traditional Keynesian IS-LM-FF model of internal-external balance (essentially the Mundell-Fleming model) with the global monetarist approach and include empirical evidence purporting to show the widespread integration of national goods and asset markets.

In equation (4e) relative interest rates are included as a variable assumed to influence the demand for money. If it is

further assumed, as in the global monetary model, that the law of one price operates in the international capital market then the uncovered interest rate parity condition must also apply. The law of one price in financial markets means that the return on assets (such as treasury bills or bonds) which are identical except for the currency of denomination should be the same. However, the total return comprises a known interest yield and an anticipated exchange rate component. For example, if a US bond currently yields 10% interest but investors/ speculators anticipate the dollar to depreciate 3% against sterling over the relevant investment horizon, then a sterling bond with the same maturity and risk would be priced to yield only 7% interest. The 3% premium of dollar *vis-a-vis* sterling bonds is thus required to compensate the investor for the expected depreciation of the dollar/sterling exchange rate. In equilibrium the expected change in the exchange rate must equal the observed interest rate differential so that $(\Delta s^e)_{t+1}$ can be substituted in equation (4e) for $(i - i^*)_t$ to get:

$$s_t = (m - m^*)_t - a(y - y^*)_t + b(\Delta s^e)_{t+1} \quad \ldots\ldots\ldots (4g)$$

The issue then is how to model expectations. If economic agents are efficient in the sense of using the information of the model itself to make the relevant predictions of the exchange rate then a rational expectations approach is appropriate. This involves a forward iteration of equation (4g) which gives:

$$s_t = (1 + b)^{-1} \sum_{z=0}^{\infty} [b/1+b]^z \cdot E[(m - m^*)_{t+z} - a(y - y^*)_{t+z} \mid_t] \quad \ldots\ldots\ldots\ldots (4h)$$

where $E[\cdot \mid_t]$ means that expectations about the variables to the left of the vertical bar are conditional on information available at time t. Thus the current exchange rate depends on the future expected path of relative money supplies and differences in real incomes. Since it is assumed that the discount factor $b/1+b < 1$, such expectations do not have to extend literally into the infinite future. Due to the discount factor, short-term expectations have a higher weighting in equation (4h) and thus exert a stronger influence over the current exchange rate.

The above model can be used to decompose a change in the exchange rate into an expected and unexpected component (see Mussa 1984: 19-20). If X is the set of variables thought to determine the exchange rate, then the expected change in the exchange rate is proportional to the difference between the discounted sum of all expected future X's and the current X_t. The unexpected change in the exchange rate, on the other hand, is a function of the discounted sum of the change in expectations about future X's resulting from new information received between one period to the next. In the monetary model X would naturally include money supplies and demands and/or the variables hypothesized to determine these. However it is important to note that rational expectations, although originally connected with the new classical economics and the monetary approach, can be applied to any fundamental model of the exchange rate including the more traditional flow models thereof[11].

It should also be noted that this later version of the monetary model is quite different to the earlier monetary models of exchange rates which relied on PPP. In the early monetary models an expected depreciation of the exchange rate reduces the real demand for money which, via PPP, *indirectly* leads to depreciation of the spot exchange rate by raising the price level. In the asset market model expectations influence the spot exchange rate *directly*, as above:

"By contrast, in the present model the anticipation of depreciation leads directly, as of given prices and interest rates, to an equiproportionate depreciation of the spot rate" (Dornbusch 1980: 26).

Dornbusch (1976a) developed a model sharing some of the main assumptions of global monetarism - full employment, uncovered interest parity and continual stock equilibrium in the money market. Growth is ignored and the small, open economy assumption makes the foreign interest rate and price level exogenous. In the long-run, the exchange rate is determined by PPP in the narrow sense of traded goods arbitrage (Dornbusch 1976a: 34-35 and 50).

However, Dornbusch assumes that prices are sticky in the short run so that the goods market adjusts more slowly than the money market in response to monetary shocks. This causes increased volatility and "overshooting" of the exchange rate which, in the short-run, is completely determined by the conditions bringing about money market equilibrium: thus PPP may be violated in the short-run. Prices adjust to their long-run PPP equilibrium level at a rate proportional to excess demand in the goods market. Expenditure is a negative function of the domestic interest rate but is positively related to the spot exchange rate since depreciation succeeds in worsening the terms of trade in the short-run and world demand shifts to home goods.

Long-run PPP anchors exchange rate expectations. Speculators expect the exchange rate to adjust to long-run equilibrium **S** at a constant fraction *f* of the prevailing gap between them:

$$s^e = f(\bar{S} - S) \qquad (4i)$$

Combined with the uncovered interest parity condition we get:

$$S = \bar{S} + (i^* - i)/f \qquad (4j)$$

Given the above equilibrium conditions and specification of the speeds of adjustment in the goods and money markets, the effects of say an increase in the domestic money supply can be traced through. In the long-run the price level and the exchange rate must increase proportionately. But in the short-run, because goods prices are sticky, the domestic interest rate falls instead to immediately clear the money market. By equation (4j) the lower interest rate requires a higher (depreciated) exchange rate to maintain interest parity. It thus overshoots its new long-run equilibrium $\bar{S}$ by $(i^* - i)/f$. Put differently, the fall in the domestic interest rate implies an expectation of an eventual appreciation in the exchange rate from its currently weaker level. The lower interest rate and higher exchange rate generate excess demand in the goods market so that, with full employment, long-run equilibrium is achieved by an increase in the price level. Higher prices reduce the

real money supply, pushing up interest rates to clear the money market. Interest parity is now maintained by a fall (appreciation) in the exchange rate. Full equilibrium and short-run overshooting of the exchange rate are illustrated in the diagram below:

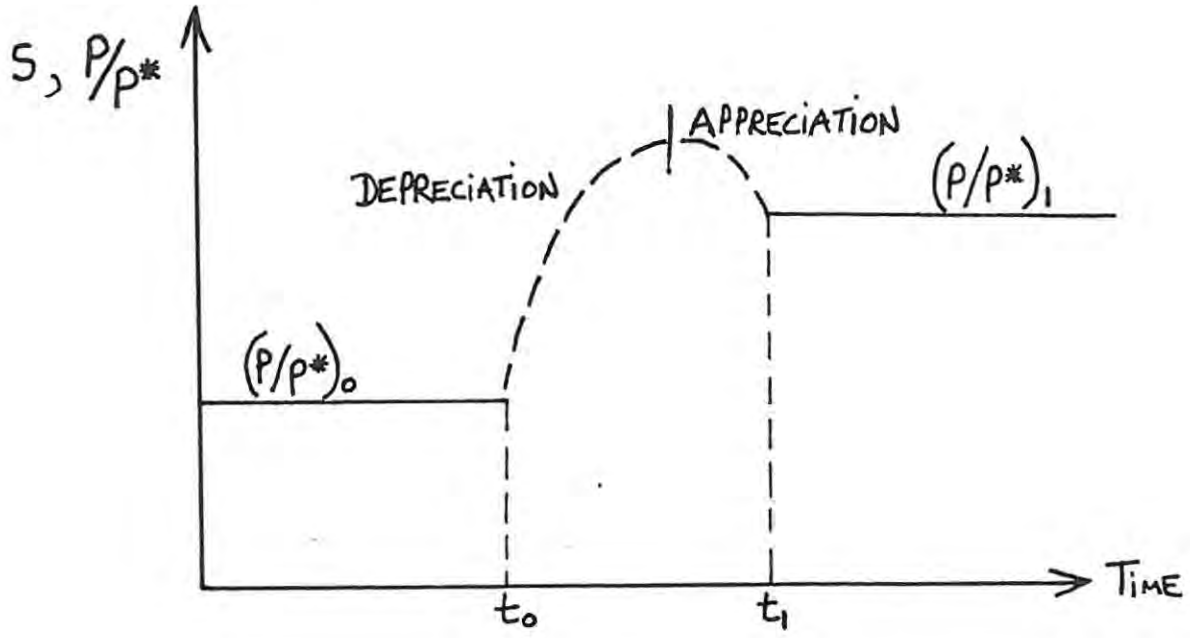

Figure 3: Typical adjustment path of the exchange rate to long-run PPP equilibrium in response to a discrete increase in the money supply

Dornbusch (1976b) extended the above model to include non-traded goods and relaxed the full employment assumption. He showed that in the short-run, changes in the real money supply are not neutral since output would partially respond to higher aggregate demand. Furthermore, changes in the relative price of traded goods may influence the real exchange rate independently of changes in money supply and demand.

An important conceptual disadvantage of the Dornbusch model is that exchange rate expectations are assumed to be independent of anticipated inflation rates. The global monetarist or 'Chicago' model, on the other hand, is unrealistic because it assumes perfectly flexible prices. Frankel (1979) proposed a hybrid model combining the assumption of sticky prices in the short-run and modelling exchange rate expectations as a

function of anticipated inflation differentials. Thus equation (4i) is modified as follows:

$$s^e = k(\bar{s} - s) + (p^e - p^{e*}) \quad \text{.....(4k)}$$

where p^e and p^{e*} are the currently anticipated, long-run domestic and foreign inflation rates respectively (again, as throughout this thesis, upper and lower case letters designate levels and rates of change of the variables respectively). In the short-run the spot exchange rate s is expected to return to its long-run equilibrium value $\bar{s}$. As in the Dornbusch model, the speed of adjustment is proportional to the gap between s and $\bar{s}$. In the long-run, when $s = \bar{s}$, the exchange rate is expected to adjust at a rate equal to $p^e - p^{e*}$. Combining equation (4k) with the uncovered interest parity condition gives:

$$s - \bar{s} = -1/k[(i - p^e) - (i^* - p^{e*})] \quad \text{.....(4l)}$$

In the short-run the gap between s and $\bar{s}$ is proportional to the real interest rate differential expressed in brackets. In the long-run, however, when $s = \bar{s}$ the real interest rate differential is zero since the nominal interest differential simply reflects the difference in anticipated inflation rates $(i - i^* = p^e - p^{e*})$. Equation (4l) also implies overshooting. For example, a discrete increase in the money supply will, as in the Dornbusch model, create excess liquidity due to sticky prices in the short-run. The nominal interest rate falls to maintain money market equilibrium. The lower interest rate, in turn, requires a depreciation in the exchange rate to maintain interest parity. If the increased money supply also leads to a higher expected inflation differential (which would be the case with rational expectations) the rate of depreciation will be even higher than in the Dornbusch model, via equation (4k).

The monetary aspects of Frankel's model, where the exchange rate is determined once again by the relative supply of and demand for two currencies, are introduced via PPP in the usual way. In contrast to the small open economy assumed by Dornbusch, Frankel assumes a two country model in which it is

necessary to consider the foreign money market explicitly. In regression equation form:

$$S_t = a + b(M - M^*)_t - c(Y - Y^*)_t + d(i - i^*)_t + e(p^e - p^{e*})_t + e_t \quad \text{.....(4m)}$$

Frankel hypothesizes that *d* is negative and *e* positive (and greater than *d* in absolute value). In the Chicago flexible price model *d* is hypothesized to be zero and *e* to be positive while in Dornbusch's sticky price model *d* should be negative and *e* zero. Frankel (1979) reports that his regressions for the mark/dollar exchange rate over the sample period July 1974 - February 1978 support the predictions of his model while refuting those of Dornbusch and the Chicago school[12].

Thus the Dornbusch and Chicago models may be regarded as limiting cases of the more general hybrid model proposed by Frankel. The Chicago model is applicable when inflation differentials are very high, as during the German hyperinflation of 1920-23 analyzed by Frenkel (1977). Inflationary expectations swamp all other variables such that full equilibrium is maintained instantly and continuously. The sticky price model is more appropriate when the inflation differential is low and stable. Frankel's model seeks to explain exchange rate variation in the environment of moderate inflation typical of the industrialized economies during the 1970s.

Two key assumptions distinguish portfolio balance models from monetary models. The monetary approach assumes that domestic and foreign nonmoney assets are perfect substitutes and can be aggregated as a homogenous "bond" market. Most monetary models (Kouri and Porter 1974 is a significant early exception) also ignore the wealth effects of current account imbalances which may arise if traded goods PPP does not hold continuously. Portfolio balance models relax these assumptions.

Perfect substitution between domestic and foreign bonds is implied by the uncovered interest parity condition which rules out changes in asset preferences as an independent variable determining the exchange rate. The aggregation of domestic and

foreign bonds means that only three markets (goods, bonds, and money) are considered in the monetary approach. By Walras' Law one of these markets is redundant, which allows explicit consideration of the bond market to be left out and for the monetary approach to focus on the remaining goods and money markets.

However, seemingly identical securities differ according to their currency of denomination and issuer. If exchange rate expectations are not held with certainty, speculators will require a higher return to compensate them for bearing perceived foreign exchange risk when trading in securities denominated in different currencies. Besides exchange rate risk, speculators and arbitragers also bear various types of political risks and costs: the risk of default by the issuing government; the imposition of new exchange controls; new taxes on repatriated earnings; and the cost of any existing controls. The nature of such risk factors and costs is explained by, amongst others, Aliber (1973), Frenkel and Levich (1975), and Dooley and Isard (1980). Their research suggests that the uncovered interest parity condition should be amended to:

$$i - i^* - \Delta s^e = R \qquad (4n)$$

where **R** is the risk premium for a particular currency. The presence of such a risk premium means that apparently identical securities are not regarded as perfect substitutes. In this case changes in asset preferences will play an independent role in determining exchange rates. The exchange rate will adjust not only to ensure that existing money stocks are willingly held, but also the stocks of domestic and foreign bonds.

Portfolio balance models also allow for the possibility of wealth effects arising from current account imbalances. If the exchange rate diverges from PPP then this has implications for the trade and current accounts of the balance of payments. In the Dornbusch overshooting model, for example, the exchange rate depreciates in response to an increase in the domestic money supply, to maintain stock equilibrium in the money market. Because goods prices are assumed to be sticky in the

short-run the real exchange rate also changes and, if the Marshall-Lerner condition is met, this leads to an improved trade balance. This implies an increase in domestic wealth as the country accumulates foreign exchange reserves and increased claims on future foreign output. Portfolio balance models specify the effects of such changes in wealth on domestic expenditure and on the demand for money and other assets.

Consideration of current account flows also introduces a new channel through which expectations may influence the exchange rate. Speculators may revise their views regarding the sustainability of current account imbalances thereby triggering changes in desired asset portfolios and thus forcing spot exchange rate adjustment to regain stock equilibrium (see endnote 8). The role of current account balances, expectations and stock-flow interaction within an asset market framework is thus an integral part of the portfolio approach.

The basic asset market equilibrium conditions and constraints characterizing portfolio balance models are:

$$W = M + B + SF \quad \text{.....(4o)}$$

$$M = M(i,i^*)W \quad \text{.....} \quad M_i < 0 \; ; \; M_{i^*} < 0 \quad \text{.....(4p)}$$

$$B = B(i,i^*)W \quad \text{.....} \quad B_i > 0 \; ; \; B_{i^*} < 0 \quad \text{.....(4q)}$$

$$SF = F(i,i^*)W \quad \text{.....} \quad F_i < 0 \; ; \; F_{i^*} > 0 \quad \text{.....(4r)}$$

$$f = T(S/P) + i^*F \quad \text{.....} \quad T_{s/p} > 0 \quad \text{.....(4s)}$$

where **W** is wealth, **F** is net holdings of foreign assets (bonds) and **B** is domestic bonds. Partial derivatives are given alongside each equation. Equations (4p-4r) are the standard asset demand functions while (4o) is the wealth constraint. It is assumed that the demand functions are homogenous in the scale variable wealth. By Walras' Law any one of these demand equations, usually the demand for foreign bonds (4r), is redundant and not considered explicitly in the analysis. Equation (4s) equates **f**, the rate of change in domestic holdings of net foreign assets (capital account), with the current account balance which is the sum of the trade balance **T(S/P)** and net interest receipts **(i*F)**.

The basic portfolio balance model should be supplemented with a description of the goods market and other auxiliary hypotheses, as in the monetary approach. It is not necessary to describe a fully specified model here. However, as summarized by Hallwood and MacDonald (1986: 126) the main features of the portfolio approach are: asset market shocks result in exchange rate overshooting but this is partly reversed over time as residents alter their stocks of foreign assets; a current account surplus (deficit) is associated with an appreciating (depreciating) exchange rate; and PPP need not hold, even in the long-run, since changes in the money supply may induce permanent changes in the real exchange rate necessary to maintain overall balance of payments equilibrium.

As a stylized example, a reduced preference for domestic bonds implies an immediate impact effect on the asset markets - the domestic interest rate falls and the exchange rate depreciates. The weaker exchange rate disturbs flow equilibrium in the balance of payments. By equation (4s) a trade surplus is realized (and the domestic money supply increases). Residents accumulate wealth in the form of greater net foreign assets. Over time, residents rebalance their portfolios by selling some of these assets and realizing their claims on foreign output. Imports thus increase but the exchange rate appreciates, further encouraging imports. Moreover, goods prices have increased in line with the increase in the domestic money supply. This further erodes competitiveness and the trade surplus. For current account equilibrium the trade balance must run into deficit because of the positive net interest income earned on the foreign asset. The exchange rate continues to appreciate until the current account balance is zero.

The above sketch assumes static exchange rate expectations, that is, the expected exchange rate for the next period is equal to the current exchange rate. Dynamic expectations may be introduced as an additional variable in each of the asset demand equations above and then deciding how to model expectations. In a rational expectations portfolio balance model, exchange rates (and interest rates) react differently depending on whether exogenous shocks are foreseen or not. With

perfect foresight, speculators are never mistaken and the expected exchange rate is always equal to the actual exchange rate that materializes in the next period. Thus the perfect foresight version of rational expectations has a stabilizing influence on the exchange rate (and on interest rates). The implications of anticipated changes in the money supply or asset preferences are immediately realized and reflected in the spot exhange rate. For example, the depreciation of the exchange rate required by an increase in the money supply will now be dampened by the anticipated appreciation resulting from the resulting current account surplus. Thus the dynamic path of the exchange rate is less volatile. However with a conditional expectations or "news" version of rational expectations, unanticipated shocks require a more pronounced initial jump in the exchange rate followed by correction to the new long-run equilibrium. As with the other models described in this section the testable implications of the model are highly sensitive to which expectations hypothesis is used.

As apparent from the above, portfolio balance models may be viewed as a conceptual hybrid of the Mundell-Fleming model and the monetary model in which both stock and flow equilibrium conditions are integrated. Krueger (1983: 82) remarks that the Mundell-Fleming model, with its explicit treatment of both current and capital accounts, is in this respect the intellectual precursor of the portfolio approach. It is worth noting that some of the basic insights of the portfolio approach had already been developed during the 1960s and early 1970s, at the same time as the monetary approach - see, for example, Oates (1966); McKinnon and Oates (1966); Branson (1968); McKinnon (1969); Willet and Forte (1969); and Kouri and Porter (1974). This raises the question, of some interest from a historical perspective, why the monetary model initially gained the ascendancy when the portfolio approach seems a more natural conceptual evolution from the then mainstream Mundell-Fleming model.

Some tentative answers may be surmized. It was perhaps the relative simplicity of the stylized monetary model, with conclusions and policy recommendations in sharp contrast to the

orthodox Keynesian models, that captured the imagination of many researchers in the field. The greater analytical complexity of the portfolio approach, although perhaps more "realistic", may have made it less accessible in the competitive market place of ideas. Moreover, initial empirical tests of the basic monetary model did seem to confirm its main predictions whereas the more complex portfolio balance models could not be tested as readily (see chapter five for more evidence in this regard). The early monetary approach to the balance of payments also benefitted from a resurgence of classical thinking generally and more specifically in the form of monetarism (see below), especially when combined with the classical assumptions of perfectly flexible prices and full employment.

4.5 Some lessons for the methodology of economics

A more complete evaluation of the monetary approach and the methodological implications thereof is postponed until the empirical issues have been discussed in chapter five. However, some preliminary observations may be made at this stage as regards the conceptual development and progress of theories of exchange rates and the balance of payments, leading up to the modern revival of the monetary approach.

The relationship between a theory and empirical observations or tests is by most accounts a defining characteristic of science but philosophers have disagreed on the nature of this relationship and thus what is to count as science as opposed to what is not. The logic of confirmation is that a good scientific theory is one which explains, or is consistent with, a great many facts and passes more empirical tests than its rivals. Popper rejected this view of empiricism and replaced it with his own view, that of falsification: theories cannot be proved simply because they agree with many facts but they can be disproved by counter evidence and active attempts should be made to do so. Good theories were simple; unambiguous criteria for refutation had been specified; and the truly scientific attitude was that which stood ready to abandon even well

established theories unequivocally if the evidence demanded it. Lakatos reinterpreted the conventionalist rules for acceptance and rejection of a theory within his methodology of scientific research programmes but also emphasized the significance of empirical tests. Arguably, by making the prediction and corroboration of novel facts the centrepiece of his approach, the empirical requirements for a genuinely scientific theory were tightened even further (Popper with a twist).

The theories of exchange rates and the balance of payments analysed in this chapter do not appear to have developed consistently in any of these ways. The development of such theories seems to bear only a tenuous and remote relationship to empirical test results, and not only during the early stages. Up until the early monetary approach to the balance of payments the research activity of most theorists in the field may best be described as 'policy science' in which empirical considerations played a minor role. There is, however, a noticeable break in the empirical approach to the monetary model compared to the theories which preceeded it.

In reading the contributions made by Alexander, Swan, Meade, Mundell, Fleming and others after the Second World War it is clear that their primary concern was the policy implications of various theories[13]. This period may thus be described as an era of economic policy science in which the empirical testing of the successive models was largely eschewed. Meade (1951), for example, tucked away his macroeconomic model of the balance of payments in a mathematical supplement. There was never any serious attempt to test his model empirically, nor was it thought necessary or desirable to do so beyond perhaps a cursory reference to some selective stylized facts. The development of the Mundell-Fleming model, in turn, did not result from any obvious empirical failings of Meade's model but from the attempt to clarify the policy implications of elastic capital flows. All the above economists were engaged in an ongoing process of conceptual puzzle solving - a series of mostly *a priori* "thought experiments" - the aim of which was to develop better theories of economic policy geared to the objectives of full employment and external equilibrium. Most of

the empirical work that took place during this period was on testing different versions of the PPP relationship as explained in chapter three and, to some extent, the testing of trade theories (see, for example, Deardorff 1985)[14].

The distinctive policy implications of the early monetary approach to the balance of payments were an important reason for its initial success against the perceived failure of Keynesian policies. Whether these could be regarded as novel predictions is unclear as no decisive empirical tests could conceivably be run to corroborate them. At best the policy implications of the monetary model were consistent with the patchy empirical evidence against those derived from the Keynesian framework. This type of evidence or 'test' is plainly unsatisfactory from a falsificationist point of view, being more in keeping with the discredited logic of confirmation instead. The evidence offered in support of the monetary approach may perhaps be better described as stylized facts. However, there was also a far greater willingness to subject the monetary models to more searching econometric tests, particularly the monetary and other asset market models of exchange rate determination, which has continued to the present day. What is of interest from a methodological perspective is that despite the acknowledged empirical failure of the monetary model against such tests it has prevailed and continues to be the subject of the latest, increasingly sophisticated econometric tests (see chapter five). By comparison, the Mundell-Fleming model was dumped unceremoniously before any serious empirical tests thereof had even been attempted. The reason for the rapid change in thinking from the Mundell-Fleming model to the monetary approach had less to do with its empirical merits or demerits than with its perceived conceptual inadequacies:

"Therefore, although the integration of asset markets and capital mobility into open economy macroeconomics was an important contribution of the Mundell-Fleming model, the model was largely rejected on *a priori* grounds as a serious contender for the explanation of exchange rate movements at the beginning of the recent float. This is because it was judged to contain a fundamental flaw - it is cast almost entirely in flow terms" (MacDonald and Taylor 1992: 3).

It was *a priori* conceptual considerations such as these which characterized theoretical advances in the study of exchange rates and the open economy, rather than the direct result of empirical tests. Neither falsification nor the corroboration of novel facts was seriously attempted (nor, it may be argued, was it appropriate to do so). An emphasis on *a priori* reasoning is a feature of the deductivist approach to the methodology of economics, a tradition going back to Senior, Mill, and Cairnes which reached its zenith in the 1930s with Lionel Robbins and which has recently enjoyed something of a revival (see Hausman 1992). The essence of deductivism is not that empiricism should be eschewed but that its role in economics is to explore the range of application and the limitations of a theory when applied to the real world. The results of empirical "tests" are not construed as either proving or disproving the theory in question. A negative test result or inconsistency with the observed data serves only to draw attention to those disturbing causes held constant by the *ceteris paribus* assumption or to logical errors of deduction. Blaug's (1980) position is that this type of weak verificationist approach to science, although not as bad as the anti-empiricism of the Austrian school and *Verstehen* doctrines, should be replaced by the more rigorous standards of falsificationism. The analysis of exchange rate and balance of payments theories in this chapter suggests otherwise.

Laudan's analysis of conceptual problem solving is consistent with the actual development of theories about the balance of payments and exchange rates reviewed in this chapter. For example, with regard to intra-scientific conflict Laudan suggests that a conflict between two competing theories in the same domain raises presumptive doubts about both theories (see chapter two). This may lead to the rejection of one of them but in many cases, and particularly it seems in economics, it is common for both theories to be retained with each theory being valid as a partial explanation for a given phenomenon.

This reasoning is evident in the development of the elasticities, absorption, monetary, and portfolio approaches to the balance of payments and exchange rates described above. It

is incorrect to think of any of these theories as being either "proved" or "falsified" by the evidence. The correct interpretation is to think of each of these theories as partial explanations for the phenomena in question, which are true over the relevant range of application of the theory, and which are not mutually exclusive. More general models may be constructed in which each of the partial approaches may be viewed as part of a larger system as attempted, for example, by Gylfason and Helliwell:

"Within the larger system...the different predictions of the partial approaches are based in part on ignoring other important parts of the system, and in part on particular assumptions about expectations and about the strength of the international linkages among national markets for goods and financial assets" (Gylfason and Helliwell 1983: 820).

A similar conclusion was reached by Michaely (1980) in a reconciliation of the PPP, elasticities, and absorption approaches to devaluation:

"I shall argue, however, that each of these approaches has a strong element of validity, in the sense that each provides a sound analysis under well-defined circumstances and that when interpreted in this way, bringing out the valid element of each approach, these three approaches constitute in fact components of one body of theory; whether to call them then one "theory" is a matter of taste, but it would certainly be legitimate to do so" (Michaely 1980: 34).

Michaely, however, regarded the monetary approach as indeed a different theory and rival to the PPP/ elasticities/ absorption model he developed from Liviatan (1979). But as noted earlier in this chapter, the monetary approach and these partial explanations are not mutually exclusive. Mussa (1984), for example, reviewed his earlier emphasis on asset market equilibrium as the critical determinants of exchange rates in a footnote commenting that:

"It is clear to me that one can arrive at an asset price expression for the exchange rate from a model that focuses on flow market equilibrium conditions. More generally, one must recognize that in any sensible model of exchange rates both asset market and flow market equilibrium conditions are important, and it is a matter of expository convenience which of them one chooses to emphasize" (Mussa 1984: 21).

It was further noted above that portfolio balance models go some way to integrate stock-flow interactions within a basically asset market framework. After comparing different asset market models of the exchange rate, including the portfolio balance model, Dornbusch concluded:

"We have reviewed a wide spectrum of exchange rate theories. There is little point in endorsing one particular formulation, since each of these models seeks to capture a special effect and thus is more or less suitable for a particular instance of policy analysis" (Dornbusch 1980: 27).

The above[15] also illustrates the general problem of exogeneity and the *ceteris paribus* assumption in economics. Unlike in the physical world, the real world of the social sciences does not often permit a relatively clean incision to extract the dominant explanatory variables in a particular case - as is sometimes pithily expressed by the phrase "everything depends on everything else", the explanatory variables in economic models are usually only weakly endogenous. It is thus common for economic theories to develop from partial to more general explanations of the phenomena under consideration, but not in the sense of the more general theory necessarily replacing (and thus falsifying) the partial explanations because it is a more successful empirical structural model. The more complex theory may indeed be inferior to the simpler partial models in predicting relevant testable facts. Under these circumstances it is not appropriate to think of the partial theories as having been falsified by negative empirical test results or other evidence which the more general model explains more readily.

Early criticism of the monetary approach well illustrates the conceptual problem of exogeneity in economics. Currie (1976), for example, argues that the distinctive policy conclusions of the monetary approach are arrived at only because the government budget constraint is ignored. As explained in section 4.3 the core assumption of the monetary approach is that, under a fixed exchange rate, a balance of payments deficit or surplus disturbs stock equilibrium in the money market and must therefore correct itself in the long-run. Currie points out that the leakage of foreign exchange reserves

base money associated with a balance of payments deficit, for example, can be neutralized by an injection of domestic base money to finance a budget deficit of similar size. In this case, private sector money stock equilibrium and the balance of payments deficit could be maintained until foreign exchange reserves and the availability of foreign credit are exhausted.

If the Keynesian assumptions of rigid prices and unemployed resources hold then Currie also shows that, contrary to the policy conclusions of the monetary approach devaluation, tariffs, import quotas, and exchange controls can permanently improve the balance of payments. Assuming an initially balanced budget and trade account, devaluation lowers the terms of trade and increases the relative price of tradable goods. If the Marshall-Lerner condition holds the trade account improves. Furthermore, the higher level of domestic aggregate demand leads to an increase in output and national income, resulting in higher tax revenues and a budget surplus. If the budget surplus is initially less than the trade surplus the inflow of reserves continues to induce increased private expenditure. This reduces the trade surplus and increases the budget surplus until they are equal. Private sector stock equilibrium in the money market is achieved and no automatic offsetting forces are generated to reduce the trade surplus any further. A similar process determines long-run equilibrium if tariffs or other measures to protect the balance of payments are imposed. If tariff revenues are retained by the government the equilibrium level of income is lower, and the trade surplus larger, than if such revenues are distributed to the private sector.

If perfect international arbitrage in goods and capital is assumed, along with the usual monetarist assumptions of flexible prices and full employment, then devaluation and trade protection measures alone cannot permanently improve the balance of payments. However, they can be effective if accompanied by expenditure reducing policies, as exlained under the absorption approach.

Currie's analysis also throws doubt on the direction of causation assumed by the monetary approach in the money base

identity. Monetarists assume domestic credit to be largely exogenous. If domestic credit and the money supply grow faster than the demand for money then, according to the monetary approach, the excess will be eliminated via a balance of payments deficit. However, reverse causation is just as plausible. Assume, for example, a large increase in the oil price (or a decline in the world price of a major export) which causes a trade deficit. The government may neutralize the deflationary implications of the trade deficit by running a budget deficit financed by increased sales of bonds to the central bank. Thus the drain in foreign exchange reserves may be offset by an increase in net domestic assets without disturbing private sector stock-flow equilibrium. Domestic credit increases, not because the central bank controls it, but passively in response to a real shock:

"This being so, analysis of the balance of payments in the longer-run should concentrate on these real factors and not on the accounting identity between reserves, domestic credit and money on which the monetary analysis focuses. In the longer-run, a real, not a monetary, analysis of the balance of payments is required" (Currie 1976: 514).

Thus the chain of causation may be the reverse to that assumed by the monetary approach. It seems that in most developed countries at least, central banks can sterilize balance of payments flows in the short-run.In the longer-run, as explained above, such flows can be neutralized by offsetting fiscal policies. This means that domestic credit is, to some extent, endogenous[16].

The monetarist response to such claims varies. For example, it is argued that a balance of payments deficit is unsustainable in the long-run because, under a fixed exchange rate, foreign exchange reserves and credit facilities will eventually be exhausted necessitating corrective action. Also, major trading partners may be unwilling to accumulate foreign exchange reserves and/or run a corresponding budget surplus. The point, however, is that the required corrective action and counter measures are discretionary and not the result of automatic offsetting forces generated by balance of payments deficits as claimed by the monetary approach[17].

From a methodological perspective it has been argued that the monetary approach may be seen as part of a broader Lakatosian research programme, that of monetarism. As pointed out in chapter two, however, the identification of research programmes in economics with a particular set of hard core propositions is problematic. Cross (1982) feels that economics does not contain indisputable hard core propositions that, *ex ante*, endure long enough to distinguish them from protective belt propositions[18]. Instead, Cross suggests that research programmes in economics can be identified more simply by means of their positive heuristics alone. Monetarism, for example, may be identified with the positive heuristic which directs members of the research programme to "explain sustained variations in the rate of inflation by sustained prior variations in the rate of monetary expansion" (Cross 1982: 335). The groupings of hypotheses thus making up the research programme can be appraised in the way proposed by Lakatos, that is, by assessing the degree of progression or degeneration in theoretical and empirical content as these hypotheses are amended.

In this context, Cross argues that the monetary approach constituted a progressive addition and amendment to the broader monetarist research programme. The monetary approach led to a progressive problem shift because it predicted more facts: under a fixed exchange rate and an endogenous national money supply, it predicts convergence in different national inflation rates; while under a flexible exchange rate it predicts changes in such exchange rates to be a function of divergent rates of national monetary expansion. Cross maintains that it is possible to detect an almost uninterrupted increase in empirical content from Friedman's (1956) reinterpretation of the quantity theory of money, to Friedman's (1968) analysis of the short-run trade-off between inflation and unemployment and the expectations augmented Phillips curve, through to Johnson's (1972) paper on the monetary approach to balance of payments theory. A further element of progressivity, according to Cross, was that most of the excess empirical content, including that of the monetary approach, was in fact corroborated up until 1973. After 1973, however, the following decade of floating exchange rates saw elements of empirical degeneration set in

such as the perverse behaviour of exchange rates like sterling in 1976 and 1980-1. These anomalies were explained by undesirable *ad hoc* auxiliary hypotheses (see chapter two on ad hocery) such as the effect of North Sea oil or changes in 'political confidence'.

Although it may be useful to interpret the monetary approach within the context of monetarism, Cross' reconstruction is not entirely true to Lakatos' MSRP. In particular, the monetary approach did not represent a theoretically progressive problem shift in the monetarist research programme in the sense of "predict(ing) some novel, hitherto unexpected fact" (Lakatos 1970: 118). It is more accurate to say that the monetary approach *post*dicted and explained some quite common and well known observations which monetarism had been previously unable to explain without resort to *ad hoc*$_3$ strategems. These observations could not be properly described as novel or unexpected facts. Rather, they were stylized facts which stood out as unexplained anomalies in the previous versions of monetarism. As such they awaited a more powerful analytic theory to account for them without making use of tacked on *ad hoc* explanations.

This distinction is significant because it illustrates a characteristic feature of many other theoretical advances in economics. For example, the attraction of Keynes' *General Theory* was not so much that it predicted novel facts. Besides clarifying a significant logical error in classical thinking its attraction lay in being able to explain, in a theoretically consistent way, the stark evidence of mass unemployment during the 1930s which could not be satisfactorily explained by classical theory at the time without resort to defensive immunizing strategems (such as defining the high rate of unemployment as 'voluntary' - see chapter two). The resurgence of monetarism during the 1970s, on the other hand, can be at least partially ascribed to its explanation of the stylized fact of stagflation which occurred at that time, which Keynesian theories could not do without bringing in factors extraneous to the theory. It would not, of course, be correct to say that monetarism predicted stagflation. Thus in strict

Lakatosian terms large parts of economics would be regarded as pseudoscience because it is not sufficient that theories merely explain 'what is there'; they must also predict unexpected facts:

"We *'accept'* problemshifts as 'scientific' only if they are at least theoretically progressive; if they are not, we *'reject'* them as 'pseudoscientific'. Progress is measured by the degree to which a problemshift is progressive, by the degree to which the series of theories leads us to the discovery of novel facts" (Lakatos 1970: 118).

Laudan's problem solving methodology appears to fit the subject matter of economics far better than the alternative falsificationist methodologies of Popper or Lakatos. This derives from Laudan's explicit account and emphasis of conceptual problem solving and from his rationalization for greater tolerance of *ad hoc*$_{1-2}$ modifications to theories to explain empirical anomalies (see chapter five for more on the latter). It is fairly clear that the progress made in theories of the balance of payments and exchange rate determination, either partially or as more general macroeconomic models, has been primarily through conceptual problem solving advances rather than by successive increases in the empirical content of the theories concerned. This point is illustrated further in chapter five where it is shown that the monetary and asset market models have largely been empirical failures. On purely empirical grounds there would be very little to choose between, say, the monetary model and the Mundell-Fleming model. A follower of either the Popperian or Lakatosian versions of falsification would be forced to concede by their own criteria that the evolution of exchange rate and/or balance of payments theories has been largely nonprogressive and irrational from a scientific point of view. Laudan's model enables us to avoid this conclusion. The problem solving model allows us to prefer one theory to another if it resolves an important conceptual problem, even if lacks any empirical advantage over its rivals.

Notes

1 This condition is represented by the Bickerdike-Robinson-Metzler formula which subsumes the Marshall-Lerner condition as a special case (see Krueger 1983: 31-40).

2 These conclusions must be amended in the large country case. A monetary expansion in the US, for example, will also significantly increase the world money supply (from an outflow of money via the balance of payments) and thus reduce world interest rates thereby shifting the FF schedule. Thus under a fixed exchange rate, monetary policy is not impotent in the large country case. Similarly it can be shown that with flexible exchange rates, fiscal policy is not ineffective in expanding output.

3 This type of development is quite common in economics. From a methodological perspective it suggests that Kuhn's distinction between 'normal' and 'revolutionary' science does not fit economics very well. Kuhn (1970) argues that scientists adhere unquestioningly to the dominant paradigm during normal science and that the jump to a new paradigm cannot be explained rationally. Since Mundell himself experimented with quite different conceptual models of the open economy, Kuhn's psychology or sociology of discovery does not appear to be a very good description of workaday economic research.

4 The key passages in Meade's book are in chapter eight, *The Meaning of Fiscal and Monetary Policy and of Internal and External Balance*. In explaining the effects of a change in relative (nominal) interest rates between two countries A and B the distinction between stock and flow demands is made as follows:

"In the first place, owners of capital in A may, as a result of the change, wish to shift a large part of their existing capital funds from A to B, either on long term in order to earn 5 per cent instead of 4 1/2 per cent or on short term in order to earn 3 per cent instead of 2 1/2 per cent. *Such a shift of capital may be very large and very short-lived*; and this is likely to be particularly true in the case of short term loans which are likely to be considered as a more mobile fund of capital by their owners than in the case of long term investments. *Once this shift of existing capital funds has taken place the strain on the balance of payments of country A which may have been very great while it was taking place will altogether cease*."

and

"But there is a second source from which the demand by residents in A for B's capital assets may proceed, namely, from the annual savings of residents in A who are seeking an outlet for this flow of savings and may be induced to invest a larger part of it in B as a result of the relative fall in yields on

capital in A. This source of demand for B's assets is as likely (perhaps, more likely) to be operative in the market for long-term assets as in that for short-term assets; it is not likely to be as great quantitatively as the first source, immediately after the reduction in A's interest rates; *but, unlike the first source, it is likely to constitute a continuing additional demand for B's assets*" (Meade 1951: 103) my emphasis.

5 More subtle considerations include the use of money to defer expenditure decisions and as a medium of search. These features arise from ignorance of the set of all goods, their characteristics and utility now or in the future (see Shackle 1972).

6 A reduction in St is defined as a lower domestic price of foreign money and thus as an appreciation.

7 The Mundell-Fleming model was developed during the Bretton Woods era with a fixed exchange rate system in mind and was thus originally applied to problems concerning balance of payments adjustment and devaluation. It can easily be adapted for application to the case of flexible exchange rates as, for example, in Kenen (1985).

8 This can be done in various ways. Speculators can take a leveraged position against a currency by selling it forward on the forward exchange market; investors/speculators may rebalance portfolios by selling assets (particularly liquid assets such as treasury bills) denominated in the currency concerned and switching to similar assets denominated in the currency(s) expected to be revalued; and importers may speed up foreign currency payments while exporters delay receipt and/or conversion of foreign exchange.

9 The failure of the Bank of England, which reportedly lost over a billion pound sterling in its attempt to support the exchange rate within the ERM, also underlines a well known disadvantage of a fixed exchange rate. If speculators are correct and the monetary authority is unable to prevent the currency from depreciating sharply, huge profits are realized. Whereas, if proved wrong, potential losses are limited due to the restricted appreciation possible with a fixed exchange rate.

10 The term 'global monetarism' is attributed to Marina Whitman (1975) from her paper titled *Global Monetarism and the Monetary Approach to the Balance of Payments*.

11 An interesting point made by Rodriguez (1980) is that with rational expectations the apparent dichotomy between stock and flow adjustment falls away. The flow implications of current account imbalances, for example, are immediately realized by

speculators and reflected in asset portfolio adjustments. Thus the spot exchange rate is completely determined by speculative capital flows so that the supposed influence of trade flows on the exchange rate is a redundant variable. The sudden adjustment in exchange rates which coincides with the release of trade data which is contrary to expectations is suggestive of this view. It is not the trade flows themselves which influence the exchange rate but rather the way in which this information is used by speculators in deciding whether to buy or sell the currency concerned.

12 Hallwood and MacDonald (1986) note that later research, where regressions were estimated for the mark/dollar exchange rate beyond this period, failed to support the real interest differential model.

13 Machlup's early attempt to apply the theory of the national income multiplier to an economy opened to international trade further illustrates the essentially non-empirical analytical approach at this time (Machlup 1943).

14 De Marchi's persuasive Lakatosian analysis of the Leontief paradox shows how critical counter-evidence was mostly ignored in the appraisal of the Samuelson-Ohlin factor-price equalisation theory. The decision to continue the pursuit of this research programme depended on analytical considerations. The programme encompassed a theoretical problem shift in which the implications of the factor proportions model were explored as an analytical special case of general price theory. Thus, where empirical testing of trade theories did occur, it was largely innocuous because such theories formed part of a broader research programme protected by the negative heuristic of its hard core (De Marchi 1976).

15 These points may be clarified and taken further by the notion of *commensurability* (see Feyerabend 1962; 1970 and Kuhn 1970). Researchers such as Michaely regard the monetary and asset market models to be incommensurable with the income-expenditure flow models of the balance of payments and exchange rate determination. Mussa, on the other hand, argues that asset models and flow models of exchange rates are not mutually exclusive - they can be combined as in the portfolio approach and are thus to some extent commensurable. However, although the two theoretical approaches may be combined it does appear to be the case that the asset market approach has established itself as the dominant or master framework for analysing exchange rates and balance of payments adjustment. Thus the (partial) Keynesian flow models, although not logically incompatible with the asset market approach (especially, as noted above, in the case of rational expectations versions thereof) are conceptually subservient to it and would require modification where necessary to make it compatible.

16 The issue whether the money supply is exogenous or endogenous is fraught with both conceptual and empirical

problems and has been the subject of intensive research. This issue and a review of the literature cannot be dealt with here. Suffice it to say that if the central bank cannot, or can but for various reasons does not, control the money supply then it becomes an endogenous variable, responding passively to changes in the demand for money. In this case the balance of payments must be viewed as exogenous and determined by real factors such as differential economic growth rates, changes in the marginal propensity to import, interest rates, fiscal policies etc. Some economists (for example, Dean 1988) argue that it is wrong to suppose that the money supply is either purely exogenous or endogenous. Although the central bank may not be able to control the money supply perfectly it can usually exert some *influence* over domestic credit and the money supply (through either direct open market operations or via the effect of interest rates on the demand for credit).

17 Darby (1980) also argues that the monetary approach is based on inaccurate assumptions about the economy.

18 Cross (1982: 331) notes that it is possible to identify candidates for hard core propositions *ex post* as those which have survived logical and empirical criticism. However, he suggests that it is doubtful whether members of a particular research programme in economics regard any propositions *ex ante* to be unchallangeable.

CHAPTER FIVE

EXCHANGE RATES AND THE MONETARY MODEL: EMPIRICAL ISSUES

In chapter four it was suggested that the succession of balance of payments and exchange rate theories could best be construed in Laudan's terms as progress in the conceptual problem solving abilities of the theories concerned. The preference for one theory over another appears to have been based on conceptual and analytical considerations which had very little to do with comparing the empirical proficiency of the different theories. Virtually no attempt was in fact made to empirically test these theories, particularly in the post Second World War policy theory phase, beyond somewhat casual references to selected institutionalized 'facts' about policy implications and the like.

With the advent of the monetary approach, however, there was a clear shift of emphasis towards empirical analysis and econometric 'tests' of the monetary model of exchange rate determination. As reviewed in section 5.3 most of the econometric test evidence has been negative and this has led many researchers to regard the monetary model as an empirical failure - yet, significantly, still worthy of pursuit within the broad asset market approach. Moreover, accompanying the disenchantment with the empirical proficiency of the monetary model, was a further shift in emphasis away from testing structural models in favour of empirical analyses of the time series properties of exchange rates and the efficiency aspects of the foreign exchange market.

This research helped uncover and clarify a further set of stylized facts about the behaviour of exchange rates, beyond the broad institutionalized facts of the policy phase era. Whatever the conceptual or analytical merits of the monetary model the empirical justification for it now turned on its 'consistency' with these new stylized facts about exchange

rates. Despite some notable exceptions the monetary and other asset market models are viewed as consistent with such facts (whereas the Keynesian flow models are not). These developments are of obvious interest from a methodological perspective. They suggest that serious attempts at falsification have not been adopted as a research strategy in this area of economics and, more pertinently, that it would not have been appropriate to do so. The empirical analysis of exchange rates has tended to be exploratory in nature, in helping to solve various empirical and conceptual puzzles, but has not 'tested' the underlying theory by providing adequate reasons for rejecting the theoretical framework of the asset market approach. As noted in section 5.5 Laudan's problem solving model of scientific progress appears to fit the evolving story of exchange rate economics better than a falsificationist approach.

This chapter examines the empirical evidence regarding the monetary model of exchange rates and related issues such as the efficiency of the foreign exchange market. The reasons for, and the response to, the acknowledged empirical failure of the monetary model are examined. Section 5.1 looks at some important stylized facts or broad empirical regularities in the behaviour of exchange rates. Although the monetary model has failed most econometric tests it is arguably better equipped to explain such stylized facts than rival theories. Section 5.2 analyses the efficiency issues more closely and suggests that because there is inadequate knowledge about the formation and operation of both expectations and risk in the foreign exchange markets, tests of the efficiency hypothesis have produced inconclusive results. The broad methodological implications of the efficient markets hypothesis for structural models of exchange rates are noted. In section 5.3 econometric evidence for and against the monetary model is then analysed and the reasons for its generally poor empirical performance are noted. This section emphasizes the evident failure to develop a structural model of exchange rates with a better than random walk forecasting capability. Some alternative research strategies that have been adopted in response to the poor performance of the monetary model are also discussed. The relatively new approach by MacDonald and Taylor (amongst

others), in which purchasing power parity and the monetary model are interpreted as long-run equilibrium conditions and tested using the more recent cointegration techniques, are singled out for closer analysis in section 5.4. Section 5.5 concludes the chapter by noting some significant implications of the foregoing for the methodology of economics, particularly as regards the issues of falsification and the question of scientific progress in economics.

5.1 Stylized facts and theories of exchange rates

The empirical acceptability of an economic theory and its success against rival theories is probably determined not so much by the evidence, positive or otherwise, from econometric tests as by its ability to explain stylized facts about the economy. Such "facts" may be thought of as broad empirical regularities[1] which are commonly accepted by researchers in the field as requiring a theoretically consistent explanation. A theory of exchange rates such as the monetary model may succeed in this regard even though it may fail most of the more discerning econometric tests that it has been subjected to. As made clear below, there is just too much that can conceivably "go wrong" in such tests, in particular the problems of expectations and model misspecification, that would lead us to reject the theory without helping us to decide on its overall explanatory value. It is perhaps partly for this reason that econometric tests are not, and cannot be, taken too seriously. What has proved more significant empirically in the evolution of thinking about exchange rates is rather the ability of the theories concerned to satisfactorily explain the various stylized facts with which they are faced. To use a metaphor, neither a square nor an oval peg may fit a round hole yet the latter may be justifiably regarded as a better model than the former, even if the average measure of deviation (the mean error) of the two pegs from the actual shape of the hole were the same.

It should be borne in mind that what comes to the fore as a stylized fact, in need of explanation, may change over time.

For example, during the Bretton Woods era of pegged exchange rates and controls over capital flows, the behaviour of floating exchange rates commanded less attention. The central stylized fact engaging the attention of economists at the time was the effect of devaluation on the trade balance. Only as exchange controls were relaxed and the size of capital account transactions grew, eventually swamping the monetary value of international trade flows, did new stylized facts emerge requiring integration with the existing theory, as in the Mundell-Fleming model of the open economy. The move to generalised floating exchange rates after the end of the Smithsonian agreement in 1973 led to the emergence of further stylized facts regarding the behaviour of exchange rates which were not, and could not easily have become, evident during the Bretton Woods era.

The recent history of stylized facts about exchange rates is regarded as broadly consistent with the general asset market approach but is unfavourable as far as both the Keynesian flow models (such as the Mundell-Fleming model) *and* the prototype monetary models are concerned. The early monetary model of the balance of payments appeared to explain the stylized fact that devaluation was as often as not unsuccessful in improving the trade balance, contrary to what was expected on the basis of flow models. The monetary model also initially appeared to receive some support from econometric tests both with regard to the balance of payments and subsequently when first applied to floating exchange rates, as documented in the Frenkel and Johnson (1976 and 1978) volumes respectively. However, as made clear below, later test results were mostly negative. With the increasing awareness of the role played by expectations in determining exchange rates came a realization of the difficulties that any empirical model, including the monetary model, would face in this regard. Papers by Mussa (1978; 1979), Dornbusch (1976a; 1976b), Frenkel (1981a), and Levich (1978) shifted attention to the full implications of the asset market view of exchange rates and to testing the efficiency aspects of the foreign exchange market. The research in this area spawned a number of new stylized facts about the behaviour of exchange rates, mostly incompatible with Keynesian flow models and the

naive monetary model. Attention was also focused on portfolio balance models which broadened the allowable range of assets to include the bond market.

The most striking characteristic of the foreign exchange markets is that freely floating exchange rates appear to follow a random walk such that:

$$s_t = as_{t-1} + e_t \ldots\ldots\ldots\ldots a = 1 \ldots\ldots\ldots\ldots\ldots\ldots (5a)$$

where s_t is the natural logarithm of the spot exchange rate and e_t is the error term. If the exchange rate contains a unit root (a=1) as in equation (5a) and the error terms are independent and identically distributed (iid) then the time series process is a random walk. The iid assumption is not usually borne out by the evidence as exchange rates often exhibit trends and are thus serially correlated. However, although iid is strictly necessary for the strong hypothesis of a pure random walk it should be noted that only a unit root is necessary for the weaker hypothesis of no mean reversion and the unpredictability of exchange rates. Also, the first differences of the time series in this case are stationary such that exchange rate changes (as opposed to the levels thereof) may be regarded as a random walk (see also the similar unit root studies on the real exchange rate in chapter three).

Models where the equilibrium exchange rate is determined by fundamental economic forces which change slowly predict similarly gradual adjustments in the exchange rate. Since actual exchange rate movements are erratic they inevitably appear to overshoot such equilibrium exchange rates. This suggests an asset market continually surprised by new information, as recognized in the paper by Frenkel (1981a). Most theories of exchange rates based on trade flows such as the relative price, factor proportions, and purchasing power parity theories thus fail to explain the essentially random movements in exchange rates. This criticism also applies to the early batch of monetary models in which the determination of exchange rates was linked to changes in relative excess money supplies via some version of PPP. Later models such as

Dornbusch's sticky price model, in which differences in the relative speeds of adjustment in goods and asset markets produce short-run overshooting of the exchange rate from its long-run PPP level, as well as the portfolio balance models, were more successful in this regard (although not in their ability to meet the more searching specifications of econometric tests).

Floating exchange rates are also often characterized by periods of relative inactivity punctuated by episodes of marked volatility, often with unidirectional exchange rate movements[2]. Although this may be a purely statistical feature of any random walk series, its occurrence in foreign exchange markets may be partially explained by the presence of central banks in the market[3]:

"Exchange rates are likely to be influenced by the actual and anticipated behaviour of central banks. As long as private market participants believe that central banks will take the actions necessary to hold an exchange rate near a given level, there will be little incentive to speculate against the authorities. But, once evidence becomes available that the authorities are unwilling or unable to hold the exchange rate, speculation will force it to change by a substantial amount" Mussa (1978: 11-12).

This also accords more closely with the asset market view of exchange rates than with naive monetary models in which the demand for currency is regarded as a function of interest rate differentials and relative growth rates of national income. Since these variables change relatively slowly over time they can only provide a partial explanation of exchange rates. Such models cannot easily explain, for example, sudden weekly deviations in exchange rates of say 5 or 10 percent whereas expectations augmented asset market models, in principle, can do so. This is not to say that a specific asset market model would have any greater success in predicting the future path of exchange rates: at present there is no satisfactory model of exchange rate expectations necessary to make this even a remote possibility. Indeed, the full implications of the asset market approach suggest that testable predictions of this kind are not possible at all in what are essentially stock markets like the foreign exchange market.

The unpredictability of exchange rates is further underscored by the observed relationship between forward and spot exchange rates. The forward exchange rate may be thought of as the market's prediction of the future spot exchange rate. If speculators expect the exchange rate to appreciate they will buy forward exchange and vice versa. Therefore, the forward exchange rate rapidly approaches the anticipated future spot rate and may be regarded as the market's prediction of the future spot rate based on all the relevant information currently available:

$$F_t x = S_{t+x} + e_t \quad \text{.....(5b)}$$

where $F_t x$ is the forward exchange rate of the currency at date t for delivery x days in the future and S_{t+x} is the spot exchange rate expected to prevail at that time. The residual error term e_t is large since the forward exchange rate is usually a poor predictor of the future spot rate[4]. Most of the variation in spot exchange rates is attributable to unexpected changes therein. Only a small proportion of such variation is explained by anticipated changes reflected in the forward rate. Again this suggests an exchange rate adjusting to new information surprising the market or "news".

A related empirical regularity is the covariance between spot and forward exchange rates. The forward exchange rate may be at a discount or a premium to the spot exchange rate depending upon the interest rates prevailing in the countries concerned. However, changes in the two exchange rates are closely contemporaneous and of similar magnitude:

$$\Delta F_t x = F_t x - F_{t-1} x = \Delta S_t = S_t - S_{t-1} \quad \text{.....(5c)}$$

This relationship is also indicative of an efficient asset market. In an efficient foreign exchange market, persistent arbitrage opportunities to make riskless profits by simultaneously buying foreign currency in the forward market and selling it spot (or vice versa) should not exist[5].

Anticipated profits from such arbitrage will only be zero, however, when the premium (or discount) on forward exchange equals the excess of the domestic nominal interest rate over the foreign interest rate. This state of equilibrium is the familiar covered interest parity condition (not to be confused with the hypothesized uncovered interest parity condition explained in chapter four):

$$F_t/S_t = (1 + i_t)/(1 + i^*_t) \quad \text{.....(5d)}$$

Together with equation (5c) this implies the observed fact that the actual variation in spot exchange rate returns is greater than both the interest differential and the forward premium. Again, this suggests an asset market in which the exchange rate rapidly adjusts to news and changes in perception about future returns. The interest rates in equation (5d) also provide a link between exchange rates and inflation. The Fisher effect suggests that nominal interest rates are comprised of a real interest rate (determined by factors such as the time preference of consumption and the productivity of capital) plus a premium reflecting the anticipated inflation rate. Thus as noted by Frenkel (1977) if the difference in expected inflation rates between two countries is very large then the factors determining real interest rates may be ignored. In this case the forward discount provides a useful observable measure of the anticipated inflation differential (as, for example, during the hyperinflation in Germany in the early 1920s).

De Vries (1994) points out that some statistical phenomena observed in the behaviour of exchange rates do not have any convincing economic explanation - none of the available theories can account for them so that various *ad hoc* explanations are resorted to. One such example is the fat tail phenomenon in which exceptionally high or low exchange rate returns occur more frequently than under a normal distribution of returns. A related observation is that periods of quiescence and turbulence in the foreign exchange market tend to cluster together. As noted by Mussa above, however, these phenomena may be partially explained by the response of speculators to the intervention activities of central banks in the market.

De Vries (1994) also lists those relationships which make sense and are expected to hold true on the basis of *a priori* economic principles but for which there is very little empirical evidence such as PPP and uncovered interest rate parity, two important assumptions of most monetary models of exchange rates. He concludes with the observation that the predictions of high frequency (typically monthly data) reduced-form exchange rate models do not outperform simple no change forecasts (see section 5.3 for a closer look at the evidence in this regard). In other words there is to date no compelling evidence for any structural model of exchange rates based on economic fundamentals, let alone a theory capable of meeting the more stringent criteria of falsification:

"The macro oriented exchange rate literature after the demise of Bretton Woods has largely been an epitaph on the fundamentals models of exchange rates" (De Vries: 365).

Thus the stylized facts presented in this section are not fully consistent with any of the structural models discussed in chapters three and four, although later versions of the monetary and asset market approach are better equipped in this regard. The available stylized facts suggest instead that the foreign exchange market behaves like an asset or stock market in which exchange rates and other interdependent financial market variables rapidly adjust in response to new information and changes in expectations. These adjustments take place to ensure (virtually) continuous stock equilibrium in the relevant markets such that the existing stocks of financial assets are willingly held. It is, of course, possible to rehabilitate the naive monetary model along these lines using rational expectations. Under this approach, news about those variables which are believed to influence the demands and supplies of monies will alter expectations about the future path of the exchange rate, thereby triggering changes in the present value of the exchange rate. As noted by Mussa in chapter four, the trick is to pin down those variables and to specify the way in which they effect expectations. The difficulty is that the variables and their functional relationships do not appear to be time invariant. If they are unstable over time then it makes the task of developing an empirical model in which exchange

rate predictions are consistently linked to the selected fundamentals a somewhat specious exercise. These difficulties are highlighted by the routine econometric failures of the monetary model.

5.2 Testing the efficiency of the foreign exchange market

The problems in the way of testing the monetary and other structural models led to greater research emphasis on the statistical properties of exchange rate behaviour and away from the direct question of exchange rate determination[6]. A productive line of research was the application of the efficient markets hypothesis (EMH) to foreign exchange markets - research which has helped uncover and clarify many of the stylized facts discussed above. The efficiency issue is also important from a methodological perspective. If the efficiency hypothesis is substantially correct then the idea that there is a "true" model out there waiting to be discovered and which can be used to consistently predict (or postdict) exchange rates does not make much sense. The best we can hope for is an explanatory theory which helps us to qualitatively understand the past behaviour of exchange rates, perhaps in the sense of explaining the kind of stylized facts referred to above.

There is a close connection between EMH and rational expectations which supports this view. In a nutshell, EMH suggests that all publicly available and relevant information is rapidly reflected (or discounted) in the prices of the assets concerned[7]. Since rational expectations are based on the information content of the best available model, the predictions of this model would similarly be reflected in the current asset price. Therefore even if the "true" fundamentals model of exchange rates were known its success in predicting the future path thereof would prove transitory. The information of the model would be quickly reflected in the spot exchange rate, preserving the efficient markets conclusion that the best predictor of the next period's exchange rate is the current spot rate.

Despite some early evidence favouring EMH there are some notable exceptions which means that the hypothesis cannot be accepted unconditionally with respect to foreign exchange markets. But before looking at the evidence in more detail below it should be pointed out that these anomalies do not necessarily invalidate the EMH. There are two reasons for this: the possibilty of a risk premium being attached to a currency and the difficulty in modelling expectations.

As regards the latter, the statement that the foreign exchange market is efficient implies an equilibrium model of expectations as well as the proposition that economic agents do not systematically ignore or misinterpret relevant information. A negative econometric test result could mean either that the market is inefficient or that the wrong model of expected prices (in this case exchange rates) was used. Consider the equation:

$$Z_{t+1} = P_{t+1} - E[P_{t+1} \mid t] \quad \text{.....(5e)}$$

Cast in this form it is apparent that we are testing a joint hypothesis. Equation (5e) states that the one period return on an asset Z_{t+1} is the difference between its actual realized price P_{t+1} and the expected price $E[P_{t+1}]$, given the information available at the start of the period ($\mid t$). In the context of the foreign exchange market the returns would be the one period change in the exchange rate while the expected price would be the forward exchange rate at time t. If the distribution of returns (the Z_t) are found to be serially uncorrelated with a zero expected value then this is strong evidence that the market is indeed efficient. No unusual profit opportunities exist - that is, no speculator can consistently make unusual gains (or losses for that matter).

If, however, the Z_t *are* found to be serially correlated this could *either* be due to the foreign exchange market being inefficient *or* to an incorrect model of the expected equilibrium exchange rate. It follows that the behaviour of observed changes in exchange rates cannot be used to judge market efficiency independently of a standard provided by such

a model. This implies, for example, that non-random changes in exchange rates do not necessarily imply inefficiency in the foreign exchange market. As Levich puts it:

"Random price movement *per se* is neither a necessary nor sufficient condition for market efficiency. If the expected equilibrium price varies considerably, market efficiency requires non-random price movements" (Levich 1979: 247-8).

To illustrate this point, consider the diagram below in the context of the foreign exchange market:

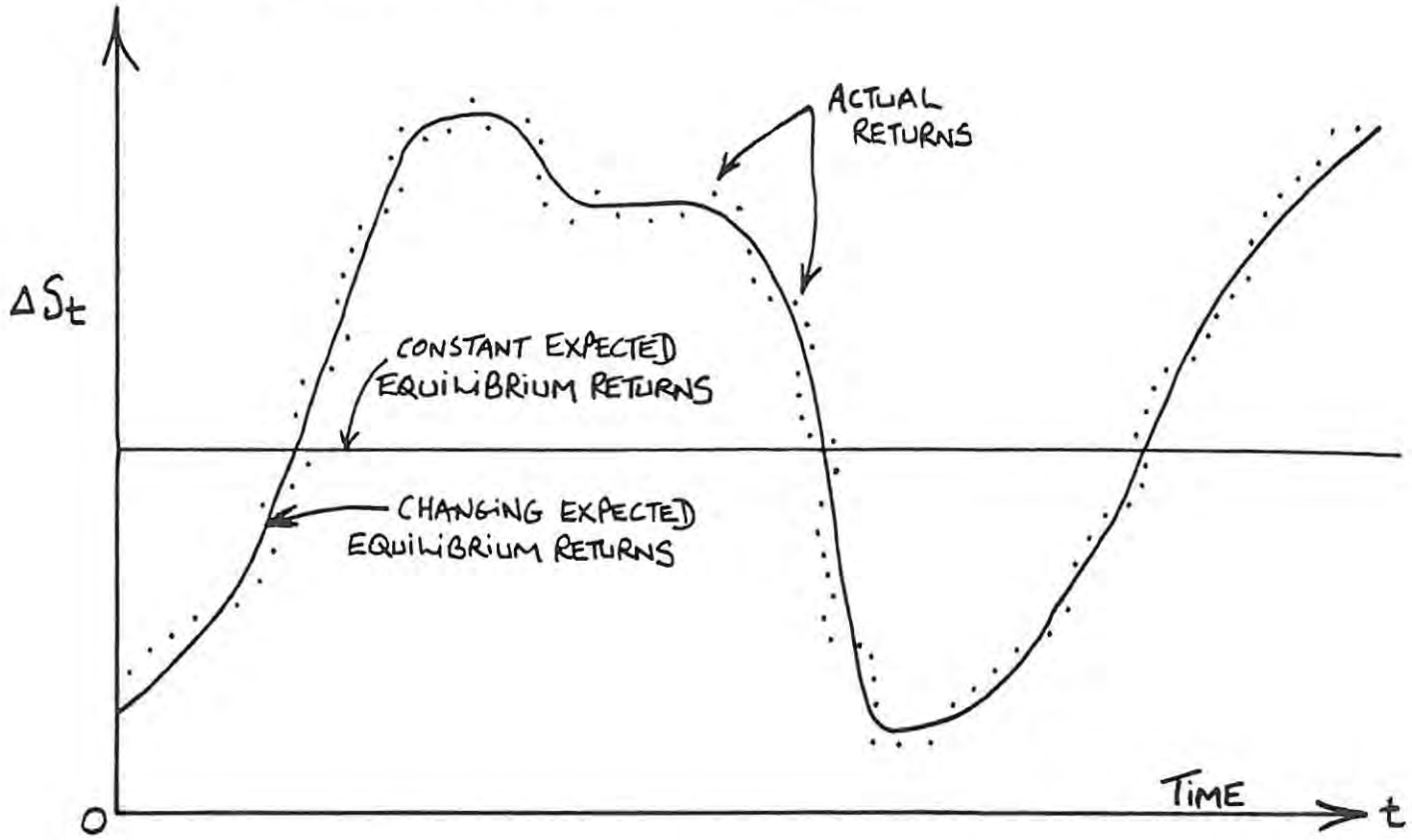

Figure 4 Efficient market behaviour dependent upon a model of expected equilibrium returns
Source: Levich (1979: 247)

In figure 5.1 actual returns (changes in the exchange rate) are serially correlated relative to a model in which expected equilibrium returns are constant. In this case the EMH would be rejected. The presence of serial correlation suggests that profit maximizing speculators have neglected information significant in determining the exchange rate. However, actual returns do follow a random walk relative to a model in which the expected equilibrium exchange rate varies. Thus the EMH

would be accepted in this case. We thus return to the basic conundrum which vitiates so many econometric test results - the absence of a satisfactory model of expectations means that there can never be any decisive test results of an economic theory which includes a joint hypothesis about expectations. Until such a model is discovered, which in the nature of the subject matter appears to be unlikely[8], all such test results can only provide inconclusive elements of circumstantial evidence.

A related problem which further hinders efficiency tests of the foreign exchange market is the possibility of a risk premium being attached to particular currencies at certain times. If, for example, it were found that the forward exchange rate systematically underpredicted the future spot exchange rate it cannot necessarily be concluded that the foreign exchange market is indeed inefficient. The forward discount may simply reflect the real or perceived risks of purchasing the foreign currency concerned. Thus the expected excess profit which can be earned by purchasing the currency at a forward discount is the required compensation for bearing such risk.

As might be expected, in the light of these problems the evidence regarding foreign exchange market efficiency is mixed. Levich (1978) and Frenkel (1981a) tested an ordinary least squares regression version of equation (5b) using logarithms of the spot and forward exchange rates of the pound, french franc, and the mark against the US dollar for the period 1973-79:

$$s_{t+1} = a + bf_t + e_{t+1} \qquad (5f)$$

They interpreted their results as highly supportive of the joint hypothesis of foreign exchange market efficiency and no risk premium. In both cases the coefficient *a* did not differ significantly from zero while *b* did not differ significantly from unity. The generally high R^2 statistics and the absence of first order serial correlation in the error terms suggests that information deemed relevant by the foreign exchange market is rapidly reflected in the forward rate f_t.

However, Hansen and Hodrick (1980) and Cumby and Obstfeld (1984) amongst others point out that such tests are inappropriate if there are trends present in exchange rate changes over time. In this case it is suggested that the data be detrended by subtracting the spot exchange rate s_t from each side of equation (5f). In this form the EMH implies that a currency which is at a forward discount (premium) should on average depreciate (appreciate) by the same percentage over the relevant time period - in other words the forward discount (or premium) should on average correctly forecast the realized change in the spot exchange rate. Most tests reject the hypothesis of exchange market efficiency in this context (see, for example, Boothe and Longworth 1986 and Taylor and MacDonald 1989). The results of these tests imply that unexploited profits exist: by simply purchasing those currencies which are at a forward discount against other currencies a speculator stands to gain from both the higher interest rate (implicit with a currency at a forward discount) *and* an appreciation in the relevant exchange rate. Further evidence against EMH was reported by Frankel (1979b) where he tested for independence between consecutive period forecast errors. Using weekly data for the franc, pound, and lira exchange rates against the dollar for July 1974 - April 1978 he rejected the hypothesis of independence and the implied orthogonality property of EMH[9].

At this point, however, the efficient markets dilemma *vis-a-vis* risk emerges: do these test results really mean that the foreign exchange market is inefficient or do they simply indicate the size of the risk premium that may be attached to certain currencies from time to time? Frankel (1986) maintains that the size of the deviations from the EMH is too great to be explained solely by the presence of a risk premium. Bilson (1979a), on the other hand, argues that since tests of the Eurocurrency markets (where foreign exchange risk is negligible) strongly suggest that they are efficiently arbitraged, extraneous factors such as risk should bear responsibility for ambiguous test results in other currency markets rather than conclude that such markets are inefficient.

To the extent that the foreign exchange market behaves like an efficient asset market, the idea that it is possible to develop an enduring structural model of exchange rates may prove illusory (the converse, however, is not necessarily true - even if the evidence of further research shows that foreign exchange markets are in fact inefficient this does not imply that better structural models with superior predictive capabilities can be developed). The efficient markets approach implies that exchange rates are inherently unpredictable whereas a structural model thereof is based on the premise that, given sufficient advance information about the relevant fundamentals, it is possible to predict (albeit imperfectly) the direction and perhaps even the magnitude of changes in the exchange rate.

5.3 Empirical exchange rate models

The empirical evidence regarding structural exchange rate models, specifically including monetary models, is mostly negative. Any falsificationist worth his salt would have long since discarded the relevant theory from which such models are derived. But to do so would be tantamount to a cure which is worse than the disease itself because it would leave economists without any explanatory theory of exchange rates, unsatisfactory though it may be. For this reason researchers have not attempted in practice to disprove the monetary model but have persisted with a less demanding confirmationist methodology - checking to see how well the model fits the data and if necessary resorting to various conventionalist strategems to preserve the underlying theory. Exchange rate economics is replete with the use of *ad hoc*$_{1-2}$ hypotheses to solve awkward empirical problems. Econometric 'tests' of the models concerned are not intended to deliver a knock out blow to the theory but to explore the possible range of application of the theory and to open up new avenues for further research. Researchers have had little choice but to adopt such a methodological position given the insurmountable empirical difficulties which stand in their way in this (and presumably many other) areas of economics. Thus the workaday research strategies adopted by such economists appear to accord more

closely with Laudan's problem solving model than with conventional falsificationist methodologies.

Since the advent of floating exchange rates in 1973, three phases may be distinguished in testing the monetary model of exchange rate determination. The period up until about 1978 was characterized by test results that were interpreted as generally supportive of the monetary approach (although even at this stage there were some important exceptions) and which contributed to the initial burst of enthusiasm that accompanied the new theory, thereby helping to establish it as a serious alternative to the orthodox Keynesian models. A second phase of tests may be discerned for the post 1978 period to around 1987. These were far less supportive of the monetary model, with most test results unable to find estimated coefficients significantly different from zero and/or correctly signed. Furthermore, out-of-sample tests were unable to find that popular monetary models were any better than naive random walk models in their ability to predict the path of exchange rates over time. From 1987 until the present there appears to have been a revival in the fortunes of the monetary model, in tandem with more favourable evidence regarding purchasing power parity theory. This has resulted from a reinterpretation of the monetary equation and of PPP as long-run equilibrium conditions and the use of more powerful econometric techniques to uncover such relationships.

The first and second phase tests of the monetary model are well summarized by MacDonald and Taylor (1992) and Taylor (1995). As regards the first phase, research by Frenkel (1976); Bilson (1978); and Hodrick (1978), all subsequently published as chapters in *The Economics of Exchange Rates* edited by Frenkel and Johnson (1978), is broadly supportive of the flexible price monetary model prior to 1978. Further evidence favouring the monetary model for this period resulted from research by Bilson (1979a), Putnam and Woodbury (1979), and Dornbusch (1980) while complementary research on the efficiency of the foreign exchange market was conducted by Levich (1978; 1979), Bilson (1981), and Frenkel and Levich (1975; 1977). Frankel (1979a) used long bond interest rate differentials as a proxy for

expected inflation and found that the coefficients on both the interest rate and expected inflation terms were significant. He interpreted his results (for the mark/dollar exchange rate between July 1974 and February 1978) as supportive of his real interest rate differential version of the monetary model and as rejecting the alternative flexible and sticky price versions thereof.

The results of second phase tests conducted after 1978, where the sample period included more recent data, generally failed to corroborate the positive earlier findings. The ability of the monetary model to track changes in the in-sample exchange rate deteriorated markedly: "few coefficients were correctly signed (many were incorrectly signed); the equations had poor explanatory power as measured by the coefficient of determination; and residual correlation was a problem" (MacDonald and Taylor 1992: 11). For example, in applying the real interest rate differential model to the mark/dollar exchange rate on sample data extended beyond 1978, Frankel (1983) found that his estimates were either not significant or of the incorrect sign and he felt compelled to label his results as a "disaster" for the monetary model, in sharp contrast to the favourable results he reported in his 1979a paper. Of particular concern were the results of tests of the mark/dollar exchange rate which were the exact opposite of that hypothesized by the monetary model: the price of the mark rose (appreciation) as its supply increased! The 'incorrect' sign on the money supply variable has been a common finding in empirical research on the monetary model.

Frankel (1982) had attempted to explain the "mystery of the multiplying marks" by introducing wealth as an independent variable in the money demand equations (with wealth defined as the sum of financial claims on the government plus cumulative current account surpluses). He argued that the large German current account surpluses during the late 1970s redistributed wealth from US to German residents, thereby increasing the demand for marks independently of the other variables. His hybrid monetary-portfolio balance equation provided a better fit to the data such that most of the estimates, including the

coefficients on home and foreign wealth (but not on the income terms), were correctly signed and statistically significant. This is an instructive example of the use of an *ad hoc* conventionalist stratagem by an eminent practitioner in the field. To digress briefly, should such responses be deplored and avoided in the interests of conforming to falsificationist criteria *a la* Popper, Lakatos, and Blaug? Or does the subject matter of economics in this area force such an approach on the researchers concerned, who are essentially making the best of a bad job? And is progress in a discipline like economics advanced or retarded thereby? Laudan's problem solving model, by way of contrast to the falsificationist approach, allows such *ad hoc* modifications if they solve important empirical problems and, if no additional conceptual problems are created thereby, Laudan's model suggests that scientific progress has been achieved (section 5.5 discusses these issues further).

In a later paper, Frankel (1984) estimated competing versions of the flexible price, sticky price, and portfolio balance models of exchange rate determination, as well as a synthesis of the portfolio balance and monetary equations. The results of the estimation for five different currencies against the dollar for the period 1974 to mid 1981 were generally negative or only weakly supportive. In particular, Frankel found some evidence favouring the sticky price model over the flexible price model although "the results must be pronounced poor for both versions" which led one "to consider possible ways of 'patching up' the monetary model" (Frankel 1984: 239). This may be seen as an improvement on the preliminary results of his 1983 research above which, as noted above, he interpreted as disastrous for the monetary model. The results for the portfolio balance models were similarly poor, particularly for Germany where the estimated coefficients on the mark and dollar assets (bonds) had the wrong sign. In the synthesized model the coefficients on the risk premiums were significant and of the right sign for most of the currencies studied, in contrast to the insignificant results from the estimation of the stand alone portfolio balance model. However, the coefficients on the variables taken from the monetary model were mostly insignificant and led Frankel to conclude that the synthesized

model was not superior to the sum of its parts. He attributed his results to possible shifts in money demand and the long-run real exchange rate and recommended further research in these areas - but still within the framework of the monetary model.

The results from other researchers testing different versions of the stand alone portfolio balance model are generally inconclusive as, for example, in Branson, Halttunen, and Masson (1977; 1979), Dooley and Isard (1982), and Boughton (1988). Although the portfolio approach is intuitively preferable to the monetary approach since it provides a conceptual framework for the role played by changes in wealth, current account balances, and risk premia in exchange rate determination, the test results are mostly no better than those of the simpler monetary model. Also, while the portfolio approach is more "realistic" it is more complex and thus more difficult to model empirically. For example, Branson *et al* identified the stock of foreign assets with the cumulative current account balance. This fails to distinguish which part of the foreign asset stock is relevant to the determination of the particular bilateral exchange rates concerned. A further complicating factor with portfolio balance models that relax the assumption of static expectations is how to model the effects of changes in wealth and/or current account balances on exchange rate expectations. Speculators take a view regarding the sustainability of current account imbalances and buy or sell the relevant currencies accordingly. In practice it has been inordinately difficult to capture the way in which such expectations are determined.

The above raises a pertinent question as regards methodological issues. On the face of the empirical evidence presently available the portfolio balance model performs no better than the monetary model in tracking the in-sample path of the exchange rate. But, as discussed above and in chapter four, there are good reasons to regard it as a superior *conceptual* model of exchange rate determination (just as the monetary model, in turn, may be regarded as a better conceptual apparatus than the Mundell-Fleming model). Falsificationist criteria alone thus cannot determine the choice between competing structural models of the exchange rate. This choice

is determined, at least in part, not by empirical evidence but by conceptual considerations which are largely *a priori* in nature. The emphasis placed by Lakatos on the successful prediction of novel facts (that is, empirical corroboration thereof) thus appears to be inapplicable in this area of economics.

A good starting point in evaluating the empirical evidence regarding *out-of-sample* predictions of exchange rates by the monetary model is the paper by Meese and Rogoff (1983). The authors compared the out-of-sample forecasting accuracy of selected structural and time series exchange rate models. Their general conclusion was that a naive random walk model performed as well or better than any of the estimated models for the time intervals tested. This conclusion was reached despite the exchange rate forecasts being based on the actual realized values of the future explanatory variables. Their conclusions devastated the conventional macroeconomic model building approach at the time and provoked both critical and supportive responses from concerned researchers in the field. Indirectly their results tended to bear out the broad implications of EMH as regards exchange rate determination as discussed above.

Meese and Rogoff compared three variants of the monetary model (chapter four describes the basic stylized models in context): the flexible price model (Frenkel 1978b and Bilson 1978; 1979a); the sticky price model of Dornbusch (1976b) and Frankel (1979a; 1981); and a hybrid sticky price portfolio balance model including the current account (Hooper and Morton 1982). Several univariate time series forecasting models were also tested using both actual and pre-filtered data with various lags. Included here was the random walk model where the current spot exchange rate is used to predict future spot exchange rates. Forecasts of each model were tested over one to twelve month time horizons for the pound, mark, yen, and trade weighted exchange rates against the US dollar commencing March 1973 through to June 1981. It is important to note that the parameters of each model were estimated using the most recent information available at the time of a given forecast. This was achieved by means of rolling regressions to re-estimate the

parameters for every forecast period. Out-of-sample accuracy was measured by three statistics: the mean error, mean absolute error, and the root mean square error. In comparing the accuracy of the forecasts generated by the various models the authors concluded[10]:

"The structural models in particular fail to improve on the random walk model in spite of the fact their forecasts are based on realized values of the explanatory variables. They predict much worse, especially at one month horizons, if serial correlation is not accounted for" (Meese and Rogoff 1983: 12).

Meese and Rogoff analysed various possible reasons for the poor out-of-sample performance of the structural models they tested including simultaneous equations bias, unstable shifts in the parameters of the reduced form equations, and possible misspecification of the equations. As regards the possibilty of unstable parameter shifts during the sample period, Meese and Rogoff (1983) argued that unless the parameters themselves followed a random walk such shifts cannot explain why the structural models fail to outperform naive random walk forecasts. They also investigated four sources of possible model misspecification: the uncovered interest parity hypothesis, the proxies for inflationary expectations, and the specifications for the goods market and the demand for money (in particular the unstable velocity of the US dollar). None of these were regarded as strong enough reasons to fully account for the inferior forecasting performance of the structural models *vis-a-vis* the random walk model.

Simultaneous equations bias was singled out for more intensive investigation in Meese and Rogoff (1984) where, in an attempt to overcome this problem, they imposed theoretical coefficient constraints on relative money suppplies, the arguments of the money demand functions, and the speeds of adjustment to PPP. None of the new coefficient estimates were found to be any better than the random walk forecasts for time horizons under twelve months. They did find some improvement at longer time horizons but concluded that in general the results were insufficient to show that simultaneous equations bias or sampling error were to blame for the poor forecasting performance of the structural models.

Most of the later (post 1978) empirical research into the monetary model of exchange rates, with the important exception of some long-run cointegration studies (see section 5.4), lend support to Meese and Rogoff's findings. Virtually all the models proposed have failed, or at best are only weakly supported by, the econometric tests to which they have been subjected. Frankel (1984: 239) noted that the field had entered "an introspective and skeptical phase, after the initial enthusiastic burst of model building and estimation that followed the beginning of floating exchange rates".

There has been no shortage of explanations for the poor empirical performance of the monetary model. Some of the more common reasons are summarized by Lane (1991):

a) Purchasing power parity does not hold in the short-run. Inflation differentials are not fully offset by changes in the exchange rate so that relative prices and the real exchange rate vary significantly in the short-run (although this can be explained to some extent by sticky price monetary models).

b) The hypothesized uncovered interest rate parity relationship has been difficult to verify empirically because of the difficulty in establishing an independent model of exchange rate expectations. Furthermore, the proposition that changes in a hypothesized risk premium are responsible for deviations from interest parity has received only limited support. For example, Cumby and Obstfeld (1981) were unable to determine whether their estimated deviations from UIP were due to the presence of a (possibly changing) risk premium attached to certain currencies or to systematic differences between actual and expected changes in the exchange rate.

c) Instability in the demand for money function in the exchange rate equation and/or misspecification of the simple functional forms hypothesized for such equations. Also, the common simplifying assumption that the parameters of the money demand function are the same for different countries may well be invalid.

d) The assumption of exogeneity in the money supply ignores the possible existence of a reaction function for the monetary authorities including the exchange rate as a target variable.

e) The possibility of bandwagon effects or "bubbles" in the foreign exchange market in which the exchange rate may stray from its fundamentals determined equilibrium. If indeed present these would further contribute to the empirical failure of the monetary model.

In evaluating these and related problems with the empirical analysis of the monetary approach, Lane (1991: 216) concluded that "it is perhaps less surprising that the model has failed empirically than that it ever appeared to succeed at all".

Some researchers have tried to isolate and measure the relative importance of some of the possible causes of failure of the monetary model. Smith and Wickens (1986) used a distinctive methodology whereby they explicitly modelled hypothesized sources of misspecification of the monetary model using pure time series techniques. They tested the flexible price, rational expectations version of the monetary model associated with Frenkel (1978b) and Mussa (1978) and estimated how much of the variance in the model's residual error term could be explained by deviations from PPP, money supply innovations, and unstable money demand functions respectively. The reduced form equation they estimated is similar to equation (4g) except that it includes the separate sources of structural errors, which in conventional reduced form equations are combined as the residual error term. Smith and Wickens also estimated how well the flexible price monetary model approximated a random walk by checking whether the lagged innovations (singly and combined) improved the model's fit to the data and thus beat a random walk in forecasting in-sample exchange rates (that is, *ex post* but not *ex ante* forecasting ability).

By comparing the extra contribution of the additional time series terms to the original restricted reduced form equation Smith and Wickens were able to measure the relative importance of the different sources of misspecification. This information

can be used to help find the optimal respecification of the relevant equations. Their approach avoids an open ended search of the whole structural model which might require the inclusion of newly specified exogenous variables. The authors note, however, that their approach is not a substitute for the ideal which is a properly specified structural model.

Using quarterly data for the pound/dollar and mark/dollar exchange rates over the period March 1973 to March 1982, Smith and Wickens concluded that the poor fit of the monetary model was due to both structural and dynamic misspecification of the model. They attributed the main source of misspecification to short-run deviations from PPP (accounting for roughly 64 per cent of the total unexplained variance in the residuals), followed by money supply innovations and unstable money demand functions. Their results also showed that although a random walk model seems a good description of the behaviour of the two exchange rates (as suggested by the absence of serial correlation) the lagged structural errors improved the fit of the monetary model which could thus be used to beat random walk forecasts of the in-sample exchange rates. This implies that the monetary model is misspecifed dynamically as well as structurally.

As in closed macroeconomic models, instability in the demand for money has been suggested as a confounding factor in empirical tests of the monetary model of exchange rates. Boothe and Poloz (1988) investigated this possibilty and the effect of possible dynamic misspecification using an autoregressive, distributed lag generalization of Frankel's (1979a) real interest rate differential model. Boothe and Poloz hypothesized that for both the US and Canada, shifts in the demand for money due to non-trivial financial innovations (such as the introduction of interest bearing cheque accounts) may have influenced the statistical significance of the estimated coefficients. For both countries they accordingly adjusted the money supply data ($1971Q_1$ - $1984Q_4$) to account for these innovations. They estimated three versions of the basic reduced form equation: a simple version with no lag dynamics; a best fit lagged equation; and the latter with a linear relative

money supply homogeneity restriction imposed (HOD1). In each case they compared the relevant test statistics before and after the shift adjustment of the monetary aggregates.

Before adjustment the simple equation coefficients on foreign money, domestic and foreign inflation were incorrectly signed, that is, contrary to the theory. The lagged version without restrictions significantly improved the fit of the model but the coefficients were still incorrectly signed. In the third version the HOD1 restriction was rejected by the data. Substituting the shift adjusted measures for the official monetary aggregates and re-estimating the three equations helped to improve the fit to the data in each case[11] (reduced standard errors and higher R^2). However, the adjustments did not correct the contrary signs on key parameters of the model and the homogeneity restrictions were still rejected.

Boothe and Poloz concluded that none of the models was satisfactory as in each case at least one of the theoretical restrictions was rejected by the data. Thus the adjustments made to account for possible unstable money demands could not fully explain the poor performance of the monetary model. However, their results showed that misspecification of the equations' dynamics was significant in this regard. In the light of their findings, Boothe and Poloz suggested the need for a richer monetary model, specifically including the influence of real shocks such as changes in the price of oil and prolonged fiscal imbalances.

The research strategies that have developed in response to these problems vary but in most cases they have attempted to resolve such problems rather than adopt a wholly new model or approach (an exception is Lane 1991 who resorts to a rival open economy version of the IS/LM model). One gets the impression that the limited empirical support for the monetary and portfolio balance models has been sufficient to confirm a decision to accept the models taken on largely *a priori* grounds. The mass of potentially falsifying evidence and test results has been ignored or at least accommodated in the sense

that it has not led to the outright rejection of the models but only to modifications thereof.

A strategy to deal with the problem of misspecification has been to estimate a structural model of the whole macroeconomy using a system of simultaneous equations. This is believed to be superior to estimating single equation reduced form models in capturing the greater complexity of interactions between the exchange rate and other macroeconomoic variables - that it is more realistic. There appears to have been some success with this approach [see Isard (1988); Papell (1988); and Masson (1988)]. However, it is questionable whether the parameters of such models remain stable over time. Moreover this approach also raises a host of other problems regarding general equilibrium modelling.

An alternative strategy has been to analyse the possibility that exchange rates have deviated from their fundamentals determined levels due to the presence of rational bubbles in the foreign exchange market. Such deviations are deemed rational in the sense that a speculator realizes that the exchange rate is, say, overvalued relative to a given set of fundamentals but perceives that there is yet a non-zero probability of the gap widening further in the short-run. Thus, depending on the risk profile of the speculator, the expectation of further appreciation may be sufficient inducement to further speculative purchases and resale of the currency before the exchange rate collapses to (or even through) its fundamental equilibrium level. As the exchange rate strays ever further from fundamental equilibrium, the probability that it may collapse increases. Hence the expected appreciation of the currency would need to accelerate to "reward" the speculator for the risk attached to further purchases. This "spiking" phenomenon is a common observation in the foreign exchange markets, as in other stock markets. Testing for such bubbles in the foreign exchange market has been conducted by Shiller (1979), Huang (1981), Evans (1986), and Flood and Hodrick (1989) amongst others. Most of these tests have rejected the hypothesis of no-bubbles.

The idea that nonfundamentalist considerations may exert a significant short-run effect on the exchange rate also receives support from research into the investment advice given by chartists. For example, Allen and Taylor (1990) found that in the short-run (up to one week) some chartists' exchange rate forecasts outperformed alternative forecasts, including those based on a random walk model and various univariate and vector autoregressions. They also surveyed chief foreign exchange dealers in the London foreign exchange market and found that over 90% of respondents reported using charting techniques for time horizons of up to one week. The proportion fell as the time horizon was extended, however, with more emphasis then being placed on fundamentals. A problem with this approach is that not all chartists interpret the past pattern of exchange rate movements in the same way and, furthermore, the consensus opinion may shift over time. Thus although this strategy may be useful for explaining the past behaviour of exchange rates it may not be much help in consistently predicting future changes in the exchange rate.

The erratic behaviour of exchange rates in the short-run and the idea that fundamentals may only be significant in the longer-run has prompted a new avenue of research which some believe may prove to be the salvation of the monetary approach. MacDonald and Taylor (1992: 26-27), for example, sum up the position thus:

"Given this evidence, it is hardly surprising that empirical models based on pure, fundamental economic theory fail to provide an adequate explanation of short-term movements in exchange rates. However, the revelation that foreign exchange participants focus more on fundamentals at longer time horizons suggests that more attention might fruitfully be paid to modelling the fundamental determinants of long-term exchange rates. This is consistent with evidence in favor of the monetary model as a long-run equilibrium condition...".

"In addition, the development of econometric techniques that aid in the identification of long-run relationships using short-run data...is likely to provide a further impetus in this direction".

These and related issues concerning the third phase of tests of the monetary model are discussed in section 5.4 below.

5.4 The monetary model as a long-run equilibrium constraint

As with PPP (see chapter three) the failure of most tests to support the monetary model has been ascribed to deficiencies in the standard regression techniques used. Negative test results, in the light of more recent techniques, have been interpreted only as evidence against the monetary model as a static short-run relationship between the exchange rate and the relevant forcing variables. The basic argument is that the more powerful econometric testing procedures, specifically those relating to cointegration and unit root testing, allow the coefficients on these variables to be estimated without the restrictions implied by earlier empirical models. The proponents of this view claim there is now sufficient evidence to conclude that: "in contrast to the findings of other researchers...an unrestricted monetary model does provide a valid explanation of the long-run nominal exchange rate" (MacDonald and Taylor 1991: 184).

This claim should be put in context. As was the case with PPP (see chapter three), early cointegration tests of the monetary model using the two-step Engle and Granger (1987) methodology proved disappointing. For example, Meese (1986), Baillie and Selover (1987), Meese and Rogoff (1988), and Kearney and MacDonald (1990) could not find cointegration between the exchange rate, relative money suppplies, or relative price levels which cast doubt on the monetary model even as a long-run equilibrium relationship. However, this was later attributed to deficiencies in the two-step Engle and Granger procedure, in particular the emphasis on the Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) techniques which were used to test for both stationarity in the time-series data (step 1) and then, if the variables were found to be integrated of the same order, for cointegration between them (step 2).

The response to these initial negative results has been the adoption of yet more sophisticated testing procedures that overcome the deficiencies of the Engle and Granger approach. Specifically this involves the use of the maximum likelihood multivariate technique proposed by Johansen (1988) and Johansen

and Juselius (1992) outlined in chapter three[12]. It is the results from these tests that are proclaimed to support the monetary model as a long-run equilibrium constraint.

Using the Johansen technique, MacDonald and Taylor (1993) tested the standard present value equation of the forward-looking monetary approach to the exchange rate (FMAER) for monthly dollar/mark exchange rates, January 1976 - December 1990. Interestingly, their results rejected the hypothesized rational expectations restrictions but did support the standard flexible price monetary model of exchange rates as a long-run equilibrium condition. A bubble term was added to the FMAER equation but the results were also inconsistent with the speculative bubbles hypothesis. Furthermore, in contrast to the findings of Meese and Rogoff and others, MacDonald and Taylor were able to beat random walk forecasts of the exchange rate using out-of-sample forecasts generated by a dynamic error correction model of the data, for all the forecast horizons tested. These results complement similar findings in MacDonald and Taylor (1991).

Moosa (1994) tested a variant of the flexible price monetary model in which he distinguished between traded and non-traded goods, with PPP applying only to the former. His data set comprised monthly bilateral exchange rates for the UK, Germany, and Japan against the US dollar for the period January 1975-December 1986 to include episodes of both "success" and "failure" of empirical monetary models. He found two cointegrating vectors each for the UK and Germany and three for Japan. All the restrictions imposed by the monetary model for the UK and Japan were rejected (at the 5% level of significance). For Germany, however, the restrictions on the coefficients on the domestic and foreign money supplies could not be rejected (this was also the case in MacDonald and Taylor 1991 who could not reject any of the restrictions for Germany). The higher the number of cointegrating vectors the more stable the relationship, so the test results were most favourable for the yen/dollar but less so for the mark and the pound sterling.

In interpreting his findings Moosa points out two difficulties with the Johansen technique. Firstly, there is an identification problem in selecting that vector which "makes the most economic sense" as a long-run determinant of the exchange rate. Many of the estimated coefficients are either of the wrong sign or are too large compared to the predictions of the monetary model. Thus some of the significant cointegrating vectors do suggest a long-run equilibrium relationship between the exchange rate and the relevant explanatory variables but not of the hypothesized direction or size. Secondly, the Johansen technique only allows the same restrictions to be tested across all the cointegrating vectors simultaneously. It is not possible to test them selectively. Thus a cointegrating vector which does satisfy the restrictions imposed by the monetary model may exist but would not necessarily be identified as such by Johansen's multivariate approach. Moosa notes that the identification problem may reflect the operation of variables missing from the simple monetary equations that have been tested. This would entail the testing of more complex models such as the portfolio balance model. Moosa's own variation, as noted above, was to limit the assumption of PPP to traded goods only.

Wickens (1996) places further doubt on the usefulness of Johansen's cointegration technique. He shows that, in general, the cointegration vectors (CVs) derived from the estimation of unrestricted vector error correction models (VECM) cannot be given an economic interpretation without the addition of *a priori* information (such as the theoretical restrictions of a properly specified structural model). He also shows that the common stochastic trends (CSTs) estimated using vector autoregressive models (VARs), which are the source of the permanent shocks determining economic variables in the long-run, generally cannot be identified (or uniquely obtained from the CVs) either. This puts a big question mark over the economic meaning of the estimated CSTs (and the coefficients on the CVs) as distinct from the pure statistical significance of the estimates.

Wickens notes that these problems result from the basic approach of cointegration analysis in which a structural model is inferred from the estimated CVs rather than proceeding from a fully specified structural model to be estimated. Wickens shows where the cointegration technique breaks down by first specifying a complete structural model in terms of the integrated variables and then representing this as a VECM so that the connections between the structural model and the coefficients of the CVs and the CSTs are readily apparent. With regard to the cointegration technique and VAR analysis he concludes that:

"This is a return to a form of measurement without theory in which *a priori* considerations are normally confined to the selection of variables. The conclusion of this paper, that it is only by taking account of the *a priori* restrictions that an economic interpretation can be given to the results, therefore raises new doubts about the usefulness of standard VAR analysis" (Wickens 1996: 256).

The upshot of these and other studies is that although the short-run dynamics of the exchange rate may be very complex - and perhaps even inherently unpredictable as suggested by the "news" and efficient asset market approaches to exchange rate determination - the latest available evidence does appear to be consistent (with some exceptions) with the view that the exchange rate converges to long-run equilibrium as determined by fundamentals such as PPP, relative money supplies, income, and interest rate differentials. However, given the reservations by Wickens and others, there is insufficient evidence at this stage to support anything but renewed research in the area[13]. Taylor, himself one of the main protagonists of the approach, comments:

"The usefulness of the cointegration approach suggested by these studies should, moreover, be taken as at most tentative: their robustness across different data periods and exchange rates has yet to be demonstrated" (Taylor 1995: 29).

5.5 Methodological issues

As should be apparent from the analysis of empirical exchange rate models in this chapter, negative test results have not

been interpreted as falsifying the underlying monetary and asset market theory of exchange rates. But rather than condemn this as malpractice on the part of the researchers concerned, various explanations were adduced to account for the failure of the monetary model. A stout defender of falsificationism such as Blaug would probably dismiss these as undesirable *ad hoc* conventionalist strategems to protect a discredited theory. However a case is made in this chapter, building on the evidence of the previous chapters, that the peculiar subject matter of exchange rate economics (and presumably many other areas of economics) prohibits any serious test of the theory along falsificationist lines whether of a Popperian or Lakatosian persuasion. In the first place tests of structural exchange rate models, that is models based on a set of fundamentals, are nearly always tests of a joint hypothesis invloving expectations or risk premia (or both). Test results under these circumstances are thus unavoidably ambiguous: the failures of the model could *either* be due to specification and identification problems with the equation(s) being tested *or* it could be due to incorrect hypotheses about expectations and risk. Thus it does not appear possible, at this stage, to lay a firm enough foundation (the "precise, checkable initial conditions" referred to by Hutchinson) upon which to erect a genuinely falsifiable model of exchange rates. Misspecification and identification problems aside, only if satisfactory models of expectations and risk are developed would it be possible even in principle to falsify structural models of the exchange rate.

Given the unavoidable ambiguity surrounding the test results of empirical exchange rate models the researcher is forced to adopt alternative strategies as outlined in this chapter. Initially this will usually involve an interpretation of the test results and an explanation for the poor performance of the model (positive test results, being rarer, are normally not vigorously contested by the researcher and are gratefully proffered as confirmatory evidence in support of the theory being tested - witness the early tests of the monetary model). Following the interpretation and explanation phase there is typically a modification of the model, where possible,

including new auxiliary hypotheses designed to account for the failure of the earlier version. A fresh round of tests of the modified model then ensues which unearths further anomalies, and so on. Hence the essentially *exploratory* nature of empirical research in exchange rate economics (as in other areas of economics).

It should be noted that this *modus operandi* also does not conform to Lakatos' more sophisticated version of falsification since the successive theories or models do not usually predict novel facts, nor are they intended to. The most that is expected and can be hoped from them is that they will help 'patch up' deficiencies in the earlier models and provide a basis for further research into new anomalies that may arise. The emphasis is on explaining what has happened rather than the discovery of new facts which were not predicted by the older model. Moreover, a pattern emerging from the successive theories of exchange rate determination is that any initial empirical successes enjoyed by the new model are seldom corroborated by subsequent test results. Thus the normal research practice of economists in this area, in Lakatosian terms, often involves neither theoretically nor empirically progressive problem shifts.

The *ad hoc* nature of such economic research, frowned upon by falsificationists like Blaug, is nevertheless quite permissible in terms of Laudan's more pragmatic problem solving methodology. If *ad* hoc_{1-2} modifications to the basic model solve more empirical problems than a rival version and, importantly, if no additional conceptual problems are created by such an adjustment then the new model constitutes a scientific advance and 'progress'. What is important here is that the modifications should not violate the negative heuristics of the basic theoretical framework - in other words, they should not be *ad* hoc_3. Thus the requirements of independent testability (and the falsificationist demand of 'saving the phenomena') are not forced upon us in the Laudan model (although, of course, if a new theory does meet these reqirements as well then this would be regarded as an added bonus). It is for this reason that Laudan's approach appears to

fit the workaday research practice of economists more readily than alternative methodological prescriptions.

The research strategies adopted in the area of exchange rate economics also accord loosely with the contention by Boland and Morgan that although specific empirical models may be "falsified", the underlying theory or conceptual framework remains impervious to negative test results. There are literally thousands of different empirical models which may be constructed as permutations of the same theoretical approach[14]. Only if all these models were exhaustively tested for all possible time periods and data sets and found wanting could the underlying theory be said to be falsified. This train of thought also links up with current empirical research into PPP and the monetary model of exchange rates using the cointegration techniques proposed by Engle and Granger and Johansen. By letting the data "speak for themselves" some favourable evidence was found validating PPP and the monetary model as long-run equilibrium conditions. This is arguably due to the ability of cointegration techniques to allow the models to be tested in unrestricted form, that is without imposing specific constraints on the coefficients in the equations concerned. In other words the cointegration procedure "searches" the time series data for the best (significant) goodness-of-fit with the basic model and then estimates the associated coefficients accordingly. However, as made clear in section 5.4, although estimated parameters of the basic model may be found which are statistically significant it remains to be seen whether they can be given a meaningful economic interpretation.

The reluctance to accept any finality as regards test results may also be due to the perceived limitations of econometric techniques. The test results, positive or negative, may well be reversed with the development of new, more powerful techniques. This is clearly seen in the recent use of cointegration tests applied to PPP and the monetary model. Whereas the earlier standard regression approach mostly produced negative test results, the new approach has discovered evidence more favourable to these models as long-run equilibrium conditions.

A similar development has been observed regarding the efficiency hypothesis, in this case in the opposite direction. Greater sophistication in testing procedures has cast doubt on the (no risk premium) efficiency hypothesis, in contrast to the generally positive results interpreted from earlier tests.

Partly because of the difficulties in testing the various exchange rate models and hypotheses, the acceptance of a new theoretical approach (or "paradigm" or "research programme") seems to rely more heavily on a combination of *a priori* considerations and consistency with broad empirical regularities or stylized facts (as explained in section 5.1). As pointed out in chapter four, significant conceptual breakthroughs and the development of a new theoretical framework do not appear to be closely linked with the empirical evidence of specific econometric tests. The monetary model did not supersede the Mundell-Fleming model because the latter "failed" econometric tests. It was subjected in fact to very little serious testing, whereas the formal monetary model has been tested intensively for more than a quarter of a century since the early 1970s (and despite repeated empirical failures during this time). The decision to accept and to persist with the monetary model seems rather to have been based on (a) conceptual considerations concerning the nature of speculative capital flows, the demand for different currencies as the demand for a stock rather than a flow, and the ensuing characterization of the foreign exchange market as an efficient asset market including the crucial role played by expectations and (b) the consistency of the new approach with emerging stylized facts about the behaviour of exchange rates which could not easily be explained by the Mundell-Fleming model or other flow models of the economy. The reliance on stylized facts as the empirical base for the appraisal of theories in economics means that such theories should properly be regarded as historically circumscribed explanatory theories as opposed to genuine predictive theories open to falsification. It is also the case that since many stylized facts are a product of a specific time, place, and institutional arrangements that external historical factors (in contrast to internal

theoretical factors) play a much greater role in the development of economic theories than in the natural sciences.

Conceptual considerations were also shown to be important as regards the choice between the portfolio balance model *vis-a-vis* the monetary model. The reasons given for preferring the portfolio balance model have little to do with the results of econometric tests as it fares as poorly as the monetary model in this regard. This does not deter proponents of the portfolio approach as they believe it to be a superior conceptual model compared to the available alternatives. The model may thus only be rejected and superseded if researchers can be persuaded that a better conceptual framework exists. Such arguments are necessarily along intuitive *a priori* lines.

Again, as suggested in chapter four, Laudan's problem solving model appears to be a better explanation of these tendencies in economics. The grounds for preferring one theory to another need not be defined exclusively by empirical considerations. Laudan's model allows one to say that 'progress' has been achieved if important conceptual problems have been resolved even if a new theory does not have excess empirical content over its predecessor. It is entirely conceivable that the new theory may be preferred even if it solves *fewer* empirical problems than its rivals, depending on the number and importance of the conceptual problems it is capable of solving. The main difficulty with Laudan's approach, as pointed out in chapter two, is in deciding how to weight the different problem gains and losses from one theory to another.

Both the research into the efficiency of foreign exchange markets and tests of exchange rate models as long-run equilibrium conditions also have some noteworthy implications as regards the symmetry thesis in economics: the thesis that there is a methodological equivalence between explanation and prediction in the sciences. Although an increasing number of more sophisticated recent tests technically reject the speculative efficiency hypothesis, research in this area shows how difficult it is to distinguish changes in the exchange rate from a random walk. The evidence suggests that asset market

prices like exchange rates are inherently *unpredictable*, at least in the short-run. However, this does not mean that the historic behaviour of exchange rates cannot be *explained*. It has already been argued why some explanatory theories or models of exchange rates are preferable to others with regard to conceptual considerations and stylized facts about macroeconomic variables. More or less satisfactory explanations of the past behaviour of exchange rates may be developed even though such theories may be equally useless in predicting exchange rates out-of-sample for given data about the relevant forcing variables. In particular, the way in which expectations are determined in the future may be very different to the way they have been determined in the past. To believe otherwise is to commit the basic inductivist fallacy writ large.

But, it may be argued, doesn't recent research using more powerful econometric tests show that the exchange rate can be predicted on the basis of long-run PPP and basic monetary models? Take, for example, the conclusion by MacDonald and Taylor in comparing the forecasting performance of their monetary model against a random walk:

"The results of this exercise...are very interesting indeed: the dynamic error-correction model outperforms the random walk forecast at every forecast horizon. The results of this section thus suggest that, treated as a long-run equilibrium condition, the monetary model of the exchange rate may still be useful in forecasting the exchange rate" (MacDonald and Taylor 1993: 103).

The nature of such prediction or forecasting should, however, be qualified. It is clear that we are not dealing with anything approaching the laws and universal constants of the natural sciences. The out-of-sample forecasts referred to here are based on the coefficients of the cointegrating vectors discovered from the relevant time-series data. The coefficients are estimated from the data "by letting the data speak for themselves" without imposing any prior restrictions on the monetary model. The sign and the magnitude of these coefficients may not even make much economic sense in terms of the chosen structural model. As Moosa (1994: 285) points out, the magnitudes of such coefficients are often excessively high

or of the wrong sign compared to the predictions of the monetary model. Thus the cointegration procedures may capture statistically significant relationships but their economic significance is questionable. Such procedures can be helpful in explaining "what is there" and in extrapolating these relationships or patterns in the data. However, for different exchange rates and for different time periods the data may speak very differently and the old extrapolations break down. The parameters of the monetary model under this approach are like ciphers, dependent upon particular data sets to give them meaning. In comparison to, say, the classical laws of motion it would be as if the gravitational constant *g* had always to be re-estimated on the basis of repeated experiments at different times and places!

A further question arising from the results of cointegration studies and dynamic error-correction models of the exchange rate is whether, if the models become generally accepted as a forecasting technique by speculators, the currently better than random walk forecasts may break down in the future. This is an implication of the efficient markets hypothesis. Any mechanical procedure such as cointegration modelling which appears to offer superior forecasts would not be ignored by competitive markets. The procedures would be treated as a valuable new source of information which, if acted upon, would bring forward the long-run implications for the exchange rate thereby rapidly eroding the competitive advantage afforded by the new technique. This would preserve at least a weak form of the EMH - that the analysis of readily available historic data cannot indefinitely produce superior price forecasts. Thus the recent better than random walk forecasts reported by MacDonald and Taylor may well prove to be fleeting - just as the initially positive empirical 'tests' of the monetary model in the early 1970s could not be replicated later in the decade.

Notes

1 Since short-term changes in exchange rates are erratic it may seem to be a misuse of language to refer to their behaviour as empirical "regularities". However, this is the term used by some researchers in the field - such as Mussa (1979). The rest of this chapter sticks to the preferred term "stylized facts".

2 The jargon term for this type of behaviour is leptokurtosis. The leptokurtic nature of asset prices is well documented and may be problematic for some econometric testing procedures. For example, the Johansen cointegration technique (see section 5.4) assumes normal distribution Gaussian errors.

3 Another explanation for time varying volatility in foreign exchange markets is 'political turmoil' - specified political events are a determinant of the conditional variance of the currency concerned. Recent events in the South African foreign exchange market provide a good example of this. For about a year after the abolition of the financial rand the unified currency was relatively stable. Since the first quarter of 1996, however, the rand exchange rate has been characterized by episodes of sharp unidirectional depreciation. This has been ascribed to speculation driven by political considerations and the perceived inability of the SARB to prevent such speculation.

4 A more recent and perhaps better way to express this relationship is to say that the spot and forward exchange rates are cointegrated. Cointegration is present if a stationary linear (logarithmic) combination $s_{t+1} + af_t = e_{t+1}$ can be found, that is, where the residual term is not significantly different to zero. Many studies show that the spot and forward rates are cointegrated with cointegrating vector $a = -1$ (see Baillie and McMahon 1989).

5 A similar relationship exists in the spot market for cross-currency transactions, the well known triangular arbitrage condition. The exchange rates (and changes therein) between three or more currencies should be consistent. If, say, the dollar appreciates by 5% against the pound but only 2% against the franc, arbitrage should ensure that the pound depreciates by 3% against the franc. If not, a risk free round-trip profit could be made by simultaneously selling dollars for pounds, pounds for francs, and francs back to dollars.

6 An interesting issue here is whether empirical problems with the monetary model led to more intensive research of the statistical time series properties of exchange rates, as suggested here, or whether practitioners of a new type of econometrics were looking for new areas of application to exchange rate economics. Probably both factors were at work to some extent. Tests for unit roots and mean reversion in

exchange rates certainly provided a fruitful new avenue for applied econometricians to ply their trade. As noted earlier, however, the repeated empirical failure of structural exchange rate models also led to the feeling that continued econometric tests thereof would be a case of flogging a dead horse. But the horse was not quite dead and buried! The revival of the monetary model from the late 1980s with the advent of modern cointegration techniques bears witness to this. I am grateful to Mr G Farrell, a colleague in the Department of Economics at Rhodes University (presently at Birmingham University), for raising this point.

7 There are three different versions of EMH. The weak version asserts that historical asset price data cannot be used to improve forecasts of such prices and thus provide abnormal returns. This rules out forecasts based on chartism, for example. The semi-strong version asserts that, in addition, any relevant public information will already be reflected in asset prices and thus also cannot be used to generate market beating returns. This would rule out the use of public economic and market data or trends to improve exchange rate forecasts. The strong version suggests that all information, both public and private or "inside information", is similarly quickly reflected in prices thereby eliminating any abnormal profit opportunities.

8 The problems in constructing an adequate model of expectations are well documented. Frankel and Froot (1987), for example, suggest that expectations about exchange rates do not concur with either the rational expectations hypothesis or alternative hypotheses such as perfect foresight and regressive expectations (see also chapter four).

9 The orthogonality property of EMH is that no economic variables (such as current or lagged spot or forward exchange rates) can be used to predict the forecast errors using information available at the time such forecasts were made. If the error terms are not independent, as in Frankel (1979b), this implies that relevant information has been ignored which, in principle, could have reduced the forecast error.

10 Meese and Rogoff's test procedures and results have been replicated by Pilbeam (1991) for the dollar-pound exchange rate using a quarterly data set for the period 1/1979 to 3/1988. However, some critics of the Meese and Rogoff approach argue that their somewhat harsh conclusions result from their use of a one-step testing procedure. Schinasi and Swami (1989), for example, argue that the comparison of the structural model forecasts with those of the naive random walk model should be between multi-step rather than one-step forecasts.

11 In post-sample ($1984Q_1$ - $1885Q_4$) dynamic simulations, four of the six alternative equations (before and after adjustment) were found to beat a naive random walk model (with and without

a drift parameter). The best performance was from the adjusted lag equation with homogeneity imposed.

12 But see endnote 2 for possible deficiencies in the Johansen cointegration technique. Size of sample problems may also be important. Phillips, with various co-authors, has developed an alternative way of estimating cointegrating regressions in which the variables are adjusted to remove the large sample effects of endogeneity, residual serial correlation and heteroskedasticity (see, for example, Phillips 1991a and 1991b; Phillips and Loretan 1991; and Phillips 1993).

13 It is difficult to comment further on these issues without a more thorough examination of the latest thinking in econometrics (that is, methodology with a small 'm' as opposed to a capital 'M' which is the main concern here). This is beyond the scope of this thesis.

14 Despite the criticisms made here of Lakatosian MSRP it may yet be useful as a historical metaphor in describing the evolution and classification of theories in economics - but bearing in mind Cross' *ex ante/ex post* distinction regarding hard core presuppositions and processes. For example, the idea that *ex post* (but not *ex ante*) certain 'hard core' assumptions appear to have been guarded by a negative heuristic may be useful metaphorically, as may the idea that different empirical modifications to a basic theoretical model are motivated by a positive heuristic guiding research within the 'protective belt' (as is the case with other ideas and terms borrowed from the history and philosophy of science, such as Kuhn's 'paradigms' and 'revolutions' which have become an enduring part of the academic discourse of economics even though it may be argued that the literal methodological models themselves have been discredited).

CHAPTER SIX

CONCLUSIONS FOR THE METHODOLOGY OF ECONOMICS

This thesis has analyzed the main developments in theories of exchange rates and the balance of payments that have taken place since Cassel's formulation of purchasing power parity in 1916. The various attempts at integrating the balance of payments and the determination of exchange rates within a macroeconomic model were also examined. Throughout this thesis, these developments have been appraised with some basic methodological issues and questions in mind.

The central concern of this appraisal has been the relationship of empirical evidence to the development of theory and what this reveals about the methods typically adopted in economics. As outlined in chapter two, both Popper and Lakatos emphasized the role played by empirical evidence as the ultimate arbiter of good theories and what should count as scientific progress. Both argued that empirical considerations were crucial as regards the logic of justification of theoretical explanations. In particular they held that although the verification of theories was an irretrievably flawed methodology, it was yet possible to use empirical evidence to falsify them. For Popper a good theory is one that is open to falsification because it excludes many possible observations or empirical outcomes and has (thus far) failed to be disproven by rigorous empirical tests. For Lakatos, regardless of the evidence for or against a theory, it need not be falsified unless there is a better alternative theory. To qualify as "better", however, the empirical requirements for the new theory are as demanding as Popper's:

a) it must explain all the unrefuted content of the old theory
b) it must have excess content by predicting novel facts, that is, it must be independently testable

c) at least some of these predictions or testable implications must also be corroborated.

Only if these threefold empirical requirements are met can a new theory claim an unambiguous victory over an established rival and only in this sense can the older theory be regarded as falsified (and even then the predecessor theory may stage a comeback). A research programme characterized by (a) and (b) is regarded as theoretically progressive and may be tolerated as such for a time, but ultimately only if condition (c) is met as well is the programme also empirically progressive and worthy of continued pursuit. Sequences of related theories that do not meet all of these criteria are regarded as a degenerating research programme.

What is typical here is the almost exclusive emphasis on empirical criteria as opposed to purely analytical or conceptual considerations. Intuitive *a priori* considerations cannot provide independent grounds for accepting one theory and rejecting another. The ultimate test is empirical and the proof of the pudding is in the eating thereof. Although both Popper and Lakatos regard theoretical issues as a matter of concern (Popper's preference for 'deep' theories that can explain a lot from a little; Lakatos' injunction that a new theory or theoretical modification should not be *ad hoc*$_3$), their concerns do not form an integral part of their methodology - at least not in the sense that such theoretical considerations can actually override their stated empirical criteria.

The basic conclusion of this thesis, with reference to specific theories of exchange rates and models of the open economy, is that falsification as construed by either Popper or Lakatos is inappropriate in economics. As examined in this thesis, the insurmountable difficulties in applying falsificationism explains why economists have rarely taken it seriously despite their high regard for falsification as an ideal. A number of supporting or subsidiary claims are examined below.

6.1 Econometric tests and the appraisal of economic theories

In most areas of economics, particularly macroeconomics, the evidence from econometric tests seems to be largely ignored in the appraisal of economic theories. This has clearly been the case with theories of purchasing power parity and the monetary model of exchange rates. The vast amount of empirical research in this regard, as outlined in chapters three and five respectively, has had very little significance in either proving or disproving the theories concerned. If anything, the cumulative weight of evidence from econometric tests would suggest that both PPP and the monetary model should have been falsified. The empirical failure of these models of exchange rate determination has been openly acknowledged (see chapters three and five). Despite these recognized failures, ongoing empirical research and 'tests' of both PPP and the monetary model have persisted up to the present day.

Thus in actual practice it does not appear that the econometric 'tests' of these theories are really tests at all. A genuine test must set a standard against which a theory can be appraised and judged to either pass or fail, unless we are to engage in a semantic contortion of the meaning of the word 'test'. It is true that specific empirical models of exchange rates have been rejected (or, in some cases, accepted) but such rejections have not had serious consequences for the general theory or conceptual framework from which the models were derived. Of course, a committed falsificationist like Blaug would simply criticise this practice as an example of "innocuous falsification" or "playing tennis with the net down" and say that this is exactly what is wrong with economics - that we don't take falsification seriously enough. But taking falsification seriously in economics would, as argued in this thesis, be a case of the cure being far worse than the disease, at least on the basis of econometric test results.

Even if we disregard naive versions of falsification and the idea that we can produce crucial test evidence that will deliver a knockout blow to an economic theory (which even Blaug does not demand) the fact of the matter is that most of the

cumulative evidence from empirical tests of PPP and the monetary model has been negative. Blaug would thus surely want economists to concede that these theories have been refuted. It will also not help to fall back on the Lakatosian defence and say that the theories need not be falsified unless there is a better theory waiting in the wings. In this area of economics (and most others) there are a number of competing theories or models. The problem is that the Lakatosian version of falsification defines "better" in empirical terms, as explained above. Since none of the competing models of exchange rate determination surveyed in this thesis can lay claim, on the basis of econometric tests, to superior empirical proficiency there is presumably nothing to choose between them on this account. It seems clear that the choice between competing models has at least as much to do with *a priori* conceptual considerations than with their empirical record.

A further noteworthy feature of econometric tests in this area of economics is that *ex post* it has been possible to rationalize the poor empirical performance of PPP and the monetary model with the advent of new, more sophisticated econometric techniques. This is most clearly seen in the evolution of tests of the PPP relationship (although the same can be said for tests of the monetary model): from the early OLS regressions of PPP; to initial unit root tests on real exchange rates; to more sophisticated Dickey-Fuller (1979) unit root tests and other means of distinguishing the real exchange rate from a random walk; to the initial Engle and Granger (1987) cointegration tests of PPP; to the most recent tests based on the Johansen (1988) maximum likelihood cointegration technique. At each stage the primarily negative empirical results of previous tests were explained in terms of the latest innovations in econometric techniques. More subtle changes were also apparent at certain stages - for example, the reinterpretation of PPP as an equilibrium long-run relationship (rather than a short-run equation in which the direction of causation is specified, from relative price levels to the exchange rate) using unit root and cointegration tests. Thus it appears at times as if the econometric cart is put in front of the theoretical horse and suggests that econometric 'tests'

have been used more in exploring existing empirical anomalies rather than as genuine tests of the basic theory. Such tests are used to explore the range and application of the theory in particular instances and not to either "prove" or "disprove" the relevant theory. This practice has much in common with the methodology of deductivism of Mill and Robbins - which went out of fashion after the Second World War but in reality has never been totally abandoned by economists.

6.2 Novel facts *vesus* stylized facts in economics

Although the evidence from econometric tests may not be particularly relevant to the appraisal of economic theories, indirect evidence in the form of broad stylized facts about the economy might be significant in this regard. For example, the early monetary approach had policy implications which were the direct opposite of models derived from the Keynesian framework (such as the Mundell-Fleming model). This seems to fit the Lakatosian requirement that a rival theory yield testable facts that are either unexpected, expressly forbidden, or merely fortuitous in terms of the established theory. Also, the general asset market model which evolved during the 1970s was consistent with certain empirical observations that could not easily be explained by the Mundell-Fleming model. Specifically these stylized 'facts' might include (i) the observed devaluation policy failures of the 1960s (ii) the convergence of international inflation rates under the Bretton Woods fixed exchange rate system and (iii) the essentially random walk behaviour of changes in floating exchange rates since the early 1970s and some of the other statistical features of the behaviour of exchange rates described in chapter five.

However, the concept of a stylized fact is not entirely clear and has not been given a precise definition - economists have used this term in different ways. For example, as regards the Bretton Woods era economists such as McKinnon (1981) and Kenen (1985) used the term to describe the typically insular

economies that emerged after the Second World War, and the balance of payments and exchange rate policy experiences before and after the War (as in i and ii above). Other economists (such as Mussa 1979 and De Vries 1994) use the term in the sense of describing general statistical features of the empirical behaviour of floating exchange rates (as in iii above). Either of these definitions raises a number of questions from a methodological perspective, especially as regards falsification. In the first place, were the contrasting policy implications of the monetary approach *vis-a-vis* the then orthodox Meade-Mundell-Fleming models genuinely testable novel facts? A Keynesian macroeconomist could either maintain that the policy predictions of the monetary approach were based on spurious tacit assumptions about other parts of the economy (as, for example, does Currie 1976 as regards the government budget balance) or that the so-called policy failures harped on about by monetarists were simply due to the influence of exogenous variables which neither model, Keynesian nor monetary, could account for (meaning that the selected instances of policy failure were insufficient to count as a stylized fact).

Moreover, a diehard Keynesian could argue that the Mundell-Fleming model was quite capable of explaining statistical features such as the near random walk behaviour of changes in the exchange rate, by reflecting the influence of expectations as shifts in the relevant flow demand or supply curves. Of course, as argued by Mussa (1979), the problem with this defence is that the inclusion of expectations becomes a tacked on auxiliary hypothesis which undoes the theoretical unity or spirit of the model. In particular, if capital flows are perfectly elastic with respect to expectations then the exchange rate is fully determined by speculative capital flows. The theory of exchange rate determination then becomes purely a matter of explaining movements in the speculative supply of international capital and, therefore, how expectations about the exchange rate are determined. Thus instead of a model designed to explain the exchange rate in terms of income and expenditure flows, trade elasticities, and the sensitivity of capital flows to changes in interest rate differentials we have

an essentially monetary or asset market view of exchange rate determination. This is a classic example of Lakatos' injunction against *ad hoc*$_3$ amendments to a theory. The point though is that this criticism of the Mundell-Fleming model is non-empirical. It is not that the Mundell-Fleming model cannot, in principle, explain certain facts about exchange rates which the asset market models can, but rather that it can only do so by making such *ad hoc*$_3$ modifications to the theory.

If we take the definition of a stylized fact as a statistical empirical regularity this raises the question: when does such an empirical regularity first become noticed and why? In both the Popperian and Lakatosian versions of falsification, a new or modified theory must have excess empirical content which is independently testable from its predecessor - otherwise it may be criticised as an undesirable *ad hoc*$_1$ modification made to save the theory from falsification. Hence the emphasis is on the prediction of novel facts, unexpected or even expressly forbidden in terms of the older theory. If economics has to rely on stylized facts as the decisive element in the logic of justification (which seems to be inevitable given the even greater empirical difficulties with econometric tests) it is thus clear that most economic theories must necessarily develop in an *ad hoc*$_1$ manner, since a fact recognized as an empirical regularity cannot at the same time be an unexpected or novel fact. Accounting for unexplained stylized facts that appear to be inconsistent with existing theories is common in economics. By strict Popperian or Lakatosian versions of falsification this practice would be unacceptable. However, to insist on such criteria of appraisal would evidently rule out most theoretical advances in economics without providing a viable alternative. In short, there is an important difference between the type of facts required for Popperian/Lakatosian falsification to work and the type of facts which economics, in practice, has to rely on in the appraisal of economic theories: novel facts are forward looking predictions of events or observations while stylized facts are backward looking historical events or time series observations.

6.3 Theory choice and conceptual-analytical considerations

The analysis of macroeconomic theories of exchange rates and the balance of payments makes it apparent that theory choice had much more to do with *a priori* conceptual considerations than with the empirical proficiencies of the theories concerned. A good example of this was the quite dramatic conceptual shift in favour of the monetary approach over the Mundell-Fleming model which took place in the early 1970s. With the articulation of the monetary approach to the balance of payments, and subsequently the monetary and asset market models of exchange rates, the conceptual flaws and theoretical inconsistencies of the earlier Keynesian flow models became evident. Henceforth any macroeconomic model which failed to account for stock equilibrium processes in the relevant asset markets (money, bonds, and foreign exchange) was disqualified *a priori* as a serious contender in explaining the determination of exchange rates and/or the balance of payments within the context of the open economy. This was not to say that flow considerations were necessarily irrelevant in such models but that they had to be linked in a theoretically consistent way to the process of stock equilibrium in these asset markets. The justification for preferring and persevering with the monetary and other asset market models, as opposed to a modified Keynesian flow model, had at best a tenuous link to empirical test results or other evidence in the form of stylized facts.

As explained in chapter four, the Mundell-Fleming model was deemed to be analytically defective on two counts. Firstly, it could not explain why a change in net capital flows induced by a discrete change in interest rate differentials should be permanent rather than a temporary flow reflecting a once-off desired adjustment in international asset portfolios. Also, the full implications for the conduct of monetary policy if such capital flows were perfectly interest rate elastic were not seriously considered. Secondly, as noted in section 6.2 above, the model had great difficulty in accounting for the effects of speculative capital flows motivated by changes in expectations in a theoretically consistent fashion. These conceptual problems with the Mundell-Fleming model were, on the other

hand, a natural and integral part of the alternative asset market approach.

It was not as if the monetary approach and asset market models enjoyed any striking competitive empirical advantage. As documented in chapter five, besides some debatable early successes the monetary and other asset market models have been a dismal empirical failure. On the other hand, the older flow approaches to exchange rates and the balance of payments were not subjected to anything like the same degree of empirical scrutiny. The Meade-Mundell-Fleming synthesis, for example, represented the zenith of Keynesian open economy macroeconomics during the 1960s. It integrated current and capital account flows in a theoretically consistent model which provided policy guidelines *vis-a-vis* the questions of internal and external balance and devaluation that preoccupied the Bretton Woods era of quasi-fixed exchange rates. However, the Mundell-Fleming model lasted only a few years, largely untouched by direct empirical tests, before being superseded by the monetary approach.

Similar comments can be made about rival models within the general asset market approach. For example, the reasons why some economists express a preference for portfolio balance models rather than the simple monetary model have little to do with the results of econometric tests. Such models fare as badly as the monetary model with regard to econometric tests. At an empirical level, proponents of the portfolio approach usually point to the stylized fact that countries like Germany and Japan which experienced large and sustained current accounts in the past also had appreciating exchange rates against countries with a less robust trade performance (but note that other theories could account for the same 'fact' and some countries like Australia with trade deficits also had relatively strong currencies). The main reason why some economists prefer the portfolio approach is that they believe it to be a superior conceptual model compared to the alternatives. The model may thus only be rejected and superseded if such researchers can be persuaded that a better conceptual-analytical model exists. Such arguments are

necessarily along intuitive *a priori* lines - with stylized facts perhaps being offered as an afterthought in support of a preconceived theory.

6.4 Ceteris paribus, expectations, and the Duhem-Quine thesis

Besides the issues raised above it has been argued throughout this thesis that the problems raised by the *ceteris paribus* assumption and the Duhem-Quine thesis contribute significantly to making genuine falsificationism an unworkable methodology in economics. The standard response of falsificationists in this regard is that the differences between a social science like economics and, say, physics are simply a "matter of degree". Accordingly there is no substantive difference in scientific method between the two disciplines and thus that the essential unity of science is preserved.

In addition to what *are* significant substantive differences as explained in sections 6.1-6.3 it may also be argued, with regard to the *ceteris paribus* clause and the difficulties posed by the Duhem-Quine thesis, that the differences as a "matter of degree" are so great that they also tend to make falsification an unworkable and inappropriate methodology in economics. To use an analogy, the difference between a man who has had only two brandies to drink and a man who has had ten is also a matter of degree. However, the effect of this difference on the driving ability of the two men is such that we can reasonably expect many more collisions from men in the latter category than in the former.

The problems raised by the Duhem-Quine thesis are readily apparent in 'tests' of the law of one price and PPP. Superficially these theories appear to be good candidates for falsification - they can be stated simply as mathematical formulas with a limited number of key explanatory variables which resemble a physical law. However, this apparent similarity soon disappears once the theory is subjected to empirical tests. What appears to be a firm foundation upon which to erect potentially falsifying tests is soon shown to be

a treacherous quicksand. The empirical results of such tests generate a seemingly endless process of qualifications, interpretations, and amendments - anything but a clear rejection of the theory.

The main reason for this is that tests of PPP are, in the Duhem-Quine sense, tests of an interconnected web of hypotheses and we can never be certain which part of the complex has been falsified by a specific test result. For example the proposition that the common currency price of identical goods (or of the appropriate comparative price indices) must be equal is necessarily subject to a range of if-then qualifiers or auxiliary hypotheses such as: transport costs are zero; there are no tariffs or other barriers to trade; no currency or arbitrage risks; perfect information; no measurement errors *etcetera*. Thus a negative test result does not logically imply that the PPP theory has been falsified since the rejection of any one or a combination of these hypotheses could have accounted for the same result. Only if, for the specific test situation in question, the effect of each and every one of this nonexhaustive list of auxiliary hypotheses could be demonstrated and accounted for would it begin to make sense to infer that negative test results had falsified the theory.

At the level of the law of one price for identical internationally traded goods the theory of PPP is axiomatic - in the same way that the statement "the shortest distance between two points is a straight line" is an axiomatic self-evident truth. The response to claims that this proposition had been tested and a number of exceptions to it had been found would not imply that the statement had been falsified. Instead questions would be asked about possible interfering causes such as the physical conditions under which the tests were conducted, the reliability of the measuring instruments used, and even the motivations of the researcher.

Moving from the law of one price for individual goods to PPP theories involving collections of goods raises similar problems. Here it is necessary to distinguish between the key

explanatory variables which are endogenous to the theory and the interfering effects of exogenous variables. The operation of the *ceteris paribus* clause in economics suggests that PPP should not be regarded as a deterministic law but rather as a stochastic functional relationship, where the relative price levels are but two independent variables hypothesized to systematically influence the exchange rate. If the error terms from this estimated relationship are shown to be serially correlated and the deviations from the PPP relationship are large, then it implies that variables other than national price levels exert a systematic influence on the exchange rate. Only if these exogenous variables can be identified, their effects on the exchange rate specified and shown to have greater explanatory power than relative price levels is the inference that PPP has been falsified by negative test results valid. Indeed despite such a conclusion PPP may yet be retained as a *partial* explanation of the exchange rate rather than discarded altogether, as in the monetary model.

These problems are not unique to a social science like economics. The *ceteris paribus* clause is similarly attached to theories in the natural sciences and may thus also disallow the unambiguous falsification of such theories. However, the issue here is once again a judgement as to the degree or force with which the *ceteris paribus* clause applies in economics compared to, say, physics. In the natural world of the physical sciences it is far easier to isolate the relevant explanatory variables and to decide whether a negative experimental test result is due to an incorrect theory of the selected endogenous variables or whether some other set of exogenous variables is to blame. In economics the problems posed by 'weak endogeneity' usually do not permit such a decision to be made unequivocally.

One aspect of the Duhem-Quine thesis which may be unique in economics, and *not* merely a matter of degree compared to the physical sciences, is the role played by expectations in economic theories. Most macroeconomic theories, for example, include a hypothesis about expectations - either tacitly or explicitly - which is inescapable when financial markets are concerned. This is clearly seen in the monetary and other asset

market models of exchange rates. Such models are comprised of two distinct parts: the identification and specification of the endogenous fundamental variables; and a hypothesis about the determination of exchange rate expectations.

If expectations are treated exogenously (as in Keynes' *General Theory*) this makes serious attempts at falsification a specious exercise since any negative test result could always be explained away with reference to *ad hoc* changes in such expectations. Thus economists have tried various ways of endogenizing expectations. The best example of this is the rational expectations hypothesis (REH) where expectations are made dependent on the fundamental variables and structural parameters of the chosen model. REH may be thought of as a logical implication of the efficient markets hypothesis, which asserts that pertinent information is not systematically ignored in determining expectations. With REH this includes the information available from the chosen structural model. In the monetary model of exchange rate determination, for example, the current spot exchange rate depends on the future expected path of relative money supplies and differences in real incomes.

From the above description of the way in which expectations are included in economic theories, as in the monetary model of exchange rates, it is evident that any test of such a model is necessarily a test of a joint hypothesis. In keeping with the Duhem-Quine thesis, it can thus usually not be reliably ascertained whether a negative test result falsifies the structural model itself or the embedded hypothesis about the determination of expectations. Without a satisfactory model of expectations, which at this stage appears to be unobtainable, neither can tests of the relevant theory produce any conclusive results as regards falsification.

Nor does REH release us from this basic dilemma. The hypothesis is useful because it allows the model builder to endogenize expectations in a theoretically consistent way and it avoids the implication of alternative hypotheses (such as adaptive and regressive expectations) that economic agents systematically ignore possibly relevant information and do not learn to

anticipate such information. However, REH assumes that all economic agents base their expectations on the same model as that of the model builder, which is held to be the "true" structural model of the economy. To the extent that economic agents base their expectations on an alternative structural model or to the extent that different groups consider different models (and thus have divergent expectations) REH is fundamentally flawed.

6.5 Asset markets, prediction, and explanation in economics

The evidence surveyed in this thesis has shown how difficult it is to predict the path of a significant macroeconomic variable like the exchange rate. This is also true of other variables determined in financial markets like bonds and shares. The empirical work on PPP and the monetary model shows how difficult it is to distinguish the real exchange rate and out-of-sample exchange rate forecasts from a random walk. Even predictions of more stable macroeconomic variables such as real GDP growth and national inflation rates are not markedly superior in this regard (and at a microeconomic level similar conclusions could be drawn from various commodity markets such as gold, silver, and oil and for sectors of the economy like the property market). This raises the question whether genuine falsification is a workable or appropriate methodology for economic research under these circumstances.

Drawing on the distinction between in-sample and out-of-sample tests of economic models, as in Meese and Rogoff's tests of some popular models of exchange rate determination, a useful distinction can perhaps also be made between *ex post* and *ex ante* versions of falsification. *Ex post*, structural economic models are tested by seeing how well they *post*dict historic in-sample time series data. Such models explain the past behaviour of the dependent variables (such as exchange rates in the monetary model) to the extent that they are a good fit to the data. *Ex ante* tests of the out-of-sample data *pre*dictions of the model, however, are much stronger and thus more in keeping

with the demand for "serious" attempts at falsification in economics.

Out-of-sample predictions of exchange rates often fail to outperform naive random walk forecasts (in which the next period forecast of the exchange rate is the same as the current period exchange rate). The reasons for the poor predictive performance of, for example, the monetary and portfolio balance models is partly attributable to unstable structural parameters of the model which vary over time. Therefore although a structural model may be found which fits the relevant time series data *ex post* it may not continue to fit the data *ex ante* for very long. Thus whatever problems falsification may have in *ex post* tests - and there are many, as suggested above - the difficulties are even greater *ex ante*. To this extent it may help to think of economics as an explanatory postdictive science rather than a genuine predictive science, whereby falsification is more appropriate to the latter than the former.

The near random walk behaviour of exchange rates in the short-run is not surprising given that the foreign exchange market is a speculative asset market *par excellence*. Like any stock market, pertinent new information is quickly reflected in the prices of the assets concerned. In the foreign exchange market stock equilibrium, in the sense of currencies being willingly held, is maintained by rapid adjustment of exchange rates in response to the release of such news. Since newsworthy items are unexpected, by definition, changes in exchange rates appear to be largely unpredictable.

As shown in chapter five, tests of the efficiency of the foreign exchange market have produced mixed results. It is not clear whether economic agents are inefficient in systematically ignoring (or misinterpreting) relevant information, or whether mispricing of exchange rates reflects a risk premium of some kind. Moreover, many tests of foreign exchange market efficiency are ambiguous because, as in tests of structural exchange rate models, a joint hypothesis including an equilibrium model of expectations is tested. So, for example,

changes in the exchange rate may be serially correlated with respect to a model in which expected equilibrium returns are constant - suggesting that profit maximizing speculators have neglected significant information relevant to the determination of the exchange rate. In this case the efficiency hypothesis is rejected. However, changes in the exchange rate may not be serially correlated compared to a model in which the expected equilibrium exchange rate varies. In this case the efficiency hypothesis cannot be rejected. In keeping with the Duhem-Quine thesis it thus cannot be decided whether a negative test result falsifies the efficiency hypothesis itself or the auxiliary expectations hypothesis.

Although most tests suggest that PPP and the monetary model of exchange rates does not hold in the short-run, some recent tests using cointegration techniques purport to find empirical support for these relationships as long-run equilibrium constraints. However, the evidence is mixed and it is too early to judge whether these more positive results are more than a statistical fluke. As with the monetary model, initial test results appeared to be favourable only to be followed by a run of mostly negative results. More importantly the meaning of the estimated cointegration coefficients has been questioned since, although they may be statistically significant, in many cases they do not have a sensible economic interpretation.

6.6 Laudan's problem solving approach and the methodology of economics

Laudan's problem solving model appears to provide a much better methodology for the appraisal of economic theories than either the Popperian or Lakatosian versions of empiricism. The basic reason for this is that Laudan defines scientific progress simply as an increase in the number of problems solved by a new theory compared to a rival. Instead of defining progress in terms of rationality, Laudan reverses this reasoning and argues that scientists are following rational research strategies just when they make progress in devising theories which solve more problems. His approach is thus a pragmatic one which is not

preoccupied with whether theoretical explanations are either "true" or "false" - although it should be mentioned that his approach is not necessarily inconsistent with this aim since theories that solve more problems are presumably closer approximations to the truth than those that solve fewer such problems.

In some respects Laudan's model is not a radical departure from Lakatosian MSRP. For example, Laudan's idea of a 'research tradition' is similar to Lakatos' idea of a 'research programme' where a positive heuristic guides modifications to new theories. The difference is that Laudan's model does not include an inviolable hard core. Any so-called hard core assumptions can be challenged and, if a new or modified theory successfully solves important problems by doing so, scientific progress may be achieved. Also in similar vein to Lakatos, unsolved empirical problems do not threaten an established theory as much as anomalous empirical problems which are successfully resolved by a competing theory but which remain unsolved by existing theories. Laudan's approach also regards modifications to a theory as problematic conceptually if they do not conform with the spirit or positive heuristic of the research tradition of which they are a part. This is similar to Lakatos' criticism of *ad hoc*$_3$ changes to a theory.

But there are significant differences between Laudan and Lakatos and these differences are important for the methodology of economics. For falsificationists like Popper and Lakatos the ultimate test of a theory is empirical. Theoretically progressive problem shifts must eventually prove their worth in the corroboration of new facts. Laudan, however, places at least as much emphasis on solving conceptual problems as he does anomalous empirical problems (see chapter two). The attraction of this for economics is that scientific progress may be achieved by solving significant conceptual problems without any increase in the scope of its solved empirical problems being necessary. Such progress is possible even if fewer empirical problems are solved, depending on the number and significance of the conceptual problems solved by the new theory. Laudan does not demand that a new or modified theory

explain all the unrefuted empirical content of an established theory *or* require that it predict novel facts. It is quite legitimate for theory modifications to be *ad hoc*, in the sense of not being independently testable, if such adjustments help solve more empirical problems.

A priori conceptual considerations were shown to be significant in the evolution of theories about exchange rates and the open macroeconomy. For example, the succession of more general theories from narrower partial explanations of the balance of payments and exchange rates (such as the elasticities, PPP, absorption, and Mundell-Fleming models) cannot be explained easily on the basis of falsifying test results or the corroboration of novel facts. The progress was almost purely theoretical in that the successive theories provided a superior conceptual and analytical framework to guide further research. None of these theories were superior at an empirical level on the basis of any test results. Indeed to some extent they became less open to falsification because, being more complex, there is more room to explain away potentially refuting test evidence using conventionalist immunizing stratagems. Until the advent of the monetary and asset market models from the early 1970s economists felt very little need to actually subject these theories to intensive empirical tests (although some attempt was made to test competing *trade* theories - inconvenient results were mostly ignored or were too ambiguous to draw any firm conclusions). The intention was to provide a theoretical framework for macroeconomic policy analysis, not to empirically test whether the theories were either "true" or "false".

Conceptual considerations were also at the fore of the shift from the Mundell-Fleming model to the monetary model in the early 1970s. At least three important conceptual problems were solved by this move: the problem of treating changes in net capital flows in response to changes in interest differentials as permanent rather than temporary flows reflecting international asset portfolio adjustments; the different effects on the national balance sheet between capital and trade flows; the difficulty in making expectations induced changes in

speculative capital flows an integral part of the theory, as opposed to accounting for them in an *ad hoc*3 manner. These conceptual advances are crucial in explaining the revealed preference for the monetary approach to the Keynesian flow models of the open economy. Whether the monetary or asset market models are more successful empirically is debatable. As explained earlier these models are generally regarded as empirical failures with regard to econometric tests. Perhaps a case can be made that they were consitent with certain stylized facts which the Keynesian models were not but this is uncertain as there does not appear to be a clear meaning or definition of a stylized fact in economics.

Laudan's problem solving model allows us to make sense of the conceptual developments above in a way that falsificationism does not. With Laudan's approach we can say that scientific progress had been achieved because significant conceptual problems had been resolved and thus that the research strategies used by economists in this field have been perfectly rational. Falsificationist methodologies would be unable to do so because they emphasize the prediction of empirical facts at the expense of purely conceptual or analytical advances.

As noted above, Laudan's approach is also much softer on *ad hoc* modifications to a theory that are designed to explain negative test results - as long as such modifications or amendments are not *ad hoc*$_3$ and create more conceptual or empirical problems than they solve. In Laudan's view such research strategies foster progress by helping to solve empirical problems as they arise. This also accords more closely with general research practice in economics. Usually, once a theory has been expressed as an empirical model amenable to econometric tests, there are many nontheoretical reasons why it may fail such tests. This is clearly evident in the many tests of PPP and the monetary model surveyed in this thesis and helps explain why economists are observed to be so reluctant to accept *prima facie* negative test results as necessarily falsifying the underlying theory. Economists should not feel that they are pursuing a theory irrationally by adopting certain *ad hoc* 'immunizing strategems' under these conditions and, according

to Laudan's approach, they don't have to. In a recent book where he revives Mill's idea of economics as an 'inexact and separate science' and argues for a sympathetic reconsideration of *a priori* methods and deductivism in economics, Hausman comments on the research practices of economists as follows:

"Their response to anomalous market data, which mimics the inexact method *a priori*, is not illegitimately dogmatic. It is, on the contrary, fully consistent with standard views of theory assessment, once one takes into account of how bad these data are. The problem is not a moral failing among economists - their inability to live up to their Popperian convictions - but a reflection of how hard it is to learn about complex phenomena if one does not know a great deal already and cannot do controlled experiments" (Hausman 1992: 226).

As one might expect, no methodological approach is without its own peculiar complications. Laudan's approach is no exception. The main problem in applying Laudan's problem solving model is in weighting the different problem gains and losses that may occur from one theory to another especially if, for example, a conceptual advance is accompanied by an empirical setback or *vice versa*. This may be an acute problem in economics where progress is mainly conceptual and analytical rather than empirical. The weighting of different conceptual gains or losses is a more subjective task and a matter of interpretation compared to empirical problem solving. However, if progress in economics is to be regarded as problem solving, and as suggested above there is much to recommend this approach, then this problem will in any event have to be addressed one way or another.

Perhaps a final word can be said about the position taken by Blaug and other proponents of falsification who feel that it is not taken seriously enough and should be applied more strictly in economics. Their fears about economics failing to live up to rigorous and honest scientific standards are, however, unfounded. Economics will not be any the poorer for abandoning an inappropriate and unworkable methodology such as falsification - indeed we may well be better off by not having to look over our shoulder every time we set about fine tuning our theories or inventing new ones. It should be noted that the call to drop falsificationism in economics does not mean that

all forms of empiricism should also be abandoned. It is only that the type of empirical methods used will be far more pragmatic, looser, *ad hoc*, and inductive than Blaug and others sympathetic to falsificationism might want - but given the subject matter of economics, as outlined in previous chapters, this is really the best we can hope for. Perhaps it is indeed time to stop paying lip service to falsificationist precepts - but to create a new self-image for economists instead.

REFERENCES

ABUAF N AND JORION P
1990. Purchasing power parity in the long-run. *Journal of Finance* 45: 157-174.

ADLER M AND LEHMAN B
1983. Deviations from purchasing power parity in the long run. *Journal of Finance* 38 (4): 1471-1481.

AGASSI J
1979. The legacy of Lakatos. *Philosophy of the Social Sciences* 9: 316-326.

ALEXANDER SS
1952. Effects of devaluation on a trade balance. *IMF Staff Papers* 2: 263-278.

ALIBER RZ
1973. The interest parity theorem: a reinterpretation. *Journal of Political Economy* 81 (6): 1451-1459.

ALLEN HL AND TAYLOR MP
1990. Charts, noise and fundamentals in the foreign exchange market. *Economic Journal* 100 (Supplement): 49-59.

ARCHIBALD GC
1979. Method and appraisal in economics. *Philosophy of the Social Sciences* 9: 304-15.

BACKHOUSE RE
1991. The neo-Walrasian research program in macroeconomics. In de Marchi N and Blaug M (eds): *Appraising economic theories: studies in the methodology of scientific research programs*, 403-427, Hants: Edward Elgar Publishing Ltd.

BAILLIE RT AND McMAHON PC
1989. *The foreign exchange market: theory and econometric evidence*. Cambridge: Cambridge University Press.

BAILLIE RT AND SELOVER DD
1987. Cointegration and models of exchange rate determination. *International Journal of Forecasting* 3: 43-52.

BALASSA B
1964. The purchasing power parity doctrine: a reappraisal. *Journal of Political Economy* 72: 584-596.

BAUMOL WJ
1957. Speculation, profitability and stability. *Review of Economics and Statistics* 39: 263-271.

BENJAMIN DK AND KOCHIN LA
1979. Unemployment in interwar Britain. *Journal of Political Economy* 87 (June): 441-478.

BILSON JFO
1978. Rational expectations and the exchange rate. In Frenkel JA and Johnson HG (eds): *The Economics of Exchange Rates*. Reading: Addison-Wesley Press.

1979a. On the efficiency of markets for foreign exchange: comment. In Dornbusch R and Frenkel JA (eds): *International Economic Policy: Theory and Evidence*. Baltimore: John Hopkins University Press.

1979b. The deutsche mark/dollar rate - a monetary analysis. In Brunner K and Meltzer AH (eds): *Policies for Employment, Prices and Exchange Rates*. Amsterdam: North-Holland Publishing Company.

1981. The 'speculative efficiency' hypothesis. *Journal of Business* 54 (July): 435-451.

BLAUG M
1976. Kuhn versus Lakatos or paradigms versus research programs in the history of economics. In Latsis SJ (ed): *Method and appraisal in economics*, 149-180, New York: Cambridge University Press.

1980. *The methodology of economics (or how economists explain)*. New York: Cambridge University Press.

1985. *Economic theory in retrospect*. 4th ed. New York: Cambridge University Press.

1988. John Hicks and the methodology of economics. In de Marchi N (ed): *The Popperian legacy in economics*, 183-195, Cambridge: Cambridge University Press.

BOLAND LA
1982. *The foundations of economic method*. London: George Allen & Unwin.

1988. *The methodology of economic model building*. Boston: George Allen & Unwin

BOOTHE P AND LONGWORTH D
1986. Foreign exchange market efficiency tests: implications of recent empirical findings. *Journal of International Money and Finance* 5: 135-152.

BOOTHE PM AND POLOZ SS
1988. Unstable money demand and the monetary model of the exchange rate. *Canadian Journal of Economics* 21 (4): 785-798.

BOUGHTON JM
1988. Exchange rates and the term structure of interest rates. *IMF Staff Papers* 35 (March): 36-62.

BRANSON WH
1968. *Financial capital flows in the US balance of payments.* Amsterdam: North Holland.

BRANSON WH, HALTTUNEN H, AND MASSON P
1977. Exchange rates in the short run: the dollar-deutschemark rate. *European Economic Review* 10 (December): 303-324.

1979. Exchange rates in the short run: some further results. *European Economic Review* 12 (October): 395-402.

BRYANT M AND CECCHETTI S
1993. The consumer price index as a measure of inflation. *Economic Review of the Federal Reserve Bank of Cleveland* Q4: 15-24.

CAGAN P
1956. The monetary dynamics of hyperinflation. In Friedman M (ed) 1956: *Studies in the quantity theory of money*, 25-117, Chicago: University of Chicago Press.

CALDWELL BJ
1982. *Beyond positivism: economic methodology in the twentieth century.* London: George Allen & Unwin.

CASSEL G
1916. The present situation of the foreign exchanges. *The Economic Journal* 26 (101): 62-65 and (103): 319-323.

1918. Abnormal deviations in international exchanges. *The Economic Journal* 28 (112): 413-415.

1921. *The world's monetary problems.* London: Constable & Co.

1922. *Money and foreign exchange after 1914.* London: Constable & Co.

1928. *Post-war monetary stabilization*. New York: Columbia University Press.

1936. *The downfall of the gold standard*. Oxford: Oxford University Press.

CHEUNG YW AND LAI KS
1993. Long-run purchasing power parity during the recent float. *Journal of International Economics* 34 (2): 181-192.

1994. Mean reversion in real exchange rates. *Economics Letters* 46 (3): 251-256.

CHOWDHURY A AND SDOGATI F
1993. Purchasing power parity in the major EMS countries: the role of price and exchange rate adjustment. *Journal of Macroeconomics* 15: 25-45.

COPI I
1972. *Introduction to logic*. New York: Macmillan Company.

CROSS R
1982. The Duhem-Quine thesis, Lakatos and the appraisal of theories in macroeconomics. *The Economic Journal* 92: 320-40.

CUMBY RE AND OBSTFELD M
1981. A note on exchange rate expectations and nominal interest differentials: a test of the Fisher hypothesis. *Journal of Finance* 36 (June): 697-703.

1984. International interest rate and price level linkages under flexible exchange rates: a review of recent evidence. In Bilson JFO and Marston RC (eds): *Exchange Rate Theory and Practice*. Chicago: University of Chicago Press.

CURRIE DA
1976. Some criticisms of the monetary analysis of balance of payments correction. *The Economic Journal* 86: 508-522.

CURTIS JM
1971. Direct foreign influences on Canadian prices and wages. In Caves RE and Reuber GL (eds): *Capital transfers and economic policy: Canada 1951-1962*, 168-197, Cambridge, Mass: Harvard University Press

DARBY M
1980. The monetary approach to the balance of payments: two specious assumptions. *Economic Inquiry* 18 (2): 321-326.

1983. Movements in purchasing power parity: the short and long-runs. In Darby M and Lothian J (eds): *The international transmission of inflation*, 462-477, Chicago: University of Chicago Press.

DEAN JW
1988. Does money supply endogeneity matter? *The South African Journal of Economics* 56 (1): 39-45.

DEARDORFF AV
1984. Testing trade theories and predicting trade flows. In Jones RW and Kenen PB (eds): *Handbook of international economics (vol. 1)*, 467-517, Amsterdam: Elsevier Science B.V.

DE MARCHI N
1976. Anomaly and the development of economics. In Latsis SJ (ed): *Method and appraisal in economics*, 109-128, New York: Cambridge University Press.

1991. Introduction: rethinking Lakatos. In de Marchi and Blaug M (eds): *Appraising economic theories: studies in the methodology of scientific research programs*, 1-28, Hants: Edward Elgar Publishing Ltd.

DE MARCHI N (ED)
1988. *The Popperian legacy in economics*. Cambridge: Cambridge University Press.

DE VRIES CG
1994. Stylized facts of nominal exchange rate returns. In van der Ploeg F (ed): *The Handbook of International Macroeconomics*, 349- 389, Oxford: Basil Blackwell.

DICKEY DA AND FULLER WA
1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 14 (June): 427-431.

DIEBOLD F, HUSTED S, AND RUSH M
1991. Real exchange rates under the gold standard. *Journal of Political Economy* 99: 1252-1271.

DOOLEY MP AND ISARD P
1980. Capital controls, political risk, and deviations from interest rate parity. *Journal of Political Economy* 88: 370-384.

1982. A portfolio balance rational expectations model of the dollar-mark exchange rate. *Journal of International Economics* 12 (May): 257-276.

DORNBUSCH R

1975. Exchange rates and fiscal policy in a popular model of international trade. *American Economic Review* 65: 859-871.

1976a. The theory of flexible exchange rate regimes and macroeconomic policy. *Scandinavian Journal of Economics* 78 (2): 255-275.

1976b. Expectations and exchange rate dynamics. *Journal of Political Economy* 84: 1161-1176.

1980. Monetary policy under exchange rate flexibility. In Bigman D and Taya T (eds): *The functioning of floating exchange rates: theory, evidence, and policy implications*, 3-31, Cambridge, Mass: Ballinger Publishing Co..

1988. *Exchange rates and inflation*. Cambridge, Mass: MIT Press.

DUHEM P

1914,translated 1954. *The aim and structure of physical theory*. Princeton: Princeton University Press.

DUNN RM

1970. Flexible exchange rates and oligopoly pricing: a study of Canadian markets. *Journal of Political Economy* 78: 140-151.

1972. Flexible exchange rates and traded goods prices: the role of oligopoly pricing in the Canadian experience. In Johnson HG and Swoboda AK (eds): *The economics of common currencies*, 8-113, London: George Allen & Unwin Ltd.

ENDERS W

1988. ARIMA and cointegration tests of purchasing power parity under fixed and flexible exchange rates. *The Review of Economics and Statistics* 70: 505-508.

ENGEL C AND ROGERS J

1995. How wide is the border? National Bureau of Economic Research: Working Paper 4829.

ENGLE RF AND GRANGER CW

1987. Cointegration and error correction: representation, estimation and testing. *Econometrica* 55: 251-276.

ENGLE RF AND YOO BS

1987. Forecasting and testing in cointegrated systems. *Journal of Econometrics* 35: 143-159.

EVANS GW
1986. A test for speculative bubbles in the sterling-dollar exchange rate: 1981-84. *American Economic Review* 76 (September): 621-636.

FEYERABEND P
1962. Explanation, reduction and empiricism. In Fiegl-Maxwell (eds) *Minnesota Studies in the Philosophy of Science* 3: 28-97.

1970. Consolations for the specialist. In Lakatos I and Musgrave A (eds): *Criticism and the growth of knowledge*, 197-230, New York: Cambridge University Press.

FLEMING JM
1962. Domestic financial policies under fixed and floating exchange rates. *IMF Staff Papers* 9: 369-379.

FLOOD RP AND HODRICK RJ
1989. Testable implications of indeterminacies in models with rational expectations. *NBER Working Paper (2903)*:Cambridge,Mass.

FRANKEL JA
1979a. On the mark: a theory of floating exchange rates based on real interest differentials. *American Economic Review* 69 (September): 610-622.

1979b. Tests of rational expectations in the forward exchange market. *Southern Economic Journal*: 1083-1101.

1981. On the mark: reply. *American Economic Review* 71: 1075-82.

1982. The mystery of the multiplying marks: a modification of the monetary model. *Review of Economics and Statistics* 64 (August): 515-519.

1983. Monetary and portfolio balance models of exchange rate determination. In Bhandari JS and Putnam BH (eds): *Economic interdependence and flexible exchange rates*. Cambridge, Mass: MIT Press.

1984. Tests of monetary and portfolio balance models of exchange rate determination. In Bilson JFO and Marston RC (eds): *Exchange rate theory and practice*. Chicago: Chicago University Press.

1986a. International capital mobility and crowding-out in the US economy: imperfect integration of financial markets or goods markets? In Hafer RW (ed): *How open is the US economy?* ,33-67, Lexington: Lexington Books.

1986b. The implications of mean variance optimization for four questions in international macroeconomics. *Journal of International Money and Finance* 5 (supplement): 53-75.

1990. Zen and the art of modern macroeconomics: a commentary. In Haraf WS and Willet TD (eds): *Monetary policy for a volatile global economy*, 117-123, Washington, DC: American Enterprise Institute for Public Policy Research.

FRANKEL JA AND FROOT KA

1987. Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review* 77: 133-153.

FRANKEL JA AND ROSE AK

1995. Empirical research on nominal exchange rates. In Grossman G and Rogoff K (eds): *Handbook of international economics, vol.3*, 1689-1723, Amsterdam: Elsevier Science B.V.

FRENKEL JA

1976. A monetary approach to the exchange rate: doctrinal aspects and empirical evidence. *Scandinavian Journal of Economics* 78 (2): 200-224.

1977. The forward exchange rate, expectations, and the demand for money. *American Economic Review* 67 (4): 651-663.

1978a. Purchasing power parity: doctrinal perspective and evidence from the 1920s. *Journal of International Economics* 8: 169-191.

1978b. A monetary approach to the exchange rate: doctrinal aspects and empirical evidence. In Frenkel JA and Johnson HG (eds): *The Economics of Exchange Rates*. Reading: Addison-Wesley Press.

1978c. Quantifying international capital mobility in the 1980s. In Bernheim D and Shoven J (eds): *National saving and economic performance*, 227-260, Chicago: University of Chicago Press.

1981a. Flexible exchange rates, prices and the role of "news": lessons from the 1970s. *Journal of Political Economy* 89 (4): 665-705.

1981b. The collapse of purchasing power parities during the 1970s. *European Economic Review* 16: 145-165.

FRENKEL JA AND LEVICH RM

1975. Covered interest arbitrage: unexploited profits? *Journal of Political Economy* 83 (April): 325-328.

1977. Transactions costs and interest arbitrage: tranquil versus turbulent periods. *Journal of Political Economy* 85 (December): 1209-1226.

FRENKEL JA AND JOHNSON HG
1976. *The monetary approach to the balance of payments*. London: George Allen & Unwin Ltd.

1978. *The economics of exchange rates*. Reading, Mass: Addison-Wesley Publishing Company.

FRIEDMAN M
1953. The case for flexible exchange rates, in *Essays in positive economics*. Chicago: University of Chicago Press.

1956. The quantity theory of money - a restatement. In Friedman (ed): *Studies in the quantity theory of money*, 3-21, Chicago: University of Chicago Press.

1968. The role of monetary policy. *American Economic Review* 58 (March): 1-17.

FROOT KA AND ROGOFF K
1995. Perspectives on PPP and long-run real exchange rates. In Grossman G and Rogoff K (eds): *Handbook of international economics, vol.3*, 1647-1688, Amsterdam: Elsevier Science B.V.

FULTON G
1984. Research programs in economics. *History of Political Economy* 16 (2): 187-206.

GENBERG AH
1975. *World inflation and the small open economy*. Stockholm: Swedish Industrial Publications.

1976. Aspects of the monetary approach to balance of payments theory: an empirical study of Sweden. In Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 298-325, London: George Allen & Unwin Ltd.

GEWEKE J, MEESE R, AND DENT W
1982. Comparing alternative tests of causality in temporal systems. *Journal of Econometrics* 77: 161-194.

GIOVANNINI A
1988. Exchange rates and traded goods prices. *Journal of International Economics* 24 (1): 45-68.

GRANGER CWJ
1969. Investigating causal relations by econometric models and cross-spectral models. *Econometrica* 37: 424-438.

GREGORY TE
1930. *Foreign exchange before, during, and after the war.* London: Oxford University Press.

GYLFASON T AND HELLIWELL JF
1983. A synthesis of Keynesian, monetary, and portfolio approaches to flexible exchange rates. *The Economic Journal* 93: 820-831.

HABERLER G
1945. The choice of exchange rates after the war. *American Economic Review* 35 (3): 308-318.

HAKKIO CS
1984. A re-examination of purchasing power parity. *Journal of International Economics* 17: 265-277.

HALLWOOD P AND MACDONALD R
1986. *International money: theory, evidence, and institutions.* Oxford: Basil Blackwell Ltd.

HANDS DW
1988. Ad hocness in economics and the Popperian tradition. In de Marchi N (ed) 1988: *The Popperian legacy in economics*, 121-137, New York: Cambridge University Press.

HANSEN LP AND HODRICK RJ
1980. Forward exchange rates as optimal predictors of future spot rates: an econometric analysis. *Journal of Political Economy* 88: 829-853.

HAUSMAN D
1992. *The inexact and separate science of economics*. Cambridge: Cambridge University Press.

HAWTREY RG
1919. *Currency and credit*. London: Longmans, Green and Co.

HICKS J
1983. *Classics and moderns, collected essays on economic theory vol. 3*. Oxford: Basil Blackwell.

HODRICK RJ
1978. An empirical analysis of the monetary approach to the determination of the exchange rate. In Frenkel JA and Johnson HG (eds): *Exchange rate economics*, 97-128, Reading, Mass: Addison-Wesley Publishing Company.

HOOPER P AND MORTON JE
1982. Fluctuations in the dollar: a model of nominal and real exchange rate determination. *Journal of International Money and Finance* 1: 39-56.

HSIEH DA
1982. The determination of the real exchange rate: the productivity approach. *Journal of International Economics* 12: 355-362.

HUANG RD
1981. The monetary approach to exchange rates in an efficient foreign exchange market: tests based on volatility. *The Journal of Finance* 36 (March): 31-41.

HUME D
1752. Political discourses - 1752. In Rotwein E (ed): *David Hume: writings on economics*. London: Nelson Press.

1777. *Enquiries concerning human understanding and concerning the principles of morals*. Oxford: Clarendon Press (reprinted 1975).

HUTCHISON TW
1935. *The significance and basic postulates of economic theory*. London: Macmillan. New York: Augustus M Kelley (reprinted 1965).
1956. Professor Machlup on verification in economics. *Southern Economic Journal* 22 (April): 476-83.

1976. On the history and philosophy of science and economics. In Latsis SJ (ed): *Method and appraisal in economics*, 181-206, New York: Cambridge University Press.

1977. *Knowledge and ignorance in economics*. Oxford: Basil Blackwell.

ISARD P
1977. How far can we push the law of one price? *American Economic Review* 67: 942-948.

1988. Exchange rate modeling: an assessment of alternate approaches. In Bryant RC, Henderson DW, Holtham G, Hooper P, and Symansky SA (eds): *Empirical Macroeconomics for Interdependent Economies*. Washington: The Brookings Institution.

JOHANNSON J
1975. *A critique of Karl Popper's methodology*. Stockholm: Skandinavian University Books.

JOHANSEN S
1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12: 231-254.

1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector and autoregressive models. *Econometrica* 59: 1551-81.

JOHANSEN S AND JUSELIUS K
1992. Testing structural hypotheses in a multivariate co-integration analysis of the purchasing power parity and the uncovered interest parity for the UK. *Journal of Econometrics* 53: 211-244.

JOHNSON HG
1961. Towards a general theory of the balance of payments. In Johnson HG: *International trade and economic growth*, 153-168, Cambridge: Harvard University Press. Reprinted in Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 46-63, London: George Allen & Unwin Ltd.

1972. The monetary approach to balance of payments theory. *Journal of Financial and Quantitative Analysis* 7: 1555-1572. Reprinted in Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 147-167, London: George Allen & Unwin Ltd.

1976a. Towards a general theory of the balance of payments. In Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 46-63, London: George Allen & Unwin Ltd.

1976b. The monetary approach to balance-of-payments theory. In Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 147-167, London: George Allen & Unwin Ltd.

1977. The monetary approach to balance of payments theory and policy: explanation and policy implications. *Economica* 44: 217-229.

JOHNSON P
1983. *A history of the modern world: from 1917 to the 1980s.* London: Weidenfeld and Nicolson.

KEARNEY CT AND MACDONALD R
1990. Exchange rate volatility, news and bubbles. *Australian Economic Papers* (June): 1-20.

KENEN PB
1985. Macroeconomic theory and policy: how the closed economy was opened. In Jones RW and Kenen PB (eds): *Handbook of international economics, vol.2*, 625-677, Amsterdam: Elsevier Science B.V.

KEYNES JM
1925. *The economic consequences of Mr Churchill*. London: Macmillan

1930. *A treatise on money (volume II)*. London: Macmillan.

1936. *The general theory of employment, interest and money*. London: The Macmillan Press Ltd.

1937. The general theory of employment. *Quarterly Journal of Economics* IV (2): 27-46.

KIM Y
1990. Purchasing power parity in the long run: a cointegration approach. *Journal of Money, Credit and Banking* 12: 491-503.

KNETTER MM
1989. Price discrimination by US and German exporters. *American Economic Review* 79 (1): 198-210.

1993. International comparisons of price-to-market behaviour. *American Economic Review* 83 (3): 473-486.

KOURI PJK AND PORTER MG
1974. International capital flows and portfolio equilibrium. *Journal of Political Economy* 82 (May/June): 443-467.

KRAVIS IB, HESTON A, AND SUMMERS R
1978. Real GDP per capita for more than one hundred countries. *Economic Journal* 88 (June): 215-242.

KRAVIS IB AND LIPSEY RE
1978. Price behaviour in the light of balance of payments theories. *Journal of International Economics* 8 (2): 193-245.

1983. Toward an explanation of national price levels. *Princeton Studies in International Finance* 52 (November): 1-32.

KRUEGER AO
1983. *Exchange-rate determination*. Cambridge: Cambridge University Press

KUGLER P AND LENZ C
1993. Multivariate cointegration analysis and the long-run validity of purchasing power parity. *Review of Economics and Statistics* 75: 180-184.

KUHN TS
1970. *The structure of scientific revolutions*. Chicago: Chicago University Press.

LAFFER AB AND MILES M
1982. *International economics in an integrated world*. Glenview, Illinois: Scott, Foresman and Company, 218-256 and 299-311.

LAKATOS I
1968. Changes in the problem of inductive logic. In Lakatos I (ed): *The problem of inductive logic*, 315-417, Amsterdam: North-Holland.

1970. Falsification and the methodology of scientific research programmes. In Lakatos I and Musgrave A (eds): *Criticism and the growth of knowledge*, 91-196, New York: Cambridge University Press.

1976. *Proofs and refutations: the logic of mathematical discovery*. Cambridge: Cambridge University Press.

1978. In Currie G and Worrall J (eds): *Philosophical papers (vol.*1). Cambridge: Cambridge University Press.

LANE TD
1991. Empirical models of exchange rate determination: picking up the pieces. *Economia Internazionale* 44: 210-226.

LATSIS SJ
1976. A research programme in economics. In Latsis SJ (ed): *Method and appraisal in economics*, 1-42, New York: Cambridge University Press.

LAUDAN L
1977. *Progress and its problems: towards a theory of scientific growth*. Berkeley: University of California Press.

1996. *Beyond positivism and relativism*. Los Angeles: University of California Press.

LEVICH RM
1978. Tests of forecasting models and market efficiency in the international money market. In Frenkel JA and Johnson HG (eds): *The Economics of Exchange Rates*, 129-158, Reading, Mass: Addison-Wesley.

1979. On the efficiency of markets for foreign exchange. In Dornbusch R and Frenkel JA (eds): *International Economic Policy: Theory and Evidence*. Baltimore: John Hopkins University Press.

LIPSEY RG, COURANT PN, PURVIS DD AND STEINER PO
1993. *Economics*. 10th ed. New York: Harper Collins College Publishers Inc.: 29-31.

LIVIATAN N
1979. A diagrammatic exposition of the monetary trade model. *Economic Record* 55: 222-229.

MACDONALD R
1993. Long-run purchasing power parity: is it for real? *Review of Economics and Statistics* 4 (November): 690-695.

1995. Long-run exchange rate modeling: a survey of the recent evidence. *IMF Staff Papers* 42 (3): 437-489.

MACDONALD R AND TAYLOR MP
1991. The monetary approach to the exchange rate: long-run relationships and coefficient restrictions. *Economics Letters* 37 (October): 179-185.

1992. Exchange rate economics: a survey. *IMF Staff Papers* 39 (1): 1-57.

1993. The monetary approach to the exchange rate: rational expectations, long-run equilibrium, and forecasting. *IMF Staff Papers* 40 (1): 89-107.

MACHLUP F
1943. *International trade and the national income multiplier.* Philadelphia: Blakiston.

1978. *Methodology of economics and other social sciences.* New York: Academic Press.

MAGEE B
1973. *Popper.* London: Fontana Paperbacks

MAGEE SP
1978. Contracting and spurious deviations from purchasing power parity. In Frenkel JA and Johnson HG (eds): *The economics of exchange rates*, 67-83, Reading, Mass: Addison-Wesley Publishing Company

MARK NC
1990. Real and nominal exchange rates in the long run: an empirical investigation. *Journal of International Economics* 28: 115-136.

MARSHALL A
1923. *Money, credit, and commerce.* London: Macmillan.

MASSON PR
1988. Strategies for modeling exchange rates and capital flows in multi-country models. *Journal of Policy Modeling* 10: 209-228.

MAYER T
1980. Economics as hard science: realistic goal or wishful thinking? *Economic Inquiry* 18: 165-78.

McCLOSKEY DN AND ZECHER JR
1976. How the gold standard worked, 1880-1913. In Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 359-385, London: George Allen & Unwin Ltd.

McKINNON RT
1969. Portfolio balance and international payments adjustment. In Swoboda AK and Mundell RA (eds): *Monetary problems of the international economy*. Chicago: University of Chicago Press

1981. The exchange rate and macroeconomic policy: changing postwar perceptions. *Journal of Economic Literature* 19: 531-557.

McKINNON RT AND OATES WE
1966. *The implications of international economic integration for monetary, fiscal, and exchange-rate policy*. Princeton Studies in International Finance (16) Princeton: Princeton University Press.

MEADE JE
1951. *The theory of international economic policy (Volume 1: The balance of payments)*. London: Oxford University Press (reprinted 1965).

MEESE RA
1986. Testing for bubbles in exchange markets: a case of sparkling rates? *Journal of Political Economy* 94 (April): 345-373.

MEESE RA AND ROGOFF K
1983. Empirical exchange rate models of the seventies: do they fit out of sample? *Journal of International Economics* 14: 3-24.

1984. The out-of-sample failure of empirical exchange rate models: sampling error or misspecification? In Frenkel JA (ed): *Exchange rates and international macroeconomics*. Chicago: National Bureau of Economic Research, 67-109.

1988. Was it real? The exchange rate - interest differential relationship over the modern floating rate period. *Journal of Finance* 43 (September): 933-948.

MICHAELY M
1960. Relative-prices and income-absorption approaches to devaluation: a partial reconciliation. *American Economic Review* 50: 144-147.

1980. Analyses of devaluation: purchasing power parity, elasticities, and absorption. In Bigman D and Taya T (eds): *The functioning of floating exchange rates: theory, evidence, and policy implications*, 33-54, Cambridge, Mass: Ballinger Publishing.

MOGGRIDGE DE
1969. *The return to gold, 1925: the formulation of economic policy and its critics*. London: Cambridge University Press.

MOOSA IA
1994. The monetary model of exchange rates revisited. *Applied Financial Economics* 26: 279-287.

MORGAN MS
1984. *The history of econometric ideas*. Cambridge: Cambridge University Press.

1988. Finding a satisfactory empirical model. In de Marchi N (ed): *The Popperian legacy in economics*, 199-212, Cambridge: Cambridge University Press.

MUNDELL RA
1960. The monetary dynamics of international adjustment under fixed and flexible exchange rates. *Quarterly Journal of Economics* 74: 227-257.

1961. Flexible exchange rates and employment policy. *Canadian Journal of Economics and Political Science* 27 (November): 509-517.

1962. The appropriate use of monetary and fiscal policy for internal and external stability. *IMF Staff Papers* 9 (March): 70-79.

1963. Capital mobility and stabilization policy under fixed and flexible exchange rates. *Canadian Journal of Economics and Political Science* 29 (November): 475-485.

1968. *International economics*. New York: MacMillan.

1971. *Monetary economics*. Pacific Palisades: Goodyear.

1976. Barter theory and the monetary mechanism of adjustment. In Frenkel JA and Johnson HG (eds): *The monetary approach to the balance of payments*, 64-91, London: George Allen & Unwin Ltd.

MUSSA M
1978. The exchange rate, the balance of payments, and monetary and fiscal policy under a regime of controlled floating. In Frenkel JA and Johnson HG (eds): *The economics of exchange rates*, 47-65, Reading, Mass: Addison Wesley Publishing Co.

1979. Empirical regularities in the behaviour of exchange rates and theories of the foreign exchange market. In Brunner K and Meltzer AH (eds): *Policies for Employment, Prices, and Exchange Rates: Carnegie-Rochester Conference Series on Public Policy* 11. Amsterdam: North-Holland Publishing Company.

1984. The theory of exchange rate determination.In Bilson JFO and Marston RC (eds): *Exchange rate theory and practice*, 13-78, Chicago: University of Chicago Press.

1986. Nominal exchange rate regimes and the behaviour of real exchange rates. In Brunner K and Meltzer A (eds): *Carnegie-Rochester Series on Public Policy* 25: 117-214.

NURKSE R
1944. *International currency experience: lessons of the inter-war period*. Princeton: League of Nations.

OATES WE
1966. Budget balance and equilibrium income: a comment on the efficacy of fiscal and monetary policy in an open economy. *Journal of Finance* 21 (September): 489-498.

OFFICER LH
1974: Purchasing power parity and factor price equalization. *Kyklos* 27: 868-878.

1976a. The purchasing power parity theory of exchange rates: a review article. *IMF Staff Papers* 23: 545-579.

1976b. Productivity bias and purchasing power parity: an econometric investigation. *IMF Staff Papers* 23: 1-60.

1984. *Purchasing power parity theory and exchange rates*. Greenwich, Conn: JAI Press.

OHLIN B
1967. *Interregional and international trade*. Cambridge, Mass: Harvard University Press.

PAPELL DH
1988. Expectations and exchange rate dynamics after a decade of floating. *Journal of International Economics* 25 (November): 307-317.

PATEL J
1990. Purchasing power parity as a long-run relation. *Journal of Applied Econometrics* 5: 367-379.

PHILLIPS PCB
1991. Optimal inference in cointegrated regression. *Econometrica* 59: 238-306.

1993. Fully modified least squares and vector autoregression. *Cowles Foundation Discussion Paper* No. 1047. New Haven, Yale University.

PHILLIPS PCB AND LORETAN M
1991. Estimating long-run economic equilibria. *Review of Economic Studies* 59: 407-436.

PILBEAM KS
1991. *Exchange rate management: theory and evidence*. London: Macmillan.

PIPPINGER MK
1993. Cointegration tests of purchasing power parity: the case of Swiss exchange rates. *Journal of International Money and Finance* 12: 46-61.

POPPER K
1957. *The poverty of historicism*. London: Routledge & Kegan Paul.

1959, translation from 1934. *The logic of scientific discovery*. London: Hutchinson.

1963. *Conjectures and refutations: the growth of scientific knowledge*. London: Routledge & Kegan Paul.

POTERBA J AND SUMMERS L
1986. The persistence and volatility of stock market fluctuations. *American Economic Review* 76: 1142-1151.

PUTNAM BH AND WOODBURY JR
1979. Exchange rate stability and monetary policy. *Review of Business and Economic Research* 15 (2): 1-10.

QUINE WVO
1980. *From a logical point of view*. 2nd ed. Harvard: Harvard University Press.

REDMAN DA
1992. *A reader's guide to rational expectations*. Aldershot: Edward Elgar Publishing Company.

RICARDO D
1821. *Principles of political economy and taxation*. London: EC Connor, G Bell and Sons (eds) 1911.

RICHARDSON JD
1978. Some empirical evidence on commodity arbitrage and the law of one price. *Journal of International Economics* 8: 341-51.

RODRIGUEZ CA
1980. The role of trade flows in exchange rate determination: a rational expectations approach. *Journal of Political Economy* 88 (6): 1148-1157.

ROGERS C AND JENKINS MA
1995. Haircuts or hysteresis? Sources of movements in real exchange rates. *Journal of International Economics* 38 (3/4): 339-360.

ROGOFF K
1996. The purchasing power parity puzzle. *Journal of Economic Literature* 34 (June): 647-668.

ROLL R
1979. Violations of purchasing power parity and their implications for efficient international commodity markets. In Sarnat M and Szego GP (eds): *International finance and trade (volume 2)*, 96-125, Cambridge, Mass: Ballinger Publishing Co.

RUNDE J AND TORR C
1985. Divergent expectations and rational expectations. *South African Journal of Economics* 53: 217-25.

SALTER WEG
1959. Internal and external balance: the role of price and expenditure effects. *Economic Record* 35: 226-238.

SAMUELSON PA
1964. Theoretical notes on trade problems. *Review of Economics and Statistics* 46 (2): 145-154.

1974. Analytical notes on international real-income measures. *Economic Journal* 84: 595-608.

SERLETIS A
1995. Maximum likelihood cointegration tests of purchasing power parity: evidence from seventeen OECD countries. *Weltwirtschaftliches Archiv* 22: 476-493.

SHACKLE GLS
1972. *Epistemics and economics: a critique of economic doctrines*. London: Cambridge University Press, 206-207.

SHEARMUR J
1991. Theoretical progress and excess content. In de Marchi N and Blaug M (eds) 1991: *Appraising Economic Theories: Studies in the Methodology of Research Programs*, 35-53, Aldershot: Edward Elgar Publishing Company.

SHEFFRIN SM
1996. *Rational expectations*. Cambridge Surveys of Economic Literature. New York: Cambridge University Press.

SHILLER RJ
1979. The volatility of long-term intertest rates and expectations models of the term structure. *Journal of Political Economy* 87 (December): 1190-1219.

SHILLER RJ AND PERRON P
1985. Testing the random walk hypothesis: power versus frequency of observation. *Economics Letters* 18: 381-386.

SHINASI G AND SWAMY PAVB
1989. The out-of-sample forecasting performance of exchange rate models when coefficients are allowed to change. *Journal of International Money and Finance* 8 (September): 375-390.

SIMS C
1972. Money, income, and causality. *American Economic Review* 62 (4): 540-552.

SMITH VL
1994. Economics in the laboratory. *Journal of Economic Perpectives* 8 (1): 113-131.

SOHMEN E
1969. *Flexible exchange rates*. Chicago: University of Chicago Press.

SMITH PN AND WICKENS MR
1986. An empirical investigation into the causes of failure of the monetary model of the exchange rate. *Journal of Applied Econometrics* 1: 143-162.

SWAN T
1960. Economic control in a dependent economy. *Economic Record* (March): 156-167.

TAYLOR MP
1988a. An empirical analysis of long-run purchasing power parity using cointegration techniques. *Applied Econometrics* 20: 1369-1381.

1988b. A dynamic model of forward foreign exchange risk, with estimates for three major exchange rates. *The Manchester School* 56: 55-68.

1995. The economics of exchange rates. *Journal of Economic Literature* 33 (March): 13-47.

TAYLOR MP AND MACDONALD R
1989. Exchange rate economics: an expository survey. In Taylor MP and MacDonald R (eds): *Exchange rates and open economy macroeconomic models*, 4-38, Oxford: Basil Blackwell.

TAYLOR MP AND MACMAHON PC
1988. Long-run purchasing power parity in the 1920s. *European Economic Review* 32: 179-197.

THUROW LC
1983. *Dangerous currents: the state of economics*. New York: Oxford University Press.

VINER J
1937. *Studies in the theory of international trade*. New York: Harper & Bros.

WEINTRAUB ER
1985a. Appraising general equilibrium analysis. *Economics and Philosophy* 1: 23-37.

1985b. *General equilibrium analysis: studies in appraisal*. New York: Cambridge University Press.

WHEATLEY J
1803. *Remarks on currency and commerce*. London: Burton.

WHITMAN MvN
1975. Global monetarism and the monetary approach to the balance of payments. *Brookings Papers on Economic Activity* 3: 121-166.

WICKENS MR
1996. Interpreting cointegrating vectors and common stochastic trends. *Journal of Econometrics* 74: 255-271.

WICKSELL K
1919. The riddle of foreign exchanges. In Lindahl E (ed) 1969: *Knut Wicksell: selected papers on economic theory*, 229-249, New York: Augustus M Kelley

WILLET TD AND FORTE F
1969. Interest rate policy and external balance. *Quarterly Journal of Economics* 83 (May): 242-262.

WONG S
1978. *The foundations of Paul Samuelson's revealed preference theory*. London: Routledge & Kegan Paul.