Essays in Applied Time Series Econometric Methodology: The Role of Random Field Regression

by

Edward J. O'Brien

Submitted to the Department of Economics
in fulfilment of the requirements for the degree of

Doctor of Philosophy

at the

University of Dublin

March 2007

Declaration

I declare that this thesis submitted to the University of Dublin, Trinity College, for the degree of Doctor of Philosophy, has not been submitted as an exercise for a degree at any other university. The large majority of research contained herein is entirely my own; several chapters of this thesis have, however, benefited from co-authorship with my thesis supervisor, Michael J. Harrison, and our colleague Derek Bond. I authorise the Library of the University of Dublin, Trinity College to lend or copy this thesis upon request.

Date: 1st March 2007

Signature: Edward J. O'Brien

Summary

This thesis is a collection of essays in applied time series econometrics, focusing particularly on the role of random field regression. The thesis is introduced in Chapter 1, which discusses the motivation for the thesis and outlines its structure.

The second chapter discusses nonlinear econometric modelling. It introduces the concept of nonlinearity and discusses its importance in economics and econometrics. It also provides a treatment of several approaches to modelling nonlinearity in economics, before giving an account of the approach to nonlinear econometric modelling proposed by Hamilton (2001). It then describes some of the methods of nonlinear optimisation that may be used in the GAUSS computer program provided by Hamilton for the implementation of his methodology. The performance of this program is investigated using data relating to Hamilton's examples, two versions of the GAUSS software and a range of alternative numerical optimisation options. The impact of changes in initial parameter estimates and the use of pairs of optimisation algorithms are also examined. The effects of changes in the sample data on the results produced by Hamilton's procedure are explored. Finally, a discussion of the published comments made by Hamilton (2005) on this work are included. The results of this study suggest some clear conclusions, which will be of value to those contemplating working with Hamilton's method.

Chapter 2 also highlights Hamilton's (2001) LM-type test, which is likely to be little known, due to the technical nature of the original paper detailing the methodology and also because the test is embedded within that methodology. Chapter 3, continuing on the theme of nonlinearity in economics and econometrics, and the approach put forward by Hamilton, investigates the properties of several tests for neglected nonlinearity in time series, using Monte Carlo simulation methods. This study is motivated by the LM-type test proposed by Hamilton. The comparative properties of this test, as an integral part of Hamilton's framework and indeed as a stand-alone test for nonlinearity, have yet to be fully explored. Therefore, this chapter investigates the comparative properties of the Hamilton test and some well-known alternatives, by applying them to some model specifications commonly encountered in empirical research. Hamilton's LM-type test is evaluated across a range of parameters and data, and compared with the Durbin-Watson (1950) bounds test, Ramsey's (1969) RESET test, the Harvey-Collier (1977) ψ-test and three tests put forward by Dahl and González-Rivera (2003). The results from this chapter confirm the powerful nature of the Hamilton test and its variants, particularly the λ_{OP}^{A} test of Dahl and González-Rivera. Interestingly, however, it is also shown that Ramsey's (1969) RESET test is powerful by comparison to the random field-based tests.

Chapter 4 provides the background theory required for the remaining chapters of the thesis. As the remaining chapters compare and contrast the results of modelling time-series relationships using the Hamilton (2001) methodology with a variety of alternative methods, they are introduced here. These alternative methods share the common trait that they exploit the concept of (co)integration in modelling economic relationships. Three approaches to modelling economic relationships are compared with Hamilton's approach: the Engle-Granger (1987) 2-Step procedure, Johansen's (1988, 1991) vector autoregressive approach and common factor (COMFAC) analysis (Hendry and Mizon, 1978). Chapter 4 reviews each

of these methods in turn. Attention is also given to the work of Dolado, Gonzalo, and Mayoral (2002) on fractional integration and Johansen's (2002) small sample correction, which may offer further insight into the implementation, application and results of some of the above methods.

Chapters 5, 6 and 7 draw attention to the limitations of the standard unit root-cointegration approach to economic and financial modelling, and to some of the alternatives based on the idea of fractional integration, long memory models and the random field regression approach to nonlinearity. Chapter 5 examines a well-known demand for money dataset for Denmark and Finland, which relates to an area of economics where cointegration is commonly employed. Chapter 6 explores purchasing power parity for Irish data by investigating the behaviour of the Irish exchange rate in respect to Germany and the United Kingdom. Purchasing power parity is another area of economic theory which has often been empirically investigated using the $I(1)/I(0)$ framework. In each chapter, standard unit root testing procedures and tests of cointegration are employed to explore the relationships outlined. This is followed by consideration of the possibility of fractional integration and nonlinearity in each case.

The findings of standard unit root testing and cointegration analyses prove somewhat confusing in both cases, suggesting perhaps that the $I(1)/I(0)$ framework is not the most appropriate in either case. Fractional integration analyses confirm this finding. Strong evidence of nonlinearity is found in both these applications, using Hamilton's (2001) test. That methodology suggests some clear conclusions regarding the variables which contribute to that nonlinearity.

Chapter 7 examines issues surrounding the testing of fractional integration and nonlinearity in relation to the forward exchange rate anomaly of Fama (1984). The behaviour of three exchange rates and premiums is investigated in terms of fractional integration and nonlinearity. The findings provide some support for $I(1)$ exchange rates but suggest fractionality for premiums, mixed evidence on cointegration, and a strong possibility of time-wise nonlinearity. Significantly, when the nonlinearity is modelled using a random field regression, the forward anomaly disappears.

The results from these illustrative case studies not only offer interesting insights into the specific areas of money demand, purchasing power parity and the forward exchange rate anomaly, but they also offer conclusions that should aid practitioners in applied time series econometrics. Specifically, these studies draw attention to the issues of stationarity, nonstationarity, structural breaks and nonlinearity in economics and econometrics, and outline methodologies and modelling approaches, at the frontier of current econometric thinking, that can be used by practitioners to explore and better understand these issues.

Chapter 8 concludes the thesis by offering a summary of the findings and by outlining avenues for future research.

Acknowledgements

There are many people who deserve thanks and recognition for the support and assistance they rendered during the completion of this thesis.

I gratefully acknowledge the valuable assistance of David A. Belsley, Boston University, Christian M. Dahl, Purdue University, James D. Hamilton, University of California at San Diego, Søren Johansen, University of Copenhagen, Laura Mayoral, Universitat Pompeu Fabra, Bruce Mizrach, Rutger's University and Editor of *Studies in Nonlinear Dynamics and Econometrics*, Theodore Panagiotidis, Loughborough University, Curt Wells, University of Lund, Jonathan H. Wright, The Federal Reserve Board and three anonymous referees. Participants at the 3^{rd} International Association for Statistical Computing World Conference on Computational Statistics and Data Analysis, Limassol, Cyprus, the 20^{th} Irish Economic Association Annual Conference, Bunclody, Co. Wexford, the 26^{th} Conference on Applied Statistics in Ireland, Killarney, Co. Kerry, the 4^{th} INFINITI Conference on International Finance, Dublin and various seminars at the Department of Economics, University of Dublin, Trinity College, are also acknowledged for their questions, comments and suggestions.

A Government of Ireland Scholarship from the Irish Research Council for the Humanities and Social Sciences, a Trinity Postgraduate Award and financial assistance from the Royal Economic Society, the Economic and Social Research Council, the International Association for Statistical Computing and the Department of Economics, University of Dublin, Trinity College, are gratefully recognised.

At Trinity College, I sincerely acknowledge R. Andrew Somerville, Alan H. Matthews and Patrick Paul Walsh for their support throughout my studies, as both an undergraduate and postgraduate student. I wish to thank those who shared my workspace, Ali Uğur, Michael Gallagher, Liam Delaney and Simon Cullen, for their advice, assistance and good humour, and all at the Department of Economics, both staff and students. I thank also the administrative staff of the Department of Economics, too numerous to name.

My family and friends have provided continued encouragement throughout my studies, for which I am very grateful. In particular, I wish to thank my parents, Richard and Yvonne, and my siblings for their unwavering support throughout my education.

Derek Bond has provided seemingly boundless energy and inspiration to this enterprise. He has contributed substantially to this work as a co-author in related research and has willingly given his opinions, advice and encouragement. For that I am sincerely grateful. His visits to Dublin were much anticipated and much appreciated.

To Michael J. Harrison, my supervisor for the duration of this thesis, an acknowledgement seems insufficient. Not only has he guided the research presented herein, but he has guided my education in econometrics and beyond, for many years. From sparking my interest in statistics and econometrics as an undergraduate student, to encouraging independence in my research, he has remained a primary source of inspiration. The insight, understanding and wisdom that he has imparted, has and will continue to make a lasting impression on my life. With his guidance and accompaniment, this project has never felt a burden. It has, rather, been a thoroughly enjoyable journey. It has been an enriching experience, tainted with sadness at journey's end. I thank him sincerely for his attention-to-detail, enthusiasm, inspiration, patience and, dare-I-say, friendship. Being his student has been a privilege.

Finally, to A. Katarina Bengtsson, my fiancée, who has known me only as a graduate student. For your companionship, your tireless patience, your good sense and endless encouragement, I will be forever grateful. Without your loving support, this work would likely remain incomplete. For this, and so much more, I will be eternally thankful and look forward to a summers day at Västra Sallerup.

Models are, for the most part, caricatures of reality, but if they are good, then, like good caricatures, they portray, though perhaps in a distorted manner, some of the features of the real world. The main role of models is not so much to explain and to predict - though ultimately these are the main functions of science - as to polarize thinking and to pose sharp questions.

Mark Kac, *Some Mathematical Models in Science.*

It must be emphasized that a cointegration analysis cannot be the final aim of an econometric investigation, but it is our impression that as an intermediate step a cointegration analysis is a useful tool in the process of gaining understanding of the relation between data and theory, which should help in building a relevant econometric model.

Søren Johansen,
Likelihood Based Inference in Cointegrated Vector Autoregressive Models.

In short, the paper proposes a single encompassing framework for nonlinear modelling, offering a new test for nonlinearity, methods to infer what the nonlinear function looks like, and checks of whether it is adequately described by some particular model.

James D. Hamilton,
A Parametric Approach to Flexible Nonlinear Inference.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Foreword

This thesis, a collection of related essays in applied time series econometrics, has modest aims. It sets out to explain and explore Hamilton's (2001) nonlinear modelling framework, a relatively new and potentially important approach to applied econometric modelling, including his proposed test for nonlinearity. It will also consider the concept of long memory time series and attempt to establish what these alternative modelling approaches have to offer, by comparison to a standard analysis of nonstationary data. This is achieved by considering three illustrative case studies where such standard analysis has been routinely employed. These aims are further elaborated upon in the subsequent section and in each respective chapter. The structure of the thesis will also be outlined briefly here.

1.2 Motivation

Nonstationary data have played a key role in economic modelling for the past two decades. The seminal works of Nelson and Plosser (1982) and Engle and Granger (1987) highlighted the prevalence of nonstationarity, particularly in macroeconomic time series. The autoregressive-integrated-moving average, or ARIMA, approach put forward by Box and Jenkins (1976), provided an early framework for modelling nonstationary data. With further developments in this area, including an understanding of spurious regression and its likely causes (see, for example, Granger and Newbold, 1974), a large theoretical and applied literature grew up. Standard approaches to dealing with nonstationary time series were introduced and the concept of cointegration and error-correction became staples of macroeconomic modelling.

Despite the attractiveness of this approach, its ability to combine long-run and short-run dynamics, and its ability of overcome the so-called spurious regression, the methodologies that built up in this area are not without their potential pitfalls. A common concern in all applied research is sample size. As cointegration is concerned with long-run relationships, that concern is particularly relevant. This issue has recently been addressed by a small sample correction to the popular Johansen (1988, 1991) procedure and has shown the potential for incorrect inferences where sample sizes are small. The primary concern, however, is that these methods are strictly applicable to unit root processes, as opposed to the closely related nonstationary processes. To classify series requires suitable testing. Two immediate problems are evident. An early test for unit roots, the augmented Dickey-Fuller (1979, 1981) test, is known to have low power. This procedures tests the null hypothesis that the series under examination is nonstationary against the alternative that it is stationary. The low power, therefore, often manifests itself in failing to reject nonstationarity, when in fact the series in question is stationary. These issues have been addressed to some degree by new testing procedures, but generally, test power has remained low. One reason for this low power is that it is difficult to statistically distinguish between nonstationary series and nonlinear but stationary series. Structural change in a series may also lead to incorrect inferences regarding nonstationarity.

A further issue relates to the definition of nonstationarity and unit roots themselves, as so often used in the applied literature. Unit root and cointegration tests offer a framework to test for nonstationarity when series are integrated precisely to integer order; unity for example. That is, the series is generated by a unit root process. A series can exhibit nonstationarity, however, if its order of integration lies between 0.5 and 1.5. This introduces the concept of fractional integration, so far little used in applied research. A series can, however, be considered to be long memory and stationary if the order of integration lies between zero and 0.5. These important issues remain less well known than the standard $I(1)/I(0)$ analysis and they can make modelling nonstationary data far more challenging than it may at first appear.

Another concern, briefly alluded to already, is nonlinearity. It can be very difficult to distinguish between a nonlinear series and a nonstationary series. That is, a series may be stationary with a nonlinear data generating process, but testing procedures cannot distinguish this from nonstationarity. The presence of structural breaks adds further complications, making what may otherwise be stationary series nonstationary.[1] The result can be a modelling effort directed towards nonstationary data, when in fact it should be directed to modelling the nonlinearity or structural breaks. The relative attractiveness of the cointegration framework, against the potentially more complex nonlinear or stationary with breaks case, goes some way to explaining why this approach is often neglected. This is reinforced by the availability of pre-programmed computing routines which make the former approach very accessible.

Recent developments in time series econometrics may offer a solution to some of these problems. The field of fractional integration has developed greatly, with several procedures available to test whether a series is integrated to integer order or fractionally. The field of fractional cointegration, although not touched upon in this thesis, is also growing.[2] Finally, developments have also been made in the area of modelling nonlinearity, a notable example being that of Hamilton (2001). This offers a potentially powerful alternative to the standard approach, when the appearance of unit roots is actually due to structural change or nonlinearity.

The aims of the thesis are to explore these issues and to highlight the possible deficiencies in the standard $I(1)/I(0)$ framework, by considering the possible alternatives of fractional integration, small samples and nonlinearity. The thesis will initially explore the alternative methods, particularly that of Hamilton (2001), before proceeding to illustrative case studies to evaluate these recent developments.

1.3 Structure

After the brief introduction of Chapter 1, which discussed the motivation for the thesis, Chapter 2 will begin with a discussion of nonlinear econometric modelling, outlining some

[1]See, for example, Perron (1989).

[2]A growing literature on fractional cointegration exists, encompassing both theoretical aspects and empirical applications. Much of the theoretical literature considers issues relating to testing and inference in fractionally cointegrated systems; see, for example, Kim and Phillips (2000), Martin (2001), Davidson (2002), published in special edition of the *Journal of Econometrics* dedicated to long memory and nonlinear time series, Hualde and Robinson (2002), Chen and Hurvich (2003a, 2003b), Gil-Alana (2003), Robinson and Hualde (2003), Velasco (2003), Dittmann (2004), and Robinson and Iacone (2005). The concept of fractional cointegration has been applied to a wide range of topics; see, for example, Cheung and Lai (1993b), Baillie and Bollerslev (1994), Lien and Tse (1999), Liu and Chou (2003), Masih and Masih (2004), and Caporale and Gil-Alana (2005).

approaches that are popular in the applied literature. It will also introduce the concept of nonlinearity and will discuss its importance in economics and econometrics. It will then give an account of the approach to nonlinear econometric modelling proposed by Hamilton (2001), the focus of this chapter. The theory behind this framework will be outlined, as will its test for nonlinearity. The issue of the applicability of this approach in the presence of nonstationarity will also be discussed. The implementation of this methodology will then be discussed, as will some of the methods of nonlinear optimisation that may be used in the GAUSS computer program provided by Hamilton. The performance of this program will be reported on, in terms of computational issues and data sensitivity: the former detailing the impact of changes in initial parameter estimates and the use of pairs of optimisation algorithms and the latter the effects of changes in the sample data on the results produced by Hamilton's procedure. Finally, the published comments made by Hamilton (2005) on this work will be discussed. The chapter will conclude with recommendations for those contemplating working with the method.

Continuing on the theme of nonlinearity in economics and econometrics, and the approach put forward by Hamilton (2001), Chapter 3 will investigate the properties of several tests for neglected nonlinearity in time series, using Monte Carlo simulation methods. This study is motivated by the LM-type test proposed by Hamilton and previously introduced in Chapter 2. This chapter will begin with a description of each of the seven tests to be used in the comparative study, with an emphasis being placed on those less well-known tests. The design of the Monte Carlo experiments, which set out to investigate the comparative properties of these tests by applying them to some model specifications commonly encountered in empirical research, will be outlined. Some basic but necessary material on hypothesis testing will also be included for completeness. The results from the simulation study will then be discussed at some length. The conclusion will outline the findings of the study and offer a comparative evaluation of the Hamilton test for nonlinearity.

Chapter 4 will serve only to provide the background theory required for the remaining chapters of this thesis. As the remaining chapters compare and contrast the results of modelling time-series relationships using the Hamilton (2001) methodology with a variety of alternative methods, the alternatives will be introduced here, although a brief treatment of stationarity, nonstationarity and unit root processes will be offered initially. The alternative methods to be introduced share the common trait that they exploit the concept of (co)integration in modelling economic relationships. Three approaches to modelling these relationships will be compared with Hamilton's: the Engle-Granger (1987) 2-Step approach, Johansen's (1988, 1991) vector autoregressive approach and common factor (COMFAC) analysis, as described by Hendry and Mizon (1978). Chapter 4 will review each of these methods in turn. Attention will also be given to the work of Dolado, Gonzalo, and Mayoral (2002), on fractional integration and Johansen's (2002) small sample correction, which may offer further insight into the implementation, application and results of some of the above methods.

Chapters 5, 6 and 7 will draw attention to the limitations of the standard unit root-cointegration approach to economic and financial modelling, and to the potential of some of the alternative methods based on the idea of fractional integration, long memory models and the random field regression approach to nonlinearity; i.e., the methods introduced in previous chapters. Chapter 5 examines a well-known demand for money dataset for Denmark and Finland, which relates to an area of economics in which cointegration is commonly employed. Chapter 6 explores purchasing power parity for Irish data by investigating the behaviour of the Irish exchange rate in respect to Germany and the United Kingdom. Purchasing power parity is another area of economic theory which has often been empirically investigated using the $I(1)/I(0)$ framework. The structure of these chapters will be very similar. In each chapter, a brief account of the demand for money and purchasing power parity, respectively, will be offered. Standard unit root testing procedures and tests of cointegration will then be employed to explore the relationships outlined. Finally, the possibility of fractional integration and nonlinearity in each case will be explored. Each chapter will conclude with a summary of the findings.

Chapter 7 will examine the issues surrounding the testing of fractional integration and nonlinearity in relation to the forward exchange rate anomaly of Fama (1984), for three exchange rates and premiums. The structure will be similar to that of chapters 5 and 6. After a brief introduction to the anomaly, standard unit root and cointegration tests will be applied. Tests for fractional integration will also be carried out, before exploring the relationship with Hamilton's (2001) methodology, to consider the possibility of nonlinearity.

It is hoped that the results from these illustrative case studies may offer interesting insights into the specific areas of money demand, purchasing power parity and forward exchange rate anomaly, but they may also offer guidance that could aid practitioners in using time series in applied economic research. Chapter 8 will conclude the thesis by offering a summary of the findings and by outlining avenues for future research.[3]

[3]The analysis carried out in this thesis uses a number of software environments, namely, EViews v. 5.1, Gauss v. 3 and v. 5, JMulTi v. 4.12, Microfit v. 4.11, Ox v. 3.4 and Rats v. 5. It has been typeset in LaTeX using MiKTeX v. 2.4; see, for example, Kopka and Daly (2004), Lamport (1994) and Mittelbach, Goosens, Braams, Carlisle, and Rowley (2004).

Chapter 2

The Hamilton Random Field Regression Methodology and its Implementation

Research carried out in the course of writing this chapter has led to the publication of 'Investigating Non-linearity: A Note on the Implementation of Hamilton's Random Field Regression Model' by Bond, D., M.J. Harrison, and E.J. O'Brien, in *Studies in Nonlinear Dynamics and Econometrics*, 9, Article 2, September 2005, and also Trinity Economic Papers, Nos. 12 (2003) and 4 (2005). See also the response from J.D. Hamilton, 'Comments on 'Investigating Nonlinearity" in *Studies in Nonlinear Dynamics and Econometrics*, 9, Article 3, September 2005. The Bond, et al. (2005a) paper was presented at the 3^{rd} International Association for Statistical Computing World Conference on Computational Statistics and Data Analysis, Limassol, Cyprus, 28^{th}-31^{st} October 2005.

2.1 Introduction

As discussed in Chapter 1, this thesis is concerned with applied time-series modelling of macroeconomic and financial data, and some of the issues involved therein. One of these issues is the presence of nonlinear relationships among economic variables. This chapter will introduce the importance of nonlinear models in economics and econometrics, and will briefly discuss some tests and methodologies used in practical applications. Those methods include approaches based on time-varying parameters, threshold and smooth transition autoregressive processes, regime-switching models and smoothing splines. The issue of a nonlinear modelling strategy is also touched upon.

Attention is then turned to a new and potentially important technique, which is the main focus of this chapter. Hamilton (2001) proposed an approach to nonlinear modelling of economic relationships that provides a single flexible parametric framework for testing for nonlinearity, drawing inference about the form of nonlinearity and assessing the adequacy of the description of nonlinearity provided by specific models. This approach treats functional form as the outcome of a latent stochastic process. This latent process is modelled using a relatively new Gaussian random field concept that generalises Brownian motion to k dimensions. From the practicing economist's viewpoint, the importance of Hamilton's approach lies in the valuable insights it can provide for model construction and the resulting enhancement of the forecasting ability of economic models.

However, the new methodology has been little used to date and its full potential remains to be established. As Hamilton (2001, p. 552) pointed out, its usefulness for particular sample sizes and nonlinearities is a matter for empirical investigation. Yet, citing his own three examples and the Monte Carlo studies by Dahl (2002), he suggested that the method holds much promise.

The main purpose of this chapter is to address a number of practical issues that arise when using the Hamilton (2001) approach. The first of these concerns computation and reports on experience gained with Hamilton's software to implement the method. It appears that the numerical optimisation involved is not an entirely straightforward matter, either when using Hamilton's dataset or alternative samples. Thus while the Hamilton case study is the primary focus, this finding may have more general relevance for procedures that employ similar optimisation techniques. The second issue concerns the sensitivity of the method to changes in data. Experiments suggest that minor data changes can have implications for computation and big effects on the results.

The structure of the chapter is as follows. Section 2.2 discusses the importance of nonlinear models in economics and Section 2.3 introduces some techniques commonly used for modelling them. Section 2.4 introduces the Hamilton (2001) method, while sections 2.5, 2.6 and 2.7 consider the findings of this research. Specifically, issues relating to the computational and data sensitivity matters are reported. Section 2.8 contains comments on the reply to this work offered by Hamilton (2005), while Section 2.9 concludes.

2.2 Nonlinear Models in Economics and Econometrics

The main goal of this chapter is to consider the method of Hamilton's (2001) flexible nonlinear inference, its theory and implementation. Before doing so, however, the importance of nonlinearity in economics is very briefly explored. It is generally accepted that nonlinearity occurs naturally in economics and that many economic theories suggest a nonlinear relationship. Many relationships of interest, be they from the field of macroeconomics, microeconomics or financial economics, have naturally occurring thresholds, constraints and boundaries, which may result in nonlinearities. A plausible nonlinear specification may sometimes be suggested by economic theory. These are often incomplete, however, and may not fit the data well. Also, even when theory may indicate nonlinearity in a given relationship, it may offer no insight into the form of that nonlinearity. Finally, theory may say nothing about a nonlinearity whatsoever. As a result of these difficulties, linear models are often assumed to be adequate. A known nonlinear specification may be linearised, thereby avoiding further complications in analytical or empirical estimation, or as has often been the case, a nonlinear model is approximated by a linear model.[1] Modelling in this way has proven to be quite successful. Problems, however, have been known to arise when this approach is adopted. Tests of such approximated specifications frequently reject parameter constancy, pointing perhaps to structural breaks. A common, but unsatisfactory solution has been to build a dummy variable into the specification to allow for such a break. It is preferable, however, in such cases to specify an equation that allows for nonlinearities in the parameters.

Advances in recent decades have been made which allow for the use of nonlinear specifications. Much work has been done in the field of nonlinear time series analysis. In tandem with this, low cost computational power has become widely available to the applied researcher, making such methods accessible. As nonparametric and semiparametric estimation methods have become computationally feasible, they have also become increasingly popular, despite the criticisms made against them, some of which will be outlined briefly in the next section. Such developments have inspired further advances in statistical techniques and have been applied throughout economics, econometrics and finance.[2]

Considering this, when should a nonlinear specification be used? From an economic and

[1] An excellent example is the Phillips curve, which is perhaps one of the most well-known, if not controversial, examples of nonlinearity in economics. Phillips (1958) proposed the following form of the relationship between wage rates and unemployment

$$y + a = bx^c,$$

which can be linearised easily to become

$$\log(y + a) = \log b + c \log x.$$

The preferred specification,

$$y + a = bx^c + k\left(\frac{1}{x^m}\frac{dx}{dt}\right),$$

was not used. It is widely assumed that this was due to a lack of available computing power. Instead, the loglinear equation above was estimated by OLS with just four observations. The original 52 observations were aggregated over intervals to produce just six observations, of which four were used.

[2] See, for example, Lütkepohl, Teräsvirta, and Wolters (1999), who examined the stability of German money demand, Lee, Kim, and Newbold (2005), who explored spurious regression in the context of nonlinearity and Taylor, Peel, and Sarno (2002), who re-assessed purchasing power parity in light of nonlinear behaviour in exchange rates.

econometric theory point of view, the choice is clear. A linear specification should only be used if theory suggests a linear relationship, or a relationship that could be reasonably approximated by a linear model, without eliminating essential elements of that model. This of course assumes that theory provides some indication of the nature of the relationship. It is far more likely that the nonlinear relationships suggested by economic theory will be vague at best. Granger and Teräsvirta (1993) suggested a strategy for nonlinear modelling. In deciding on the variables to include, they advocate both the simple-to-general approach, suggested by Box and Jenkins (1976) in the context of linear modelling, and also the general-to-specific approach of Hendry.[3] Having tested for, and found nonlinearity, 'there appears to be no simple answer' to the question of which nonlinear model should be used.[4] In fact, it is recommended that a variety of models be considered and that economic theory, vague as it may be, must play an important role, where possible. In addition, the decision-making process can be guided by post-estimation misspecification testing. Such post-estimation testing is useful for inferring whether a given nonlinear form is appropriate. Note that although some tests for nonlinearity test a null of linearity against a specific nonlinear alternative, the majority test against a nonspecific alternative.[5] The issue of testing will be discussed in greater depth in Chapter 3. A linear model may also be considered if the nonlinear model is difficult to work with and it is felt that the linear alternative may be useful or instructive. Finally, an exploratory data analysis approach may be taken, allowing the data to guide the final specification of the model. The role played by economic theory cannot be understated.

Despite the obvious attraction of using nonlinear models, as outlined above, difficulties may still arise. The chosen specification will, therefore, be one of the vast array of alternative nonlinear models, selected perhaps more for analytical convenience than prior theoretical considerations. Of course, as with all model specifications, any nonlinear model remains a simplification and can be expected to fail from time to time.

Given this brief overview of why nonlinear modelling may be preferable, the next section will give consideration to some of those methods successfully employed in economics and econometrics.

2.3 Methods for Modelling Nonlinearity

This section reviews some of the methods available in the field of nonlinear econometric modelling, including both parametric and nonparametric techniques. The methods to be considered are time-varying parameters, threshold autoregressions, models of regime switching, smooth transition autoregressions and smoothing splines. Only brief consideration is given to the nonparametric approaches. These methods, as Hamilton (2001) pointed out, sacrifice many of the benefits of parametric methods.[6] Nonparametric methods present problems in inference, be it Bayesian or classical in nature, and are not readily adaptable to the

[3] See, for example, Gilbert (1986).

[4] Granger and Teräsvirta (1993), p. 166.

[5] Teräsvirta (1994), for example, proposed a test for nonlinearity which tests the null of linearity against the alternative of a smooth transition autoregression (STAR) model.

[6] Hamilton (2001), p. 537. In light of this, it is felt nonparametric approaches such as kernels, wavelets and nearest neighbour are beyond the scope of this review.

hypothesis testing or model simplification required for multivariate modelling. What follows is a brief description of each method.

2.3.1 Time varying parameters

Sims (1993) outlined a model with time-varying parameters. His motivation for such a study stemmed from the work of Litterman (1986), who forecast macroeconomic variables using a small Bayesian vector autoregressive model. Sims updated Litterman's model, allowing for the nonnormality of forecast errors, and more importantly, for time-varying variances and time-varying autoregressive coefficients.[7] His model takes the form

$$\mathbf{X}_i(t) = \sum_{j=1}^{k} \left[\sum_{s=1}^{m} \mathbf{X}_j(t-s)\boldsymbol{\beta}(t;i,j,s) + \boldsymbol{\beta}(t;i,j+1,1) \right] + u(t;i), \tag{2.1}$$

given a time series of k-vectors, $\mathbf{X}_i(t)$, determined by a state vector $\boldsymbol{\beta}(t;i,j,s)$, an equation disturbance $u(t;i)$ and where $i = 1,\ 2, \ldots,$ the number of equations in the vector autoregression. The $\boldsymbol{\beta}$'s and u's are stochastic processes, with distributions that are conditional on the initial values of the **X**s, for the other observed **X**s. The model has substantial time variation in its coefficients, which as previously outlined, and as illustrated by Sims, may be important in modelling aggregate macroeconomic variables.

Time-varying parameter models, in many respects encompass a broad range of approaches, some of which will be considered in the following subsections. As a result, these methods are widely used, for example, Koopman and Ooms (2003), who used them for modelling tax revenues.

2.3.2 Threshold autoregressive processes

The threshold autoregressive (TAR) model was first proposed by Tong (1978). This class of model encompasses features such as limit cycles, amplitude dependent frequencies and jump phenomena.[8] A time series Y_t is a self-exciting TAR process if it follows the model

$$Y_t = \Phi_0^{(j)} + \sum_{i=1}^{p} \Phi_i^{(j)} Y_{t-i} + a_t^{(j)}, \tag{2.2}$$

$$r_{j-1} \leq Y_{t-d} < r_j, \tag{2.3}$$

where $j = 1, \ldots, k$ and d is a positive integer. The thresholds are

$$-\infty \leftarrow r_0 < r_1 < \ldots < r_k \rightarrow \infty. \tag{2.4}$$

[7] Sims (1993), p. 179.

[8] A limit cycle can be defined as an attracting set to which orbits or trajectories converge and upon which trajectories are periodic (Guckenheimer and Holmes, 1997, p. 150-154). Amplitude dependent frequencies relate to asymmetric cyclical behaviour, where the period of the cycle is dependent upon the amplitude of the oscillation, i.e., the period is lower when the amplitude of the oscillation is lower; see, for example, Tong (1990). A jump is defined as a point of discontinuity (Jeffreys and Jeffreys, 1988, p. 26). Jump phenomena encompass, among others, models of regime switching.

For each j, $\left\{a_t^{(j)}\right\}$ is a sequence of martingale differences satisfying[9]

$$E\left(a_t^{(j)}|F_{t-1}\right) = 0, \tag{2.5}$$

$$\sup_t E\left(|a_t^{(j)}|^\delta|F_{t-1}\right) < \infty, \tag{2.6}$$

for some $\delta > 2$, with F_{t-1} the σ-field[10] generated by $\left\{a_{t-i}^{(j)}|i = 1, 2, \ldots; j = 1, \ldots, k\right\}$.

Such a process partitions the one-dimensional Euclidean space into k regimes and follows a linear autoregressive model in each regime. The overall process Y_t is nonlinear when there are at least two regimes with different linear models. As previously stated, the model can contain certain features that cannot be captured by a linear time series, such as limit cycles, amplitude dependent frequencies and jump phenomena. Outline approaches to modelling such TAR processes can be found in Tsay (1989) and Tong and Lim (1980).

TAR models have been used widely in empirical economics. These models have been used by Enders and Granger (1998), to examine the term structure of interest rates, by Caner and Hansen (1998), to explore unemployment, and Pesaran and Potter (1997), to model US output.

2.3.3 Regime switching models

Many variables undergo periods in which the behaviour of the series seems to be quite dramatic. This can be seen by examining any sufficiently long macroeconomic or financial time series. Variables can be considered to go through behavioural phases or regimes. A change in such a regime cannot be seen as the outcome of a perfectly foreseeable, deterministic event. The change itself must be viewed as the outcome of a random process.

It can be said that the outcome of this random process is influenced by the unobserved random variable s_t^*, which is the state or regime that the process was in at time t. Since s_t^* can only take on discrete values, a suitable model must be selected. The simplest time-series model for a discrete-valued random variable is a Markov chain.[11]

Briefly, Markov chains can be defined as follows. Let s_t be a random variable that

---

[9] A stochastic process $\{Y(t),\ t \in \mathcal{T}\}$ is said to be a martingale difference process relative to the increasing sequence of σ-fields $\mathcal{D}_1 \subset \mathcal{D}_2 \subset \cdots \subset \mathcal{D}_t \subset \cdots$, if

- $Y(t)$ is a random variable relative to $\mathcal{D}_t$;
- $E\left(|Y(t)|^2\right) < \infty$; and
- $E\left(Y(t)|\mathcal{D}_{t-1}\right) = 0,\ t \in \mathcal{T}$. (Spanos, 1986, p. 147).

[10] Given the conditions

- $A \in \mathcal{F}$, $\tilde{A} = \Omega - A \in \mathcal{F}$ and
- $A,\ B \in \mathcal{F}$, $A \cup B \in \mathcal{F}$,

a collection $\mathcal{F}$ of subsets of a nonempty set Ω satisfying these conditions is called an algebra or field. Further, a collection $\mathcal{F}$ of subsets of a nonempty set Ω satisfying

- $A \in \mathcal{F}$, $\tilde{A} = \Omega - A \in \mathcal{F}$ and
- $A_j \in \mathcal{F}$, $j = 1, 2, 3, \ldots$, $\cup_{j=1}^{\infty} A_j \in \mathcal{F}$

is defined as a σ-algebra or σ-field. (Bierens, 2004, p. 4).

[11] See Hamilton (1994), Chapter 22, for an excellent treatment of this topic.

can assume only an integer value $1, 2, \ldots, N$. Suppose the probability that s_t equals some particular value j depends on the past only, through to the most recent value, s_{t-1}, such that

$$P\{s_t = j | s_{t-1} = i, s_{t-2} = k, \ldots\} = P\{s_t = j | s_{t-1} = i\} = p_{ij}. \tag{2.7}$$

Such a process is called an N-state Markov chain with transition probabilities $\{p_{ij}\}$, $i, j = 1, 2, \ldots, N$.[12] The transition probability p_{ij} gives the probability that state i will be followed by state j. For example, a simple first-order autoregressive time-series model of a regime switching specification takes the form

$$Y_t = c_{s_t} + \phi_{s_t} Y_{t-1} + \varepsilon_t, \tag{2.8}$$

where ε_t is normally and identically distributed (n.i.d.) with a zero mean and variance σ^2, and where both the constant and autoregressive terms may vary with different states, s_t.

Markov chains are useful for several reasons. A permanent regime change can be modelled by a two-state Markov chain. On the other hand, unusual, short-lived events can also be accounted for. Furthermore, the Markov chain is a flexible tool, as the specification is consistent with a broad range of outcomes. Granger and Teräsvirta (1993) discussed the use of regime switching models and Markov chains and outlined several studies that have employed these methods.[13]

2.3.4 Smooth-transition regressions

Granger, Teräsvirta, and Anderson (1993) proposed a class of models known as smooth transition regressions (STR), which are further discussed by Teräsvirta (1994).[14] Their motivation follows the arguments presented previously. They acknowledge that economic relationships may be nonlinear, but recognise that there 'is no generally accepted class of nonlinear models that can be applied to explore relations' i.e., to employ in exploratory or specification search forms of modelling.[15] This, they believe, is due to the wide variety of alternative nonlinear models available, and a lack of experience in deciding which of these models is most appropriate.

Their class of STR models take the form

$$Y_t = \alpha_1 + \boldsymbol{\beta}_1' \mathbf{X}_t + \phi(Z_t)(\alpha_2 + \boldsymbol{\beta}_2' \mathbf{X}_t) + \varepsilon_t, \tag{2.9}$$

where $0 \leq \phi(Z) \leq 1$ and Z is the 'indicator' variable, a linear combination of the components of $\mathbf{X}_t$. The model is seen to be a smooth transition between the model

$$Y_t = \alpha_1 + \boldsymbol{\beta}_1' \mathbf{X}_t + \varepsilon_t, \tag{2.10}$$

[12] Hamilton (1994), p. 678.

[13] Granger and Teräsvirta (1993), p. 141-145.

[14] Attention here is initially limited to the class of models termed smooth transition regressions (STR). Teräsvirta (1994) refers to the more general smooth transition autoregressive (STAR) models.

[15] Granger, et al. (1993), p. 311.

when $\phi(Z) = 0$, and the alternative linear model

$$Y_t = (\alpha_1 + \alpha_2) + (\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2)'\mathbf{X}_t + \varepsilon_t, \tag{2.11}$$

when $\phi(Z) = 1$.

These two models can be viewed as having very different properties. One could be stationary while the other is nonstationary, for example. These models could be considered to represent two distinct regimes. The models, therefore, 'thus represent a smooth regime-switching situation'.[16] The STR representation can be written as:

$$Y_t = \varphi'\mathbf{Z}_t + \theta'\mathbf{Z}_t G(\gamma, c, s_t) + \varepsilon_t, \tag{2.12}$$

where ε_t is independently and identically distributed, $\varepsilon_t \sim$ i.i.d.$(0, \sigma^2)$, $\mathbf{Z}_t = (\mathbf{w}_t', \mathbf{X}_t')$ and $\mathbf{w}_t = (1, Y_{t-1}, \ldots, Y_{t-p})'$. The transition function $G(\gamma, c, s_t)$ determines the degree of mean reversion and is itself a function of γ, the slope coefficient, c the location parameter and s_t the transition variable. Normally s_t is set to be a lagged value of Y_t. Several choices of specification for the transition function, G, are available. Its form is often taken to be exponential:

$$G(\gamma, c, s_t) = 1 - \exp\left[-\gamma (s_t - c)^2\right], \tag{2.13}$$

and the resultant model is known as the Exponential Smooth Transition Autoregressive (ESTAR) model. This results in a symmetrical transition function. An asymmetric logistic function, and hence the LSTAR model, could also be considered:

$$G(\gamma, c, s_t) = [1 + \exp[-\gamma (s_t - c)]]^{-1}. \tag{2.14}$$

A more general alternative to the ESTAR model is the LSTAR2 model:

$$G(\gamma, c, s_t) = \left[1 + \exp\left[-\gamma \prod_{k=1}^{2} (s_t - c_k)\right]\right]^{-1}. \tag{2.15}$$

The use of the LSTAR2 model overcomes the problem that as $\gamma \to \infty$, Equation (2.12), where $G(\gamma, c, s_t)$ is described by Equation (2.13), 'becomes practically linear, for the transition function equals zero at $s_t = c$ and unity elsewhere'.[17]

As with the other methods considered in this brief review, STR models have been widely used in the empirical literature, not just within the fields of economics and econometrics. For example, Bacon and Watts (1971) used STR techniques to model chemical data; Lütkepohl, et al. (1999) applied them in the context of German money demand.

2.3.5 Smoothing splines

'Spline smoothing is a natural solution to the regression problem when one is given a set of regression functions, but one also wants to hedge against the possibility that the true model

[16] Granger, et al. (1993), p. 313.
[17] Teräsvirta (2004), p. 224.

is not exactly in the span of the given regression functions'.[18] Deviations of the true model from the span of the functions given are easily derived from spline theory.

Following Wahba (1978), consider the equation

$$\mathbf{Y}(t_i) = g(t_i) + \varepsilon_i \qquad i = 1, 2, \dots n, \quad t_i \in \mathcal{T}, \tag{2.16}$$

where $\varepsilon_i \sim$ n.i.d.$(0, \sigma^2 \mathbf{I}_n)$ and $g(\cdot)$ is some 'smooth' function defined on some index set $\mathcal{T}$. The function $g(\cdot)$ can be satisfactorily estimated by cubic polynominal smoothing splines, when $\mathcal{T}$ is an interval of the real line, and given the realisation $\mathbf{y} = (y_1, \dots, y_n)'$ of $\mathbf{Y} = (Y(t_1), \dots, Y(t_n))'$.[19]

Splines can be considered to be an alternative to fitting a specified set of regression functions, when the true nature of $g(\cdot)$ is actually in the range of the specified regression function. It can be shown that spline smoothing is an extension of Gauss-Markov regression, given the specified regression functions. Wahba (1978) claimed 'that spline smoothing is an appropriate solution to the problem arising when one wants to fit a given set of regression functions to the data, but in so doing, also hedge against model errors, that is, against the possibility that the true model g is not exactly in the span of the given set of regression functions'.[20] Wahba showed that spline smoothing may also lead to a measure of deviation of the true function g, from the range of the regression functions. This deviation can be estimated from the data. This measure of deviation can also be viewed as a bandwidth parameter, controlling the smoothness of the estimated function. Using this approach to nonparametric regression, the bandwidth parameter may be estimated from the data. Poirier (1976) discussed smoothing splines and their applications in economics.

2.4 The Hamilton Methodology

2.4.1 Introduction

In an important paper, Hamilton (2001) proposed an approach to nonlinear modelling of economic relationships that provides a single flexible parametric framework for testing for nonlinearity, drawing inference about the form of nonlinearity and assessing the adequacy of the description of nonlinearity provided by specific models. Following Wecker and Ansley (1983), the approach treats functional form as the outcome of a latent stochastic process that is part of the data-generating process; that is, the conditional expectation function associated with a regression model is thought of as being generated randomly prior to the generation of the data. This latent process is modelled using a Gaussian random field concept that generalises Brownian motion to k dimensions, and the parameters of the process are estimated by maximum likelihood.

[18] Wahba (1978), p. 364.

[19] A cubic polynomial smoothing spline is defined as a smoothly joined piecewise polynomial of degree n. For example, if $t_1, t_2, \dots, t_n$ are a set of n values in the interval a, b, such that $a < t_1 \leq t_2 \leq \dots \leq t_n \leq b$, then a cubic spline is a function g such that on each of the intervals $(a, t_1), (t_1, t_2), \dots, (t_n, b)$, g is a cubic polynomial and, secondly, the polynomial pieces fit together at the points t_i in such a way that g itself and its first and second derivatives are continuous at each t_i and hence on the whole of a, b. (Everitt, 2002, p. 356.)

[20] Wahba (1978), p. 364.

The method is a good deal more than an exploratory data or data-smoothing device, although in this regard alone, it may prove very useful. From the practicing economist's viewpoint, its importance lies in the valuable insights it can provide for model construction and the resulting enhancement of the forecasting ability of economic models. However, this modelling framework has not been widely used to date and its full potential remains to be established.[21] As Hamilton (2001, p. 552) pointed out, its usefulness for particular sample sizes and nonlinearities is a matter for empirical investigation. Yet, citing his own three examples and the Monte Carlo studies by Dahl (2002), he suggested that the method holds much promise.

The aims of the remainder of this chapter are modest and the focus intentionally narrow. The main purpose is to address a number of practical issues that arise when using the Hamilton (2001) framework. The first of these concerns computation: issues concerning Hamilton's software, which implements the method, are documented.[22] It appears that the numerical optimisation involved is not an entirely straightforward matter, either when using Hamilton's data set or alternative samples. It should be stressed at the outset that this is a result of difficulties with numerical optimisation and not the Hamilton method itself. Thus while the Hamilton case study is the primary focus, this finding may have more general relevance for procedures that employ similar optimisation techniques. The second issue concerns the sensitivity of the method to changes in data. Experiments suggest that minor data changes can have implications for computation and big effects on the results. Another aim, given the length and difficulty of the original paper, is to provide a concise and reasonably accessible account of Hamilton's methodology for nonspecialist practitioners, though the nature of the subject matter is such that it is not possible to avoid some technicalities.

2.4.2 The model

Consider the nonlinear regression denoted by[23]

$$y_t = \mu(\mathbf{x}_t) + \varepsilon_t, \qquad t = 1, 2, \dots, T, \tag{2.17}$$

where y_t is scalar, $\mathbf{x}_t = [x_{it}]$ is a k-vector of observations on the explanatory variables at time t, ε_t is a stochastic disturbance with zero mean and constant variance, independent of lagged values of $\mathbf{x}_t$ and y_t, and $\mu(\mathbf{x})$ denotes the conditional expectation function $E(y|\mathbf{x})$. The nature of $\mu(\mathbf{x})$ is fundamental to Hamilton's (2001) approach and is considered to be

[21] There are, however, some notable exceptions. Dahl and González-Rivera (2003) further developed and evaluated the test for nonlinearity proposed by Hamilton (2001), which will be discussed in some detail in Chapter 3. Dahl and Hylleberg (2004) investigated the forecasting performance of flexible nonlinear regression models for US unemployment and industrial production. Dahl, González-Rivera, and Qin (2005) studied the performance of nonlinear models when additive random fields are employed. Hamilton (2003) used his proposed framework to model the relationship between oil-price changes and GDP growth. Finally, Kim, Osborn, and Sensier (2005), used this flexible approach to explore nonlinearities in the monetary policy rule of the US Federal Reserve.

[22] The program is written in GAUSS and can be freely downloaded from `http://weber.ucsd.edu/~jhamilto/`.

[23] To facilitate cross-reference to the original paper, the notation used in this section is similar to that used by Hamilton (2001).

determined by

$$\mu(\mathbf{x}) = \alpha_0 + \boldsymbol{\alpha}'\mathbf{x} + \lambda m(\boldsymbol{g} \odot \mathbf{x}), \tag{2.18}$$

where α_0 and λ are scalar parameters, $\boldsymbol{\alpha} = [\alpha_i]$ and $\boldsymbol{g} = [g_i]$ are k-vectors of parameters, $m(\cdot)$ is a realisation of a stochastic process with a continuous path called a random field and $\odot$ denotes the Hadamard product, i.e., element-by-element multiplication. The realisation of $m(\cdot)$, and hence $\mu(\mathbf{x})$, is assumed to be generated by nature, prior to and independently of, all of the observations. Given this fixed $\mu(\mathbf{x})$, the values for ε_t and $\mathbf{x}_t$ are then generated and y_t, is determined according to the Equation (2.17).

The interpretation of the parameters in Equation (2.18) is particularly important for the understanding and application of Hamilton's (2001) method. In particular, the scalars λ and g_i, $i = 1, 2, \ldots, k$, characterise the relationship between $m(\cdot)$ and the conditional expectation function $\mu(x_1, x_2, \ldots, x_k)$. Specifically, λ is a measure of the overall 'weight' of the process $m(\cdot)$ in the conditional expectation, while the magnitudes of the g_i indicate the degree of nonlinearity associated with their respective x_i. Thus $\lambda = 0$ indicates that $m(\cdot)$ makes no contribution and the conditional expectation is linear, in which case Equation (2.17) is the familiar general linear model. Similarly, $g_i = 0$ implies that the conditional expectation is linear in x_i, while $g_i \neq 0$ signifies that it is nonlinear in x_i. If all of the $g_i \to 0$, the contribution of $m(\cdot)$ to the conditional expectation, hence to y_t, becomes indistinguishable from that of α_0; if all of the $g_i \to \infty$, the contribution to y_t is indistinguishable from that of ε_t. The standard interpretation applies to α_i, $i = 0, 1, \ldots, k$.

The key component in Equation (2.18), on which the interpretation of the g_i depends, is the random realisation $m(\cdot)$. Its nature and role require explanation before the practical matters of estimation and testing are considered. First, consider a uniform orthogonal grid in $\mathbb{R}^k$, bounded in the direction of each of the k standard basis vectors or Cartesian co-ordinates by some lower value a_j and some upper value b_j, $j = 1, 2, ..., k$.[24] Let the set of all nodes in the grid be A_N, where $N - 1$ is the number of grid intervals in each direction and N^k is therefore the number of distinct points in A_N. For each point $\mathbf{x} \in A_N$, let $e(\mathbf{x}) \sim N(0, 1)$ and be independent of $e(\mathbf{z})$ for all $\mathbf{x} \neq \mathbf{z}$; let $B_N(\mathbf{x}) = \{\mathbf{z} \in A_N : (\mathbf{x} - \mathbf{z})'(\mathbf{x} - \mathbf{z}) \leq 1\}$, i.e., the set of all points in A_N whose distance from $\mathbf{x}$ is less than or equal to unity; and let $n_N(\mathbf{x})$ denote the number of points in $B_N(\mathbf{x})$. Hamilton (2001, p. 540) then defines the scalar process $m_N(\mathbf{x})$ as[25]

$$m_N(\mathbf{x}) = [n_N(\mathbf{x})]^{-\frac{1}{2}} \sum_{\mathbf{z} \in B_N(\mathbf{x})} e(\mathbf{z}). \tag{2.19}$$

Taking the limit of Equation (2.19) as the grid partition becomes finer, i.e., as $N \to \infty$ and the interval length in each direction of the grid tends to zero, the notion of a continuous scalar-valued k-dimensional random field emerges. The stochastic nature of this is such that

[24] By *uniform* it is meant that the intervals defined by the grid are of equal length in the direction of each of the k co-ordinates, and the number of intervals in each direction is the same. Note that this does not imply that the intervals in different directions have to be the same length unless the a_j are equal and the b_j are equal for all j.

[25] This processes is illustrated for $k = 2, a_1 = a_2 = 0, b_1 = 5, b_2 = 3$, and equal interval lengths in Hamilton (2001, p. 541), so that the number of intervals in each direction is not the same, as required by the definition of A_N given here.

for any $\mathbf{x} \in A_N, m(\mathbf{x}) \sim N(0,1)$. The similarity to standard Brownian motion is apparent.

For arbitrary points $\mathbf{x}$ and $\mathbf{z}$ in $\mathbb{R}^k$, the correlation between $m(\mathbf{x})$ and $m(\mathbf{z})$ is zero if the distance between $\mathbf{x}$ and $\mathbf{z}$ is greater than 2. If this distance is not greater than 2, it can be shown, though the proofs are difficult,[26] that

$$H_k(h) = Cov_k\left(m(\mathbf{x}), m(\mathbf{z})\right) = \frac{G_{k-1}(h,1)}{G_{k-1}(0,1)}, \tag{2.20}$$

where $G_{k-1}(h,1) = -\frac{h}{k}(1-h)^{\frac{k-1}{2}} + \frac{k-1}{k}G_{k-2}(h,1)$, h is one-half the distance between $\mathbf{x}$ and $\mathbf{z}$, $k = 2, 3, \ldots$ and the initial values are $G_0(h,1) = 1-h$ and $G_1(h,1) = \frac{\pi}{4} - \frac{1}{2}h(1-h^2)^{\frac{1}{2}} - \frac{1}{2}\sin(h)$. Equation (2.20) can be calculated recursively, but fortunately its values for $k = 1$ to 5 inclusive are provided in Table I of Hamilton (2001, p. 542) and are included in his program, although the option to calculate recursively is available. It is this covariance that provides the means by which the g_i govern the curvature of $\mu(\mathbf{x})$ in Equation (2.18); see the illustrative case of $k = 1$ in Hamilton (2001, p. 540).

2.4.3 Estimation

Assuming normality of the ε_t, it follows from equations (2.17), (2.18) and (2.20) that

$$\mathbf{y} \sim N\left(\mathbf{X}\boldsymbol{\beta}, \mathbf{C} + \sigma^2\mathbf{I}_T\right), \tag{2.21}$$

where $\mathbf{y}$ is the T-vector of observations on the dependent variable in Equation (2.17), $\mathbf{X}$ is the $T \times (k+1)$ matrix of observations on the k explanatory variables and a column of ones associated with the intercept, $\boldsymbol{\beta} = [\alpha_0, \boldsymbol{\alpha}']'$ is the $(k+1)$-vector of parameters of the linear component of the conditional expectation, $\mathbf{C} = \left[\lambda^2 H_k(h_{ts})\right]$ is a $T \times T$ variance-covariance matrix whose typical element is $\lambda^2 Cov_k\left(m(\boldsymbol{g} \odot \mathbf{x}_t), m(\boldsymbol{g} \odot \mathbf{x}_s)\right)$, and h_{ts} is one-half the distance between $\boldsymbol{g} \odot \mathbf{x}_t$ and $\boldsymbol{g} \odot \mathbf{x}_s$. The likelihood function follows straightforwardly from Equation (2.21) as

$$\ln f(\mathbf{y}; \boldsymbol{\beta}, \boldsymbol{g}, \lambda, \sigma^2) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\mathbf{C} + \sigma^2\mathbf{I}_T\right| - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{C} + \sigma^2\mathbf{I}_T)^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{2.22}$$

Maximum likelihood provides the basis for inference concerning the parameters $\boldsymbol{\beta}, \boldsymbol{g}, \lambda$ and σ^2. Hamilton (2001) showed that the procedure is valid for regressors that are deterministic or lagged values of the dependent variable. However, in the interests of simplifying the calculations, Equation (2.22) is rewritten. Defining $\zeta = \frac{\lambda}{\sigma}$, letting $\boldsymbol{\psi} = \left[\boldsymbol{\beta}', \sigma^2\right]'$ be the $(k+2)$-vector of parameters of the linear part of the model and $\boldsymbol{\theta} = [\boldsymbol{g}', \zeta]'$ be the $(k+1)$-vector of parameters of the nonlinear component, and setting $\mathbf{W}(\mathbf{X}; \boldsymbol{\theta}) = \zeta^2\mathbf{C}^* + \mathbf{I}_T$, where $\mathbf{C}^* = \lambda^{-2}\mathbf{C}$, the right-hand side of Equation (2.22) can be written as

$$-\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2}\ln\left|\mathbf{W}(\mathbf{X}; \boldsymbol{\theta})\right| - \frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{W}(\mathbf{X}; \boldsymbol{\theta})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{2.23}$$

The values of the elements of $\boldsymbol{\psi}$ that maximise Equation (2.23) for given $\boldsymbol{\theta}$ can then be

[26]See Lemma 2.1 and Theorem 2.2 in Hamilton (2001, p. 541). Note also that the details relating to Equation (2.20) are expressed slightly differently than in Hamilton's (2001) lemma and theorem.

calculated analytically as

$$\widetilde{\beta}(\boldsymbol{\theta}) = [\mathbf{X}'\mathbf{W}(\mathbf{X};\boldsymbol{\theta})^{-1}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{W}(\mathbf{X};\boldsymbol{\theta})^{-1}\mathbf{y}, \tag{2.24}$$

and

$$\widetilde{\sigma}^2(\boldsymbol{\theta}) = \frac{1}{T}[\mathbf{y} - \mathbf{X}\widetilde{\beta}(\boldsymbol{\theta})]'\mathbf{W}(\mathbf{X};\boldsymbol{\theta})^{-1}[\mathbf{y} - \mathbf{X}\widetilde{\beta}(\boldsymbol{\theta})]. \tag{2.25}$$

Thus Equation (2.22) may be concentrated as

$$\phi(\boldsymbol{\theta};\mathbf{y},\mathbf{X}) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\widetilde{\sigma}^2(\boldsymbol{\theta}) - \frac{1}{2}\ln|\mathbf{W}(\mathbf{X};\boldsymbol{\theta})| - \frac{T}{2}. \tag{2.26}$$

The numerical maximisation of Equation (2.26), therefore, gives the maximum likelihood estimate of $\boldsymbol{\theta}$, which through equations (2.24) and (2.25) yields the estimate of $\boldsymbol{\psi}$.

2.4.4 Bayesian analysis

This subsection introduces the Bayesian aspects of the Hamilton (2001) methodology. Before doing so, however, a brief review of the Bayesian approach to inference is offered.

Consider the vector of parameters $\boldsymbol{\vartheta}$, estimated for a sample of observations, such that $\boldsymbol{\vartheta} = (\mu, \sigma^2)'$ and, for example, $z_t \sim$ n.i.d. $(\mu.\sigma^2)$ and $\mathbf{Z} = (z_1, z_2, \ldots, z_T)$.[27] The classical statistical approach assumes that a true value of $\boldsymbol{\vartheta}$ exists. This is unknown and invariant, and $\widehat{\boldsymbol{\vartheta}}$ represents its sample estimate, and is a random variable.

The Bayesian viewpoint considers $\boldsymbol{\vartheta}$ to be a random variable, implying a degree of uncertainty about $\boldsymbol{\vartheta}$. This uncertainty can be described in terms of a probability distribution. Information held before the observation of data forms a prior density, $f(\boldsymbol{\vartheta})$, and probability statements regarding $\boldsymbol{\vartheta}$, made prior to observation, can be made in terms of this prior density. The sample likelihood can be defined as $f(\mathbf{Z}|\boldsymbol{\vartheta})$. The joint density of $\mathbf{Z}$ and $\boldsymbol{\vartheta}$ is $f(\mathbf{Z}, \boldsymbol{\vartheta}) = f(\mathbf{Z}|\boldsymbol{\vartheta}) f(\boldsymbol{\vartheta})$. The posterior density $f(\boldsymbol{\vartheta}|\mathbf{Z}) = f(\mathbf{Z}, \boldsymbol{\vartheta}) f(\mathbf{Z})^{-1}$, relates to statements about $\boldsymbol{\vartheta}$ after $\mathbf{Z}$ has been observed.

Returning to the Hamilton (2001) approach, recall that $\boldsymbol{\psi} = [\boldsymbol{\beta}', \sigma^{-2}]'$ and $\boldsymbol{\theta} = [\mathbf{g}', \zeta]'$ are vectors of the linear and nonlinear parameters, respectively. Since the elements of $\boldsymbol{\theta}$ become unidentified as $g_i \to \infty$, the use of a nondiffuse prior is necessary in order for the posterior distribution to be well defined.[28]

[27] This exposition closely follows Hamilton (1994), Chapter 12.

[28] Consider $y_t \sim$ n.i.d. (μ, σ^2), the sample likelihood of which is

$$f(\mathbf{y} \mid \mu;\, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y} - \mu\mathbf{1})'(\mathbf{y} - \mu\mathbf{1})\right\},$$

where $\mathbf{1}$ is a vector of 1s. Prior information about μ is contained within its prior distribution, $\mu \sim N(m,\, \sigma^2/\nu)$,

$$f(\mu :\, \sigma^2) = \frac{1}{(2\pi\sigma^2/\nu)^{1/2}} \exp\left[\frac{-(\mu - m)^2}{2\sigma^2/\nu}\right],$$

where m and ν are parameters that describe the nature and quality of the prior information. As $\nu \to 0$, the quality of prior information becomes ever poorer and the Bayesian estimate of $\bar{y}$, where $\bar{y} = (1/T)\sum_{t=1}^{T} y_t$, approaches its classical counterpart. As $\nu \to \infty$, the prior is known as a diffuse prior, in which case the prior information is disregarded (Hamilton, 1994, Chapter 12).

A prior for each of σ^{-2}, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is required.[29] The prior distribution of σ^{-2} is given by

$$p\left(\sigma^{-2}\right)=\frac{\xi^{\nu}}{\Gamma(\nu)}\sigma^{-2(\nu-1)}\exp\left[-\xi\sigma^{-2}\right], \tag{2.27}$$

where $\nu = 0.25$ and $\xi = (\nu s_y^2)/2$ for s_y^2, the sample variance of y. Similarly, the prior distribution of $\boldsymbol{\beta}$ is defined as

$$p\left(\boldsymbol{\beta}\mid\sigma^{-2}\right)=\frac{1}{\left(2\pi\sigma^{2}\right)^{\frac{k+1}{2}}}|\mathbf{M}|^{-\frac{1}{2}}\exp\left[\frac{-1}{2\sigma^{2}}\left(\boldsymbol{\beta}-\mathbf{m}\right)'\mathbf{M}^{-1}\left(\boldsymbol{\beta}-\mathbf{m}\right)\right], \tag{2.28}$$

where $\mathbf{M} = T(\mathbf{X}'\mathbf{X})^{-1}$, a $(k+1)$-square matrix, and $\mathbf{m}$ is a matrix whose first element is the sample mean of y_t. Finally, the prior for $\boldsymbol{\theta}$ is lognormal and given by

$$p\left(\boldsymbol{\theta}\right)=\prod_{i=1}^{k+1}\frac{1}{\sqrt{2\pi}\theta_{i}}\exp\left[\frac{-\left[\ln\left(\theta_{i}\right)+\ln\left(\sqrt{ks_{i}^{2}}\right)\right]^{2}}{2}\right]. \tag{2.29}$$

Given these prior distributions, the posterior conditional distribution for $\boldsymbol{\beta}$ is

$$f\left(\boldsymbol{\beta}\mid\sigma^{-2},\mathbf{Y}_{T},\boldsymbol{\theta}\right)=\frac{1}{\left(2\pi\sigma^{2}\right)^{\frac{(k+1)}{2}}}|\mathbf{M}^{*}|^{-\frac{1}{2}}\exp\left[\frac{-1}{2\sigma^{2}}\left(\boldsymbol{\beta}-\mathbf{m}^{*}\right)'\mathbf{M}^{*-1}\left(\boldsymbol{\beta}-\mathbf{m}^{*}\right)\right], \tag{2.30}$$

where $\mathbf{Y}_T = \left(y_T,\ \mathbf{x}_T',\ y_{T-1},\ \mathbf{x}_{T-1}',\ldots,y_1,\ \mathbf{x}_1'\right)$, $\mathbf{M}^* = \left(\mathbf{M}^{-1}+\mathbf{X}'\mathbf{W}^{-1}\mathbf{X}\right)^{-1}$ and $\mathbf{m}^* = \mathbf{M}^*\left(\mathbf{M}^{-1}\mathbf{m}+\mathbf{X}'\mathbf{W}^{-1}\mathbf{y}\right)$.. Also, the posterior conditional distribution for σ^{-2} is

$$f\left(\sigma^{-2}\mid\mathbf{Y}_{T},\boldsymbol{\theta}\right)=\frac{\xi^{*\nu^{*}}}{\Gamma\left(\nu^{*}\right)}\sigma^{-2(\nu^{*}-1)}\exp\left[-\xi^{*}\sigma^{-2}\right], \tag{2.31}$$

where $\nu^* = \nu + (T/2)$ and $\xi^* = \xi + \frac{1}{2}(\mathbf{y}-\mathbf{X}\mathbf{m})'[\mathbf{W}(\mathbf{X};\boldsymbol{\theta})+\mathbf{X}\mathbf{M}\mathbf{X}']^{-1}(\mathbf{y}-\mathbf{X}\mathbf{m})$. Finally, it can be shown that the joint distribution of $\mathbf{y}$ and $\boldsymbol{\theta}$ is given by

$$f\left(\boldsymbol{\theta},\mathbf{y}\mid\mathbf{X}\right)=f\left(\mathbf{y}\mid\boldsymbol{\theta},\mathbf{X}\right)\cdot p(\boldsymbol{\theta}), \tag{2.32}$$

where

$$f\left(\mathbf{y}\mid\boldsymbol{\theta},\mathbf{X}\right)=\frac{\Gamma\left(\nu^{*}\right)\xi^{\nu}}{(2\pi)^{T/2}\Gamma(\nu)\xi^{*\nu^{*}}}|\mathbf{W}\left(\mathbf{X};\boldsymbol{\theta}\right)+\mathbf{X}\mathbf{M}\mathbf{X}'|^{-1/2}. \tag{2.33}$$

How then can the conditional expectation function $\widetilde{E}\left[\mu(\mathbf{x})|\mathbf{Y}_T\right]$ be estimated? Consider the random vector ζ whose distribution, conditional on $\boldsymbol{\theta}$ and $\mathbf{Y}_T$, is $f\left(\zeta|\boldsymbol{\theta},\mathbf{Y}_T\right)$. It can be shown that the posterior probability that ζ falls into a region C is as follows. An independently and identically distributed sample of $\boldsymbol{\theta}^{(1)},\ldots,\boldsymbol{\theta}^{(N)}$ is drawn from an arbitrary density function, $I(\boldsymbol{\theta})$. For each $\boldsymbol{\theta}^{(j)}$, generate $\zeta^{(j)}$ from $f\left(\zeta|\boldsymbol{\theta}^{(j)},\mathbf{Y}_T\right)$, and calculate

$$\widetilde{Pr}\left(\zeta\in C\mid\mathbf{Y}_{T}\right)=\frac{\sum_{j=1}^{N}\delta_{[\zeta^{(j)}\in C]}w\left(\boldsymbol{\theta}^{(j)},\mathbf{Y}_{T}\right)}{\sum_{j=1}^{N}w\left(\boldsymbol{\theta}^{(j)},\mathbf{Y}_{T}\right)}, \tag{2.34}$$

[29] A standard prior is used for σ^{-2}. See DeGroot (1970), p. 251.

where C is some region containing ζ, $\delta_{[\zeta^{(j)} \in C]} = 0$ when $\zeta^{(j)} \notin C$ and 1 otherwise, and

$$w\left(\boldsymbol{\theta}^{(j)}, \mathbf{Y}_T\right) = \frac{f\left(\boldsymbol{\theta}^{(j)}, \mathbf{y}|\mathbf{X}\right)}{\mathbf{I}\left(\boldsymbol{\theta}^{(j)}\right)}, \tag{2.35}$$

where $\mathbf{I}\left(\boldsymbol{\theta}^{(j)}\right)$ is an arbitrary importance density. Finally, it can be shown, conditional on $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$, that

$$\mu(\mathbf{x}) \mid \boldsymbol{\psi}, \boldsymbol{\theta}, \mathbf{Y}_T \sim N\left(\xi_T\left(\mathbf{x} \mid \boldsymbol{\psi}, \boldsymbol{\theta}\right), p_T\left(\mathbf{x}, \mathbf{x} \mid \boldsymbol{\psi}, \boldsymbol{\theta}\right)\right), \tag{2.36}$$

where $\xi_T(\mathbf{x}) = E\left[\mu(\mathbf{x}|\mathbf{Y}_T)\right]$ and $p_T(\mathbf{z}, \mathbf{w}) = E\left[\xi_T(\mathbf{z}) - \mu(\mathbf{z})\right]\left[\xi_T(\mathbf{w}) - \mu(\mathbf{w})\right]$. The posterior mean can be calculated, therefore, from

$$\widetilde{E}\left[\mu(\mathbf{x}) \mid \mathbf{Y}_T\right] = \frac{\sum_{j=1}^{N} \xi_T\left(\mathbf{x} \mid \boldsymbol{\psi}^{(j)}, \boldsymbol{\theta}^{(j)}\right) w\left(\boldsymbol{\theta}^{(j)}, \mathbf{Y}_T\right)}{\sum_{j=1}^{N} w\left(\boldsymbol{\theta}^{(j)}, \mathbf{Y}_T\right)}. \tag{2.37}$$

2.4.5 Testing for nonlinearity

The form of the model used in Hamilton's (2001) approach, Equation (2.18), suggests that a simple method of testing for nonlinearity is to test the hypothesis that $\lambda = 0$, or $\lambda^2 = 0$. Hamilton shows that if $\lambda^2 = 0$, and the nonlinear model is estimated, then for fixed $\boldsymbol{g}$, the maximum likelihood estimator $\widetilde{\lambda}^2$ is consistent for the true value of zero and asymptotically normal. Thus a test based on the use of standard normal tables is suggested. However, given the maximum likelihood approach to estimation and the linearity of Equation (2.17) under the null hypothesis that $\lambda^2 = 0$, an obvious and perhaps more appealing way of testing is to use the Lagrange multiplier principle, which requires only a simple linear regression to be estimated. Under the assumption of normality, Hamilton derives the appropriate score vector of first derivatives and the associated information matrix and proposes a form of LM test for practical application. The procedure has four steps. Set $g_i = 2/\sqrt{ks_i^2}$, excluding the constant term whose variance is zero. This g_i is approximately the mean of the lognormal Bayesian prior used by Hamilton as the initial value for θ_i, $i = 1, 2, \ldots, k$.[30] Calculate the $T \times T$ matrix, $\mathbf{H}$, whose typical element is $H_k\left(\frac{1}{2}||\boldsymbol{g} \odot \mathbf{x}_t - \boldsymbol{g} \odot \mathbf{x}_s||\right)$, i.e., the function $H_k(h_{ts})$ defined in Equation (2.20). Use OLS to estimate the standard linear regression $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ and obtain the usual residuals, $\widehat{\boldsymbol{\epsilon}}$, and the standard error of estimate, $\widehat{\sigma} = (T - k - 1)^{-\frac{1}{2}}\sqrt{\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}}}$. Finally, compute the statistic

$$\lambda_H^E(\boldsymbol{g}) = \frac{\left[\widehat{\boldsymbol{\epsilon}}'\mathbf{H}\widehat{\boldsymbol{\epsilon}} - \widehat{\sigma}^2 tr(\mathbf{M}_T\mathbf{H})\right]^2}{\widehat{\sigma}^4\left[2tr\left(\left[\mathbf{M}_T\mathbf{H}\mathbf{M}_T - (T - k - 1)^{-1}\mathbf{M}_T tr(\mathbf{M}_T\mathbf{H})\right]^2\right)\right]}, \tag{2.38}$$

where $\mathbf{M}_T = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is the familiar symmetric idempotent matrix.[31]

As $\lambda_H^E(\boldsymbol{g}) \overset{A}{\sim} \chi_1^2$ under the null hypothesis, linearity ($\lambda^2 = 0$) would be rejected if $\lambda_H^E(\boldsymbol{g})$ exceeded the critical value, $\chi_{1,\alpha}^2$, for the chosen level of significance, α.[32] In the former case the alternative nonlinear specification given by equations (2.17) and (2.18) would be preferred. The identification of a specific form of nonlinearity is greatly aided by the estimate of the

[30] See Subsection 2.4.4.

[31] Hamilton (2001) denoted his LM-type test statistic by ν^2; in this and subsequent chapters, this test is denoted by $\lambda_H^E(\boldsymbol{g})$, following Dahl and González-Rivera (2003).

[32] For example, at the 5 per cent significance level, the null would be rejected if $\lambda_H^E(\boldsymbol{g}) > 3.84$.

conditional expectation $\mu(\mathbf{x})$ and, specifically, the $\widetilde{g}_i$ and $\widetilde{\zeta}$. The matter is explained in Hamilton (2001, Section 5) and to some extent in the following subsection. It is illustrated in the three examples in his Section 7 and at several points in the remainder of this thesis.

2.4.6 Inference and modelling

Thus far, consideration has been given to the theoretical aspects of the Hamilton (2001) methodology. The remainder of this chapter is dedicated to the issues arising from the practical application of the procedure. Hamilton's suggested new approach, a framework for modelling nonlinear relationships, is a three-step procedure that 'proposes a single encompassing framework for nonlinear modelling, offering a new test for nonlinearity, methods to infer what the nonlinear function looks like and checks of whether it is adequately described by some particular model'.[33]

The first stage of the three-step procedure is to test for nonlinearity in the data. Testing for nonlinearity has been described in Subsection 2.4.5. As has been previously noted in Subsection 2.4.2, if the true relationship is linear, then $\boldsymbol{g}$ is unidentified. At this point, the implementation of the procedure, using the software provided, may be a source of some confusion. If the true relationship is indeed linear, the test statistic of the Lagrange Multiplier (LM) test for nonlinearity should, for a given level of significance, be statistically insignificant.[34] Even if this is the case, however, the GAUSS procedure continues to attempt to estimate $\boldsymbol{\theta}$. Experience has shown that in such cases, the iterative process required to estimate $\boldsymbol{\theta}$ often fails. While this may lead to the belief that the program has encountered one of the many implementation issues to be outlined later in this chapter, this is clearly not the case. Should the procedure not fail, the estimates obtained for $\boldsymbol{\theta}$ may not necessarily be statistically insignificant. Without due care and attention, therefore, the applied researcher may infer a nonlinear specification when in fact the data should be modelled with a linear form. The result of the nonlinearity test must, therefore, be evaluated before turning attention to the results of the optimisation.

The estimation of $\boldsymbol{\theta}$ from the concentrated likelihood function in Equation (2.26), discussed in Subsection 2.4.3, provides estimates for $\boldsymbol{g}$ and ζ. These are important elements of the assumed underlying nonlinear relationship, and are the elements which contribute to the nonlinearity. Given the estimation of $\boldsymbol{\theta}$, estimates for $\boldsymbol{\psi}$, whose elements, α_i, describe the linear aspects of the relationship, can easily be found from equations (2.24) and (2.25). As noted in Subsection 2.4.2, the estimates of g_i and ζ describe the relationship between $m(\cdot)$ and the conditional mean function $\mu(\mathbf{x})$. Not only does the GAUSS procedure return estimates for g_i, ζ and α_i, it also returns the estimates' asymptotic standard errors and the square roots of the posterior Bayesian variances, given by $E\left\{[\theta_i - E(\theta_i|\mathbf{Y}_T)]^2|\mathbf{Y}_T\right\}$. The Bayesian analysis was discussed previously in Subsection 2.4.4.

The estimation of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ is necessary for the second stage of the procedure. Those estimates can only assist, however, in inferring a suitable nonlinear specification to model

[33]Hamilton (2001), p. 538. It should be noted that the use of the same data sample for testing and estimation calls into question the underlying assumption of independence.

[34]This conditionality stems from the probability of the test making a Type I error, i.e., the chosen level of significance for the test.

the relationship in question. Consider the case of a model with two explanatory variables, and assume that the LM test for nonlinearity rejects the hypothesis of a linear relationship. Further, assume that at least one of the explanatory variables contributes to the underlying nonlinearity. Having obtained estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$, and where, for example, g_1 is found to be statistically significant and g_2 is not, consider the conditional expectation function with respect to x_1, holding x_2 constant, given by

$$\widetilde{E}\left[\mu\left(x_1, \overline{x}_2\right)|\mathbf{Y}_T\right], \tag{2.39}$$

where $\overline{x}_2$ is the sample mean of the explanatory variable x_2. Plotting this function may allow the researcher to infer the correct nonlinear specification. For more complex examples, a similar procedure is followed. First, the LM test is used to examine whether the relationship between the dependent and independent variables is nonlinear. Second, having estimated $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$, a conventional parametric specification is inferred from the flexible nonlinear inference procedure. The third and final stage of this process is to re-test the inferred specification using the LM test as a misspecification test, to confirm that the correct form has been chosen. Where more than one g_i is found to be statistically significant, plotting the conditional expectation function may not be straightforward. For two significant g_i, Hamilton (2001) used contour plots, although surface plots may be equally useful here. For three significant g_i, the use of the conditional expectation function is abandoned altogether.[35] In some circumstances, it may be straightforward for the researcher to infer a conventional specification from the plot of the conditional expectation, should that plot be sensible. It may be, however, that this process will be far from trivial!

2.4.7 Nonlinear inference and nonstationarity

An understanding of the general characteristics of any methodology is fundamental. Only by understanding the applicability and shortcomings of any technique, may it be appropriately utilised. Hamilton's (2001) proposed framework is indeed flexible; it has been shown that 'the class of nonlinear functions estimated consistently is quite flexible', regardless of whether the explanatory variables are discrete or continuous values.[36] It has also been shown that it is possible to obtain consistent estimators of the conditional mean, even under fairly general misspecification. By using an alternative covariance specification, which will be discussed further in Chapter 3, Dahl and González-Rivera (2003) showed that the problem of unidentified nuisance parameters can be overcome. However, 'given a particular sample size or function, the usefulness of our approach is a matter for empirical investigation'.[37]

While it is the aim of this chapter and indeed this thesis to empirically investigate this technique, one final consideration regarding its applicability is necessary. This thesis is concerned with modelling data which may be (fractionally) integrated and possibly cointegrated, but also data which may contain structural breaks or latent nonlinearity, and possibly both.

[35] For details, see Hamilton's (2001) three examples, p. 559-564.

[36] Hamilton (2001), p. 551. It is assumed that the term 'consistently' is used in the usual sense. That is, an estimator is said to be consistent when the probability limit of a sequence of estimators is equal to the true population parameter (Hamilton, 1994, p. 181).

[37] Hamilton (2001), p. 552.

The issues of structural breaks and latent nonlinearity are not troubling. Exploratory data analysis should highlight the existence of breaks and the practicing economist should not have difficulty detecting them in many cases. Latent nonlinearity is even less of a concern, as the Hamilton (2001) framework is designed to detect it, and studies to date suggest the technique is powerful.[38]

What may be troubling, however, is the performance of the Hamilton methodology where the data are integrated to some degree. Lee, et al. (2005) addressed the issue of spurious nonlinear regression in econometrics. Specifically, they assessed the performance of a variety of nonlinearity tests, including Hamilton's (2001) LM-type test, introduced in Subsection 2.4.5. Lee, et al. used the following two independent random walk variables,

$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_{yt}, \\ x_t &= x_{t-1} + \varepsilon_{xt}, \end{aligned} \tag{2.40}$$

where $\varepsilon_{yt} \sim \text{i.i.d.}(0, \sigma_y^2)$, $\varepsilon_{xt} \sim \text{i.i.d.}(0, \sigma_x^2)$. The tests under examination were employed to test whether y_t and x_t were nonlinearly related for various simulated samples, ranging in size from $T = 50$ to $T = 10,000$. The results present startling evidence of spurious nonlinear regression. Of particular interest here, the Hamilton LM-type test performs poorly. For the smallest sample size of $T = 50$, the test rejects linearity in over 50 per cent of cases against a nominal size of 5 per cent. Worryingly, rather than converging to its χ^2 distribution, the statistic $\lambda_H^E(\boldsymbol{g})$ diverges to infinity as the sample size increases, i.e., its rate of rejection of the null of linearity converges to 100 per cent. They concluded by warning that 'when interpreting nonlinear test results in favour of nonlinearity, applied economists should weigh the evidence against the persistence structure of the time series under investigation'.[39] These results suggest that care should be taken applying the Hamilton methodology to persistent or integrated data. This may lead to incorrect inferences regarding nonlinearity and possibly result in misdirected modelling efforts.

This is a pertinent warning that highlights some of the potential difficulties faced in applied research. Tests for nonlinearity appear to have low power against spurious nonlinearity. It is well known that tests for unit roots perform badly in the presence of nonlinearities and or structural breaks, i.e., unit root tests may incorrectly infer that stationary series with nonlinearities or breaks are nonstationary. Together, these results suggest that attempts to untangle the issues of stationarity, nonlinearity and structural instability are less than straightforward.

2.5 Computational issues

The implementation of Hamilton's (2001) methodology is straightforward, in principle, given the on-line availability of Hamilton's program code. In practice, however, difficulties beyond those outlined in Subsection 2.4.6, may await the unwary. These difficulties relate to the nonlinear optimisation algorithms in the OPTMUM procedure of GAUSS, which is at the heart

[38]See Chapter 3 for further details.

[39]Lee, et al. (2005), p. 306.

of Hamilton's program. Indeed, initial attempts to run the program using GAUSS 5 failed completely and no nonlinear estimates were obtained. It was this experience that motivated the research contained in this chapter.

The methods of nonlinear optimisation are familiar to many econometricians. The particular algorithms available in the GAUSS OPTMUM procedure are some of the most well known. However, to understand fully the issues raised in the following sections, a brief description of numerical optimisation and the relevant algorithms used in GAUSS is provided in the following subsection.[40]

2.5.1 Nonlinear optimisation

The OPTMUM procedure in GAUSS maximises a function, for example Equation (2.26), $\phi(\boldsymbol{\theta}; \mathbf{y}, \mathbf{X})$, by minimising the negative of the function with respect to its vector of parameters, in this case $\boldsymbol{\theta}$. Given the derivatives of this objective function with respect to $\boldsymbol{\theta}$, i.e., the gradient vector, which is numerically computed, and initial values for $\boldsymbol{\theta}$, obtained as described in Subsection 2.4.4, the OPTMUM procedure advances iteratively, computing a direction, $\mathbf{d}$, and a step length, s, at each iteration. The quantity $s\mathbf{d}$ is a vector of values that is added to the current estimate of $\boldsymbol{\theta}$, and therefore has the same dimension as $\boldsymbol{\theta}$, and s is a scalar. Thus, given a value for $\mathbf{d}$ the current estimate, $\widetilde{\boldsymbol{\theta}}$, is updated as

$$\widetilde{\boldsymbol{\theta}}_{+} = \widetilde{\boldsymbol{\theta}} + s\mathbf{d}; \tag{2.41}$$

hence s may be interpreted as changing the rate of descent of the objective function in the given direction. How $\mathbf{d}$ and s are computed will be described in turn, concentrating on the former.

Defining $\mathcal{G}$ to be the $k+1$ gradient vector and $\mathcal{H}$ to be a $(k+1) \times (k+1)$ symmetric matrix, a standard method of calculating $\mathbf{d}$ is from

$$\mathbf{d} = \mathcal{H}^{-1}\mathcal{G}. \tag{2.42}$$

However, as numerical matrix inversion may be challenging, the OPTMUM procedure avoids it by computing $\mathbf{d}$ as the solution of the equation

$$\mathcal{H}\mathbf{d} = \mathcal{G}, \tag{2.43}$$

which is thought to be numerically more reliable. While $\mathcal{G}$ is calculated in a standard manner, $\mathcal{H}$ may be calculated in different ways depending on which algorithm is selected. Several approaches are available in OPTMUM.

The STEEPEST DESCENT algorithm simply sets $\mathcal{H} = \mathbf{I}_{k+1}$, the identity matrix of order $k+1$. While this is computationally undemanding and therefore attractive, the descent may be slow and require many iterations before convergence. The PRCG or Polak and Ribiere (1969) conjugate gradient method is a development of the STEEPEST DESCENT method that

[40]Further details on numerical optimisation may be found in the texts by Brent (1973), Greene (2003), Murray (1972) and the GAUSS reference manual *Optimization*, Aptech Systems, Inc., 2001, especially chapters 2 and 3.

also uses only the gradient but updates the direction as

$$\mathbf{d}_+ = \mathcal{G}_+ + r\mathbf{d}, \text{ where r} = \frac{(\mathcal{G}_+ - \mathcal{G})'\mathcal{G}_+}{\mathcal{G}'\mathcal{G}}. \tag{2.44}$$

There are several more complex methods. The NEWTON algorithm equates $\mathcal{H}$ to the Hessian of the objective function, which may be computed numerically as the gradient of the gradient. Unfortunately, this computation is generally a formidable numerical problem and, as it is required at each iteration, makes the algorithm slow and possibly unreliable. However, when it works smoothly, the NEWTON algorithm may converge in fewer iterations than other methods.

The large computational problems associated with the calculation of the Hessian in the NEWTON method are avoided by certain so-called quasi-Newton algorithms. These start with an initial estimate of the Hessian and employ updates that add information at each iteration without requiring the calculation of second derivatives. Although they generally need more iterations to achieve convergence than the NEWTON method, their numerical efficiency means that they are usually faster and, furthermore, tend to be more robust to the condition of the model and data. The OPTMUM procedure contains three such algorithms: the BFGS method due to Broyden (1967), Fletcher (1970), Goldfarb (1970) and Shanno (1970), the DFP method of Davidon (1968) and Fletcher and Powell (1963) and BFGS-SC, which is a modified BFGS algorithm in which the formula for the computation of the update of the Hessian estimate has been changed to make it scale free. In all three cases, the OPTMUM implementation of the algorithm uses the Cholesky factorisation of the approximation to the Hessian in Equation (2.43), i.e., $\mathcal{H} = \mathcal{C}'\mathcal{C}$, before solution for $\mathbf{d}$. The BFGS algorithm is the default choice in OPTMUM, while the other five are available as options.[41]

The OPTMUM procedure in GAUSS 5 also includes a number of methods for computing the step length, s. The default method is called STEPBT, which is described in Dennis and Schnabel (1983). It first attempts to fit a quadratic to the objective function and computes an estimate of s that minimises the quadratic. If that fails, it tries a cubic function, which is rather more versatile in cases where the objective function is not well approximated by a quadratic.

If STEPBT fails, then BRENT is used, a technique due to Brent (1972) that evaluates the objective function at a sequence of test values for s, determined by extrapolation and interpolation using the inverse of the 'golden ratio', namely, the constant $\frac{(\sqrt{5}-1)}{2} = 0.61803$. This method is generally more efficient than STEPBT but requires significantly more function evaluations.

If, in turn, BRENT fails, then a procedure called HALF is used. Denoting the objective function by $F\left(\widetilde{\boldsymbol{\theta}} + s\mathbf{d}\right)$, this method first sets $s = 1$. If $F\left(\widetilde{\boldsymbol{\theta}} + s\mathbf{d}\right) < F\left(\widetilde{\boldsymbol{\theta}}\right)$, then s is set to 1; if not, then $s = 0.5$ and $F\left(\widetilde{\boldsymbol{\theta}} + s\mathbf{d}\right)$ is tried. The attempted step length is halved each time the objective function fails to decrease. When the function does decrease, s is set to

[41] It should be noted that all of the methods discussed here are locally but not globally convergent. While not considered here, a range of globally convergent optimisation techniques is available. These include smoothing (homotopy) methods, response surface techniques, simulated annealing and genetic algorithms. Further details can be found in Horst and Pardalos (1995), Pinter (1996) and Neumaier (2004). For econometric applications see, for example, Maddala and Nelson (1974), Goffe (1996), Jerrell and Campione (2001), and Tucci (2002).

its current value. This method usually requires the fewest function evaluations but is most likely to fail to find the step length that decreases the objective function.

Finally, if HALF fails, a final search for a random direction that decreases the objective function is implemented. The radius of the random search is fixed via an important global variable in OPTMUM called `_oprteps`, the default value of which is 0.01. It is, however, possible to specify any positive value for `_oprteps`.

2.5.2 Computations and results

The computations and results in this subsection and the following section relate to Hamilton's (2001) Example 3 concerning the post-war US Phillips curve. No results for the test statistic in Equation (2.38), the LM-type test for nonlinearity, are given here as they derive from a simple ordinary least squares regression, which is unproblematical. However, they were checked for all of the cases considered and, without exception, the null of linearity was rejected.

Noting that an OLS regression of inflation (π_t) on unemployment (u_t), lagged inflation (π_{t-1}) and a time trend (t) reveals statistically insignificant evidence of an inflation-unemployment trade-off using annual data for the period $t = 1949$ to 1997, Hamilton (2001, Section 7) investigated the use of a nonlinear relation like that defined in equations (2.17) and (2.18), of the specific form

$$\pi_t = \mu\left(u_t, \pi_{t-1}, t\right) + \varepsilon_t. \tag{2.45}$$

As previously noted, however, it was initially not possible to replicate Hamilton's results using GAUSS 5 and Hamilton's original data: the numerical optimisation associated with the maximisation of Equation (2.26) failed. The original data is included here, for reference, as Table A.1.[42]

Alternative algorithms and step lengths

Examination of Hamilton's (2001) GAUSS program revealed that it employs the BFGS algorithm and relies on the default value of `_oprteps`. It had also been implemented by Hamilton using an earlier version of GAUSS. Indeed, it was possible to reproduce his results using GAUSS 3 without any modification of the program, although for his reported value of $\widetilde{g}_2$ of 0.16, a value of -0.16 was recorded, when rounded to two decimal places, like all of his results.[43] However, when using GAUSS 3 with different algorithms, by suitable and straightforward adjustments of the program code, the results, when they were produced using Hamilton's data, were not always similar to those reported by Hamilton. Table A.4 shows the results of all six optimisation procedures available in GAUSS. In this table, algorithms 1, 2, 3, 4, 5 and 6 refer to the STEEPEST DESCENT, BFGS, BFGS-SC, DFP, NEWTON and PRCG methods, respectively; and the g_i and α_i refer to the parameters in the nonlinear and the linear components of the conditional expectation function, respectively. The values of $i = 1, 2, 3$ relate to u_t, π_{t-1} and t, respectively, while α_0 is the constant.

[42] All tables and figures referred to in this and subsequent chapters can be found in the relevant appendices.

[43] This difference in sign is of no consequence for the nonlinear inference, as both positive and negative values imply the identical value of g_i^2, and therefore of the likelihood function.

The results for BFGS (algorithm 2) in Table A.4, obtained using the program in unmodified form, are those corresponding to Hamilton's (2001) published results and are included as a convenient reference. Apart from the one difference in sign for $\widetilde{g}_2$, they are identical to his. However, it was found that BFGS-SC (algorithm 3) and PRCG (algorithm 6) fail for Hamilton's dataset, STEEPEST DESCENT (algorithm 1) produces noticeably different numerical results from Hamilton's, NEWTON (algorithm 5) produces very similar results except for the sign on the nonlinear parameter estimate $\widetilde{g}_2$, and DFP (algorithm 4) replicates the results of BFGS, the Hamilton case. Despite the big numerical differences in the results produced using algorithm 1, the high statistical significance of $\widetilde{g}_3$ remains and the inference concerning nonlinearity would be basically the same as that drawn by Hamilton.

Following an amount of experimentation involving straightforward modification of the program code to utilise different values of `_oprteps`, Hamilton's (2001) results were eventually reproduced using GAUSS 5. It should be noted that not all values of `_oprteps` proved successful and led to results. Table A.5 contains the results for what was the most successful value for this parameter, namely, `_oprteps` = 0.00001, while Table A.6 gives results for certain other `_oprteps` values, namely, 0.001, 0.1 and 1.0. As can be seen from Table A.5, the results from GAUSS 5, algorithm 2, are identical to those of Hamilton produced by the same algorithm in GAUSS 3. The results from GAUSS 5, algorithms 1 and 5 are similar in absolute terms to those given by algorithm 2, but there are some sign changes on $\widetilde{g}_2$ and $\widetilde{g}_3$. In contrast to what was found using GAUSS 3, there are surprisingly large numerical differences between the results from algorithms 4 and 6 and those from algorithm 2 when using GAUSS 5. Despite these various changes across some algorithms and the two versions of GAUSS, $\widetilde{g}_3$ remains the most statistically significant of the nonlinear parameter estimates, though $\widetilde{g}_2$ is marginally significant for most of the algorithms and `_oprteps` values. Algorithm 3 failed in all experiments due to a problem with the Cholesky decomposition,[44] and it was concluded that this may be due to a program error in the GAUSS software, which remains to be investigated. The numbers of iterations used by the alternative algorithms are, in relative terms, broadly in line with what was said about relative efficiencies in Subsection 2.5.1.

Initial parameter estimates

It is well known that the initial estimates of parameters are often critical for convergence of numerical optimisation procedures. The initial estimates for the g_i parameters in Hamilton's (2001) program, as have been used previously in testing for nonlinearity, are $g_i = \frac{2}{\sqrt{ks_i^2}}$, where s_i^2 is the sample variance of explanatory variable x_i.[45] While there is no information in Hamilton's paper concerning the initial value of ζ, examination of his program revealed the start value $\zeta = 0.5$. It was decided to investigate the effect of changes in the initial value of ζ, using the STEEPEST DECENT algorithm, which, according to Schoenberg (2001, p. 14), is the least affected of the algorithms by choice of starting values, and which was the only algorithm not to have failed in our earlier experiments.

Using `_oprteps` = 0.00001 and a range of start values for ζ from 0.1 to 1.5, inclusive, in

[44] The GAUSS diagnostic message produced was *Cholesky downdate failed.*

[45] See Subsection 2.4.5.

steps of 0.1, it was found that failures occurred for the lower values of 0.1, 0.2 and 0.3, but that all other values of ζ produced results. In all successful cases, the number of iterations of algorithm 1 was 150, the default maximum in the code. There was considerable variability in the final estimates of the g_i parameters and their standard errors, although $\widetilde{g}_3$ was always statistically significant. The results for these cases are given in Table A.7.

Algorithm switching

In order to capitalise on the characteristics of the various optimisation algorithms, and thereby economise on the number of iterations, and to examine whether the variability in parameter estimates noted in the previous subsection could be reduced, the investigation of the effects of different start values was extended to consider the use of an algorithm switching procedure involving two different methods, as described by Schoenberg (2001). This procedure begins with algorithm 1, STEEPEST DESCENT, which is robust to initial conditions, i.e., initial parameter values and a step length of unity.[46] If the function fails to improve by the default amount of 0.01, then a switch to algorithm 2, BFGS with step length BRENT takes place.[47] The use of algorithm 5 (NEWTON) was also explored as the second of the algorithms.[48]

Again, the procedure fails for initial values of ζ of 0.1, 0.2 and 0.3, but produces estimates for all other start values of ζ examined. Considerably fewer iterations are involved in obtaining the final estimates. A maximum of 29 iterations for initial $\zeta = 0.6$, 0.7 and 0.8, and a minimum of 11 iterations for initial $\zeta = 1.1$ were required. Fewer iterations are required in all cases but one when the second algorithm is chosen to be NEWTON. Greatly reduced variability is observed in the final estimates and their standard errors across all cases, and these estimates are very close to Hamilton's (2001) original results as given in Table A.5. Indeed, to three decimal places, the absolute values of the estimates and their standard errors are identical, except for the case of $\widetilde{\alpha}_0$, where the differences are none the less very small. Details of the results for the two versions of the algorithm switching experiment for the odd start values of ζ are presented in tables A.8 and A.9. Results for the even start values, which are similar to those contained in the tables, were omitted for brevity.

2.6 Sensitivity to Data

This section reports on the performance of the program, the GAUSS algorithms, and the results produced, when various small changes are made to the dataset used in Hamilton's (2001) Example 3. Three types of change were considered. The first deleted observations at the start

[46] Although in this case, both the algorithm and step-length method switch, OPTMUM can switch just the algorithm, allowing the step-length to be chosen in the usual way (see Subsection 2.5.1).

[47] This switching is controlled by the value assigned to the global variable `_opdfct`. If the function fails to improve by the percentage `_opdfct`, OPTMUM switches to the secondary search method. Its default = 0.01. This methodology can pair any two algorithms together, although algorithms 1, 2 and 5 are usually recommended, given their respective characteristics.

[48] The advantages of such an approach are apparent. The OPTMUM optimisation process can be started without strong demands on either the condition or starting points of the model, given the characteristics of the STEEPEST DESCENT method, and can then switch to more efficient methods, in this case BFGS or NEWTON, that are more demanding, after the function is closer to the minimum. This switching algorithm technique offers the potential to overcome various problems encountered en route to the optimisation of $\boldsymbol{\theta}$.

of the dataset and the second deleted observations at the end, thus giving successively smaller samples. The third added new observations to create successively larger, somewhat updated samples. The additional observations on the US unemployment data and the US consumer price index were obtained from the US Bureau of Labor Statistics.[49] Checks confirmed that the observations for the period 1949 to 1997, also available from this website, were identical to those in Hamilton's dataset. In line with Hamilton's treatment, no re-scaling of the data was undertaken.

In all, ten alternative samples were created. Hamilton's (2001) original dataset is designated as dataset 1. Deleting the first observation ($t = 1949$) from Hamilton's data gives dataset 2; deleting the first and second observations gives dataset 3; deleting the first, second and third observations gives dataset 4. Similarly, deleting the last observation ($t = 1997$) gives dataset 5; deleting the last two observations gives dataset 6; deleting the last three observations gives dataset 7. Finally, adding the observation for 1998 gives dataset 8; adding the two observations for 1998 and 1999 gives dataset 9; adding the three observations for 1998, 1999 and 2000 gives dataset 10; and adding the four observations for 1998 to 2001, inclusive, gives dataset 11.

For each of the 10 alternative samples, Hamilton's (2001) program was implemented using GAUSS 5 and the values for `_oprteps` of 0.00001, 0.001, 0.1 and 1.0, which were referred to in Subsection 2.5.2. The GAUSS 3 implementation was also used with datasets 2, 5, 7, 8 and 11. A large volume of results was therefore produced and the relevant details are tabulated in tables A.10 to A.25. Table A.10 summarises the nonlinear estimates given by the GAUSS 3 implementation using dataset 2. These results are typical; three of the six nonlinear optimisation algorithms fail, and of those that did not fail, there are considerable differences in the results produced. The results from GAUSS 3 for datasets 5, 7, 8 and 11 are contained in tables A.11, A.12, A.13 and A.14, respectively. The level of program failure in each case is at least as great as that observed using dataset 2, and the variation in the nonlinear estimates remain considerable in some cases.

As shown by Table A.10, algorithm 2, which is Hamilton's (2001) default method, as well as algorithms 3 and 4, fail in GAUSS 3. The apparent reason is that after one or several iterations, the algorithm encountered a nonpositive definite matrix.[50] Of the methods that worked, algorithm 1 and 6 give similar results but algorithm 5 gives very different results from these, including different signs for all of the $\widetilde{g}_i$ coefficients.[51] These differences are noteworthy, as is the fact that algorithm 2 fails for all of the modified data sets examined, as can be seen in tables A.11 to A.14.

The results of the GAUSS 5 implementation using datasets 2 to 11, inclusive, are contained in tables A.16 to A.25, respectively. The information from these tables on the success and failure of the algorithms is summarised in Table A.15, for convenience. This Table includes similar information for Hamilton's (2001) data (dataset 1), for comparison. From a total of 264 program runs, 102 or 39 per cent failed to produce nonlinear estimates. At the

[49]`http://www.bls.gov/`.

[50]One or other of two GAUSS diagnostic messages were obtained in this event. The first was *Negative of Hessian is not positive definite*; the second was *Matrix not positive definite.*

[51]Although as previously noted, the sign on the g_i parameters is unimportant, as it is g_i^2 that is relevant.

extremes, algorithm 3 (BFGS-SC) failed in all cases, in line with previous findings, while algorithm 1 (STEEPEST DESCENT) was successful in all cases. These latter cases all required the maximum number of iterations permitted by the program.[52] Algorithm 5 (NEWTON) was the most efficient, converging after the least number of iterations, but it failed in 10 out of 44 runs, i.e., in 23 per cent of cases. Worryingly, algorithm 2 (BFGS), the default in Hamilton's program, failed in 28 out of 44 runs or 64 per cent of cases.

As found in the case of GAUSS 3, there are many differences in the nonlinear estimates obtained from a given dataset when different algorithms converge using GAUSS 5, including some sign changes. Furthermore, there are also some big changes in numerical parameter estimates, again including some sign changes, when marginal changes in the original dataset, such as the addition or deletion of just one observation, are made. Of particular interest is the case of the results obtained using dataset 11 and algorithm 2, where the addition of four extra observations leads to radically different conclusions concerning nonlinearity than those in Hamilton's example. In this case, all three $\widetilde{g}_i$ are found to be significant, in comparison to just $\widetilde{g}_3$ using the original data with this algorithm. The implications for inference regarding the data in this case are obvious, implying that π_t is now nonlinear in not just t, as before, but in u_t and π_{t-1} also. This appears to be an exception, however, as the relatively high statistical significance of $\widetilde{g}_3$ is generally maintained across the range of experiments that have been conducted.

Finally, the algorithm switching procedures, outlined previously, were also applied to datasets 1 to 11, inclusive. Interestingly, while there were a few cases of failure, this procedure appears to offer increased efficiency and somewhat improved numerical stability in the final numerical parameter estimates. These results can be found in tables A.26 and A.27. Note that the algorithm switching procedure has been shown in Subsection 2.5.2 to be somewhat robust to initial parameter estimates. Given the increased efficiency and stability noted here, these methods seem to offer the best approach to numerical optimisation in this context.

2.7 Further Results: Hamilton's Example 1 and 2

The computations and results reviewed thus far have related solely to Hamilton's (2001) Example 3, concerning the US Phillips curve. It is believed, however, that the issues outlined in Subsection 2.5.2 may be common to any implementation of the Hamilton approach. To examine if this was indeed the case, Hamilton's Examples 1 and 2 were reconsidered in light of the issues raised on the implementation of the methodology, namely, alternative algorithms, initial parameter estimates, algorithm switching and sensitivity to data.

Hamilton's (2001) Example 1 and Example 2 rely on randomly generated data. Example 1 makes use of a simple threshold model, defined by

$$y_t = 0.6x_{1t}1_{[x_{1t}>0]} + 0.2x_{2t} + \varepsilon_t, \tag{2.46}$$

[52]The number of permissible iterations is controlled by the GAUSS parameter `_opmiter`, which defaults to 150 in Hamilton's (2001) program. It was found through suitable experimentation that increasing the maximum number of iterations to 250, for those algorithms that reached the original maximum of 150, did not alter the results obtained to three places of decimals.

where $x_{it} \sim$ n.i.d.$(0, 100)$, $1_{[x_{it}>0]} = 0$ when $x_{it} < 0$ and 1 otherwise, $\varepsilon_t \sim$ n.i.d.$(0, 1)$ $i = 1, 2$, $t = 1, \ldots, T$, and the number of observations, $T = 100$. For Example 2, a more elaborate specification is used, namely,

$$y_t = 5 + 2x_{1t}x_{2t}1_{[x_{1t}>0]}1_{[x_{2t}>0]} + 0.7x_{3t} + \varepsilon_t, \tag{2.47}$$

where $x_{it} \sim$ n.i.d.$(0, 4)$, $\varepsilon_t \sim$ n.i.d.$(0, 1)$ $i = 1, 2, 3$, $t = 1, \ldots, T$, and again, $T = 100$. In both cases, the null of linearity is correctly rejected by Hamilton's LM test, with test statistics of $\lambda_H^E(\boldsymbol{g}) = 232.93$ and $\lambda_H^E(\boldsymbol{g}) = 64.39$, respectively, far in excess of the 5 per cent critical value of $\lambda_H^E(\boldsymbol{g}) = 3.84$. In Example 1, $\widetilde{g}_1$ is found to be significant, and in Example 2, $\widetilde{g}_1$ and $\widetilde{g}_2$ are found to contribute to the nonlinearity. It should be made clear at this point that as the x_{it} were randomly generated for both examples, the results presented for the remainder of this section relate to a fixed sample of x_{it} for each example, generated exactly as in Hamilton. In this way, the results presented here are directly comparable. For this reason, the results for Hamilton's Example 1 and Example 2, found using algorithm 2, GAUSS 3 and shown in tables A.28 and A.29, differ slightly from those reported in his paper. These differences stem solely from the random generation of the observations on the x_{it}. The particular samples of data used here can be found in tables A.2 and A.3.

Tables A.28 and A.29 present the results for the fixed sample of data using GAUSS 3 and the six available algorithms. As with previous findings, the various algorithms return some small numerical differences in the parameter estimates and some sign changes. As before, the algorithm BFGS-SC fails completely for both examples. The high level of significance for $\widetilde{g}_1$ and therefore, the inference concerning nonlinearity remains the same however, regardless of which of the five functioning algorithms were chosen. Tables A.30 and A.31 show the results for the various algorithms in GAUSS 5. A similar pattern emerges here; despite using a range of values for `_oprteps` and the six algorithms, the results remain largely unchanged, apart from some small numerical differences and some sign changes. Again, BFGS-SC fails completely. Also, this experimentation gives further confirmation of the relative efficiencies of the algorithms, given the number of iterations required for convergence. These findings confirm those of Subsection 2.5.2, where the choice of algorithm was found not to affect the statistical inference of the methodology.

The effect of altering the initial value of the parameter ζ has almost no impact, again beyond creating small numerical differences and sign changes, unlike the case of Example 3, mentioned in Subsection 2.5.2. These results can be seen for Examples 1 and 2 in tables A.32 and A.33, respectively, where once again, STEEPEST DESCENT, the algorithm most robust to initial conditions is employed, and in each case, converges after the maximum 150 iterations.[53] It should be noted that the program failed to produce estimates in both

[53] In these cases, the optimisation algorithm gave results after 150 iterations, the default maximum in the program. To investigate if the algorithm had actually converged to an optimum or simply stopped at the maximum, the optimisation was repeated with a maximum of 300 iterations. In every case, the optimisation converged before reaching this maximum and interestingly, the estimates were very similar to those produced after 150 iterations. Observing the optimisation iteration-by-iteration in each case confirmed that before 150 iterations were complete, the algorithm had converged to a small neighbourhood. As the iterations progressed, the algorithm simply moved around that neighbourhood. The numerical values obtained from any of these iterations were very close, and so the results for 150 iterations are reported here. These results were confirmed

examples for $\zeta = 0.1$ and 0.2, and in the case of Example 1, the greatest numerical differences in parameter estimates were observed for $\zeta = 0.3$. As was the case with Example 3, where estimates were produced, the inference regarding the nonlinearity is unchanged by the initial values of ζ.

Tables A.34 to A.37 report on the method of algorithm switching, previously introduced, for Example 1 and 2, in the context of varying the initial parameter estimate of ζ. It should be noted that for Example 1 and 2, for $\zeta = 0.5$, the 'default' setting for ζ, the algorithm switching techniques produce very similar parameter estimates to those produced using GAUSS 3 in tables A.28 and A.29, although in all cases, fewer iterations were required. Also, regardless of the algorithm pairing used in the switching procedure, or the initial value of ζ chosen, the parameter estimates are very similar to the original results. Although there was no instability in the results found in Table A.28, these results confirm that where algorithm switching is used, estimates converge to the same results. This confirms the previous findings that the algorithm switching procedure is the most suitable for this optimisation problem.

Finally the issue of data sensitivity was addressed, looking just at Example 1. Six datasets were created for this purpose, with dataset 1 denoting the original fixed sample for Example 1. Deleting the first two observations from x_{it} gives dataset 2; deleting the last two observations gives dataset 3; deleting the first five observations gives dataset 4; deleting the last five observations gives dataset 5; deleting the first ten observations gives dataset 6; deleting the last ten observations gives dataset 7. For the purposes of this analysis, $\zeta = 0.5$ and `_oprteps`= 0.00001. By contrast to the results of Example 3, the results are remarkably similar across all datasets, and are not noticeably different from the original results. These results can be found in tables A.38 to A.43. These results must be viewed with caution, however, and the concerns outlined in considering Example 3, which is based on real data, should be bourne in mind, as the applied researcher is unlikely to be confronted with a dataset and model specification such as that found in Example 1.[54]

2.8 Comments on 'Investigating Nonlinearity'

As noted previously, research carried out in the course of writing this chapter appears as 'Investigating Nonlinearity: A Note on the Estimation of Hamilton's Random Field Regression Model'.[55] In his accompanying reply, 'Comments on 'Investigating Nonlinearity'', Hamilton

by Hamilton (2005).

[54]Although they do not arise in the context of the examples considered here, two further potential difficulties should be noted when using the OPTMUM optimisation procedures. The first relates to the scaling of the data. Several of the optimisation algorithms calculate the Hessian matrix. A scaling problem may arise if this matrix is not balanced, i.e., when the sum of the columns, or rows, are quite unequal. As the elements of the principal diagonal of the Hessian matrix determine to a large extent these sums, OPTMUM may fail to converge if the elements of the diagonal are unequal in magnitude. As a method for scaling the elements of the Hessian may not be apparent, it is often sufficient to ensure that the data used in the model are of about the same magnitude. A second problem relates to conditioning. A matrix may be ill-conditioned if it is poorly scaled or if the model in question is misspecified. If the Hessian matrix is poorly conditioned, its elements may be small. When its inverse is calculated, as may be necessary for a given algorithm, its elements may be very large. The search direction may, therefore, fail to function appropriately in the presence of such large numbers. Model re-specification may be necessary in this case.

[55]Bond, D. M.J. Harrison, and E.J. O'Brien (2005a): "Investigating Nonlinearity: A Note on the Estimation of Hamilton's Random Field Regression Model," *Studies in Nonlinear Dynamics and Econometrics*, 9, Article

(2005) addressed some of the issues raised here. This section briefly discusses these very useful comments.[56]

Hamilton begins by discussing numerical optimisation. While this is discussed in Section 2.5, it is worthwhile bearing in mind the 'inherently bumbling nature'.[57] Numerical optimisation algorithms may only be expected to find local extrema, and under conditions of flat, near-flat or bimodal likelihood surfaces, among others, the algorithms may run into serious difficulties in converging.

Table A.4 is highlighted to illustrate some of the potential difficulties with the optimisation in this context. As Hamilton (2005) pointed out, it is not appropriate to accept the estimates of the optimisation if its convergence has been halted by reaching the maximum permissible number of iterations. He considers the case of algorithms 2 (BFGS) and 4 (DFP). In the later case, it is noted that it takes 172 iterations to reach the very same results as algorithm 2. The estimates contained in Table A.4, however, are only marginally different from those of algorithm 2. They certainly would not be a source of concern. For both algorithms, it has been noted previously that extending the number of permissible iterations to 250 made little difference, as seen in the case of algorithm 4.

Algorithm 1 (STEEPEST DESCENT), however, does not converge to the results found by the other 'successful' algorithms, even when 250 iterations are permitted. Given what is known about the relative strengths and weaknesses of this method, and the potentially difficult likelihood surface, these results are not surprising. In fact, by studying the behaviour of algorithm 1 through its numerical search, it is apparent that it converges to a neighbourhood, before wandering aimlessly around it, failing to converge to a point. Hamilton (2005) acknowledged that by controlling the operation of algorithm 1 more carefully, through the parameter `_oprteps`, great improvements can be made, and convergence can be achieved.

Hamilton (2005) repeated the starting value experiment for algorithm 2, reporting the results in Table 1.[58] He noted that 'this exercise uncovers that this was but one of several local maxima' and not the 'single maximum' suggested previously, and that the 'nonconcave slope of the likelihood surface in the valleys and saddles separating the local hills may account for the difficulties some of the algorithms had in making their way to the top of any hill'.[59] This could explain the observed behaviour of algorithm 1, noted in the previous paragraph.[60]

An issue that has not been raised thus far in this thesis, but will feature in later chapters, is the difficulty concerning estimates of the parameter σ^2 that are close or equal to zero. Recall that $\zeta = \frac{\lambda}{\sigma}$. Hamilton (2005) found this situation analogous to the 'pile-up' phenomenon

2.

[56]It should be noted that Hamilton (2005) is based on earlier versions of the work contained in Chapter 2 of this thesis and Bond, et al. (2005a). Hamilton has not, therefore, had the opportunity to respond to the experiments regarding algorithm switching, for instance. There are also some discrepancies in cross-referencing between the articles, as a result.

[57]Hamilton (2005), p. 1.

[58]Hamilton (2005), p. 4.

[59]Hamilton (2005), p. 3.

[60]The potential for the failure of optimisation procedures given certain likelihood surface characteristics is documented in the literature. For example, useful insights into two possible difficulties are provided by Warnes and Ripley (1987) and Mardia and Watkins (1989). For instance, the likelihood function may be multimodal and the global maximum may not correspond to sensible values of the parameters, or the likelihood function may have a very flat ridge with a number of local maxima (Warnes and Ripley, 1987, p. 640-641).

seen in the time-series literature (see, for example, DeJong and Whiteman, 1993).[61] This matter does not arise with Hamilton's (2001) original data, but with the derivative samples generated. It is also apparent in the applications to be found in later chapters, where it will be given further consideration.

A further issue, again to be raised later in this thesis, concerns the alternative covariance function suggested by Dahl and González-Rivera (2003), which replaces Hamilton's (2001) specification

$$H_k(h) = Cov_k\left(m(\mathbf{x}), m(\mathbf{z})\right) = \frac{G_{k-1}(h,1)}{G_{k-1}(0,1)}, \tag{2.20}$$

with

$$\mathbf{C}_k^*(h^*) = \begin{cases} (1-h^*)^{2k} & \text{if } h^* \leq 1, \\ 0 & \text{if } h^* > 1. \end{cases} \tag{2.48}$$

Given the focus of this chapter, and the vast array of results gathered and presented, this alternative specification was not considered. Its advantages are noted, however, and it is introduced more carefully in Chapter 3. It is utilised routinely thereafter. Hamilton recognises that this specification 'may avoid at least some of the problems identified'.[62]

The final comments concern data sensitivity. No mention is made of the most striking result: that with the addition of four extra observations, inference regarding the form of nonlinearity is radically changed. Rather, the case where just one observation (the first) is deleted, referenced previously as dataset 2, is considered. Hamilton does not 'deny that it shakes one's confidence in the inference that emerges', although the suggestion that this anomaly may be due to that observation's influential nature remains unconvincing.[63] The fact remains that this method, as with all methods, can be sensitive to the data.

2.9 Conclusion

This chapter has given a brief review of nonlinear economic modelling. The motivation for, and several methods of modelling nonlinear economic relationships have been introduced and discussed. Most importantly, a new approach to nonlinear econometric modelling, proposed by Hamilton (2001), has been described.

An account of this new approach has been given, as has a brief description of some of the methods of nonlinear optimisation that may be used in the GAUSS computer program

[61] The pile-up problem results from the fact that sampling distributions of moving average parameters can be observed to pile-up around unity when the true parameter is close to that value. DeJong and Whiteman (1993), citing Ansley and Newbold (1980), Cooper and Thompson (1977), Cryer and Ledolter (1981), Harvey (1981), and Sargan and Bhargava (1983) note that 'sampling distributions of maximum likelihood estimators of dominant moving average parameters "pile-up" at unity when the true parameters are less than unity but lie near the boundary of the invertiblility region'. In the context of the Hamilton (2001) methodology, the random field can be seen as a generalised moving average process. Hamilton (2005) pointed out that for a moving average process with an order of unity, the maximum autocorrelation of the process is 0.5. If the sample autocorrelation of a dataset is greater than 0.5, maximum likelihood estimates of the moving average parameters can pile-up around unity, implying that $\sigma^2 = 0$. In practical terms, the pile-up phenomenon 'clearly undermines the parametric interpretation of any inference. Viewed as a parametric model of the data-generating process, $\sigma = 0$ would imply that there is no error term in Equation (2.17), which is hardly a defensible position' (Hamilton, 2005, p. 5).

[62] Hamilton (2005), p. 5.

[63] Hamilton (2005), p. 7.

provided by Hamilton (2001) for the implementation of his methodology. The performance of this program has been investigated using data relating to Hamilton's three examples, using not only randomly generated data, but also data concerning the US Phillips curve, two versions of the GAUSS software, a range of alternative numerical optimisation options and alternative values for the GAUSS parameter `_oprteps` and model parameter starting values. The performance of algorithm switching procedures has also been examined. Finally, the effects of changes in the sample data on the results produced by Hamilton's procedure have been explored. The focus has designedly been on the GAUSS implementation of the procedure and, while several changes in the GAUSS programs have been investigated, no attempt has been made at modification of Hamilton's methodology *per se*.

The results presented suggest some clear conclusions, which will hopefully be of value to those contemplating working with Hamilton's (2001) method. First, different algorithms used for the numerical optimisation have different chances of success.[64] Hamilton's choice of the BFGS algorithm fails in over 60 per cent of the cases examined in the study of Example 3, while the less computationally efficient STEEPEST DESCENT method succeeds in all cases. Secondly, when different algorithms work, they may produce significantly different numerical results, including different signs for parameter estimates. Thirdly, the use of procedures that employ two algorithms and a switching criterion appears to produce far more consistent estimates than any one algorithm used on its own. Algorithm switching is also less sensitive to the choice of initial parameter estimates. Fourthly, minor changes in data can have significant effects, both in terms of whether an algorithm operates or not and, in the case of it operating, the numerical results it produces. For example, it is interesting to note that if Hamilton's Example 3 data had just one less observation at either end of the sample, or one more observation at the end, his version of the program would have failed to produce nonlinear estimates, not only with GAUSS 3 but also with GAUSS 5 and all of the values of `_oprteps` used in this study. Moreover, if his dataset had contained the four additional observations for 1998 to 2001 (dataset 11), while the program would have produced results, all three nonlinear parameter estimates would have been significant, in contrast to just $\widetilde{g}_3$ as found in his original study (dataset 1). Thus his inferences concerning the form of nonlinearity would also have been different. However, despite the sensitivity of results to choice of algorithm, initial values and data changes, the statistical significance of the nonlinear parameter estimates, hence the inference about the form of nonlinearity, generally seems to be little affected according to the findings that have been reported here. For the simulated datasets used in Examples 1 and 2, the sensitivity of results to choice of algorithm and size of sample is less pronounced than has been found using the real data of Example 3. Given the on-line availability of Hamilton's program and of the data used, it is a straightforward matter to replicate all of these findings.[65] This chapter is only the beginning of the work advocated by Hamilton to

[64] As previously noted, only locally convergent numerical optimisation algorithms have be applied here. The use of globally convergent procedures may offer some relief to the difficulties encountered with numerical optimisation in the context of random field regression, as documented in this chapter and Hamilton (2005). The use of such methods represents and interesting and potentially fruitful agenda for future research.

[65] In addition to Hamilton's (2001) original code, which as previously mentioned is available from his website, an annotated version of the code and all datasets used in this chapter and in Bond, et al. (2005a) are available from the website of *Studies in Nonlinear Dynamics and Econometrics*: `http://www.bepress.com/snde/`.

establish the usefulness of his new methodology, and empirical investigations continue, in this thesis and beyond.

Hamilton (2005) offered some insight into the difficulties which may have caused the numerical instability of the algorithms revealed in this study. His findings seem to confirm that 'difficult' likelihood surfaces make convergence challenging. Another possibility is that difficulties exist within Hamilton's (2001) methodology, and that the use of the alternative covariance functions for specifying the random field regression model may help to reduce numerical instability,[66] as may different procedures for estimating the parameters of Hamilton's random field regression model, such as the recently proposed two-step method of Dahl and Qin (2004), which promises unique identification and good convergence properties. The latter constitutes an interesting agenda for future research, while the former is introduced more formally in Chapter 3, and routinely taken into consideration thereafter. Finally, Hamilton (2005) introduced the phenomenon known as 'pile-up'. While issues relating to this have not been discussed in this chapter, they are of considerable importance, and will be raised again in Chapter 5.

With regard to the implementation of the Hamilton (2001) random field regression methodology, which this chapter addresses, the main recommendation is to employ an algorithm switching approach supplemented as necessary by changes of the GAUSS `_oprteps` parameter, the starting value of ζ, and using both the Hamilton and the Dahl and González-Rivera (2003) covariance specifications. On the basis of the evidence provided here and supplemented in later chapters, such an approach is more efficient than the use of single algorithms and appears to be less susceptible to numerical instability and failure.

[66]See Dahl and González-Rivera (2003).

Chapter 3

Testing for Nonlinearity: A Note on the Power of Random Field LM-Type Approaches

Research carried out in the course of writing this chapter is forthcoming as 'Testing for Nonlinearity: A Note on the Power of Random Field LM-Type Approaches' by M.J. Harrison and E.J. O'Brien, Trinity Economic Papers.

3.1 Introduction

In light of the discussion in Chapter 2 on the importance of nonlinearity in economics and econometrics, the focus here switches to testing for nonlinearity. In most practical applications, testing for linearity/nonlinearity serves two purposes. The first is to establish the presence of nonlinearity, i.e., to decide if a particular dataset is best modelled with a nonlinear specification. If such testing fails to reject linearity, subject to the power of the test(s) employed, the pursuit of a nonlinear model would be futile. Secondly, post-estimation testing can determine whether or not the chosen nonlinear model adequately represents the data generating process. In both cases, such testing can be an integral part of model building.

This chapter investigates the properties of several tests for neglected nonlinearity in time series, using Monte Carlo simulation methods.[1] This study is motivated by the LM-type test proposed by Hamilton (2001), introduced in Chapter 2. The comparative properties of this test, as an integral part of Hamilton's framework, and indeed as a stand-alone test for nonlinearity, have yet to be fully understood. The goal of this chapter is to further explore them.

The properties of this test have been studied to some degree, but remain to be fully established. Dahl (2002), for example, found that the Hamilton (2001) test was powerful by comparison to existing tests based on spline smoothing and neural networks. The aim of this chapter is to compare the Hamilton test with some well-known alternative tests, by applying them to some model specifications commonly encountered in empirical research. Not only are these properties of considerable importance for the overall Hamilton methodology, having implications for the applicability of this 'complete tool' for nonlinear modelling, but they are also important as they assess the worth of this LM-type test as a stand alone test for neglected nonlinearity in time series. Chapter 2 has highlighted the test, which is likely to be little known, due to the technical nature of the original paper detailing the methodology, and also because the test is embedded within that methodology. Many applied researchers may not be familiar with the test for these reasons.

Here, the Hamilton (2001) LM-type test will be evaluated across a range of model specifications, parameters and data, and will be compared with several other tests, namely, the Durbin-Watson (1950) bounds test, Ramsey's (1969) RESET test, the Harvey-Collier (1977) ψ-test and three tests put forward by Dahl and González-Rivera (2003). The Durbin-Watson and RESET tests are widely known, although more so as tests for autocorrelation and functional misspecification, respectively. These tests, however, may also be effective in testing for nonlinearity, as autocorrelation and misspecification may indeed result from latent nonlinearity. The Harvey-Collier test, based on recursive residuals, may be less well known and certainly seems little used.

The remaining three tests, Dahl and González-Rivera's (2003) $\lambda^E_{OP}(\boldsymbol{g})$, λ^A_{OP} and g_{OP} tests, are derivatives of Hamilton's (2001) LM-type test, denoted $\lambda^E_H(\boldsymbol{g})$.[2] Each will be introduced

[1] For an excellent introduction to Monte Carlo simulation in the social sciences, see Mooney (1997).

[2] The notation used here is that of Dahl and González-Rivera (2003) and follows that used in Chapter 2. The ^{E} signifies that full knowledge of the parametric nature of the covariance function is assumed. The alternative is ^{A}, signifying that no assumption about the covariance function is made. The subscript $_H$ signifies that the Hessian of the information matrix is used. The alternative is $_{OP}$ signifying that the outer product of

in Section 3.2. These tests exploit the features of the Hamilton framework and may offer a superior performance.

This chapter proceeds as follows. Section 3.2 introduces the various tests in some detail, particularly where those tests may be unfamiliar. Section 3.3 outlines the design of the Monte Carlo experiment, including details of the model specifications, parameters and data employed. Section 3.4 provides the necessary theoretical background. Section 3.5 presents the results and offers a discussion of these findings and Section 3.6 concludes.

3.2 The Tests

This section will introduce the various tests that will be investigated in this study. Some will be very familiar to the practicing economist: specifically, Ramsey's (1969) RESET test and Durbin and Watson's (1950) bounds test. The Durbin-Watson procedure is designed to detect serial correlation in the disturbances of a regression model, based on the residuals of that model. Any serial correlation detected in the residuals, however, may result from serially correlated disturbances or 'they may also be symptomatic of some other type of misspecification. In particular, it should be clear ... that positive serial correlation can be expected when the functional form is inappropriate'.[3] The remaining tests are likely to be unfamiliar and so will be introduced here. The intention here is to just outline the nature of each test. This section proceeds by looking at the more familiar tests of Durbin and Watson, RESET and Harvey and Collier (1977), before reviewing the newer tests: Hamilton's $\lambda_H^E(\boldsymbol{g})$ test, introduced in Chapter 2, and the tests derived from this by Dahl and González-Rivera (2003), namely, the $\lambda_{OP}^E(\boldsymbol{g})$, λ_{OP}^A and g_{OP} tests.

There are many more tests that could have been considered here.[4] The motivation for including those chosen, however, is straightforward. RESET, the d-statistic and the Harvey-Collier (1977) approach are all available in commonly used econometric software packages. With the exception of the Harvey-Collier test, these procedures are well known and routinely reported, making them attractive to the practicing economist. It seems appropriate, therefore, to include them here. The remaining tests, the $\lambda_H^E(\boldsymbol{g})$ test and its derivatives, are themselves the motivation for this chapter, the aim of which is to empirically evaluate them. Comparing the performance of each of these tests will certainly provide results that will be of interest to the practitioner.[5]

---

the score function is used.

[3]Harvey (1990), p. 157.

[4]Other tests of note include that of Harrison and Keogh (1985), who proposed a test based on disturbance estimators; McLeod and Li (1983), who applied the Ljung-Box statistic to the squared residuals of an ARMA model; the Bispectral test, which exploits the result that a properly normalised bispectrum of a linear time series is constant and zero under normality, over all frequencies (Priestley, 1988); the Brock, Dechert, and Scheinkman (1987) test, based on the notion of the correlation integral, which is known to have power against a range of nonlinear stochastic process, despite being a test of independence (Granger and Teräsvirta, 1993, p. 36); Keenan's (1985) adaptation of the RESET test; Tsay's (1986) F-test, also based on RESET; and White's (1989) neural network test of nonlinearity.

[5]Many others have compared the relative performance of tests of nonlinearity. Lee, White, and Granger (1993), for example, examined the performance of several of the tests mentioned here. More recently, Dahl and González-Rivera (2003) and Lee, Kim, and Newbold (2005) have undertaken similar studies, more detailed discussions of which will follow in later sections.

3.2.1 The Durbin-Watson bounds test

The Durbin-Watson (1950) bounds test does not test directly a hypothesis of linearity against an alternative of nonlinearity, unlike the other tests considered in this study. The test can, however, be used in this regard, for reasons previously outlined. This study not only reveals the potential for testing in the presence of nonlinearity, it should also reveal something about the behaviour of the test across a wide range of specifications. The test is based on OLS residuals, which can be very useful in detecting misspecification that results from a latent nonlinear functional form or otherwise. Variation in observations not captured by some linear model will be captured by the residuals. Those residuals may then depart from their expected random distribution. The d-statistic exploits this characteristic of the residuals, although it is better known as a test for first-order autoregression in the disturbances. Given the model

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t, \tag{3.1}$$

where $\varepsilon_t \sim \text{n.i.d.}(0, \sigma^2)$ and $t = 1, \ldots, T$, the d-statistic is defined as

$$d = \frac{\sum_{t=2}^{T} (\widehat{\varepsilon}_t - \widehat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T} \widehat{\varepsilon}_t^2}, \tag{3.2}$$

where $\widehat{\varepsilon}_t$ are the OLS residuals from Equation (3.1). When Equation (3.1) is correctly specified, the distribution of d is centred around 2, and d tends towards 0 in the presence of positive autocorrelation, and towards 4 in the presence of negative autocorrelation. It can be shown that $d \overset{A}{\sim} N(2, 4/T)$ with a range of $[0, 4]$, which results from the fact that $d \approx 2(1-\rho)$, where ρ is the autoregressive parameter. The distribution of d depends on the explanatory variables, $\mathbf{x}_t$, which results in exact points of significance being unavailable. Upper and lower bounds, however, have been provided by Durbin and Watson. These bounds allow the null hypothesis of no autocorrelation be tested against the alternative of either positive or negative autocorrelation, as follows. The null of positive autocorrelation is rejected if $d < d_L$, where d_L is the lower bound, and the null is not rejected if $d > d_U$, where d_U is the upper bound, for a given level of significance. If $d_L \leq d \leq d_U$, the test is inconclusive. For the null of negative autocorrelation, the same procedure is followed, although the decision criteria in this case are $d > 4 - d_L$, $d < 4 - d_U$, and $4 - d_U \leq d \leq 4 - d_L$, respectively, for a given level of significance.

As Harvey (1990) pointed out, serial correlation in the residuals does not imply serially correlated disturbances. So while values of d close to 0 may suggest serial correlation in the disturbance terms, it may also result from misspecification and possibly neglected nonlinearity. That is, an inappropriate or inadequate functional form may lead to serially correlated residuals. It should be noted that this test is not valid if the model to be tested includes lagged dependent variables.

3.2.2 The Harvey-Collier test

The ψ-test proposed by Harvey and Collier (1977) is another test based on residual analysis, although it employs recursive residuals, which as the authors note may 'exhibit a very different pattern of behaviour to the OLS residuals under functional misspecification'.[6] In some cases, this test is more powerful than the previously considered Durbin-Watson (1950) bounds test.[7] Before considering the test, consider first the recursive residuals. Given the model

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t, \tag{3.1}$$

where $t = 1, \ldots, T$, $\mathbf{x}_t$ is a k-vector of observations, $\boldsymbol{\beta}$ is a vector of coefficients and ε_t is the disturbance term where $\varepsilon_t \sim$ n.i.d.$(0, \sigma^2)$.[8] The $T - k$ recursive residuals are defined by

$$\nu_t = \frac{y_t - \mathbf{x}_t'\widehat{\boldsymbol{\beta}}_{t-1}}{\left(1 + \mathbf{x}_t'\left(\mathbf{X}_{t-1}'\mathbf{X}_{t-1}\right)^{-1}\mathbf{x}_t\right)^{\frac{1}{2}}}, \tag{3.3}$$

where $t = k+1, \ldots, T$, $\widehat{\boldsymbol{\beta}}_t$ is the OLS estimate of $\boldsymbol{\beta}$ obtained from the first t observations and $\mathbf{X}_t$ is a $t \times k$ matrix of full rank consisting of the first t sets of observations on the independent variables. Under the null hypothesis of Equation (3.1), i.e., of no functional misspecification, the recursive residuals share the properties of the true disturbances. Given $\bar{\nu}$, the arithmetic mean of the recursive residuals, the statistic

$$\psi = \left[(T - k - 1)^{-1} \sum_{t=k+1}^{T} (\nu_t - \bar{\nu})^2\right] (T - k)^{-\frac{1}{2}} \sum_{t=k+1}^{T} \nu_t, \tag{3.4}$$

follows a t-distribution with $(T - k - 1)$ degrees of freedom under the null hypothesis. If the null hypothesis is true, ψ should be close to zero. A two-sided t test is then normally carried out, although a one-sided test may be sometimes appropriate.[9]

3.2.3 Ramsey's RESET test

Ramsey's (1969) regression specification error test, or RESET, is as well known as the Durbin-Watson (1950) test and is widely used as a specification test. Its choice here as a test for neglected nonlinearity is clear: a nonlinear data generating process modelled linearly is a misspecification. In principle, the procedure tests the null of linearity in the independent variables against an unspecified alternative of nonlinearity. The RESET test is valid asymptotically, and may even be so in finite samples, depending on the model in question. Given a model like Equation (3.1),

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t, \tag{3.1}$$

[6]Harvey and Collier (1977), p. 104.

[7]Harvey and Collier (1977), p. 104, suggested that when the 'form of misspecification is such that the 'correct' functional form of the misspecified variable is a concave or a convex function of the variable actually included in the regression', that this test is 'more powerful than the Durbin-Watson and von Neumann tests'.

[8]The Harvey and Collier (1977) test is limited in practice to cases with either one explanatory variable, or where multiple explanatory variables are assumed to have the same form of nonlinearity. While this may not be problematical in some cases, it is not a desirable quality.

[9]Thanks to Curt Wells, University of Lund, for making available his recursive residual GAUSS code, without which, the task of programming the Harvey and Collier (1977) test would have been considerably more arduous. His website can be found at `http://www.nek.lu.se/nekcwe/`.

the test is carried out by estimating the OLS parameters $\widehat{\boldsymbol{\beta}}$ and $\widehat{\varepsilon}_t$ and regressing

$$\widehat{\varepsilon}_t = \mathbf{x}_t'\boldsymbol{\beta} + \sum_{j=2}^{h} \psi_j\, \widehat{y}_t^{\,j} + \nu_t, \tag{3.5}$$

where $\widehat{y}_t = \mathbf{x}_t'\widehat{\boldsymbol{\beta}}$. The null hypothesis of no misspecification is

$$H_0 : \psi_2 = \ldots = \psi_h = 0, \tag{3.6}$$

and the test statistic is

$$R_h = \frac{\left(\sum_{t=1}^{T} \widehat{\varepsilon}_t^2 - \sum_{t=1}^{T} \widehat{\nu}_t^2\right)/(h-1)}{\sum_{t=1}^{T} \widehat{\nu}_t^2/(T-k-h+1)}, \tag{3.7}$$

which has an approximate $F(h-1, T)$-distribution, where k is the dimension of $\mathbf{x}_t$. In practice $h = 2$ is generally found to be sufficient and it is this value that is employed in this study.

Keenan (1985) proposed an adaptation of the RESET test that attempts to avoid multicollinearity between $\widehat{y}_t^{\,j}$ and $\mathbf{x}_t$.[10] The proposed test uses just $\widehat{y}_t^2$. Rather than regressing Equation (3.5), the equation

$$\widehat{y}_t^2 = \mathbf{x}_t'\boldsymbol{\phi} + u_t, \tag{3.8}$$

is regressed to obtain the residuals $\widehat{u}_t = \widehat{y}_t^2 - \mathbf{x}_t'\widehat{\boldsymbol{\phi}}$. The regression

$$\widehat{\varepsilon}_t = \widehat{u}_t\alpha + \nu_t, \tag{3.9}$$

provides the sum of squared residuals, $\sum_{t=k+1}^{T} (\widehat{\varepsilon}_t - \widehat{u}_t\widehat{\alpha})^2 = \sum_{t=k+1}^{T} \widehat{\nu}_t^2$, and allows for testing the null hypothesis, $\alpha = 0$.

Tsay (1986) proposed a further adaptation of RESET to increase its power, suggesting that $\text{vech}(\mathbf{x}_t\mathbf{x}_t')$ be used instead of $\sum_{j=2}^{h} \psi_j\, \widehat{y}_t^{\,j}$, where $\text{vech}(\mathbf{A})$ denotes the half-stacking vector of the matrix $\mathbf{A}$, using elements on or below the diagonal only.[11] In practice, the hypothesis to be tested is equivalent to that found in Equation (3.6).[12] These revised versions of RESET are not considered in this simulation study. As noted, their comparative performance versus RESET has already been assessed. Also, a major attraction of the standard RESET test is the fact that it is widespread and routinely reported, making it easily accessible. Any comparative study of these variant tests and the random field methods, pending the results of this study, is left for future research.

[10]This multicollinearity arises from the fact that $\psi_j\, \widehat{y}_t^{\,j}$ for $j = 2, \ldots, h$ tends to be highly correlated with $\mathbf{x}_t$.

[11]Tsay (2002), p. 157.

[12]Luukkonen, Saikkonen, and Teräsvirta (1988) extended this test by augmenting $\text{vech}(\mathbf{x}_t\mathbf{x}_t')$ with cubic terms.

3.2.4 Hamilton's LM test

The $\lambda_H^E(\boldsymbol{g})$ test was introduced in Chapter 2. It warrants, therefore, but a brief mention here. Recall the form of the model used in Hamilton's (2001) approach,

$$y_t = \alpha_0 + \boldsymbol{\alpha}'\mathbf{x}_t + \lambda m(\boldsymbol{g} \odot \mathbf{x}_t) + \varepsilon_t, \tag{3.10}$$

where $t = 1, 2, \ldots, T$. This suggests that a simple method of testing for nonlinearity is to check if λ is zero or not. Hamilton shows that if $\lambda^2 = 0$, and the nonlinear model is estimated, then for fixed $\boldsymbol{g}$, the maximum likelihood estimator $\widetilde{\lambda}^2$ is consistent for the true value of zero and asymptotically normal. Hamilton proposes a form of LM test for practical application and the procedure has four steps, as follows:

- Set $g_i = \frac{2}{\sqrt{ks_i^2}}$, where s_i^2 is the variance of explanatory variable x_i, excluding the constant term whose variance is zero.
- Calculate the $T \times T$ matrix, $\mathbf{H}$, whose typical element is $H_k\left(\frac{1}{2}\left\|\bar{\mathbf{x}}_t - \bar{\mathbf{x}}_s\right\|\right)$, i.e., the function $H_k(h)$ defined in Equation (2.20).
- Use OLS to estimate the standard linear regression $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ and obtain the usual residuals, $\widehat{\boldsymbol{\epsilon}}$, and standard error of estimate, $\widehat{\sigma}^2 = (T-k-1)^{-1}\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}}$.
- Finally, compute the statistic

$$\lambda_H^E(\boldsymbol{g}) = \frac{[\widehat{\boldsymbol{\epsilon}}'\mathbf{H}\widehat{\boldsymbol{\epsilon}} - \widehat{\sigma}^2 tr(\mathbf{M}_T\mathbf{H})]^2}{\widehat{\sigma}^4\left[2tr\left([\mathbf{M}_T\mathbf{H}\mathbf{M}_T - (T-k-1)^{-1}\mathbf{M}_T tr(\mathbf{M}_T\mathbf{H})]^2\right)\right]}, \tag{2.38}$$

  where $\mathbf{M}_T = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is the familiar symmetric idempotent matrix.

As $\lambda_H^E(\boldsymbol{g}) \overset{A}{\sim} \chi_1^2$ under the null hypothesis, linearity ($\lambda^2 = 0$) would be rejected if $\lambda_H^E(\boldsymbol{g})$ exceeded the critical value, for the chosen level of significance. In this case the alternative nonlinear specification given by Equation (3.10) would be preferred.

The performance of the $\lambda_H^E(\boldsymbol{g})$ test has been assessed by Dahl (2002). In a simulation study, this test was compared to those of Cox and Koh (1989), Cox, Koh, Wahba, and Yandell (1988) and the neural network tests of Lee, et al. (1993) and Teräsvirta, Lin, and Granger (1993), which like the $\lambda_H^E(\boldsymbol{g})$ test, require no prior knowledge of the functional form. The Hamilton (2001) test was found to perform very well in finite samples. Its size and power properties where good when compared to the 'most popular and powerful tests in the literature', including the neural network tests, mentioned above.[13]

Dahl and González-Rivera (2003), however, identified several areas where the $\lambda_H^E(\boldsymbol{g})$ test's performance may suffer. These are discussed in detail in the next section. More recently, Lee, et al. (2005) assessed the performance of a variety of nonlinearity tests, including the $\lambda_H^E(\boldsymbol{g})$ test. They present worrying evidence of spurious nonlinear regression, when the variables in question are random walks. The $\lambda_H^E(\boldsymbol{g})$ test performs poorly. They find that the test rejects the true null of linearity in over 50 per cent of cases examined, against a nominal size of 5 per cent, for a sample size of $T = 50$. Rather than converging to its χ^2 distribution, the test

[13] Dahl (2002), p. 282.

statistic diverges to infinity as the sample size increases. Although none of the simulated data used in this chapter can be characterised as a random walk, the findings of Lee, et al. (2005) have serious implications for the practical application of the test. This issue will be addressed further in later chapters.

3.2.5 The $\lambda^E_{OP}(\boldsymbol{g})$, λ^A_{OP} and g_{OP} tests

Studies have shown the excellent properties of the $\lambda^E_H(\boldsymbol{g})$ test. Dahl (2002), for example, found that the test 'performs well in finite samples' and 'has good size and power properties when compared to some of the most popular and powerful tests in the literature'.[14] Several questions may be raised about its performance in certain circumstances. Specifically, as the number of variables in the model under consideration increases, so too does the number of nuisance parameters. As Hansen (1996) showed, dealing with unidentified nuisance parameters by assuming full knowledge of the parameterised stochastic process that determines the random field may have adverse effects on the power of the test. Also, the test relies upon a specific specification of the variance-covariance function of the random field. Dahl and González-Rivera (2003) noted that where alternative variance-covariance functions are more appropriate, the performance of the $\lambda^E_H(\boldsymbol{g})$ test may suffer.

To offset the potential for decreased performance in the situations outlined above, Dahl and González-Rivera (2003) proposed a troika of tests that are robust to the specification of the variance-covariance function of the random field and that encompass a broad class of variance-covariance functions. Their tests are also unaffected by increasing dimensionality, as they are free of the problem of unidentified nuisance parameters. Their tests, therefore, 'aim to generalise and complement the Hamilton (2001) statistic'.[15]

The issue of nuisance parameters was touched upon very briefly in Chapter 2, Section 2.8.[16] It is appropriate to expand this discussion here, as it is relevant to the $\lambda^E_{OP}(\boldsymbol{g})$, λ^A_{OP} and g_{OP} tests, and to the applications undertaken in subsequent chapters. Recall the fundamental model used in Hamilton's (2001) framework, as shown in Equation (3.10). Dahl and González-Rivera (2003) recognised the potential for nuisance parameters in this model. That is, under the null of $H_0\colon \lambda^2 = 0$, the parameter vector $\boldsymbol{g}$ becomes unidentified, while the number of nuisance parameters increases with the dimensionality of $\mathbf{x}$. Alternatively, consider the null of $H_0\colon \boldsymbol{g} = \mathbf{0}_k$, the motivation for which will be introduced shortly. In this case, λ becomes unidentified and y_t will be nonergodic.[17] Under the null, therefore, the test may not have a well-defined asymptotic distribution.

Two solutions to this problem are proposed by Dahl and González-Rivera (2003). The first considers the random field based on the L_1 norm, or Minkowski distance, as opposed to the L_2 norm, or Euclidean distance, used by Hamilton (2001). This has the advantage of

[14] Dahl (2002), p. 282.

[15] Dahl and González-Rivera (2003), p. 162.

[16] Davies (1977, 1987) was one of the first to analyse the problem of unidentified nuisance parameters. He suggested viewing test statistics as functions of unidentified parameters. Andrews and Ploberger (1994) derived asymptotic tests for problems where unidentified nuisance parameters are present under the alternative hypothesis but not under the null.

[17] Ergodicity is the property that the eventual distribution of the states of a system is independent of the initial state (Everitt, 2002, p. 132).

simplifying tests for neglected nonlinearity, as this measure is a linear function of the nuisance parameters, unlike the L_2 measure, which is nonlinear in the nuisance parameters.[18]

The disadvantage of this approach is that such a random field is nonisotrophic, i.e., the covariance function may be unwieldy to evaluate in models of large dimension. The approach does suggest, however, an alternative covariance function, itself based on an isotrophic random field. That function is

$$\mathbf{C}_k^*(h^*) = \begin{cases} (1-h^*)^{2k} & \text{if } h^* \leq 1, \\ 0 & \text{if } h^* > 1, \end{cases} \tag{2.48}$$

where $h^* \equiv \frac{1}{2} d_{L_1}(\mathbf{x}, \mathbf{z})$.

As Hamilton (2005) confirmed, the issue of nuisance parameters may be overcome by using this alternative covariance function.[19] He experiments with this new specification and finds that the pile-up problem, noted briefly in Chapter 2, Section 2.8, may be avoided by employing the Dahl and González-Rivera (2003) approach.[20]

The troika of tests, developed from the $\lambda_H^E(\boldsymbol{g})$ test, are based on this approach. Dahl and González-Rivera conduct an extensive Monte Carlo study to investigate the properties of their tests and find that they are indeed effective in detecting neglected nonlinearity. They out-perform the $\lambda_H^E(\boldsymbol{g})$ statistic's power in a range of models, for both small and moderate samples sizes.[21] In addition to the simulation study, an empirical application suggests that by avoiding the problem of nuisance parameters, the test's performances are comparatively better than others. Interestingly, Dahl and González-Rivera (2003) reported that these tests out-perform the well-known Tsay (1986) test for nonlinearity, suggesting perhaps that it is better to view an unobserved nonlinear function as random and not deterministic, the approach taken by Tsay (1986).

Recall from Subsection 3.2.4 the model being used by Hamilton (2001), as shown in Equation (3.10). The parameters λ and $\boldsymbol{g}$ are fundamentally important, as they characterise

---

[18]Recall from Subsection 2.4.2 and above, the model and framework of Hamilton (2001). Dahl and González-Rivera (2003) replaced the set $B_N(\mathbf{x})$ with the set $B_N^*(\mathbf{x})$. For each point $\mathbf{x} \in A_N$, $B_N^*(\mathbf{x}) = \{\mathbf{z} \in A_N : |\mathbf{x}-\mathbf{z}|'\iota \leq 1\}$, and m_N^* is the random field with a moving average representation:

$$m_N^*(\mathbf{x}) = [n_N^*(\mathbf{x})]^{-\frac{1}{2}} \sum_{\mathbf{z} \in B_N^*(\mathbf{x})} e(\mathbf{z}),$$

where $n_N^*(\mathbf{x})$ is the number of points in $B_N^*(\mathbf{x})$. It can be shown that the correlation between $m^*(\mathbf{x})$ and $m^*(\mathbf{z})$ equals zero if the L_1 distance between $\mathbf{x}$ and $\mathbf{z}$ is greater than or equal to two. This covariance function, illustrated for the cases $k = 1, 2$ is:

$$\begin{aligned} \mathbf{H}_1^*(\mathbf{x}_t, \mathbf{x}_s) &= \begin{cases} 1-h_{ts}^* & \text{if } h_{ts}^* \leq 1, \\ 0 & \text{if } h_{ts}^* > 1, \end{cases} \\ \mathbf{H}_2^*(\mathbf{x}_t, \mathbf{x}_s) &= \begin{cases} (1-h_{ts}^*)^2 + (1-h_{ts}^*)\, \min\{|x_{1t}-x_{1s}|, |x_{2t}-x_{2s}|\} & \text{if } h_{ts}^* \leq 1, \\ 0 & \text{if } h_{ts}^* > 1, \end{cases} \end{aligned}$$

where $h_{ts}^* = \frac{1}{2} d_{L_1}(\mathbf{x}_t, \mathbf{x}_s)$.

[19]Hamilton (2005), p. 6.

[20]Given the range of issues highlighted in Chapter 2, and very much in keeping with the findings therein, it would seem prudent that the applied researcher consider both the Hamilton (2001) and Dahl and González-Rivera (2003) covariance functions in empirical applications. It adds another element to the prescribed actions, detailed in Chapter 2, for those contemplating use of the Hamilton framework: use the algorithm switching methodology, pairing STEEPEST DESCENT with both BFGS and NEWTON, for a range of values for ζ and `_oprteps`, if required, and also for both covariance functions. Such investigations should allow for greater understanding of the results obtained.

[21]The models considered had bilinear, neural network and smooth transition autoregressive specifications.

the relationship between $m(\cdot)$ and the conditional expectation function $\mu(x_1, x_2, \ldots, x_k)$. This relationship is exploited by Hamilton, who uses it to test for nonlinearity. Should $\lambda = 0$, $m(\cdot)$ makes no contribution to Equation (3.10), and the conditional expectation is linear. The parameter λ^2 can, therefore, be used as a test for nonlinearity. Dahl and González-Rivera (2003) used this feature, but also make use of another. If $\boldsymbol{g} = \mathbf{0}$, the contribution of $m(\cdot)$ to the conditional expectation is indistinguishable from α_0, and hence, once again, the model is linear. Therefore, $\boldsymbol{g} = \mathbf{0}$ also serves as a test for nonlinearity. What follows is a more detailed outline of each of these tests.[22]

The $\lambda_{OP}^{E}(\boldsymbol{g})$ test

The $\lambda_{H}^{E}(\boldsymbol{g})$ test for nonlinearity is based on the Hessian of the log likelihood function, with $\mathbf{C}_k(\mathbf{x}_t, \mathbf{x}_s) = \mathbf{H}_k(h)$, where $h = \frac{1}{2} d_{L_2}(\mathbf{x}_t, \mathbf{x}_s)$. The g_i are fixed, as mentioned in the previous subsections.[23] The $\lambda_{OP}^{E}(\boldsymbol{g})$ test is derived from the loglikelihood function

$$
\begin{aligned}
& l\left(\boldsymbol{\beta}, \lambda^2, \boldsymbol{g}, \sigma^2\right) = \\
& \quad -\frac{T}{2}\log(2\pi) - \frac{1}{2}\log\left|\lambda^2 \mathbf{C}_k + \sigma^2 \mathbf{I}_T\right| - \frac{1}{2}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)'\left(\lambda^2 \mathbf{C}_k + \sigma^2 \mathbf{I}_T\right)^{-1}\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right),
\end{aligned}
\tag{3.11}
$$

where $\boldsymbol{\Omega}_{DGR} = \lambda^2 \mathbf{C}_k + \sigma^2 \mathbf{I}_T$. Dahl and González-Rivera (2003) showed that the score functions of Equation (3.11) are

$$
s\left(\lambda^2\right)|_{\lambda^2=0,\,\boldsymbol{g}} = -\frac{1}{2\sigma^2}\widetilde{\mathbf{x}}_1'\mathbf{u}, \tag{3.12}
$$

$$
s\left(\sigma^2\right)|_{\lambda^2=0,\,\boldsymbol{g}} = -\frac{1}{2\sigma^2}\widetilde{\mathbf{x}}_2'\mathbf{u}, \tag{3.13}
$$

where $\widetilde{\mathbf{x}}_1 = \mathrm{vec}(\mathbf{C}_k)$, $\widetilde{\mathbf{x}}_2 = \mathrm{vec}(\mathbf{I}_T)$ and $\mathbf{u} = \mathrm{vec}\left(\mathbf{I}_T - \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'/\sigma^2\right)$. Letting $\widetilde{\mathbf{x}} = (\widetilde{\mathbf{x}}_1 : \widetilde{\mathbf{x}}_2)$, the LM test statistic is

$$
\lambda_{OP}^{E}(\boldsymbol{g}) = \frac{T^2}{2}\frac{\mathbf{u}'\widetilde{\mathbf{x}}\left(\widetilde{\mathbf{x}}'\widetilde{\mathbf{x}}\right)^{-1}\widetilde{\mathbf{x}}'\mathbf{u}}{\mathbf{u}'\mathbf{u}} \sim \chi^2(1). \tag{3.14}
$$

As with the $\lambda_{H}^{E}(\boldsymbol{g})$ test, the $\lambda_{OP}^{E}(\boldsymbol{g})$ statistic is obtained easily. The model is estimated under the null, providing estimates of $\widehat{\boldsymbol{\varepsilon}} = \mathbf{y} - \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}$ and $\widehat{\sigma}^2 = T^{-1}\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}}$. Then, $\widehat{\boldsymbol{\nu}}$ is obtained from the least squares auxiliary regression of $\widehat{\mathbf{u}} = \phi_1\widetilde{\mathbf{x}}_1 + \phi_2\widetilde{\mathbf{x}}_2 + \boldsymbol{\nu}$, using $\widehat{\mathbf{u}} = \mathrm{vec}\left(\mathbf{I}_T - \widehat{\boldsymbol{\varepsilon}}\widehat{\boldsymbol{\varepsilon}}'/\widehat{\sigma}^2\right)$. The test statistic is $\lambda_{OP}^{E}(\boldsymbol{g}) = \frac{1}{2}T^2 R^2$, where R^2 is the uncentred R^2, defined as $R^2 = 1 - \widehat{\boldsymbol{\nu}}'\widehat{\boldsymbol{\nu}}/\widehat{\mathbf{u}}'\widehat{\mathbf{u}}$. The $\lambda_{OP}^{E}(\boldsymbol{g})$ is not asymptotically distributed. As there are T^2 observations in the auxiliary regression, the $\lambda_{OP}^{E}(\boldsymbol{g})$ test can be transformed to an asymptotically distributed $\chi^2(1)$ by multiplying the R^2 by T^2 instead of T.

The λ_{OP}^{A} test

Fixing $\boldsymbol{g}$ in the previous test is equivalent to assuming complete knowledge of the covariance function of the random field. This test does not depend on the parameterisation of the covariance function and avoids the problem of nuisance parameters. This is achieved by

---

[22] Thanks to Christian M. Dahl and Gloria González-Rivera for making the GAUSS code for their $\lambda_{OP}^{E}(\boldsymbol{g})$, λ_{OP}^{A} and g_{OP} tests available. It can be found at `http://www.krannert.purdue.edu/faculty/dahlc/`.

[23] For further details, see Chapter 2, Subsection 2.4.5.

replacing $\widetilde{\mathbf{x}}_1 = \text{vec}(\mathbf{C}_k)$ in Equation (3.12) with a Taylor approximation of $\widetilde{\mathbf{x}}_1$, as outlined below, in the auxiliary regression. 'Using the Taylor approximation captures the characteristics of a wide class of unknown but continuous and differentiable real-valued nonlinear functions, vec($\mathbf{C}_k$)'.[24] Using the covariance function

$$\mathbf{C}_k^*\left(h_{ts}^*\right) = \begin{cases} \left(1 - h_{ts}^*\right)^{2k} & \text{if } h_{ts}^* \leq 1, \\ 0 & \text{if } h_{ts}^* > 1, \end{cases} \tag{2.48}$$

where $h_{ts}^* \equiv \frac{1}{2} d_{L_1}\left(\boldsymbol{g} \odot \mathbf{x}_t, \boldsymbol{g} \odot \mathbf{x}_s\right) = \frac{1}{2}\mathbf{r}_{ts}'\boldsymbol{g}$, $\mathbf{r}_{ts} = \{|x_{t1} - x_{s1}|, |x_{t2} - x_{s2}|, \ldots, |x_{tk} - x_{sk}|\}'$, and given that $\left(1 - h_{ts}^*\right)^{2k} = \sum_{j=0}^{2k} \binom{2k}{j} h_{ts}^{*j} \left(-1\right)^{-j}$, the auxiliary regression can be written as

$$\widehat{\mathbf{u}}_{ts} = \phi_1 \widetilde{\mathbf{x}}_{ts,1} + \phi_2 \widetilde{\mathbf{x}}_{ts,2} + \boldsymbol{\nu}_{ts} = \phi \left[\sum_{j=0}^{2k} \binom{2k}{j} h_{ts}^{*j} \left(-1\right)^j \right] 1_{(h_{ts}^* \leq 1)} + \phi_2 \widetilde{\mathbf{x}}_{ts,2} + \boldsymbol{\nu}_{ts}, \tag{3.15}$$

where $1_{(h_{ts}^* \leq 1)}$ is an indicator function that can be approximated by the logistic function $1_{(h_{ts}^* \leq 1)} \approx \left(1 + \exp\left(-\gamma\left(1 - h_{ts}^*\right)\right)\right)^{-1}$, for fixed $\gamma >> 0$. Rather than proceeding as in the case of the $\lambda_{OP}^E(\boldsymbol{g})$ test, a second Taylor expansion is considered, where the norm depends on $\boldsymbol{g}$ and the logistic approximation of the indicator function. That gives

$$\begin{aligned} \widehat{u}_{ts} = \bar{\phi}_0 + \bar{\phi}_1 \sum_{i=1}^{k} g_i r_{ts,i} + \bar{\phi}_2 \sum_{i=1}^{k} \sum_{j \geq i}^{k} g_i g_j r_{ts,i} r_{ts,j} & \qquad (3.16) \\ + \bar{\phi}_3 \sum_{i=1}^{k} \sum_{j \geq i}^{k} \sum_{l \geq j}^{k} g_i g_j g_l r_{ts,i} r_{ts,j} r_{ts,l} + \ldots & \\ + \bar{\phi}_{2k+2} \sum_{i=1}^{k} \sum_{j \geq i}^{k} \cdots \sum_{m}^{k} g_i g_j \cdots g_m r_{ts,i} r_{ts,j} \cdots r_{ts,m} + \phi_2 \widetilde{x}_{2,ts} + \nu_{ts}, & \end{aligned}$$

where $\bar{\phi}_j$ is proportional to ϕ_1, and the subindex ts, relating to $\widehat{\mathbf{u}}$, $\widetilde{\mathbf{x}}_2$ and $\boldsymbol{\nu}$, refers to the ts^{th} row in the vector, for $t, s = 1, 2, \ldots, T$, and g_i and $r_{ts,i}$ are the i^{th} entries in $\boldsymbol{g}$ and $\mathbf{r}_{ts}$, respectively. Estimating this auxiliary regression by OLS, with $\sum_{j=1}^{2k+2} \binom{k+j-1}{k-1}$ regressors added, is the basis for the λ_{OP}^A test, which has the same LM test statistic as the $\lambda_{OP}^E(\boldsymbol{g})$, i.e., Equation (3.14), but using the auxiliary regression outline above. This has a χ^2 distribution of $1 + \sum_{j=1}^{2k+2} \binom{k+j-1}{k-1}$ degrees of freedom. The authors argue that if a covariance function of the form in Equation (2.48) is used, it encompasses a very wide class of covariance functions. Given that the practitioner is unlikely to know the true covariance function, this is particularly useful.

The g_{OP} test

The g_{OP} test is based on similar models, but here the random field has the variance-covariance structure

$$\widetilde{\mathbf{C}}_k\left(\mathbf{x}_t, \mathbf{x}_s\right) = \mathbf{C}_k\left(\mathbf{x}_t, \mathbf{x}_s\right) + \mathbf{C}_k\left(\mathbf{0}_k, \mathbf{0}_s\right) - \mathbf{C}_k\left(\mathbf{x}_t, \mathbf{0}_k\right) - \mathbf{C}_k\left(\mathbf{0}_k, \mathbf{x}_s\right). \tag{3.17}$$

[24] Dahl and González-Rivera (2003), p. 152.

Note the covariance function is that given by Equation (2.48) and the loglikelihood function is as before, with the exception that $\widetilde{\mathbf{C}}_k$ replaces $\mathbf{C}_k$ and $\boldsymbol{\Omega}_{DGR} = \lambda^2\widetilde{\mathbf{C}}_k + \sigma^2\mathbf{I}_T$. Proceeding as in the case of $\lambda_{OP}^E(\boldsymbol{g})$, the score functions, for fixed λ, H_0: $\mathbf{g} = \mathbf{0}$, and $\boldsymbol{\Omega}_{DGR}|_{\boldsymbol{g}=\mathbf{0}} = \sigma^2\mathbf{I}_T$, can be written as

$$s(g_i)\,|_{\lambda^2,\boldsymbol{g}=\mathbf{0}} = -\frac{\lambda^2}{2\sigma^2}\widetilde{\mathbf{x}}_i'\mathbf{u}, \qquad i = 1, 2, \dots k, \tag{3.18}$$

$$s(\sigma^2)\,|_{\lambda^2,\boldsymbol{g}=\mathbf{0}} = -\frac{1}{2\sigma^2}\widetilde{\mathbf{x}}_{k+1}'\mathbf{u}, \tag{3.19}$$

where $\widetilde{\mathbf{x}}_i = \partial\,\mathrm{vec}\left(\widetilde{\mathbf{C}}_k\right)/\partial g_i|_{\boldsymbol{g}=\mathbf{0}}$, for $i = 1, 2, \dots, k$, $\widetilde{\mathbf{x}}_{k+1} = \mathrm{vec}(\mathbf{I}_T)$ and $\mathbf{u} = \mathrm{vec}\left(\mathbf{I}_T - \varepsilon\varepsilon'/\sigma^2\right)$. The TR^2, or g_{OP} test, is $\frac{1}{2}T^2R^2 \sim \chi^2(k)$, using the R^2 from the auxiliary regression

$$\widehat{u}_{ts} = \sum_{i=1}^{k}\widetilde{\phi}_i\widetilde{r}_{ts,i} + \widetilde{\phi}_{k+1}\widetilde{x}_{k+1,ts} + \widetilde{\nu}_{ts}, \tag{3.20}$$

where $\widetilde{r}_{ts,i} = -k\left(|x_{ti} - x_{si}| - |x_{ti}| - |x_{si}|\right)$, for $t, s = 1, 2, \dots, T$, and where R^2 is uncentred.

3.3 The Design of the Monte Carlo Experiments

To assess the performance of the seven tests outlined in Section 3.2, a Monte Carlo simulation was undertaken. Five models were used. These are summarised in Table 3.1. The first was the null linear model $y_t = \alpha + \beta x_{it} + \varepsilon_{it}$. This null case calibrates each of the tests by obtaining the empirical power and size in each case. The second model is the quadratic specification, $y_t = \alpha + \beta x_{it} + \beta x_{it}^2 + \varepsilon_t$, where α, β, ε_t and x_{it} are as in the null case. The third model is the correlated error model, $y_t = \alpha + \beta x_{it} + \epsilon_{it}$, where $\epsilon_{it} = \rho\epsilon_{it-1} + \varepsilon_{it}$. For each of these three models, $\alpha = 1$, $\beta = \{0.25,\ 0.50,\ 0.75\}$, $\varepsilon_{it} \sim N(0, \sigma_i^2)$, x_{1t} is a linear time trend, $x_{1t} = 20t/T$, x_{2t} is generated from a random uniform distribution, $x_{2t} \sim U(1, 10)$ and finally, x_{3t} is generated from a random normal distribution, $x_{3t} \sim N(5.5, 2.25^2)$.[25] A similar model, here called the square specification, was also included. This was used by Dahl and González-Rivera (2003) and makes a useful comparative tool. The square model is $y_t = x_{it}^2 + \epsilon_t$, where x_{it} for $i = 1, 2, 3$ is as above, and also $x_{4t} = 0.6x_{1t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$ and $\epsilon_t \sim N(0, \sigma_i^2)$. The final model specification is adapted from Hamilton's (2001) Example 1.[26] This is a simple threshold regression, $y_t = 0.6x_{it}1_{[x_{it}>5.5]} + 0.2x_{jt} + \varepsilon_t$, where $i = 2, 3$, $j = 1, 2, 3$, $1_{[x_{it}>5.5]} = 0$ when $x_{it} < 5.5$ and 1 otherwise, and $\varepsilon_t \sim N(0, 1)$. This example is included to allow direct comparison to Hamilton, in applying the other tests for nonlinearity. Summary details of σ_i^2 for each case are provided in Table 3.1. Full details of the design of the Monte Carlo simulation can be found in Appendix B.1, Table B.1.

In almost all cases, 20,000 replications were undertaken, for sample sizes of $T = 25$,

[25] These moments and parameters were chosen for their relevance in applied econometrics, by ensuring that the models to be investigated had the reasonably high R^2-values often encountered in research, although it should be noted that the residuals of all model specifications are invariant to β. The generated data approximates well to many slowly changing economic time series, particularly the trend series and the ordered normal and uniform distributions, which are used for several of the tests.

[26] Hamilton (2001), p. 559. As noted in Subsection 3.2.2, the Harvey-Collier (1977) test is limited in practice to bivariate cases. Therefore, it is not applied to the Hamilton specification in the simulation.

75, 125, 175 and 225. In the case of Dahl and González-Rivera (2003) tests, however, both asymptotic and bootstrapped p-values are available for each test. The asymptotic values were recorded, as above, for 20,000 replications, but the bootstrapped values were also recorded. These were based on 1,000 replications, however, as the bootstrapped p-values were in turn based on 100 re-samples and the computational demands of this process are very considerable. Section 3.5 presents the results obtained from this study. Finally, consideration was also given to some features of the Durbin-Watson (1950) bounds test and the Harvey-Collier (1977) test. For the latter, observations should be ordered by the explanatory variable, prior to testing. This may also apply in the case of the Durbin-Watson test, depending on circumstances. Simulations were undertaken for these tests using both ordered and unordered data.[27]

Table 3.1: Models analysed in the Monte Carlo study.

Null	$y_t = \alpha + \beta x_{it} + \varepsilon_{it}$	$\varepsilon_{it} \sim N(0, \sigma_i^2)$
Quadratic	$y_t = \alpha + \beta x_{it} + \beta x_{it}^2 + \varepsilon_{it}$	$\varepsilon_{it} \sim N(0, \sigma_i^2)$
Correlated Error	$y_t = \alpha + \beta x_{it} + \epsilon_{it}$	
	$\epsilon_{it} = \rho\epsilon_{it-1} + \varepsilon_{it}$	$\varepsilon_{it} \sim N(0, \sigma_i^2)$
Hamilton	$y_t = 0.6x_{it}1_{[x_{it}>5.5]} + 0.2x_{jt} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1)$
	$i = 2, 3,\ j = 1, 2, 3$	
	$x_{1t} = 20t/T$	$\sigma_1^2 = 1.5$
	$x_{2t} = U\,(1, 10)$	$\sigma_2^2 = 0.4$
	$x_{3t} = N\left(5.5, 2.25^2\right)$	$\sigma_3^2 = 0.3$
Square	$y_t = x_{it}^2 + \epsilon_{it}$	$\epsilon_{it} \sim N(0, \sigma_i^2)$
	$x_{1t} = 20t/T$	$\sigma_i^2 = 1$
	$x_{2t} = U\,(1, 10)$	$\sigma_i^2 = 25$
	$x_{3t} = N\left(5.5, 2.25^2\right)$	$\sigma_i^2 = 400$
	$x_{4t} = 0.6x_{1t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1)$

In all cases, $\alpha = 1$ and $\beta = \{0.25, 0.50, 0.75\}$.

3.4 Hypothesis Testing

To explain the various concepts used in the remainder of this chapter, a brief treatment of hypothesis testing is provided.[28] Consider the random variable X defined on the probability space $(S,\ \mathcal{F},\ P(\cdot))$ and some model $\mathbf{\Phi} = \{f(x;\ \theta),\ \theta \in \mathbf{\Theta}\}$, where $\mathcal{F}$ is a σ-field generated by X and $\mathbf{x}$ is a random sample from $f(x;\ \theta)$.

For hypothesis testing, consider the null hypothesis to be $H_0\colon \theta \in \mathbf{\Theta}_0$, i.e., if θ lies in some subset $\mathbf{\Theta}_0$ of $\mathbf{\Theta}$, where the sample data $\mathbf{x} \in C_0$, H_0 is accepted. If $\mathbf{x} \in C_1$, that hypothesis is

[27]The GAUSS code for this simulation can be found in Appendix B.2.

[28]This treatment draws heavily upon and follows closely that of Spanos (1986).

rejected. The sets C_0 and C_1, also known as the acceptance and rejection regions, respectively, are defined by the given test statistic. More formally, the null and alternative hypotheses can be expressed as

$$H_0 : \theta \in \mathbf{\Theta}_0, \qquad \mathbf{\Theta}_0 \subseteq \mathbf{\Theta}, \tag{3.21}$$

$$H_1 : \theta \notin \mathbf{\Theta}_0, \qquad \text{or} \qquad \theta \in \mathbf{\Theta}_1 \equiv \mathbf{\Theta} - \mathbf{\Theta}_0. \tag{3.22}$$

In this chapter, only simple hypotheses are used, i.e., where $f(\mathbf{x};\ \theta)$ is specified completely if $\theta \in \mathbf{\Theta}_0$ or $\theta \in \mathbf{\Theta}_1$ are known. Such hypothesis testing gives rise to two possibilities for error, entitled Type I and Type II errors. The probability of a Type I error is defined as

$$Pr\left(\mathbf{x} \in C_1;\ \theta \in \mathbf{\Theta}_0\right) = \alpha, \tag{3.23}$$

whilst the probability of a Type II error is

$$Pr\left(\mathbf{x} \in C_0;\ \theta \in \mathbf{\Theta}_1\right) = \beta. \tag{3.24}$$

Arising from this framework and of particular interest here are the following definitions:[29]

Definition 1. The probability of rejecting H_0 when false at some point $\theta_1 \in \mathbf{\Theta}_1$, i.e., $Pr\left(\mathbf{x} \in C_1;\ \theta = \theta_1\right)$ is called the power of the test at $\theta = \theta_1$.

Definition 2. $\mathcal{P}\left(\theta\right) = Pr\left(\mathbf{x} \in C_1\right)$, $\theta \in \mathbf{\Theta}$ is called the power function of the test defined by the rejection region C_1.

Definition 3. $\alpha = \max_{\theta \in \mathbf{\Theta}_0} \mathcal{P}\left(\theta\right)$ is defined to be the size, or the significance level, of the test.

Definition 4. A test of $H_0 : \theta \in \mathbf{\Theta}_0$ against $H_1 : \theta \in \mathbf{\Theta}_1$ as defined by some rejection region C_1 is said to be uniformly most powerful test of size α if

1. $\max_{\theta \in \mathbf{\Theta}_0} \mathcal{P}\left(\theta\right) = \alpha$;

2. $\mathcal{P}\left(\theta\right) = \mathcal{P}^*\left(\theta\right)$ for all $\theta \in \mathbf{\Theta}_1$;

where $\mathcal{P}^*\left(\theta\right)$ is the power function of any other test of size α.

These concepts, particularly size and power, will be used throughout the remainder of this chapter.

3.5 Results and Discussion

The results of the simulation study can be found in Appendix B.1. A key to the parameter specifications, a more detailed version of Table 3.1, can be found in Table B.1. The remainder of Appendix B.1 contains the individual results for each test. The simulated sizes, in terms of percentage frequencies of rejection of the null hypothesis for the Durbin-Watson (1950)

[29] Spanos (1986), p. 290-291.

bounds test, may be found in tables B.2 to B.6 for the unordered cases and tables B.7 to B.11 for the ordered cases. Results for the Harvey-Collier (1977) test can be found in Table B.12 for the unordered case and Table B.13 for the ordered case. The results of Ramsey's (1969) RESET test can be found in Table B.14 and for Hamilton's (2001) $\lambda_H^E(\boldsymbol{g})$ test in Table B.15. Tables B.16 to B.21 contain the results of the $\lambda_{OP}^E(\boldsymbol{g})$, λ_{OP}^A and g_{OP} tests.[30] To analyse the results of the study outlined in Section 3.3, the performance of each test will be discussed before comparing relative performance. Of particular interest here is the comparative performance of the $\lambda_H^E(\boldsymbol{g})$ test to the well-known tests, but also to the more recent $\lambda_{OP}^E(\boldsymbol{g})$, λ_{OP}^A and g_{OP} tests.

3.5.1 The Durbin-Watson test

To explore the importance of ordering on the Durbin-Watson (1950) bounds test, results are reported for unordered data and also for data ordered by increasing magnitude of the explanatory variable. The empirical size and power of the test are reported for both $d < d_L$, the rejection region, but also for $d_L < d < d_U$, the so-called zone of indecision. Given the design of the experiments here, only positive autocorrelation is considered.

The size of the test with unordered data is found to be very much dependant on T, which can be seen from Table B.2. The rate of rejection of the null increases from above 1 per cent towards 5 per cent with increasing sample size, T. The size is close to or above 4 per cent in many cases. The results for the quadratic specification, which can be found in Table B.3, show that the test is powerful when x_t is a time trend, with full power observed for $d < d_L$ in every case. For $x_t \sim U$ and $x_t \sim N$, however, the power in the majority of cases is zero for both $d < d_L$ and $d < d_U$. Very similar results are seen for the square specification, as shown in Table B.4. Power is very high when x_t is trended, over 90 per cent in all cases, but very low for the other cases, with the exception of x_{4t}.

As the Durbin-Watson (1950) test is explicitly a test for serial correlation, it could be expected to perform well for the autoregressive specification.[31] The test does indeed perform well, as can be seen from the results in Table B.5. For higher values of ρ and T, power is frequently observed to be 100 per cent for $d < d_L$. For $\rho = 0.1$, however, power never exceeds 41 per cent, regardless of T. The inverse relationship between the size of $d < d_L$ and $d < d_U$, observed previously, holds true for power in this case. Finally, for the Hamilton (2001) specification, the results of which can be seen in Table B.6, the power is very low for both $d < d_L$ and $d < d_U$, regardless of T.

The size of the test with ordered data appears to be less variable, although still somewhat dependant on T. The size of $d < d_L$ generally approximates to 5 per cent, as shown by Table B.7. As with the unordered data, the quadratic specification where data is trended, reported in Table B.8, has full power for $d < d_L$, but the test now also has full power for $x_t \sim U$ and $x_t \sim N$. The effect of ordering is clear. The same can be said of the ordered square specification, the results of which can be found in Table B.9. Although the test is still sensitive

[30]Two sets of results are reported for the $\lambda_{OP}^E(\boldsymbol{g})$, λ_{OP}^A and g_{OP} tests. The first are based on asymptotic p-values, while the second are based on bootstrapped p-values.

[31]As previously noted, however, latent nonlinearity or misspecification may result in serially correlated errors.

to T and σ^2, powers of 100 per cent are observed in many cases.

Interestingly, the ordering of data for the autoregressive specification, as shown in Table B.10, appears to have a negative impact on the rate of rejection of the null hypothesis, although power still depends on ρ and T when data is ordered. For the larger values of T, this fall may not be significant enough to warrant concern. Finally, the power for the ordered Hamilton (2001) specification, as shown in Table B.11, has greatly improved and is above 94 per cent for all cases where $T \geq 75$. It appears, therefore, that ordering the data, where permissible, greatly improves the power of the Durbin-Watson (1950) test across a range of models. This is particularly noteworthy as such ordering approximates the data examined here to many economic time series, which are slowly changing. This result is unsurprising given the fact that the Durbin-Watson test is a test for first-order autoregression. Unordered normally or uniformly distributed data is not well characterised by an autoregressive process and the Durbin-Watson test statistic will, therefore, be close to 2. Ordering the data, however, may reveal a degree of autocorrelation, which may be picked up by the d-statistic.

3.5.2 The Harvey-Collier test

Despite its obvious potential, the Harvey-Collier (1977) test appears little used. Although the concept of recursive residuals, on which this test procedure is based, is familiar to many economists, the test itself remains little known. Its absence from most econometric software is evidence of this fact.[32] An obvious drawback of this test is its limited applicability to multivariate models. It is applicable in a multivariate context if all variables are assumed to have the same form of nonlinearity. Should this not be the case, it can be used in multivariate cases, but requires that a decision be made regarding which independent variable to test. This limits the tests practical applicability and also rules out the possibility of testing for multivariate nonlinearity. These draw-backs may go some way to explaining this test's under-utilisation in the literature.[33] In the context of this study, the final model specification was omitted for this test, as it is multivariate. Table 3.1 provides further details.

An integral part of the Harvey-Collier (1977) test is that the data should be ordered. That is, the observations are ordered on the explanatory variable in increasing order of magnitude.[34] To explore the importance of this step, the simulation was undertaken for both unordered and appropriately ordered data. The results for these two cases can be found in tables B.12 and B.13, respectively.

For both the unordered and ordered data, the size of the Harvey-Collier (1977) test is close to 5 per cent in all cases, as expected. There is no obvious relationship between the size of the test and the sample size, T, or the variance of the disturbance, σ^2.

The results for the quadratic specification are certainly interesting. For unordered data, where $x_t \sim U$ and $x_t \sim N$, the test has almost no power. The power is noticeably above the size in under 25 per cent of cases considered here. In general, no pattern emerges regarding

[32] A notable exception is the software `R` which includes routines to carry out the Harvey-Collier (1977) test.

[33] It should be noted that of all the tests used in this simulation, the Harvey-Collier (1977) was the fastest. While this fact may be of little value in many practical applications, in the context of a Monte Carlo study, it certainly is noteworthy.

[34] Harvey (1990), p. 157.

sample size or parameter values. By contrast, for the ordered data, the power of the test is 100 per cent in all but three cases and these are for the smallest sample size. Similar results can be observed when x_t is a time trend and hence intrinsically ordered; the test has full power, i.e., 100 per cent, regardless of sample size or parameter values. As the nonlinearity here is so stark, this is not unexpected.

For the so-called square specification, the power is very low for $x_t \sim U$ and $x_t \sim N$. Only in one case is the power notably above the size. Where x_t is a time trend and therefore ordered, however, the test generally has full power, with just one case being marginally below 100 per cent. Power is also low for the x_{4t} specification, with lower powers found for higher values of σ^2. Again, sample size appears to have little effect. This contrasts greatly with the results for ordered data, however. The power is very much greater in all cases. In some instances, it is low for small sample sizes, but improves greatly with T. Only in the case of the x_{4t} with large σ^2 is the power low for all T.

A very clear pattern emerges for the autoregressive specification with unordered data. Power increases with both sample size and the autoregressive parameter ρ, although in no case does it exceed 35 per cent and for $\rho = 0.1$ the power is not noticeably greater than the size, regardless of T. A very similar pattern emerges when x_t is a time trend. For the $x_t \sim U$ and $x_t \sim N$ cases, however, power approximates to the size in almost all cases, regardless of parameter values, although some sensitivity to sample size is evident. These results are interesting as the autoregressive specification is essentially linear. It could be argued, however, that the autoregressive error constitutes a misspecification or nonlinearity. This is certainly not picked up by the test using ordered data, with the exception of the cases where data is trended. The Harvey-Collier (1977) test appears to be robust to autoregressive residuals. On the other hand, of course, it has no power to detect them. Once again, the importance of ordering is clear.

3.5.3 Ramsey's RESET test

The results for the RESET test can be found in Table B.14. The size of this test is very close to 5 per cent for all cases, regardless of T or other parameter values. For the quadratic specification, it has full power in every case. This is unsurprising given the nature of the misspecification here. This test also performs very well for the square specification. For x_{4t}, power increases with T, but decreases with σ^2. For all but the highest value of σ^2, power is 100 per cent for $T = 225$. For the trending data, like the quadratic case, which has a very similar specification, power is 100 per cent for all parameters. A similar pattern is found for $x_t \sim U$ and $x_t \sim N$, with power being 100 per cent in many cases. Only where σ^2 is high does power fall, although it increases towards 100 per cent with increasing T.

The autoregressive specification reveals similar results to the Harvey-Collier (1977) test. For the trending data, power increases with ρ and T to a maximum of approximately 65 per cent. For lower values of ρ, power is close to size. For the remainder, where $x_t \sim U$ and $x_t \sim N$, power is close to size in most cases. Once again, if this specification is considered linear, the RESET test is robust to autoregressive errors where $x_t \sim U$ and $x_t \sim N$. If this specification is viewed as a misspecification, this test fails to pick it up under such conditions.

The reverse is true for the trending data.

Finally, the results of the Hamilton (2001) specification, the simple threshold model, also show a clear pattern. The RESET test clearly has power in detecting nonlinearity of this form. Where x_{1t} and x_{2t} are $\sim U$ and $\sim N$, respectively, the power is over 90 per cent for $T = 225$. Although low for small sample sizes, the power increases steadily with T. Only when x_{1t} is distributed either uniformly or normally and x_{2t} is trended, does the power fall. In both cases, power is less the 80 per cent for $T = 225$. Given that the RESET test is routinely available in most econometric packages, its performance here is particularly noteworthy.

3.5.4 Hamilton's LM test

Before comparing the performance of this test with the preceding results, the performance of the $\lambda_H^E(\boldsymbol{g})$ test is reported. These results can be found in Table B.15.

The size of the test is close to 5 per cent, although in the majority of cases, the size is below 5 per cent. Size appears to be effected somewhat by sample size. For the quadratic specification, the power of the test is 100 per cent in all cases. The power is generally very high for the square specification with trended data and also when $x_t \sim U$ and $x_t \sim N$. The notable exceptions are for the smaller sample sizes ($T = 25$, 75) and for the largest variance of the disturbances ($\sigma^2 = 400$). Even in these cases, the power exceeds 70 per cent for $T \geq 125$.

As with the Harvey-Collier (1977) and RESET tests, a clear pattern emerges for the autoregressive specification. For the trended data, the power approaches 100 per cent for $T = 225$ and $\rho = 0.9$. For the lowest value of ρ (0.1), the power does not exceed 10 per cent. These results do not depend on the other parameter values. For $x_t \sim U$ and $x_t \sim N$, the power is approximately equal to size. In a small number of cases, it exceeds 10 per cent, but no obvious pattern is evident. As with tests previously examined, the $\lambda_H^E(\boldsymbol{g})$ test is robust to autoregressive errors, except where the data is time trending. Finally, for the Hamilton (2001) specification, the power of the test is high. For all cases where $T \geq 75$, the power is above 90 per cent and increases with T to 100 per cent for $T = 225$. Where $T = 25$, the power ranges from 28 per cent to 39 per cent, which is quite low, although perhaps untroubling given the very small sample size and the considerable improvement as T increases. Next these results are compared with the performance of the three tests considered previously.

It is clear from the Durbin-Watson (1950) test results that the test is not very powerful for some of the models used here, when data remains unordered. The $\lambda_H^E(\boldsymbol{g})$ test clearly performs better. It is worth comparing, however, the performance of the Durbin-Watson test using ordered data with $\lambda_H^E(\boldsymbol{g})$ test. For the quadratic and square specifications, the power of the $\lambda_H^E(\boldsymbol{g})$ test is marginally better, but the differences are small. Comparisons for the autoregressive specification are difficult to make. If, on the one hand, autoregressive errors are to be viewed as a form of nonlinearity, the Durbin-Watson test out-performs the $\lambda_H^E(\boldsymbol{g})$ test by some way, perhaps unsurprisingly given that the Durbin-Watson is a test for autocorrelation. If autoregressive errors are not viewed as nonlinearity, no meaningful comparison can be made as the tests have different objectives. In the case that the autoregressive specification is viewed as a linear model with autocorrelated errors, the $\lambda_H^E(\boldsymbol{g})$ test is sensitive to high values of ρ combined with trending data, i.e., it frequently rejects the null of linearity in favour of the alternative. Finally, both tests perform very well on the Hamilton (2001) specification, with very small differences in power between them. Choosing which test is more powerful, given the above confusion is difficult, but it does appear that the $\lambda_H^E(\boldsymbol{g})$ test is slightly more powerful.

Only the results obtained using ordered data for the Harvey-Collier (1977) test are used here. For the quadratic specification, the results for the $\lambda_H^E(\boldsymbol{g})$ and Harvey-Collier tests are almost identical, with the Harvey-Collier test reporting power in several cases just below 100 per cent, while the $\lambda_H^E(\boldsymbol{g})$ test records full power in every case. Both tests return very similar results for the square specification with x_{4t}. For the remaining cases, the Harvey-Collier test appears to be marginally more powerful. This is particularly noticeable when the sample size is small and σ^2 is high, although the power of both tests tends towards 100 per cent as sample size increases, regardless of other parameters. For the autoregressive specification, the Harvey-Collier test again appears to be somewhat more powerful than the $\lambda_H^E(\boldsymbol{g})$ test. For trending data, the powers attained by the Harvey-Collier test are considerably lower then the $\lambda_H^E(\boldsymbol{g})$ test. For the remaining cases, the power of both tests is not significantly greater than their respective sizes. As previously stated, the Hamilton (2001) specification, being multivariate, was not included in the Harvey-Collier simulations, so no comparison can be made here. Given this drawback and ignoring the results of the trended autoregressive case, there is little difference between the performance of these tests in the cases explored here. The greater applicability of the $\lambda_H^E(\boldsymbol{g})$ test may be the deciding factor.

Consider next the comparative performance of the $\lambda_H^E(\boldsymbol{g})$ and RESET tests. Both tests prove effective in testing the quadratic specification, with full power observed in every case. As with the Harvey-Collier (1977) test, RESET proves to be marginally better for the square specification, although as previously noted, both tests perform very well. The differences here are small, with RESET's power increasing faster with T. Again, RESET's power is lower for the trended autoregressive case than that of the $\lambda_H^E(\boldsymbol{g})$ test and there is little difference in the remaining cases. For the Hamilton (2001) specification, the $\lambda_H^E(\boldsymbol{g})$ test is considerably better. For $T \geq 75$, its power is above 90 per cent, regardless of all other parameters. The RESET test is sensitive to the inclusion of trended data for x_{2t} and in these cases power remains below 80 per cent for maximum T. In the remaining cases, the power increases steadily with T, from just less than 60 per cent when $T = 75$, to over 90 per cent when $T = 225$. By contrast, the $\lambda_H^E(\boldsymbol{g})$ test's power is above 90 per cent for all cases when $T \geq 75$. If the results, excluding the trended autoregressive cases, are reviewed, while there is little real difference, the superior performance of the $\lambda_H^E(\boldsymbol{g})$ test for the threshold specification may once again make it preferable, as the results for the other specifications are very close. However, given that Keenan (1985), Tsay (1986) and Luukkonen, et al. (1988) all proposed potentially more powerful variants of the RESET test, it would be interesting to compare their relative performance to that of the $\lambda_H^E(\boldsymbol{g})$ test. If those tests are indeed more powerful than RESET, they may well be more powerful than the $\lambda_H^E(\boldsymbol{g})$ test for the models considered here. This remains to be explored by future research.

3.5.5 The $\lambda_{OP}^E(\boldsymbol{g})$, λ_{OP}^A and g_{OP} tests

The results for the three tests proposed by Dahl and González-Rivera (2003) are reported in tables B.16 to B.21. As with the $\lambda_H^E(\boldsymbol{g})$ test, the individual performance of each test will be first assessed before reviewing the comparative performance. These tests will be assessed relative to the $\lambda_H^E(\boldsymbol{g})$ test, which itself has been compared with the Durbin-Watson (1950), Harvey-Collier (1977) and RESET tests, but also to the simulation results obtained by Dahl and González-Rivera (2003). As noted in Section 3.3, both asymptotic and bootstrapped p-values were recorded for all three tests. In the analysis which follows, primary consideration will be given to the bootstrapped p-values. Unlike the results presented thus far, which are based on 20,000 Monte Carlo replications, the bootstrapped p-values were obtained using 1,000 Monte Carlo replications with 100 re-samples for bootstrapping.

The bootstrapped p-values for the λ_{OP}^A test can be found in Table B.16. The size is approximately 5 per cent, although values as low as 3.5 per cent are observed. The test's power is 100 per cent for all cases of the quadratic specification. For the square specification with x_{4t}, the power decreases with σ^2, but increases with T. For the remaining cases, the power is 100 per cent, except where $\sigma^2 = 400$. Power, however, is above 90 per cent in those cases where $T \geq 125$. For the autoregressive specification a familiar pattern emerges. When x_t is trended, the power approaches 100 per cent as T and ρ increase. For the remainder of the cases, power is not dissimilar to size. The power is high for the Hamilton (2001) specification. It is at its lowest, 36.9 per cent, for $T = 25$, but is above 97 per cent for all cases when $T \geq 75$. Interestingly, the results for the asymptotic p-values are not dissimilar, with the exception of the size, as seen in Table B.17. The size is generally found to be between 1 per cent and 2 per cent, but the power results are very close indeed to those found in Table B.16. This could have important implications, given the computationally intensive and time consuming nature of obtaining bootstrapped p-values.

Results for the bootstrapped p-values of the $\lambda_{OP}^E(\boldsymbol{g})$ test can be found in Table B.18. The size of this test is once again approximately 5 per cent, but with some degree of variability, shown by a size in once case of 7 per cent. For the quadratic specification, in all but one case, power is 100 per cent and in every case above 99 per cent. The power of the square specification with x_{4t} is very sensitive to σ^2; when $\sigma^2 = 1$, power is 100 per cent for $T \geq 75$, which falls to approximately 5 per cent for $\sigma^2 = 400$, for all T. For the remaining cases, power is 100 per cent for all but one where data is trended. For $x_t \sim U$ and $x_t \sim N$, power is 100 per cent for the lowest value of σ^2, but it is again sensitive where $\sigma^2 \geq 25$. In these cases, power increases to 81 per cent and 71 per cent, respectively, for maximum T and σ^2.

The quadratic specification yields similar results. Power increases with ρ to approximately 100 per cent for maximum T and ρ, for trended data. For $\rho = 0.1$, however, power approximates to size. For the remainder of cases, the power also approximates to size. Finally, for the Hamilton (2001) specification, power is above 98 per cent for all cases where $T \geq 125$ and above 75 per cent for $T \geq 75$. For $T = 25$, powers below 10 per cent are observed. Once again, the asymptotic p-values, which can be found in Table B.19 tally well with the bootstrapped results, again with the exception of size, but now also for the case of the autoregressive specification where $x_t \sim U$ and $x_t \sim N$.

The final test examined was the g_{OP} test, the bootstrapped results of which can be found in Table B.20. The size of the test was found to be approximately 5 per cent. Curiously, as shown in Table B.21, the size of the test is zero in every case, when asymptotic p-values are used. This remains to be explained and highlights the necessity for using the bootstrapped p-values. For the quadratic specification, power was 100 per cent in all but two cases and was above 67 per cent in all cases. For the square specification, the results are similar to the other tests. The power of the test is sensitive to σ^2 and T. For $T = 225$, however, power is above 90 per cent for all cases except for x_{4t} when $\sigma^2 = 400$. The pattern observed in other tests is again found here for the autoregressive specification. For the trended data, the power is high for larger values of ρ and T. For the other cases, power approximates to size. Finally, for the Hamilton (2001) specification, the power is low for small sample sizes, but increases steadily, to lie above 89 per cent in all cases where $T \geq 125$.

By contrast with the $\lambda_{OP}^{E}(\boldsymbol{g})$ and λ_{OP}^{A} tests, the asymptotic p-values of the g_{OP} test, which can be found in Table B.21, seem much less reliable. The size is found to be 0 per cent and although most cases for the quadratic specification tally with the bootstrapped results, for $T = 25$, the power is 0 per cent. The same can be said for the remaining specifications; while some results are similar, many are not. In this case, the asymptotic p-values seem very unreliable.

Of the three tests, the λ_{OP}^{A} test performs best in terms of power, across the range of specifications and data used here. It outperforms the $\lambda_{OP}^{E}(\boldsymbol{g})$ test for the quadratic specification, although as previously noted, the differences are very small. It outperforms the g_{OP} test in the square specification, where there are some notable differences in power for the small sample sizes. Even larger differences are seen between the λ_{OP}^{A} and the $\lambda_{OP}^{E}(\boldsymbol{g})$ test. The g_{OP} test appears to perform best on the trending data of the autoregressive specification, and for other data its power is close to size. Once again, however, if the autoregressive specification is considered to be a linear model, all three tests are sensitive to high values of ρ combined with trended data. Finally, for the Hamilton (2001) specification, the λ_{OP}^{A} test clearly outperforms the $\lambda_{OP}^{E}(\boldsymbol{g})$ test, again with some notable differences in power at $T = 25$. Once again, these differences are even larger when compared to the g_{OP} test. Overall, therefore, the λ_{OP}^{A} test is the most powerful, in terms of the data and specifications used here, of the three tests proposed by Dahl and González-Rivera (2003).

Having established this fact, just the λ_{OP}^{A} test performance will be compared to the $\lambda_{H}^{E}(\boldsymbol{g})$ test. It should be noted, of course, that the p-values reported in the Hamilton (2001) case were asymptotic and not bootstrapped.[35] The empirical sizes of the two tests appear to be equivalent. The performance on the quadratic specification is identical. Slight differences emerge for the square specification. While there is little difference in power with x_{4t}, for other data types, the λ_{OP}^{A} test is less sensitive to σ^2 and T. The tests perform comparably with the autoregressive specification, although the power of the $\lambda_{H}^{E}(\boldsymbol{g})$ test appears to be more in line with its size for untrended data. Finally, the λ_{OP}^{A} test performs best for the Hamilton specification, although the differences are very small and only noticeable for $T \leq 75$.

Overall, therefore, it is judged that the λ_{OP}^{A} test is the most powerful over the range of

[35] Dahl and González-Rivera (2003) have made GAUSS code available to compute bootstrapped p-values for the $\lambda_{H}^{E}(\boldsymbol{g})$ test. It was not used in this study.

specifications and data used. It outperformed the $\lambda_H^E(\boldsymbol{g})$ test, which in turn outperformed, although in some cases only just, the Durbin-Watson (1950), Harvey-Collier (1977) and RESET tests. These results very much confirm the findings of Dahl and González-Rivera (2003), that the λ_{OP}^A test performs well across a range of models. They too find that the λ_{OP}^A is more powerful than either the λ_H^E and $\lambda_{OP}^E(\boldsymbol{g})$. Interestingly, they report that the g_{OP} test is particularly successful with bilinear models. Such models have not been included here.

3.6 Conclusion

This chapter has assessed the power of several tests of nonlinearity, some well known, across a range of specifications often encountered in economics and econometrics. These results present some clear conclusions.

The well-known tests of Durbin and Watson (1950), Harvey and Collier (1977) and RESET are powerful against misspecification and nonlinearity, particularly when the former tests are applied to ordered data. This is particularly noteworthy as such ordering approximates the data examined here to many economic time series, which are slowly changing. Given the relative simplicity of these tests and their wide availability, with the exception of Harvey-Collier (1977), these results certainly endorse their use.

The $\lambda_H^E(\boldsymbol{g})$ test does offer a more powerful solution, but this increased power is small and given its more complex nature and lack of widespread availability, it may remain under-utilised. Of the three tests proposed by Dahl and González-Rivera (2003), the λ_{OP}^A test appears to be most powerful across the range of specifications examined here. This of course refers to powers based on bootstrapped p-values. Interestingly, there does not appear to be a large difference between the powers obtained from the bootstrapped p-values and the asymptotic p-values, despite the relatively small sample sizes used. This is not the case for the g_{OP} test, where asymptotic and bootstrapped p-values differ considerably. Also, the λ_{OP}^A test appears to be somewhat more powerful than $\lambda_H^E(\boldsymbol{g})$ test, although recall that the powers for this test are based on the asymptotic p-values. All of the random field-based tests appear to be sensitive to a linear model specification with autocorrelated errors, particularly in the case of trended data.

Avenues for further research would be to compare the performance of the bootstrapped $\lambda_H^E(\boldsymbol{g})$ test with its asymptotic equivalent; to consider the performance of the Keenan (1985), Tsay (1986) and Luukkonen, et al. (1988) tests, adaptations of RESET, against the random field methods discussed here; and to consider a wider range of model specifications and data, including nonnormal distributions, to gain a greater understanding of the properties of these tests.

Chapter 4

Theories and Concepts

This chapter is based on material taken from Banerjee, Dolado, Galbraith, and Hendry (1993), Davidson and MacKinnon (2004), Enders (1995), Franses (1998), Greene (2003), Hamilton (1994), Johansen (1996) and Stewart and Gill (1998).

4.1 Introduction

The aim of this chapter is to provide the background theory required for the remaining chapters of this thesis. Those remaining chapters aim to compare and contrast the results of modelling time-series relationships using the Hamilton (2001) methodology, as outlined in Chapter 2 and further discussed in Chapter 3, with a variety of alternative methods. These alternative methods share the common trait that they exploit the concept of (co)integration in modelling economic relationships. Naturally, this work follows on from the work of Chapter 2, which investigated in detail the implementation of the Hamilton procedure. Chapters 5, 6 and 7 will further this work by evaluating this procedure's potential as a modelling tool. Three approaches to modelling economic relationships will be compared with Hamilton's: the Engle-Granger (1987) 2-Step approach, Johansen's (1988, 1991) vector autoregressive (VAR) approach and common factor (COMFAC) analysis.[1] This chapter will review each of these methods in turn. Attention will also be given to the work of Dolado, Gonzalo, and Mayoral (2002) and Johansen (2002), which may offer further insight into the implementation, application and results of some of the above methods. To illustrate all of the above-mentioned methods, three case studies will be used in the remaining chapters: the demand for money in Chapter 5, the theory of purchasing power parity in Chapter 6 and the forward exchange rate anomaly in Chapter 7. Background theory and information on data used will be supplied as necessary.

The structure of this chapter is as follows. Section 4.2 will provide the background theory necessary for the subsequent chapters. The concepts and definitions outlined here will be used routinely throughout the remainder of the chapter and, indeed, the remainder of the thesis. Section 4.3 provides details on the three methodologies to be used in this study: the Engle-Granger (1987) 2-Step approach, Johansen's (1988, 1991) VAR approach and COMFAC analysis. Sections 4.4 and 4.5 outline the more recent developments of Dolado, et al. (2002) and Johansen (2002). This work has yet to be given textbook treatment and is, therefore, quite detailed, although unnecessary technicalities are avoided. It should be stressed from the outset that the intention here is to give just an overview of these new techniques.

4.2 Stationarity, Nonstationarity and Unit Root Processes

This section introduces briefly some key concepts that will be referred to routinely in the remainder of this chapter. It is felt prudent, therefore, to give these concepts a brief treatment at the outset. The section begins by introducing stationarity and nonstationarity. Tests for stationarity are then considered, before the introduction of the concept of cointegration.

4.2.1 Concepts and definitions

To begin, define a stochastic process as an ordered sequence of random variables, $\{x(s,t),\ s \in \mathcal{S},\ t \in \mathcal{T}\}$, such that for each $t \in \mathcal{T}$, $x(\cdot,t)$ is a random variable on the sample space $\mathcal{S}$, and for each $s \in \mathcal{S}$, $x(s,\cdot)$ is a realisation of a stochastic process on the index set $\mathcal{T}$. Given a

[1]See, for example, Hendry and Mizon (1978).

stochastic process, it can be said that such a process is strictly stationary if, for any subset $(t_1,\ t_2,\ \ldots,\ t_n)$ of $\mathcal{T}$ and any real number h, such that $t_i + h \in \mathcal{T}$, $i = 1,\ 2,\ \ldots,\ n$, there exists

$$F\left(x(t_1),\ x(t_2),\ \ldots,\ x(t_n)\right) = F\left(x(t_1+h),\ x(t_2+h),\ \ldots,\ x(t_n+h)\right), \tag{4.1}$$

where $F(\cdot)$ is the joint distribution function of the n-values. Strict stationarity implies that all existing moments of the process are time invariant. Also, the process is said to be weakly stationary, second order stationary, or covariance stationary, if

$$\begin{aligned} E\left[x(t_i)\right] &= E\left[x(t_i+h)\right] = \mu_1 < \infty, & (4.2)\\ E\left[(x(t_i))^2\right] &= E\left[(x(t_i+h))^2\right] = \mu_2 < \infty, & (4.3)\\ E\left[x(t_i)x(t_j)\right] &= E\left[x(t_i+h)x(t_j+h)\right] = \mu_{ij} < \infty, & (4.4) \end{aligned}$$

where μ_1, μ_2 and μ_{ij} are constant over t, for all $t \in \mathcal{T}$, and h, such that $t_r + h \in \mathcal{T}$, for $(r = i,\ j)$.[2]

The concept of stationarity is necessary for understanding integration. An integrated process can be defined as a process with no deterministic component and which has a stationary and invertible autoregressive moving average (ARMA) representation after differencing d times, but which is not stationary after differencing only $d-1$ times. Such a process is said to be integrated to order d, denoted $x_t \sim I(d)$. A time series may be integrated, or alternatively, it may be near-integrated or fractionally integrated. The concept of fractional integration was put forward by Granger and Joyeux (1980) and Hosking (1981). A simple fractionally integrated time-series model can defined as

$$(1-L)^d y_t = \varepsilon_t \qquad \text{for } 0 < d < 1, \tag{4.5}$$

where L is the lag operator, defined as

$$L^n x_t \equiv x_{t-n}, \tag{4.6}$$

and $(1-L)^d$ can be expanded infinitely as

$$\begin{aligned} (1-L)^d &= 1 - dL + \frac{1}{2!}\, d(d-1)L^2 - \frac{1}{3!}\, d(d-1)(d-2)L^3 + \ldots \\ &\quad + \frac{(-1)^j}{j!}\, d(d-1)\ldots(d-j+1)L^j + \ldots, \end{aligned} \tag{4.7}$$

which reduces to 1 for $d = 0$ and $(1-L)$ for $d = 1$. When $0 < d < 0.5$, the time series is said to be stationary with long memory, and when $0.5 \leq d < 1$, it is said to be nonstationary with long memory.[3] Baillie (1996) and Parke (1999) give an account of long memory and fractional integration in economics and econometrics.

---

[2] Banerjee, et al. (1993), Chapter 1.

[3] More formally, series can be classified as stationary with anti-persistence when $-0.5 < d < 0$, stationary with long memory when $0 < d < 0.5$, nonstationary with long memory when $0.5 \leq d < 1$ and nonstationary with strong long memory when $1 < d < 1.5$. See, for example, Tsay and Chung (2000) for further details.

4.2.2 Testing for unit roots

An $I(d)$ series that becomes stationary by differencing once contains exactly one unit root. Many economic time series may contain an exact unit root in the logarithmic transformations routinely applied to such series. Otherwise, roots very close to, but slightly greater than unity, imply nonstationary series that are not $I(d)$ for any d. Roots slightly less than unity generate near-integrated series. Such processes will tend to be difficult to distinguish from those with a root of exactly unity, particularly in moderately sized samples. Roots substantially greater than unity are characterised by explosiveness in the series. Following Banerjee, et al. (1993), consider the following data generating process

$$y_t = \rho y_{t-1} + \varepsilon_t, \tag{4.8}$$

where $y_0 = 0$ and $\varepsilon_t \sim$ i.i.d.$(0,\ \sigma_\varepsilon^2)$. The null hypothesis $H_0{:}\rho = \rho_0$ for $|\rho_0| < 1$, can be tested by regressing Equation (4.8), and by considering the t-statistic, defined by

$$t_{\mathrm{OLS}_{\widehat{\rho}}} = \frac{\widehat{\rho} - \rho_0}{\mathrm{SE}(\widehat{\rho})}, \tag{4.9}$$

which has, asymptotically, a standard normal distribution. For $\rho_0 = 1$, however, this result is no longer valid. The distribution of the t-statistic given that $\rho_0 = 1$, is not asymptotically normal or symmetric. Dickey and Fuller (1979, 1981) reported the critical values required for the following three models:

$$y_t = \rho_1 y_{t-1} + \varepsilon_t, \tag{4.10}$$

$$y_t = \mu_2 + \rho_2 y_{t-1} + \varepsilon_t, \tag{4.11}$$

$$y_t = \mu_3 + \gamma_3 t + \rho_3 y_{t-1} + \varepsilon_t. \tag{4.12}$$

Clearly, the first model contains no trend or intercept term, the second contains an intercept, μ_2, and the third contains both an intercept, μ_3, and a trend, t. Introducing the difference operator Δ, where $\Delta x_t \equiv x_t - x_{t-1}$, equations (4.10), (4.11) and (4.12) can be rearranged as,

$$\Delta y_t = \phi_1 y_{t-1} + \varepsilon_t, \tag{4.13}$$

$$\Delta y_t = \mu_2 + \phi_2 y_{t-1} + \varepsilon_t, \tag{4.14}$$

$$\Delta y_t = \mu_3 + \phi_3 y_{t-1} + \gamma_3 t + \varepsilon_t, \tag{4.15}$$

where $\phi_i = \rho_i - 1$ and $\varepsilon_t \sim$ i.i.d. $\left(0, \sigma_\varepsilon^2\right)$. The Dickey-Fuller (DF) test consists of regressing one of these equations and considering the resultant t-statistic

$$t_{\mathrm{DF}} = \frac{\widehat{\phi_i} - 1}{\mathrm{SE}(\widehat{\phi_i})}, \tag{4.16}$$

where the null and alternative hypotheses are $H_0{:}\phi_i = 0$ and $H_1{:}\phi_i < 0$, respectively. The DF test, however, assumes white noise disturbances. The test can be extended to allow for some forms of serial correlation, thereby becoming the augmented Dickey-Fuller (ADF) test.

The model

$$y_t = \mu + \gamma y_{t-1} + \gamma_1 \Delta y_{t-1} + \ldots + \gamma_p \Delta y_{t-p} + \varepsilon_t, \tag{4.17}$$

is transformed by first-differencing to give

$$\Delta y_t = \mu + \gamma^* y_{t-1} + \sum_{j=1}^{p-1} \varphi_j \Delta y_{t-j} + \varepsilon_t, \tag{4.18}$$

where $\varphi_j = -\sum_{k=j+1}^{p} \gamma_k$ and $\gamma^* = (\sum_{i=1}^{p} \gamma_i) - 1$. The limiting distribution of the DF and ADF tests is nonstandard and depends on the deterministic terms in the model, but does not depend on φ_j. The same critical values are, therefore, applicable to both tests. Those critical values have been simulated by Fuller (1976), Davidson and MacKinnon (1993) and MacKinnon (1996). In these tests, the number of lagged differences of y_t may be based on model selection criteria, such as the Akaike (1973), Hannan and Quinn (1979) or Schwarz (1978) information criteria. For series with more than one unit root, the Pantula (1989) principle may be applied.[4]

In testing for unit roots with the DF procedure, the question arises as to which of the equations (4.13), (4.14) or (4.15) should be used. Although estimating Equation (4.15), the most general model, might seem appropriate, it reduces the power of the test unnecessarily if either μ_3 or t are not present in the data generating process. This may lead to nonrejection of the null of a unit root, when in fact the process is stationary. Also, different critical values for the DF procedure apply, depending on the model under consideration. It is crucial, therefore, that the data generating process being considered is suitably modelled by the choice of DF equation.

Dolado, Jenkinson, and Sosvilla-Rivero (1990) suggested a method for overcoming the potential for misspecification in the DF regression, when the form of the data generating process is unknown.[5] Firstly, estimate Equation (4.15). If this rejects the null of a unit root, conclude that the series is trend stationary. If the test cannot reject the null, test the significance of the trend term, by imposing the unit root in the DF equation. If the trend is significant, re-test for the presence of a unit root using the standardised normal distribution.[6] If the trend is not found to be significant, estimate Equation (4.14) and re-run the DF unit root test. If the null is rejected, again, it can be concluded that there is no unit root. Otherwise, test the significance of the intercept. As before, if it is significant, re-test using the standardised normal critical values. If it is not significant, estimate Equation (4.13) and test for a unit root in the normal manner.

Consideration here has been limited to just the DF-type testing of unit roots. Alternative test procedures do exist. A popular alternative is the method suggested by Phillips (1987) and Phillips and Perron (1988), whose tests generalise the DF framework. Whereas the DF test

[4]Pantula (1989) suggested a technique for testing unit roots with the Dickey-Fuller test, where the series in question is $I(d)$ and d is some positive unknown integer. In other words, the series may contain one or more unit roots. The procedure suggests differencing the series as many times as is necessary to make it stationary (this will of course imply differencing the series d times, but recall d is unknown!). If the hypothesis of a unit root is rejected, as expected if in fact d is some positive integer, then the $d - n$-differenced series are sequentially tested for unit roots, where $n = 1, 2, \ldots,$ until the hypothesis of a unit root cannot be rejected.

[5]This exposition follows closely that which can be found in Enders (1995).

[6]This result is due to Savin and Nankervis (1985).

assumes that errors are statistically uncorrelated and have constant variance, the Phillips and Perron test relaxes these assumptions. The same critical values apply to both procedures, however. The ADF generalised least squares test, attributable to Elliot, Rothenberg and Stock (ERS) (1996), allows for very general formulations of the error term. Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) (1992) proposed a test in which the null is stationarity, in contrast to the other methods mentioned here. Leybourne and McCabe (1994, 1999) proposed a variant of this test that corrects for any serial correlation in the data generating process. Hall (1989) suggested an approach based upon an instrumental variables estimation. Bhargava's (1986) testing methodology employs von Neumann-type ratios. Ng and Perron (NP) (2001) proposed a battery of tests based on those of Bhargava (1986) and Phillips and Perron.[7] Excellent reviews of some of these methods may be found in Dolado, et al. (1990), Phillips and Xiao (1998) and Bierens (2001).

A consistent estimate of the zero frequency residual spectrum is required to carry out each of the KPSS, ERS and NP tests. Two classes of estimators are typically available, namely, spectral and moment estimators. Several estimates are available within each class. For example, the moment, or kernel sum-of-covariance class of estimators, as they are also known, are made up of the Bartlett (1950), Parzen (1961) and Quadratic spectral kernel estimators.[8] For spectral estimation, either the Newey and West (1994) or Andrews (1991) automatic bandwidth parameter selection methods may be used. Several autoregressive spectral density estimators are available within the spectral class, including OLS- and GLS-based estimators. The usual selection criteria can be used to choose the appropriate lag length in these cases.

Seasonal Unit Roots

The illustrative examples that are contained within the subsequent chapters of this thesis employ a variety of data, many of which are quarterly in periodicity. This raises the possibility that the variables therein could be seasonally integrated. This concept is directly analogous to that of integration and can be formally defined as[9]

> The nonstationary stochastic process y_t, observed at S equally spaced time intervals per year, is said to be seasonally integrated of order d, denoted $y_t \sim SI(d)$, if $\Delta_S^d y_t$ is a stationary, invertible ARMA process.[10]

Considering the quarterly case, as this will be the most relevant here, such seasonally integrated data may have standard, nonstandard and complex unit roots. This can be seen by considering the Dickey-Fuller type test for the quarterly case:

$$\left(1 - L^4\right) y_t = \pi y_{t-4} + \varepsilon_t, \tag{4.19}$$

where L is the usual lag operator. $\left(1 - L^4\right)$ can be rewritten as $(1 - L)(1 + L)\left(1 + L^2\right)$. The roots of $\left(1 - L^4\right)$ are 1, -1 and $\pm i$, and represent the standard, nonstandard and complex

[7] A fuller description of these tests has not be included here for compactness. Some of the tests are implemented in Chapter 5, but only as supplementary procedures, thus warranting this omission.

[8] See, for example, Andrews (1991) or Priestley (1981).

[9] Following the definition of Ghysels and Osborn (2001).

[10] Δ_S is the seasonal differencing operator, defined as $\Delta_s \equiv \left(1 - L^S\right)$, where L is the lag operator.

unit roots, respectively. Alternatively, those roots can be view as representing the zero, π and $\pi/2$ spectral frequencies.

Several tests have been developed to examine data for seasonal unit roots for various values of S. Tests include those by Dickey, Hasza, and Fuller (1984), Osborn, Chui, Smith, and Birchenhall (1988) and Hylleberg, Engle, Granger, and Yoo (1990), widely known as the HEGY test. Testing in this thesis will be undertaken with the HEGY test, not least because the test can be easily carried out using the software `JMulTi`, further details of which can be found in Lütkepohl and Krätzig (2004). What follows is a brief description of that method.

The HEGY test, for the case of quarterly observations, is based on the model

$$\Delta_4 y_t = \pi_1 z_{1,t-1} + \pi_2 z_{2,t-1} + \pi_3 z_{3,t-1} + \pi_4 z_{3,t-2} + \sum_{j=1}^{p} \alpha_j \Delta_4 y_{t-j} + \varepsilon_t, \tag{4.20}$$

where

$$\begin{aligned} z_{1t} &= \left(1 + L + L^2 + L^3\right) y_t, \\ z_{2t} &= -\left(1 - L + L^2 - L^3\right) y_t, \\ z_{3t} &= -\left(1 - L^2\right) y_t. \end{aligned} \tag{4.21}$$

Testing for regular, semi-annual and quarterly roots involves testing the null hypotheses $H_0{:}\pi_1 = 0$, $H_0{:}\pi_2 = 0$ and $H_0{:}\pi_3 = \pi_4 = 0$, respectively, using t- and F-type tests. `JMulTi` tests these hypotheses along with $H_0{:}\pi_2 = \pi_3 = \pi_4 = 0$ and $H_0{:}\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$. Critical values for these tests are provided by Franses and Hobijn (1997). Equation (4.20) can be estimated by OLS and p, the number of lagged differences, can be determined by the usual information criteria.

Deterministic terms can be included in Equation (4.20). In fact, Ghysels and Osborn (2001) recommend that both a linear trend and seasonal dummy variables should be included when testing the $SI(1)$ null hypothesis. Caution must be taken, however, when using this methodology. The F_{1234} test, in the notation used by Lütkepohl and Krätzig (2004), that which tests the hypothesis $H_0{:}\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$, should not be viewed as a test for seasonal integration against an alternative of stationarity at all frequencies. Rejection of the null may be due to stationarity at one or more, but not all, of the frequencies. Also, da Silva Lopes and Montanes (1999) showed that 'asymptotically the HEGY test statistics can distinguish between a structural break in the deterministic seasonal component and seasonal unit roots, but that empirically the presence of a structural break will reduce the power of these tests'.[11]

Finally, the concept of fractional integration or near integration, also discussed in this thesis, can be applied to seasonal processes; see, for example, Rodrigues (2001) and Rodrigues and Osborn (1999). As this is deemed beyond the scope of this thesis, no further consideration is warranted. Next, consider the concept of cointegration.

[11] Ghysels and Osborn (2001), p. 74.

4.2.3 Cointegration

An n-vector time series $\mathbf{x}_t$ is said to be cointegrated if each of the series taken individually is $I(1)$, that is, nonstationary with a unit root, while some linear combination of all the series, $\boldsymbol{\alpha}'\mathbf{x}_t$, is stationary, or $I(0)$, for some nonzero n-vector $\boldsymbol{\alpha}$. Cointegration implies that although many developments can cause permanent changes in individual elements of $\mathbf{x}_t$, there is some long-run equilibrium relation tying the individual components together, represented by the linear combination $\boldsymbol{\alpha}'\mathbf{x}_t$. The cointegrating vector is not unique, for if $\boldsymbol{\alpha}'\mathbf{x}_t$ is stationary, then so is $b\boldsymbol{\alpha}'\mathbf{x}_t$ for any nonzero scalar b; if $\boldsymbol{\alpha}$ is a cointegrating vector, then so too is $b\boldsymbol{\alpha}$. When referring to the cointegrating vector, an arbitrary normalisation must be made. Usually, the first element of $\boldsymbol{\alpha}$ is transformed to unity. If there are more than two variables contained in $\mathbf{x}_t$, then there may be two nonzero n-vectors $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, such that $\boldsymbol{\alpha}_1'\mathbf{x}_t$ and $\boldsymbol{\alpha}_2'\mathbf{x}_t$ are both stationary, where $\boldsymbol{\alpha}_1$ and $\mathbf{a}_2$ are linearly independent. Indeed there may be $r < n$ linearly independent n-vectors, $(\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ \ldots,\ \boldsymbol{\alpha}_k)$ such that $\mathbf{A}'\mathbf{x}_t$ is a stationary n-vector $\mathbf{A} \equiv [\boldsymbol{\alpha}_1'\ \boldsymbol{\alpha}_2'\ \ldots\ \boldsymbol{\alpha}_k']$.

More formally, consider the definition offered by Banerjee, et al. (1993), adapted from Engle and Granger (1987):

> The components of the vector $\mathbf{x}_t$ are said to be cointegrated of order d, b, denoted $\mathbf{x}_t \sim CI(d,\ b)$, if $\mathbf{x}_t \sim I(d)$ and there exists a nonzero vector $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}'\mathbf{x}_t \sim I(d-b)$, $d \geq b > 0$. The vector $\boldsymbol{\alpha}$ is called the cointegrating vector. If $\mathbf{x}_t$ has $n > 2$ components, then there may be more than one cointegrating vector $\boldsymbol{\alpha}$. If there exists exactly r linearly independent cointegrating vectors with $r \leq n-1$, then these can be gathered into an $n \times r$ matrix $\mathbf{A}$. The rank of $\mathbf{A}$ will be r and is called the cointegrating rank.[12]

In recent decades, the concept of cointegration has become increasingly important in economics and econometrics. Several important economic relationships have been empirically explored in a cointegration framework, including the demand for money function and purchasing power parity, both of which will be examined more closely in later chapters. Engle and Granger (1987) outlined several examples where cointegrating relationships have been or are likely to be found. Various procedures exist to test for the presence of stationary linear combinations of integrated variables. Some of these tests will be outlined later in this chapter.

4.2.4 Error-correction representation

Given that cointegration implies a long-run relationship tying individual components together, it is straightforward to view this in terms of error correction. An error-correction model is a dynamic model where the short-run dynamics are influenced by the deviations from long-run equilibrium. As such, error-correction models are very useful in representing cointegrating relationships. Consider the autoregressive distributed lag model, or ADL$(1,1)$,

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, \tag{4.22}$$

[12] Banerjee, et al. (1993), p. 145.

where $\varepsilon_t \sim$ i.i.d.$(0,\ \sigma^2)$ and $|\alpha_1| < 1$. Treating $(y_t,\ x_t)$ as jointly stationary, assume also that all change in the above model has ceased. The long-run values are given by the unconditional expectations. Let $y^* = E(y_t)$ and $x^* = E(x_t)$ for all t. Equation (4.22) then becomes

$$y^* = \alpha_0 + \alpha_1 y^* + \beta_0 x^* + \beta_1 x^*, \tag{4.23}$$

or

$$y^* = \frac{\alpha_0 + (\beta_0 + \beta_1)\, x^*}{(1 - \alpha_1)} \equiv k_0 + k_1 x^*, \tag{4.24}$$

which is equivalent to

$$E(y_t) = k_0 + k_1 E(x_t), \tag{4.25}$$

where k_1 is the long-run multiplier of y with respect to x. By further manipulation Equation (4.22) becomes

$$\Delta y_t = \alpha_0 + (\alpha_1 - 1)(y_{t-1} - k_1 x_{t-1}) + \beta_0 \Delta x_t + \varepsilon_t. \tag{4.26}$$

In general, an ADL$(m,\ n;\ p)$ can be defined as:

$$y_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i y_{t+i} + \sum_{j=1}^{p} \sum_{i=0}^{n} \beta_{ji} x_{jt-i} + \varepsilon_t. \tag{4.27}$$

Such models were used by Sargan (1964), Hendry and Anderson (1977) and Davidson, Hendry, Srba, and Yeo (1978), as a way of capturing adjustments in a dependent variable which depended not on the level of some explanatory variable, but on the extent to which the dependent variable deviated from an equilibrium relationship with the explanatory variables.

A generalised error-correction model (ECM) corresponding to the ADL$(m,\ n;\ p)$ model with p exogenous variables $x_1, \ldots, x_p$, by steps similar to those used in the specific case above, allows the direct specification of a general dynamic regression model in the form of an ECM, for $(r \leq m)$:

$$\begin{aligned}
\Delta y_t \quad = \quad & \alpha_0 + \sum_{i=1}^{r} \eta_i \left(y_{t-i} - \sum_{j=1}^{p} x_{jt-i} \right) + \sum_{j=1}^{p} \beta_{j0} \Delta x_{jt} + \sum_{j=1}^{p} \sum_{i=1}^{r} \zeta_{ji} x_{jt-i} \\
& + \sum_{j=1}^{p} \sum_{i=r+1}^{n} \beta_{ji} x_{jt-i} + \sum_{i=r+1}^{m} \alpha_i y_{t-i} + \varepsilon_t,
\end{aligned} \tag{4.28}$$

where $\eta_1 = \alpha_1 - 1$, $\eta_i = \alpha_i$ for $i = 2, \ldots, r$ with $r \equiv \min(m, n)$. Also, $\zeta_{j1} = \alpha_1 - 1 + \beta_{j0} + \beta_{j1}$, $\zeta_{ji} = \alpha_i + \beta_{ji}$, $i = 2, \ldots, r$ and $\Delta x_{jt-i} \equiv (x_{jt-i} - x_{jt-i-1})$.

The ECM is simply a linear transformation of the ADL model. In the ECM formulation, however, parameters describing the extent of short-run adjustment to disequillibrium are immediately provided by the regression.

4.3 Testing for Cointegration

Having reviewed the concepts of integration and cointegration, attention now turns to testing for cointegration. An exposition of three frequently used methods now follows.

4.3.1 The Engle-Granger 2-step method

As suggested by its name, this single-equation method of testing for cointegration has two steps. The first is to test the variables of interest for unit roots, to establish that the variables are indeed integrated to the same order. The DF test discussed in Subsection 4.2.2 can be used for this purpose. If the evidence from this test suggests that the variables are integrated to different orders, or not at all, then the specification of the model should be reconsidered, perhaps by modelling in differences. The second step, given that all the series are found to have unit roots, is to consider whether or not they are cointegrated. This can be done by examining the residuals of a static OLS regression.

Following Davidson and MacKinnon (2004), consider the model

$$\mathbf{x}_t = \mathbf{D}_t\boldsymbol{\phi} + \mathbf{X}_{t2}\boldsymbol{\beta} + \nu_t, \tag{4.29}$$

where $\mathbf{D}_t$ is a deterministic row vector that may not have any elements, $\mathbf{X}_t$ is a k-vector of $I(1)$ variables, $\mathbf{X}_t = [\mathbf{x}_t\, \mathbf{X}_{t2}]$, and $\mathbf{x}_t$ is the regressand. Estimates of $\boldsymbol{\beta}$ can be obtained from Equation (4.29) by OLS. If the variables cointegrate, then this estimate will be superconsistent and the estimated residuals will be superconsistent estimators 'of the particular linear combination of the elements of $\mathbf{X}_t$ that is $I(0)$'.[13] If the variables do not cointegrate, however, the residuals will be $I(1)$.

There are two reasons why the estimates of $\boldsymbol{\beta}$ may not be consistent. First, since both sides of the Equation (4.29) contain $I(1)$ variables, the problem of spurious regression may arise.[14] Second, consistency is unlikely here, as the proposed approach involves estimating the parameters of an ordinary simultaneous equation model by OLS. It is, however, an extraordinary result that in this case, neither of these considerations is a problem. In fact, as shown by Stock (1987), not only is $\widehat{\boldsymbol{\beta}}$, the OLS estimator of $\boldsymbol{\beta}$, consistent, it is superconsistent in that its asymptotic variance is $O(1/T^2)$ rather than $O(1/T)$, as in the usual case, when the null hypothesis of no cointegration is rejected.[15] Consequently, such regressions may not be spurious. The next step, therefore, is to estimate the cointegrating vector(s) by OLS.

Under the assumption of cointegration, the residuals from Equation (4.29), $\widehat{\nu}_t$, are estimates of the equilibrium errors. As such, they should be $I(0)$. An obvious approach would be to apply the familiar DF test to these residuals. The DF critical values are inappropriate, however, for these estimated errors. Estimates of the appropriate critical values for the test are given by Engle and Granger (1987), Engle and Yoo (1987), Phillips and Ouliaris (1990) and Davidson and MacKinnon (1993). If autocorrelation in the equilibrium errors is

[13] Davidson and MacKinnon (2004), p. 636.

[14] Hamilton (1994), p. 557, offers a useful definition of spurious regression. Consider a regression of the form $y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t$, for which elements of y_t and $\mathbf{x}_t$ might be nonstationary. If there does not exist some population value for $\boldsymbol{\beta}$ for which the residual $u_t = y_t - \mathbf{x}_t'\boldsymbol{\varphi}$ is $I(0)$, then OLS may produce spurious results.

[15] See also, for example, Davidson and MacKinnon (1993).

suspected, then an augmented Engle-Granger (AEG) test can be applied. Having obtained residuals from Equation (4.29), the AEG test is performed by running the regression

$$\Delta\widehat{\nu}_t = \mathbf{D}_t\phi + \gamma'\widehat{\nu}_{t-1} + \sum_{j=1}^{p}\delta_j\Delta\widehat{\nu}_{t-j} + \varepsilon_t, \tag{4.30}$$

where p is chosen to remove serial correlation and the asymptotic distribution of the test depends on k. If the null hypothesis of $\gamma' = 0$ cannot be rejected against the alternative $\gamma' < 0$, then it is concluded that the variables are not cointegrated.[16] Two further single-equation approaches warrant a mention and due to their computational simplicity will be considered in later analyses. They are the ECM and cointegrating regression Durbin-Watson tests for cointegration. They are briefly introduced here.

The error-correction mechanism test

An alternative single-equation approach to testing for cointegration is the error-correction mechanism (ECM) test, put forward by Banerjee, Hendry, and Smith (1986). Consider again the ECM model in Equation (4.28). If the variables are not cointegrated, the coefficients of the error-correction terms $\sum_{i=1}^{r}\eta_i\left(y_{t-i} - \sum_{j=1}^{p}x_{jt-i}\right)$ must be zero. Estimating the ECM model by OLS gives a suitable test statistic in the form of the t-statistic for $\eta_i = 0$. Ericsson and MacKinnon (2002) provided suitable critical values for this test, along with programs to compute them. Davidson and MacKinnon (2004) suggested that the ECM test has greater power than the Engle-Granger approach. These test procedures share the disadvantage, however, that they both depend on the choice of the regressor.

The cointegrating regression Durbin-Watson test

The final single-equation method considered here is the cointegrating regression Durbin-Watson (CRDW) test. This test, proposed by Sargan and Bhargava (1983), is computed in the same way as the Durbin-Watson d-statistic. Given that $d \approx 2\left(1-\widehat{\rho}\right)$, and that under the null hypothesis of no cointegration $\rho = 1$, the CRDW tests the null hypothesis that H_0: $\rho = 1$, as opposed to the usual Durbin-Watson d test that H_0: $\rho = 0$, where

$$\text{CRDW} = \frac{\sum_{t=2}^{T}\left(\widehat{\nu}_{it} - \widehat{\nu}_{i,t-1}\right)^2}{\sum_{t=1}^{T}\widehat{\nu}_{it}^2} \tag{4.31}$$

and $\widehat{\nu}_{it}$ are the OLS residuals from the cointegrating regression. Although this method is made easy by the routine reporting of the d-statistic by most econometric software, it is not without its limitations. The test statistic depends on the number of regressors. The bounds of the critical values change as the number of regressors increases. Also, the null of a random walk in the residuals is tested against an alternative of a stationary first-order autoregressive process. This is rather restrictive. Further details can be found in Banerjee, et al. (1993).

[16]It should be noted that errors of Type I and II can be made, regardless of whether the null hypothesis of no cointegration is rejected or not.

4.3.2 Johansen's maximum likelihood approach

Before detailing the nature of Johansen's (1988, 1991) test for cointegration, it is necessary to introduce some further background material. Systems of multiple equation models can be considered as being a collection of single-equation models. These can be written as vector autoregressive (VAR) processes. In vector notation, the n-dimensional, autoregressive process, $\mathbf{X}_t$, where $\mathbf{X}_t = (x_{1t}, \ldots, x_{kt})'$ can be written as[17]

$$\mathbf{X}_t = \mathbf{\Pi}_1 \mathbf{X}_t + \ldots + \mathbf{\Pi}_k \mathbf{X}_{t-k} + \Phi \mathbf{D}_t + \boldsymbol{\varepsilon}_t, \tag{4.32}$$

for $t = 1, \ldots, T$ and where $\boldsymbol{\varepsilon}_t \sim$ n.i.d$(0, \mathbf{\Omega})$, $\mathbf{D}_t$ is a deterministic term containing a constant, trend and seasonal dummies, and $\mathbf{\Pi}_i$ are $k \times k$ coefficient matrices.

Equation (4.32) can be written in error-correction form, to give

$$\Delta \mathbf{X}_t = \mathbf{\Pi} \mathbf{X}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{X}_{t-i} + \Phi \mathbf{D}_t + \varepsilon_t, \tag{4.33}$$

for $t = 1, \ldots, T$, fixed values of $\mathbf{X}_{-k+1}, \ldots, \mathbf{X}_0$, and where $\boldsymbol{\varepsilon}_t \sim$ i.i.d., $\mathbf{\Pi} = \sum_{i=1}^{k} \mathbf{\Pi}_i - 1$, $\mathbf{\Gamma}_i = -\sum_{j=i+1}^{k} \mathbf{\Pi}_j$ and $\mathbf{\Gamma} = \mathbf{I} - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i$. The characteristic polynomial for this equation is

$$A(z) = (1-z)\mathbf{I} - \mathbf{\Pi}_z - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i (1-z) z^i. \tag{4.34}$$

Using this characteristic polynomial, Equation (4.33) can be expressed as

$$A(L)\mathbf{X}_t = -\mathbf{\Pi}\mathbf{X}_t + (\mathbf{\Gamma} + \mathbf{\Pi})\,\Delta\mathbf{X}_t + A^{**}(L)\Delta^2 \mathbf{X}_t = \boldsymbol{\varepsilon}_t + \Phi \mathbf{D}_t, \tag{4.35}$$

where

$$A^{**}(z) = \frac{A(z) - A(1) - \dot{A}(1)(z-1)}{(1-z)^2}, \tag{4.36}$$

and $\dot{A}(1) = -\mathbf{\Pi} - \mathbf{\Gamma}$. In the presence of unit roots, $\mathbf{\Pi}$ has to be singular and it can be represented as $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, for some $n \times r$ matrices, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Equation (4.33) can, therefore, be rewritten as

$$\Delta \mathbf{X}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{X}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{X}_{t-i} + \Phi \mathbf{D}_t + \varepsilon_t, \tag{4.37}$$

for $t = 1, \ldots, T$, and where $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{k-1}$ describe the short-term dynamics, and the effects of levels depends on $\boldsymbol{\alpha}\boldsymbol{\beta}'$. By further refinement, Equation (4.37) can be rewritten as

$$\mathbf{Z}_{0t} = \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{Z}_{1t} + \mathbf{\Psi} \mathbf{Z}_{2t} + \varepsilon_t, \tag{4.38}$$

for $t = 1, \ldots, T$, and where $\mathbf{Z}_{0t} = \Delta \mathbf{X}_t$, $\mathbf{Z}_{1t} = \mathbf{X}_{t-1}$ and $\mathbf{Z}_{2t}$, of dimension $n(k-1)+m$, is the stack of variables $\Delta \mathbf{X}_{t-1}, \ldots, \Delta \mathbf{X}_{t-k+1}$ and $\mathbf{D}_t$. $\mathbf{\Psi}$ is the $n \times (n(k-1)+m)$ matrix of parameters of $\mathbf{Z}_{2t}$, i.e., consisting of $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{k-1}$ and Φ.

[17] Following closely Johansen (1996).

The loglikelihood for this equation is

$$\log L(\boldsymbol{\Psi}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Omega}) = \tag{4.39}$$
$$-\frac{1}{2}T\log(\boldsymbol{\Omega}) - \frac{1}{2}\sum_{t=1}^{T}\left(\mathbf{Z}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{Z}_{1t} - \boldsymbol{\Psi}\mathbf{Z}_{2t}\right)'\boldsymbol{\Omega}^{-1}\left(\mathbf{Z}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{Z}_{1t} - \boldsymbol{\Psi}\mathbf{Z}_{2t}\right).$$

Finally, defining $\mathbf{M}_{ij} = T^{-1}\sum_{t=1}^{T}\mathbf{Z}_{it}\mathbf{Z}'_{jt}$, for $i, j = 0, 1, 2$, $\mathbf{R}_{0t} = \mathbf{Z}_{0t} - \mathbf{M}_{0t}\mathbf{M}_{22}^{-1}\mathbf{Z}_{2t}$, $\mathbf{R}_{1t} = \mathbf{Z}_{1t} - \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{Z}_{2t}$, and $\mathbf{S}_{ij} = T^{-1}\sum_{t=1}^{T}\mathbf{R}_{it}\mathbf{R}_{jt}^{-1} = \mathbf{M}_{ij} - \mathbf{M}_{i2}\mathbf{M}_{22}^{-1}\mathbf{M}_{2j}$, for $i, j = 0, 1$, allows for testing the hypothesis $H(r) : \boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, as it can be shown that the maximisation of the likelihood function is equivalent to solving the equation

$$|\boldsymbol{\lambda}\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}| = 0, \tag{4.40}$$

for the eigenvalues $1 > \widehat{\lambda}_1 > \ldots > \widehat{\lambda}_n > 0$ and for the eigenvectors $\widehat{\mathbf{V}} = (\widehat{\nu}_1, \ldots, \widehat{\nu}_n)$. The maximised loglikelihood function is then given by

$$L_{max}^{-2/T}(H(r)) = |\mathbf{S}_{00}|\Pi_{i=1}^{n}\left(1 - \widehat{\lambda}_i\right). \tag{4.41}$$

This leads to the so-called Trace statistic, the likelihood ratio function $Q(H(r)|H(n))$, for $H(r)$ in $H(n)$,

$$-2\log Q(H(r)|H(n)) = -T\sum_{i=r+1}^{n}\log\left(1 - \widehat{\lambda}_i\right), \tag{4.42}$$

and the Maximal eigenvalue test for $H(r)$ in $H(r+1)$,

$$-2\log Q(H(r)|H(r+1)) = -T\log\left(1 - \widehat{\lambda}_{r+1}\right). \tag{4.43}$$

It is possible that the two different, but related, test procedures will conflict in their findings.

The critical values for the Trace and the Maximal eigenvalue statistic defined in equations (4.42) and (4.43) depend upon $n - r$, whether Equation (4.37) contains an intercept or trend and whether there are restrictions imposed upon these. The distribution of the Trace statistic is derived under the hypothesis that there are r cointegrating vectors and tests H_r within H_n. The testing proceeds in sequence from $\eta_0, \eta_1, \ldots, \eta_{n-1}$. The number of cointegrating vectors selected is $r + 1$ where the last significant statistic is η_r. The test's distribution is not the conventional χ^2 distribution because $\mathbf{X}_t$ is a multivariate $I(1)$ process. The Maximal eigenvalue test, alternatively, tests H_r with H_{r+1}.

Both are distributed as functionals of multivariate Wiener processes. There are no analytical forms of the distributions, but critical values can be obtained by simulation.[18] They are available in tabulated form, from among others, Johansen (1988, 1991), Osterwald-Lenum (1992) and Pesaran, Shin, and Smith (1996). Pesaran, et al. computed these critical values using stochastic simulation techniques. They include cases with and without intercepts and

[18] An approximation for the Trace statistic has been put forward by Johansen (1988) as

$$\text{Trace} = h\chi^2\left(2m^2\right), \tag{4.44}$$

where $h = 0.85 - 0.58/\left(2m^2\right)$, for $m = n - r$.

trends, and the cases where restrictions are imposed upon them. This simulation includes up to twelve endogenous $I(1)$ variables and five exogenous $I(1)$ variables in the vector error-correction model (VECM). In large samples, the critical values do not depend on the order of the VAR or on the stochastic properties of the $I(0)$ exogenous variables. Osterwald-Lenum, following Johansen's approach, also computes the critical values for the Trace and Maximal eigenvalue statistics. In this case, up to eleven endogenous $I(1)$ variables but no exogenous $I(1)$ variables are included. These critical values, like Johansen's, differ from those offered by Pesaran, et al. as they do not consider restrictions to the intercept or trend terms.

Finally, in a cointegrating vector, for unrestricted intercepts and trends, $\mathbf{X}_t$ will be trend stationary when the rank of $\mathbf{\Pi}$ is full. But if it is rank deficient, the solution for $\mathbf{X}_t$ will contain quadratic trends. For unrestricted intercepts and no trends, a rank deficiency in $\mathbf{\Pi}$ will result in $\mathbf{X}_t$ containing linear deterministic trends. To avoid these situations, the choice of restricted intercepts and no trends or unrestricted intercepts and restricted trends is normally made. However, this results in the cointegrating vectors containing a deterministic trend in the first case and intercepts in the second case.

4.3.3 Common factor analysis

Common factor analysis (COMFAC) is perhaps less well known than the above methods. A more detailed account, therefore, is given here. For exposition, the two variable case is considered first, followed by a more generalised account.

Consider a linear regression ADL(1, 1) equation relating y_t to the variable x_t,

$$y_t = \beta_1 y_{t-1} + \gamma_0 x_t + \gamma_1 x_{t-1} + \nu_t, \tag{4.45}$$

where $|\beta_1| < 1$, $\nu_t \sim$ i.i.d.$(0, \sigma^2)$ and is serially uncorrelated. The lag operator, L, allows Equation (4.45) to be rewritten as

$$y_t = \beta_1 L y_t + \gamma_0 x_t + \gamma_1 L x_t + \nu_t, \tag{4.46}$$

or

$$(1 - \beta_1 L)\, y_t = (\gamma_0 + \gamma_1 L)\, x_t + \nu_t. \tag{4.47}$$

In this equation, both variables are multiplied by polynomials in the lag operator. If parameter $\gamma_1 = -\beta_1 \gamma_0$, then Equation (4.47) becomes

$$(1 - \beta_1 L)\, y_t = \gamma_0 (1 - \beta_1 L) x_t + \nu_t. \tag{4.48}$$

Now y_t and x_t have a common factor $(1 - \beta_1 L)$. These polynomials have a common root of $\frac{1}{\beta_1}$. β_1 corresponds to what Hendry and Mizon (1983) refered to as the latent root of equation $|\lambda_1 - \beta_1| = 0$. Dividing both sides of Equation (4.48) by their common factor gives

$$y_t = \gamma_0 x_t + \frac{\nu_t}{(1 - \beta_1 L)}, \tag{4.49}$$

or

$$y_t = \gamma_0 x_t + u_t, \tag{4.50}$$

where

$$u_t = \frac{\nu_t}{(1-\beta_1 L)}, \tag{4.51}$$

or finally,

$$u_t = \beta_1 u_{t-1} + \nu_t. \tag{4.52}$$

This shows that u_t is generated by a first order autoregressive AR(1) process. If the restriction $\gamma_1 = -\beta_1\gamma_0$ is satisfied for Equation (4.45), with one period lagged variables, then the polynomials in L have a common root. This root can be thought of as the serial correlation coefficient of a first-order autoregressive error process, when Equation (4.45) is a static model. Therefore, u_t being AR(1) is a convenience, not a nuisance.[19] In Equation (4.50), only β_1, γ_0 and σ^2 need to be estimated, as opposed to of β_1, γ_0, γ_1 and σ^2 in Equation (4.45). Estimation of this equation, however, requires Cochrane-Orcutt, maximum likelihood, or an equivalent method. The benefit of estimating from Equation (4.50) rather than Equation (4.45), when $\gamma_1 = -\beta_1\gamma_0$ is a valid assumption, is the improved efficiency of the estimates. The validity of such a common root restriction could be tested, as in Sargan (1964), by comparing the goodness of fit of Equation (4.45) with that of Equation (4.50). Equation (4.45) is more general than Equation (4.50) and it is likely, therefore, that dynamic behaviour cannot be accurately summarised by AR errors. In many cases a common factor may not be found.

When $\gamma_1 \neq -\beta_1\gamma_0$, and beginning from

$$y_t = \gamma_0 x_t + u_t, \tag{4.53}$$

a significant value of the Durbin-Watson d-statistic, calculated from the OLS residuals $\widehat{u}_t$, will often be observed. This results from the fact that Equation (4.50), with a nonzero common root, will be a better approximation to Equation (4.45) than Equation (4.53), where the root has been restricted to zero. A notable case of a model with one period lags and a common root, which has become important in empirical work is Equation (4.48) where $\beta_1 = 1$, which gives a simple regression model in first differences

$$\Delta y_t = \gamma_0 \Delta x_t + \nu_t. \tag{4.54}$$

Transforming trending series by differencing to approximately stationary series has been discussed by Box and Jenkins (1976). Granger and Newbold (1974) suggested that when dealing with the levels of trending variables, the danger of spurious regression 'is especially large when the warning of a significant d-statistic has been ignored' and to circumvent this problem they also proposed the use of differenced variables.[20]

While at times this is approach is acceptable, i.e., when there is a common unit root, in

[19] See Hendry and Mizon (1978).
[20] Hendry and Mizon (1978), p. 552.

other situations, it may cause serious problems. The model in differences

$$\Delta y_t = \gamma_0 \Delta x_t + \varpi_t, \tag{4.55}$$

where $\varpi_t = \Delta \varepsilon_t$, solves the problem of potentially spurious fits, but replaces it with the problem of an error process with a root of minus unity.[21] Over-differencing, however, a linear regression model will result in the first-order serial correlation coefficient of the errors being close to -0.5. A similar problem can arise also with seasonal differences, using $(1 - L^4)$, to remove seasonality. 'If the true model is that in Equation (4.55), and the tentative hypothesis is that of Equation (4.45), then there is a latent root with a value of zero and the COMFAC approach should detect this and allow the redundant dynamics or autocorrelation to be eliminated'.[22] This problem may be avoided by testing the hypothesis that the model should be formulated in differences. This could be done by testing the hypothesis of a unit root against the alternative of a model which contains lagged values of y_t and x_t with unrestricted coefficients.

This hypothesis can be tested by first testing for a common root in the alternative model, and if this is not rejected, by then testing that the root is unity. This is a joint hypothesis, testing first for a common root, and then that the root is unity. This approach will be valid even if the variables are 'spuriously' related in levels and unrelated in differences. Consider

$$\Delta y_t = \gamma_0 \Delta x_t + (\beta_1 - 1) y_{t-1} + (\gamma_0 + \gamma_1) x_{t-1} + \nu_t. \tag{4.56}$$

If Equation (4.54) is a valid model for any value of γ_0, the coefficients of y_{t-1} and x_{t-1} must both be zero in Equation (4.56). In terms of hypothesis testing, the common root implies $\gamma_1 = -\beta_1 \gamma_0$, and if that is a unit root, and if $\beta_1 = 1$ also, Equation (4.56) becomes Equation (4.54). Testing this composite hypothesis, therefore, of a unit root can be achieved by testing the joint significance of the coefficients of the lagged variables in Equation (4.56). It can be shown that ν_t in Equation (4.56) are white noise errors.

A difference model, like Equation (4.54) can be reformulated as a model in levels, like Equation (4.45), with parametric restrictions of $\beta_1 = 1$ and $\gamma_0 = -\gamma_1$. The error term will be unaffected by this. Therefore, 'if differencing is a valid solution to the spurious regression problem, then so must be the inclusion of lagged values of all variables'.[23] As mentioned previously, however, it is not always appropriate to take the equation in differences, as there may not be common factors. Given the restriction $\beta_1 + \gamma_0 + \gamma_1 = 1$, the autoregressive distributed lag model takes on an error-correction form. Testing this restriction is equivalent to testing for a cointegrating relationship.

For the more general case, consider

$$\beta(L) y_t = \gamma(L) x_t + \delta(L) z_t + \nu_t, \tag{4.57}$$

where $\beta(L), \gamma(L)$ and $\delta(L)$ are scalar polynomials in the lag operator, L, of orders p, q and r,

[21] Since $\varpi_t = \Delta \varepsilon_t$ is a moving average with a coefficient of -1.
[22] Hendry and Mizon (1978), p. 553.
[23] Hendry and Mizon (1978), p. 554.

respectively. $\beta(L)$, $\gamma(L)$ and $\delta(L)$ may have at most l common roots, where $l = \min(p,\, q,\, r)$. If there are $n \leq l$ common roots, 'then there exists a polynomial $\rho(L)$ of order n, common to $\beta(L), \gamma(L)$ and $\delta(L)$'.[24] Then it can be shown that

$$\beta(L) = \rho(L)\beta^*(L), \qquad \gamma(L) = \rho(L)\gamma^*(L), \qquad \delta(L) = \rho(L)\delta^*(L), \tag{4.58}$$

or

$$[\beta(L){:}\gamma(L){:}\delta(L)] = \rho(L)\left[\beta^*(L){:}\gamma^*(L){:}\delta^*(L)\right], \tag{4.59}$$

where $\beta^*(L)$, $\gamma^*(L)$ and $\delta^*(L)$ are polynomials of order $(p-n)$, $(q-n)$ and $(r-n)$ respectively. Using Equation (4.59), Equation (4.57) becomes

$$\rho(L)\beta^*(L)y_t = \rho(L)\gamma^*(L)x_t + \rho(L)\delta^*(L)z_t + \nu_t, \tag{4.60}$$

or

$$\beta^*(L)y_t = \gamma^*(L)x_t + \delta^*(L)z_t + u_t, \tag{4.61}$$

where

$$\rho(L)u_t = \nu_t. \tag{4.62}$$

The number of parameters from the general model of Equation (4.57) is reduced by a factor of $2n$ (or kn for k regressors), which greatly improves efficiency. If it can be shown that $\rho(L)$ has a factor of $\Delta = (1-L)$, where $\rho(L) = (1-L)\rho^*(L)$, then Equation (4.61) can be rewritten to give

$$\beta^*(L)\Delta y_t = \gamma^*(L)\Delta x_t + \delta^*(L)\Delta z_t + u_t, \tag{4.63}$$

where

$$\rho^*(L)u_t = \nu_t. \tag{4.64}$$

In practice, the orders of the lag polynomials will not be known. To overcome this, two approaches could be adopted 'both being two-stage decision procedures which commence from the most general unrestricted model which it seems reasonable to consider'.[25] Using these techniques allows for the a priori specification of the longest lags. The common factor technique could be applied to the model as specified a priori, followed by simplification can through testing for zero roots from the set of n common roots. The Wald test is then used to test the common factor restrictions. In summary, the COMFAC approach works as follows. The coefficients and their variance matrices in the general unrestricted form of Equation (4.57) are estimated, by OLS. These results are denoted by $\widehat{\mathbf{b}}$ and $\widehat{\mathbf{V}}$. The restrictions in Equation (4.59) are equivalent to requiring a vector nonlinear function of $\mathbf{b}$, say $\mathbf{f}(\mathbf{b})$, to be zero. The asymptotic variance matrix of $\mathbf{f}(\widehat{\mathbf{b}})$ is $\mathbf{S} = \mathbf{J}'\mathbf{V}\mathbf{J}$ where $\mathbf{J} = \partial\mathbf{f}/\partial\mathbf{b}'$, and $\mathbf{f}(\widehat{\mathbf{b}})'\widehat{\mathbf{S}}^{-1}\mathbf{f}(\widehat{\mathbf{b}})$ will, therefore, have a χ^2 distribution in large samples when the $\mathbf{f}(\mathbf{b}) = 0$ restrictions are valid, with degrees of freedom equal to the number of restrictions tested. Having completed the review of methods of testing for cointegration, attention now turns to the concept of fractional integration and the work of Dolado, et al. (2002).

[24] Hendry and Mizon (1978), p. 554.
[25] Hendry and Mizon (1978), p. 555.

4.4 A Fractional Dickey-Fuller Test for Unit Roots

Many time series, be they economic series or otherwise, are not well represented by either stationary, $I(0)$, or nonstationary, $I(1)$, processes. To overcome the potential difficulties presented by this fact, and to capture the effects of the persistence properties of long memory processes, the class of fractionally integrated processes has proven useful. A fractionally integrated process is denoted $FI(d)$, where the order of integration, d, is extended to include any real number. Although unit root tests have been shown to be consistent under the alternative hypothesis of an $FI(d)$ process, their power is quite low. This lack of power has been the motivation for various approaches to take this alternative into consideration, using Wald-type and Lagrange Multiplier (LM) tests. These approaches are attributable to Geweke and Porter-Hudak (1983), Fox and Taqqu (1986), Sowell (1992), Robinson (1994) and Tanaka (1999). Many of these tests have been unsatisfactory, however, suffering from a lack of power.[26]

Dolado, et al. (2002) proposed a simple Wald-type test in the time domain, which, unlike those methods mentioned, has acceptable power properties.[27] This method is a generalisation of the well-known Dickey-Fuller test, introduced in Subsection 4.2.2, and considers the hypothesis of $FI(d_0)$ against $FI(d_1)$, where $d_1 < d_0$. This test is referred to by the authors as a fractional Dickey-Fuller (FDF) test and they concentrate on the case of $d_0 = 1$ and $0 \leq d_1 < 1$. As the FDF test is a Wald-type test, the value of d is required under the alternative, to make the testing procedure feasible. Therefore, for general hypotheses, the pre-estimation of d under the alternative is necessary for the implementation of this test.

The FDF test has several obvious advantages. It is a simple generalisation of the well-known Dickey-Fuller test. Unlike LM tests, no assumptions are required about the form of the density function, greatly increasing the robustness of this method. Finally, it has been shown that the test fares very well in finite samples, in terms of both power and size.[28]

4.4.1 The fractional Dickey-Fuller test

As introduced in Subsection 4.2.2, the DF test statistic is based upon the statistical significance of parameter ϕ in the following model,

$$\Delta y_t = \phi y_{t-1} + \varepsilon_t. \tag{4.65}$$

If ε_t is i.i.d. and $\phi = 0$, then y_t is a random walk process. If alternatively $\phi < 0$, then y_t is a stationary AR(1) process. The regression model in Equation (4.65) can be generalised to test the null hypothesis that a series is $FI(d_0)$ against the alternative that it is $FI(d_1)$, where d_0,

[26] Dolado, et al. (2002), p. 1963-1964.

[27] Further details can also be found in Dolado, Gonzalo, and Mayoral (2005a).

[28] Dolado, Gonzalo, and Mayoral (2005b) proposed a test which considers the hypothesis of $FI(d_0)$, for $0 \leq d_0 < 1$ against $FI(d_1)$, for $d_1 = 0$, for processes that may be subject to structural breaks at known or unknown dates. This test, the SB-FDF test, may be more appropriate in certain circumstances, regardless of breaks. Ideally, both the FDF and SB-FDF test should be used to test series, as both the $d = 0$ and $d = 1$ hypotheses are tested against $0 < d < 1$. Given the very recent nature of this proposed test, it will not be included here. It is recognised, however, as an area of future research with rich potential.

$d_1 \in \mathbb{R}$, by testing the significance of ϕ in the regression

$$\Delta^{d_0} y_t = \phi \Delta^{d_1} y_{t-1} + \varepsilon_t, \tag{4.66}$$

where $\Delta^{d_1} = (1-L)^{d_1}$ and ε_t is an $I(0)$ process.[29] When $\phi = 0$, Equation (4.66) becomes $\Delta^{d_0} y_t = \varepsilon_t$, implying that y_t is $FI(d_0)$. This allows the formulation $H_0{:}\phi = 0$ and $H_1{:}\phi < 0$, where H_0 implies y_t is $FI(d_0)$, and H_1 implies y_t is $FI(d_1)$. Dolado, et al. (2002) restricted their analysis to the specific case where $d_0 = 1$, namely, y_t is $I(1)$ under the null hypothesis, and $FI(d_1)$, where $0 \leq d_1 < 1$, under the alternative. This is chosen for its empirical relevance in the literature. In principle, this framework could be extended to deal with more general cases.

4.4.2 The test and its asymptotic properties

In the case where $d_0 = 1$ and where $\{\varepsilon_t\}$ is a sequence of zero mean i.i.d. random variables with unknown variance σ^2 and finite fourth order moment, the OLS estimators of ϕ, $\widehat{\phi}_{\text{OLS}}$ and its t-ratio, $t_{\widehat{\phi}_{\text{OLS}}}$, are

$$\widehat{\phi}_{\text{OLS}} = \frac{\sum_{t=2}^{T} \Delta y_t \Delta^{d_1} y_{t-1}}{\sum_{t=2}^{T} (\Delta^{d_1} y_{t-1})^2}, \tag{4.67}$$

and

$$t_{\widehat{\phi}_{\text{OLS}}} = \frac{\sum_{t=2}^{T} \Delta y_t \Delta^{d_1} y_{t-1}}{S_T \left(\sum_{t=2}^{T} (\Delta^{d_1} y_{t-1})^2 \right)^{\frac{1}{2}}}, \tag{4.68}$$

where the variance of the residuals, S_T^2 is

$$S_T^2 = \frac{\sum (\Delta y_t - \widehat{\phi}_{\text{OLS}} \Delta^{d_1} y_{t-1})^2}{T}. \tag{4.69}$$

It can be shown under the null hypothesis that y_t is a random walk, $\widehat{\phi}_{\text{OLS}}$ is a consistent estimator of $\phi = 0$ and $\widehat{\phi}_{\text{OLS}}$ converges to its true value at a rate of T^{1-d_1} when $0 < d_1 < 0.5$, $(T \log T)^{\frac{1}{2}}$ when $d_1 = 0.5$, and at the standard rate $T^{\frac{1}{2}}$ when $0.5 < d_1 < 1$. Asymptotic distributions under this null have been derived. If the data generating process is defined as $\Delta^{d_1^*} y_t = \varepsilon_t 1_{(t>0)}$, where $d_1^* \in [0, 1)$ and $1_{(t>0)} = 0$ when $t \leq 0$, and 1 otherwise, the test statistic based upon $\widehat{\phi}_{\text{OLS}}$ or the t-ratio of $\widehat{\phi}_{\text{OLS}}$ in the regression of Δy_t on $\Delta^{d_1} y_{t-1}$ is consistent for any value of $d_1 \in [0, 1)$.[30] This guarantees the consistency of the proposed tests, even when under the alternative an incorrect value of d_1 is employed to implement the procedure, insofar as $d_1 \in [0, 1)$. The standard or nonstandard asymptotic behaviour of the previous test statistics depend on the distance between the null and alternative hypothesis.

The proposed test, simulated for a random walk and a series of FI processes, with order of integration $d_1^* \in [0, 0.9)$ and two sample sizes, was found to perform very well in both size and power, by comparison to several other procedures, including the standard DF test, the Geweke

[29] It should be noted that Equation (4.66) is 'an unbalanced regression where regressand and regressor have been differenced in agreement with their degree of integration under the null and the alternative hypothesis, respectively' (Dolado, et al., 2002, p. 1966).

[30] Dolado, et al. (2002), p. 1969-1971.

and Porter-Hudak (1983) test and the LM test proposed by Tanaka (1999). In practice, the true value of d_1, d_1^*, is unknown under the alternative hypothesis. Under misspecification of the value of d_1, it was found that the power of the FDF test decreases when values of d_1 larger than d_1^* are selected, particularly when $d_1^* > 0.7$. This is to be expected, as the alternative is now close to the null. In general, however, it is found that the procedure is robust, in finite samples, to misspecification in d_1, and that the desirable qualities of the test do not depend on an accurate choice of d_1 under the alternative.

An estimate of the memory parameter d_1, under the alternative hypothesis, is required to implement the FDF test. Generally, this value is unknown and must be estimated, therefore, particularly when a composite alternative hypothesis is being posed, which is always the case in practice. A substantial literature exists on the estimation of the order of integration, d_1, in $FI(d_1)$ models, in both the time and frequency domains. If a $T^{\frac{1}{2}}$-consistent estimator of d_1 is used, the asymptotic distribution of the t-ratio of $\widehat{\phi}_{\mathrm{OLS}}$ under the null is $N(0,1)$.

Under the null hypothesis that y_t is a random walk, the test statistic, the t-ratio of $\widehat{\phi}_{\mathrm{OLS}}$ for $\widehat{d_1}$, associated with the parameter ϕ in the regression

$$\Delta y_t = \phi \Delta^{\widehat{d_1}} y_{t-1} + a_t, \tag{4.70}$$

where $\widehat{d_1}$ has been chosen according to

$$\widehat{d_1} = \begin{cases} \widehat{d_T} & \text{if } \widehat{d_T} < 1-c \\ 1-c & \text{if } \widehat{d_T} \geq 1-c, \end{cases} \tag{4.71}$$

is asymptotically distributed as

$$t_{\widehat{\phi}_{\mathrm{OLS}}}\left(\widehat{d_1}\right) \xrightarrow{w} N(0,1), \tag{4.72}$$

where $c > 0$ and is a fixed value in a neighbourhood of zero, such that $(1-c)$ is close to unity, and $\xrightarrow{w}$ denotes weak convergence. This result suggests that when a pre-estimated value of d_1 is used to implement the FDF test, the associated critical values are those from a $N(0,1)$ distribution, given that the value of $\widehat{d_1}$ satisfies Equation (4.71). If, however, the value of d_1 is assumed known, a priori, and $d_1 \in [0, 0.5)$, the test then has a nonstandard distribution under the null. As will be seen later, nonstandard critical values are available from the authors for such cases.

Dolado, et al. (2002) recommended the use of a parametric estimator of d_1. Their reasoning is that semiparametric estimators often converge at a rate slower than $T^{\frac{1}{2}}$. In principle, a $T^{\frac{1}{2}}$ rate of convergence is required for feasible use of the FDF test. They further suggest using a time domain, as opposed to frequency domain, parametric estimator. Within that class, they recommended the use of Mayoral's (2003) general minimum distance (GMD) estimator. Briefly, consider the ARFIMA(p, d_0, q) model

$$\Phi_0(L)\Delta^{\delta}\left(\Delta^{m_0} y_t - \mu_0\right) = \Theta_0(L)\varepsilon_t, \tag{4.73}$$

where $\Phi_0(L)$ and $\Theta_0(L)$ are autoregressive and moving average polynomials of order p and

q, respectively, and ε_t is a sequence of i.i.d. random variables with zero mean and unknown variance, σ^2. Note that in this model $d_0 = m_0 + \delta$. 'The integer $m_0 = \lfloor d_0 + 1/2 \rfloor$, where $\lfloor \cdot \rfloor$ denotes integer part, is the number of times that y_t must be differenced to achieve stationarity (therefore $m_0 \geq 0$). The parameter δ, the fractional part, lies in the interval $(-0.75, 0.5)$, in such a way that, for a given d_0, $\delta = d_0 - \lfloor d_0 + 1/2 \rfloor$'.[31] Also, $\boldsymbol{\lambda} = (d_0, \boldsymbol{\psi}') \in \mathbb{R}^{p+q+1}$, where $\boldsymbol{\psi}$ is a vector of autoregressive and moving average parameters. All possible values of $\boldsymbol{\lambda}$ are contained within the set $\boldsymbol{\Lambda}$. Given that the estimated residuals of this process are $e_t(\boldsymbol{\lambda})$, let the sample i^{th} autocorrelation of the residuals be

$$\widehat{\rho}_{e(\boldsymbol{\lambda})}(i) = \frac{\sum_{t=1}^{T-i} e_t(\boldsymbol{\lambda}) e_{t+i}(\boldsymbol{\lambda})}{\sum_{t=1}^{T} e_t(\boldsymbol{\lambda})^2}. \tag{4.74}$$

The general minimum distance estimator is given by

$$\widehat{\boldsymbol{\lambda}}_k = \arg\min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} V_{ke}(\boldsymbol{\lambda}, y), \tag{4.75}$$

where

$$V_{ke}(\boldsymbol{\lambda}, y) = \sum_{i-1}^{k} \widehat{\rho}_{e(\boldsymbol{\lambda})}(i)^2. \tag{4.76}$$

The authors have made available the MATLAB program to implement this estimator, but this has yet to be generalised, and so is limited in its applicability. This may render it useless for the applications in later chapters. Alternative estimators are considered, therefore. Four estimators were chosen primarily because these estimators were freely available. These were the Geweke and Porter-Hudak (GPH) (1983) and Robinson (1994) nonparametric methods, and the parametric methods attributable to Sowell (1992) and Beran (1995).[32] These procedures will be introduced in due course.

The FDF test, using the GMD estimator of d_1, performs very well. Replacing the true value of d_1 with an estimated value, $\widehat{d}_1$, was found to have little effect on the size and power properties of the test. Also, by comparing the empirical distributions of the t-ratio of $\widehat{\phi}_{\mathrm{OLS}}$ for $\widehat{d}_1$ and a $N(0,1)$ distribution, for various sample sizes, it was found that they approximate to a standardised normal variate very well.

4.4.3 The augmented fractional Dickey-Fuller test

To broaden the applicability of the test, the framework outlined above has been expanded to allow for serial correlation in the disturbance terms. Although some series may behave as fractional white noise processes, it is desirable to consider series where this may not be so, i.e., where there may be serial correlation in the error terms. By following the ADF approach, it can be shown that 'the asymptotic distribution of the t-ratio remains valid in the presence of serial correlation, as long as a sufficient number of lags of $\Delta^{d_0} y_t$ are included in the regression'.[33]

[31] Mayoral (2003), p. 4.

[32] Estimation procedures for all of these methods are available in Doornik and Ooms (1999) Ox package, ARFIMA, while the nonparametric procedures are also available in RATS.

[33] Dolado, et al. (2002), p. 1980.

The fractional augmented Dickey-Fuller (FADF) test imitates the Dickey-Fuller approach in the context of an autoregressive integrated moving average ARFIMA process. Consider again Equation (4.65)

$$\Delta y_t = \phi y_{t-1} + u_t, \tag{4.65}$$

where now $\alpha(L)u_t = \varepsilon_t$, an autoregressive process of order p, such that $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$ has all its roots outside the unit circle. The ADF test is based on the regression

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^{p} \zeta_i \Delta y_{t-i} + \varepsilon_t. \tag{4.77}$$

Under the null, Dickey and Fuller (1981) proved that the asymptotic distribution of the t-ratio of $\widehat{\phi}_{\text{OLS}}$, is identical to that obtained in the absence of serial correlation. Imitating this process, the regression for the FADF test, where u_t is an AR(p) process is

$$\Delta y_t = \phi \Delta^{d_1} y_{t-1} + \sum_{i=1}^{p} \zeta_i \Delta y_{t-i} + \varepsilon_t. \tag{4.78}$$

As in Equation (4.66), the null and alternative hypotheses of a unit root and $FI(d_1)$ process can be considered in terms of ϕ, when $\phi = 0$ and $\phi < 0$, respectively. The t-ratio is computed from the estimation of Equation (4.78). The asymptotic distribution of the t-ratio of $\widehat{\phi}_{\text{OLS}}$ for $\widehat{d}_1$ is identical to that derived under the assumption of uncorrelated disturbances.

The performance of this more general framework of the FADF test, with a $FI(d_1)$ process with an AR(1) error structure[34] and different values of d_1 and α indicate that the test is very well-behaved in terms of size. The power of the test is, by and large, dependent on how close the polynomial $\alpha(z)$ is to having a unit root. Only small changes in the power are observed when the value of d_1 is misspecified. This indicates that the test is robust to misspecification of d_1. Also, and as in previous cases, the asymptotic distribution of the t-ratio with an estimated d_1 is identical to that obtained with a known d_1, where a $T^{\frac{1}{2}}$-consistent estimator of d_1 is employed. The distribution of the t-ratio of $\widehat{\phi}_{\text{OLS}}$ for $\widehat{d}_1$ and an $N(0,1)$ distribution are compared, when the data generating process is an ARIMA$(1,1,0)$ and d_1 has been estimated using Mayoral's (2003) GMD estimator. As before, the approximation to an $N(0,1)$ distribution works well in the finite samples considered. Also, it can be shown that the power loss due to the pre-estimation of d_1 is minimal, although it happens to be larger than that obtained under identically and independently distributed error terms. Next, brief consideration is given to the alternative estimators mentioned previously, the GPH (1983), Robinson (1994), Sowell (1992) and Beran (1995) estimators.

[34] The structure in question was

$$\Delta^{d_1^*} y_t = u_t,$$

where

$$u_t = \alpha u_{t-1} + \varepsilon_t,$$

and where $\varepsilon_t \sim$ n.i.d.$(0,1)$.

4.4.4 Estimating the order of fractional integration

As previously noted, Dolado, et al. (2002) recommended the use of the GMD estimator of d_1. While this is available from the authors, it is however, not widely applicable. For this reason, four alternative, more widely used and, therefore, more readily available, estimation methods are described in this chapter and used later in the thesis; they are the nonparametric methods of Geweke and Porter-Hudak (1983) and Robinson (1994), and the parametric methods of Sowell (1992) and Beran (1995). Each method is briefly described below.

Geweke and Porter-Hudak (1983)

The concept of fractional integration has been previously introduced in this chapter. One of the most commonly used frequency domain estimators of d, the order of fractional integration, is the method of Geweke and Porter-Hudak (1983). Consider the long memory model proposed by Granger and Joyeux (1980) and Hosking (1981). This model is of the form

$$(1-L)^d X_t = \varepsilon_t, \tag{4.79}$$

where $d \in (-0.5,\ 0.5)$ and ε_t is serially uncorrelated. Note that this model is stationary and long memory, since d lies between -0.5 and 0.5. Such a model was initially considered as some time series appeared to have unbounded spectral densities at the frequency $\lambda = 0$. The spectral density of X_t, given Equation (4.79), can be defined as[35]

$$f(\lambda; d) = \left(\frac{\sigma^2}{2\pi}\right) |1 - \exp^{i\lambda}|^{-2d} = \left(\frac{\sigma^2}{2\pi}\right) \left\{4\sin^2\left(\frac{\lambda}{2}\right)\right\}^{-d}. \tag{4.80}$$

Geweke and Porter-Hudak referred to this as a simple integrated process. This is generalised to become a general integrated series,

$$(1-L)^d X_t = u_t, \tag{4.81}$$

where u_t is a stationary linear process, with a spectral density function $f_u(\lambda)$, if the spectral density is of the form $f(\lambda; d) f_u(\lambda)$, where $f_u(\lambda)$ is a positive continuous function bounded above and away from zero on the interval $[-\pi,\ \pi]$.[36] The concern is now to estimate d for the general integrated series. The spectral density function of X_t is

$$f(\lambda) = \left(\frac{\sigma^2}{2\pi}\right) \left\{4\sin^2(\lambda)\right\}^{-d} f_u(\lambda). \tag{4.82}$$

Taking natural logarithms gives

$$\ln\{f(\lambda)\} = \ln\left\{\frac{\sigma^2}{2\pi}\right\} - d\ln\left\{4\sin^2\left(\frac{\lambda}{2}\right)\right\} + \ln\{f_u(\lambda)\}. \tag{4.83}$$

[35]The conventional spectral density notation is used here. The spectral density of the time series X_t is $f(\lambda) = \sum_{s=-\infty}^{\infty} R_x(s)\exp(-i\lambda s)$, where R_x is the autocovariance function of X_t.

[36]Geweke and Porter-Hudak (1993), p. 222.

In a sample of size T, the harmonic ordinates are given by $\lambda_{j,T} = 2\pi j/\pi$, where $j = 0, \ldots, T-1$, and the periodogram of these ordinates is $I(\lambda_{j,T})$. In evaluating Equation (4.83) at $\lambda_{j,T}$, it can be shown that

$$\ln\{I(\lambda_{j,T})\} = \ln\left\{\frac{\sigma^2}{2\pi}\right\} - d\ln\left\{4\sin^2\left(\frac{\lambda_{j,\,T}}{2}\right)\right\} + \ln\{f_u(\lambda_{j,\,T})\} + \ln\left\{\frac{I(\lambda_{j,\,T})}{f(\lambda_{j,\,T})}\right\}. \quad (4.84)$$

Ignoring the term $\ln\{f_u(\lambda_{j,T})\}$, which becomes negligible where the harmonic frequencies are close to zero, Equation (4.84) can be rewritten as

$$Y_t = \alpha_t - dX_t + \varepsilon_t, \quad (4.85)$$

where $Y_t = \ln\{I(\lambda_{j,T})\}$, $\alpha_t = \ln\left\{\frac{\sigma^2}{2\pi}\right\}$ plus the mean of ε_t, where $\varepsilon_t = \ln\left\{\frac{I(\lambda_{j,T})}{f(\lambda_{j,T})}\right\}$ and $X_t = \ln\left\{4\sin^2\left(\frac{\lambda_{j,T}}{2}\right)\right\}$. The analogy to a simple linear regression is obvious. Estimating d is equivalent to estimating the coefficient of the explanatory variable. Geweke and Porter-Hudak show that for $d < 0$, the estimator is consistent and the conventional interpretation of the standard error of the coefficient is appropriate asymptotically.[37] They provide empirical evidence that this result remains true for $d \geq 0$.

Recall that for the test procedure above, $d \in [-0.5, 0.5]$. What about the case where $d > 0.5$? By differencing a fractionally integrated series by integer values, the original series can be transformed suitably. Since in many applications $d \approx 1$, it is usual to operate the GPH procedure on the first difference of a series, although appropriate differencing in the spirit of the Pantula (1989) principle may indicate the required order of differencing.

Robinson (1994)

Robinson (1994) considered the 'discretely averaged periodogram, where the averaging is done over a neighbourhood of the origin which slowly degenerates to zero as sample size T increases'.[38] By manipulating averaged periodograms, an estimate of the parameter $H = d + \frac{1}{2}$, the parameter of interest here, can be obtained. It is shown that this estimate is consistent for the nonparametric function $L(\lambda)$, and also gives a consistent estimate of G, where $L(\lambda) = GM(\lambda)$, $G > 0$, $M(\lambda)$ is a known function and G is unknown. $L(\lambda)$ is a slowly varying function at infinity, a positive measurable function, satisfying

$$\frac{L(t\lambda)}{L(\lambda)} \to 1 \quad \text{as} \quad \lambda \to \infty, \quad \text{for all} \quad t > 0. \quad (4.86)$$

The average periodogram under consideration here is

$$F(\lambda) = \int_0^\lambda f(\theta)\, d\theta, \quad (4.87)$$

where

$$f(\lambda) \sim L(\frac{1}{\lambda})\lambda^{1-2d} \quad \text{as} \quad \lambda \to 0_+. \quad (4.88)$$

[37] Geweke and Porter-Hudak (1983), p225.

[38] Robinson (1994), p. 516.

Various estimates of $F(\lambda)$ exist, including

$$\widetilde{F}(\lambda) = \int_0^\lambda \widetilde{I}(\theta)\,d\theta \qquad\qquad 0 < \lambda \leq \pi, \tag{4.89}$$

where

$$\widetilde{I}(\lambda) = |\widetilde{w}(\lambda)|^2, \tag{4.90}$$

$$\widetilde{w}(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^{T} (x_t - \mu) \exp^{it\lambda}, \tag{4.91}$$

and

$$\widehat{F} = 2\pi/T \sum_{j=1}^{[T\lambda/2\pi]} I(\lambda_j), \tag{4.92}$$

where $[T\lambda/2\pi]$ denotes the integer part and $\lambda_j = 2\pi_j/T$, and

$$I(\lambda) = |w(\lambda)|^2, \tag{4.93}$$

$$w(\lambda) = (2\pi T)^{-1/2} \sum_{t=1}^{T} x_t \exp^{it\lambda}. \tag{4.94}$$

In the model

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - \exp^{i\lambda}\right|^{1-2H} \frac{\left|b(\exp^{i\lambda})\right|^2}{\left|a(\exp^{i\lambda})\right|^2}, \qquad -\pi < \lambda \leq \pi, \tag{4.95}$$

a semiparametric estimate of H would be useful, as it would allow for the estimation of this fractional ARIMA, or ARFIMA, model. Geweke and Porter-Hudak (1983) proposed a closed form semiparametric estimate of H, which assumed the function $L(\lambda)$ was constant. Robinson presents a new semiparametric estimator of H, which under certain conditions, is consistent. It can be shown that for any $q > 0$,

$$\frac{F(q\lambda)}{F(\lambda)} \sim q^{2(1-H)} \frac{L(1/q^\lambda)}{L(1/\lambda)} - q^{2(1-H)}, \tag{4.96}$$

as $\lambda \to 0+$. An estimate of d can be derived, therefore, from

$$\widehat{H}_{mq} = 1 - \frac{\log\left[\widehat{F}(q\lambda_m)/\widehat{F}(\lambda_m)\right]}{2\log q}, \tag{4.97}$$

where $\widehat{F}(\lambda_m)$ is a special case of the general class of weighted periodogram spectrum estimates.[39] Robinson shows that under certain conditions

$$\widehat{H}_{mq} \xrightarrow{p} H, \quad \text{as} \quad T \to \infty. \tag{4.98}$$

The proof of this result requires no knowledge of the functional form of $L(\lambda)$.

---

[39]Robinson (1994), p. 518.

Whereas the GPH estimator required $d \in [-0.5,\ 0.5]$, here d is constrained to be $d \in [-1.5,\ 0.5]$. The same approach is taken, however, to ensure the series under investigation fits this constraint.

Exact Maximum Likelihood

Sowell (1992) derived an unconditional exact likelihood function for a stationary, fractionally integrated and normally distributed time series, and provides recursive procedures to estimate that function.

To understand the procedure, consider the stationary, fractionally integrated and normally distributed time series given by

$$\Phi(L)(1-L)^d z_t = \Theta(L)\varepsilon_t, \tag{4.99}$$

where $\Phi(L)$ and $\Theta(L)$ are lag polynomials, and assume, amongst other things, that $d < \frac{1}{2}$. The autocovariance function, required for evaluation of the likelihood function, can be written as

$$\gamma(s) = Ez_t z_{t-s} = \frac{1}{2\pi}\int_0^{2\pi} f_z(\lambda)\, e^{i\lambda s} d\lambda. \tag{4.100}$$

The spectral density of z_t,

$$f_z(\lambda) = (1-\omega)^{-d}\left(1-\omega^{-1}\right)^{-d} f_u(\lambda), \tag{4.101}$$

where $\omega = \exp^{i\lambda}$, can be calculated by first calculating the spectral density of $u_t = (1-L)^d z_t$, $f_u(\lambda)$. Substitution of $f_u(\lambda)$ into $f_z(\lambda)$ gives

$$f_z(\lambda) = \sigma^2 \sum_{l=-q}^{q}\sum_{j=1}^{p} \psi(l)\,\zeta_j \left[\frac{\rho_j^{2p}}{(1-\rho_j\omega)} - \frac{1}{\left(1-\rho_j^{-1}\omega\right)}\right] \times (1-\omega)^{-d}\left(1-\omega^{-1}\right)^{-d}\omega^{p+l}. \tag{4.102}$$

The autocovariance function, $\gamma(s)$, can now be calculated by substituting in the function for $f_z(\lambda)$, to give

$$C(d,h,\rho) = \tag{4.103}$$
$$\frac{1}{2\pi}\int_0^{2\pi}\left[\frac{\rho^{2p}}{(1-\rho e^{-i\lambda})} - \frac{1}{(1-\rho^{-1}e^{-i\lambda}\omega)}\right] \times \left(1-e^{-i\lambda}\right)^{-d}\left(1-e^{i\lambda}\right)^{-d} e^{-i\lambda h}\, d\lambda.$$

Using this equation, and several others derived from it, the terms of $C(d,h,\rho)$ can be recursively estimated. Mirroring issues raised in Chapter 2, appropriate start values are required for this procedure, as it depends on a numerical optimisation algorithm. Pre-estimation of d with the method of Geweke and Porter-Hudak (1983) or a grid search method are suggested in this case.

Nonlinear Least Squares

Unlike the estimators of the fractional parameter discussed thus far, Beran (1995) proposed an estimator that is not restricted to the stationary range, i.e., $-\frac{1}{2} < d < \frac{1}{2}$, but rather to any real $d > -\frac{1}{2}$. A simple algorithm to estimate d was suggested.

Beran (1995) generalised the definition of fractional integration used here thus far, to include parameters for fractional stationarity and nonstationarity. To illustrate, consider the Gaussian time series X_t for which

$$\phi(L)(1-L)^{\delta}\left\{(1-L)^m X_t - \mu\right\} = \psi(L)\epsilon_t, \tag{4.104}$$

where L is the lag operator, μ is the expected value of X_t, $\epsilon_t \sim$ i.i.d.$(0, \sigma_\epsilon^2)$, and $\phi(x) = \sum_{j=0}^{p} \phi_j x^j$ and $\psi(x) = \sum_{j=0}^{q} \psi_j x^j$ are polynomials with unit roots outside the unit circle. Also, $m \geq 0$ is an integer and $\delta \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. The m^{th} difference of X_t is a stationary fractional autoregressive integrated moving average, or $\textsc{Arima}(p, \delta, q)$. The differencing parameter, $d = m + \delta$, is the difference required to render X_t a stationary $\textsc{Arma}(p, q)$.

Where μ is unknown, the m^{th} difference of X_t is stationary with expected value μ and therefore

$$U_t(m) = (1-L)^m X_t, \tag{4.105}$$

where U_t is a stationary fractional $\textsc{Arima}(p, \delta, q)$ with mean μ. Let

$$\boldsymbol{\vartheta} = \left(\sigma_\epsilon^2, d, \phi_1, \ldots, \phi_p, \psi_1, \ldots, \psi_q\right) = \left(\sigma_\epsilon^2, \boldsymbol{\eta}\right), \tag{4.106}$$

the unknown parameter vector. Since U_t is ergodic, the sample mean is

$$\bar{U} = \frac{1}{T-m} \sum_{t=m+1}^{T} U_t, \tag{4.107}$$

and adjusted residuals can be defined by

$$e_t(\boldsymbol{\eta}) = \sum_{j=0}^{t-1} a_j(\boldsymbol{\eta}^*)\left(U_{t-j} - \bar{U}\right), \tag{4.108}$$

where $\boldsymbol{\eta}^* = (d-m, \phi_1, \ldots, \phi_p, \psi_1, \ldots, \psi_q) = (\delta, \phi_1, \ldots, \phi_p, \psi_1, \ldots, \psi_q)$, which for the population can be written as

$$\epsilon_t(\boldsymbol{\eta}) = \sum_{j=0}^{\infty} a_j(\boldsymbol{\eta}^*)\left(U_{t-j} - \mu\right). \tag{4.109}$$

$\boldsymbol{\vartheta}$ can be estimated by maximum likelihood, using the minimised sum of squared residuals

$$S(\boldsymbol{\eta}) = \sum_{t=2}^{T} e_t^2(\boldsymbol{\eta}), \tag{4.110}$$

with respect to $\boldsymbol{\eta}$, giving

$$\widehat{\boldsymbol{\vartheta}}_1 = \frac{1}{T-1} S(\widehat{\boldsymbol{\eta}}), \tag{4.111}$$

where $e_t(\boldsymbol{\eta})$ is defined as the adjusted residual above. It can be shown that $\widehat{\boldsymbol{\vartheta}}$ converges to the true value of $\boldsymbol{\vartheta}$.

Practically, Beran (1995) suggested an approach to estimating d. For $p = q = 0$, evaluating $S(d) = \sum e_t^2$ estimates $\widehat{\boldsymbol{\vartheta}}_2 = \widehat{d}$, for a fine grid of d-values. In cases where $\min(p, q) \neq 0$, d

can be estimated by evaluating $\widetilde{e}_t = \sum_{j=0}^{t-1} b_j(\delta)\left(U_{t-j} - \bar{U}\right)$, again for a sufficiently fine grid of $d = m + \delta$. The ARMA parameters $\sigma_\epsilon^2, \phi_1, \ldots, \phi_p, \psi_1, \ldots, \psi_q$ can be estimated for the series $\widetilde{e}_1, \ldots, \widetilde{e}_n$, and $\widehat{d}$ is the value of d which minimises $\widehat{\sigma}_\epsilon^2$. Doornik and Ooms (1999) implemented this approach, which they call nonlinear least squares in their OX package ARFIMA. In fact, all four of the methods outlined above are available in the OX ARFIMA package.

4.4.5 Implementing the fractional Dickey-Fuller test

Thus far, the theoretical underpinnings and properties of the FDF and FADF tests have been reviewed. This section very briefly summarises the steps necessary to actually test for fractional integration using this method.

1. An estimate must be made of the order of fractional integration, $\widehat{d}_1$, of the series of interest. Several methods are available for estimating the parameter $\widehat{d}_1$. Dolado, et al. (2002) advocated a generalised minimum distance estimator attributable to Mayoral (2003). Since no general estimator is available here, alternatives are used, as outlined above.

2. Having estimated $\widehat{d}_1$, a regression of the form[40]

$$\Delta y_t = \mu + \phi\Delta^{\widehat{d}_1} y_{t-1} + \sum_{i=1}^{p} \zeta_i \Delta y_{t-i} + \gamma t + \varepsilon_t, \qquad (4.112)$$

   where Δ^{d_1} can be expanded as

$$\begin{aligned}\Delta^{d_1} y_t &= y_t - d_1\, y_{t-1} + \frac{1}{2!}\, d_1(d_1-1)y_{t-2} - \frac{1}{3!}\, d_1(d_1-1)(d-2)y_{t-3} + \ldots \\ &+ \frac{(-1)^j}{j!}\, d_1(d_1-1)\ldots(d_1-j+1)y_{t-j} + \ldots \qquad (4.113)\end{aligned}$$

   is estimated to obtain $\widehat{\phi}$.

3. The FADF test is carried out by examining the t-ratio of $\widehat{\phi}_{\text{OLS}}$ for $\widehat{d}_1$, which is the test statistic of interest, testing the hypothesis H_0: $FI(d_0 = 1)$ against H_0: $FI(\widehat{d}_1)$ where $0 \leq \widehat{d}_1 \leq 1$.

Recall that when a pre-estimated value of d_1 is used to implement the FDF or FADF test, the associated critical values are those from a $N(0,1)$ distribution. If, however, the value of d_1 is assumed known, a priori, and $d_1 \in [0, 0.5)$, the test then has a nonstandard distribution under the null. Critical values are available from Dolado, et al. (2002), Appendix B, Tables X, XI, XIII, p. 2003-2004.

4.5 Johansen's Small Sample Correction

As outlined previously in this chapter, Johansen's (1988, 1991) Trace test is commonly used in exploring cointegrating relationships in economic data. It has come to be a standard

[40] The inclusion of an intercept and trend term, as in the standard Dickey-Fuller procedure, depends on the data.

method of analysis, with many econometric software packages including standard routines for carrying out this procedure. In his more recent paper, however, Johansen (2002) highlighted the potentially poor small-sample properties of this method. This confirms several earlier studies, as outlined by Kennedy (2003), which highlight the potential small-sample deficiencies of Johansen's test for cointegration. Cheung and Lai (1993a) pointed to several finite-sample shortcomings. A large sample size, in the order of 100 observations, is needed for reliable inference. Even then the procedure produces outliers, particularly when the errors are not distributed independently normal, and so rejects the null of no cointegration too often. Hansen, Kim, and Mittnik (1998) found that the χ^2 statistics for testing cointegrating relationships have fat tails and suggested a correction to the critical values. Zhou (2000) suggested a bootstrapping alternative which solves the problem. Johansen (2002) recommended a correction to the asymptotic critical values, which may greatly improve the accuracy of inferences made when using the technique. In the following sections, notation closely follows that of the original paper and that introduced in Subsection 4.3.2.

4.5.1 The correction factor

Consider the vector autoregressive model

$$\Delta \mathbf{X}_t = \mathbf{\Pi X}_{t-1} + \mathbf{\Upsilon} t^{n_d} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{X}_{t-i} + \sum_{i=0}^{n_d-1} \Phi_i t^i + \varepsilon_t, \tag{4.114}$$

where $t = 1, \ \ldots \ , T$, $\mathbf{X}_t$ is an n-dimensional process, $\mathbf{\Pi}$ and $\mathbf{\Gamma}$ are matrices of coefficients, $\mathbf{\Pi} = \sum_{i=1}^{k} \mathbf{\Pi}_i - \mathbf{I}$, $\mathbf{\Gamma}_i = -\sum_{j=i+1}^{k} \mathbf{\Pi}_j$, t^{n_d} is a trend term, $\Phi_i t^i$ is a deterministic term, and ε_t has the usual properties. This model is frequently employed in the analysis of economic data. It can be shown that if $\mathbf{X}_t$ is a nonstationary process, and that if $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $n \times r$, then $\boldsymbol{\beta}'\mathbf{X}_t - E[\boldsymbol{\beta}'\mathbf{X}_t]$ is stationary. If this is in fact the case, $\mathbf{X}_t$, is said to cointegrate, with cointegrating vector $\boldsymbol{\beta}$.

To implement Johansen's (1988, 1991) Trace test, consider the null hypotheses $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ and $\mathbf{\Upsilon} = \boldsymbol{\alpha\rho}'$, where, as above, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $n \times r$, and $\boldsymbol{\rho}$ is $1 \times r$. Anderson's (1951) technique of reduced rank regression is employed to derive the likelihood ratio test, and to estimate the parameters, assuming Gaussian errors. Johansen (1988, 1996) and Ahn and Reinsel (1990) derived the asymptotic distribution of the likelihood ratio, $-2\log \textsc{Lr}$. It is found to be a nonstandard distribution, which can be expressed as a Brownian motion. It can be tabulated by simulation, as a function of $n - r$ and n_d. For further details, see Subsection 4.3.2

It is widely known that the small sample properties of the Trace test are different from its asymptotic properties. This has been confirmed by simulation.[41] Ahn and Reinsel (1990) and Reimers (1992) corrected for such small samples using a method based on degrees of freedom. However, the distribution in question, that of the likelihood ratio test statistic, depends on T and θ, where θ is a function of the parameters under the null hypotheses. As $T \to \infty$ the dependence on θ disappears, but not uniformly. So, if θ is close to a boundary of cointegrating properties, any approximation may be poor. As the Trace test is so widely used in inferring cointegrating rank, where cointegrating rank, $\rho(\mathbf{\Pi})$, may be defined as the number

[41]See Cheung and Lai (1993a), Toda (1995), Haug (1996) and Gonzalo and Pitarakis (1999).

of linearly independent cointegrating relations, improving the asymptotic approximation is important. One potential, but unattractive solution, is to simulate the exact distribution each time, using the estimated parameter values and generating identically and independently distributed Gaussian errors. Although this will give the same limiting distribution, it may not improve approximation due to its nonuniform convergence.

Johansen (2002) suggested a correction factor which will improve the finite sample properties of the likelihood ratio test. Using a Bartlett (1937) correction,[42] the expectation of the LR test statistic is found, and then corrected to have the same mean as the limiting distribution. Given $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ and $\mathbf{\Upsilon} = \boldsymbol{\alpha\rho}'$, let θ denote the parameters of Equation (4.114). The Bartlett correction can be implemented by approximating

$$E_{\theta}\left[-2\log \mathrm{LR}\left(\Pi = \boldsymbol{\alpha\beta}', \Upsilon = \boldsymbol{\alpha\rho}' | \Delta \mathbf{X}_t\right)\right], \tag{4.115}$$

a function of θ and T. This result is expressed in terms of the LR test of $\mathbf{\Pi}^* = \mathbf{0}, \mathbf{\Upsilon}^* = \mathbf{0}$ in the model for $\mathbf{X}_t$

$$\Delta \mathbf{X}_t^* = \Pi^* \mathbf{X}_{t-1}^* + \Upsilon^* t^{n_d} + \sum_{i=0}^{n_d - 1} \Phi_i^* t^i + \varepsilon_t^*. \tag{4.116}$$

Note that when $\mathbf{\Pi}^* = \mathbf{0}$, then $\mathbf{\Gamma}_i = 0$ since $\mathbf{\Gamma}_i = -\sum_{j=i+1}^{k} \mathbf{\Pi}_j$. If $\mathbf{\Pi}^* = \mathbf{0}, \mathbf{\Upsilon}^* = \mathbf{0}$, then

$$f(T,\ n_b,\ n_d) = E\left[-2\log \mathrm{LR}\left(\mathbf{\Pi}^* = \mathbf{0}, \mathbf{\Upsilon}^* = \mathbf{0} \mid \mathbf{X}_t^*\right)\right], \tag{4.117}$$

where the function $f(\cdot)$ is an approximation of Equation (4.115) when $\mathbf{\Pi}^* = \mathbf{0}, \mathbf{\Upsilon}^* = \mathbf{0}$ and $n_b = n - r$. This function can be tabulated by simulation, as it depends only on T, n_b and n_d. An approximation of Equation (4.115) can be found, and takes the form

$$f\left(T,\ n_b\ n_d\right)\left(1 + T^{-1} b(\theta)\right), \tag{4.118}$$

where $b(\theta)$ is to be defined [see Equation (4.124)]. Given this, the correction factor can be shown to be

$$\frac{f(n_b\ , n_d)}{f\left(T,\ n_b,\ n_d\right)} \frac{1}{\left(1 + T^{-1} b\left(\widehat{\theta}\right)\right)}, \tag{4.119}$$

where

$$f(n_b,\ n_d) = \lim_{T \to \infty} f(T,\ n_b,\ n_d). \tag{4.120}$$

Defining

$$a(T,\ n_b,\ n_d) = \frac{f(T,\ n_b,\ n_d)}{f(n_b,\ n_d)}, \tag{4.121}$$

the correction factor becomes

$$\frac{1}{a\left(T,\ n_b,\ n_d\right)\left(1 + T^{-1} b\left(\widehat{\theta}\right)\right)}. \tag{4.122}$$

[42] Bartlett corrections are scalar transformations of the likelihood ratio statistic. They improve the LR statistic by transforming its distribution under the null hypothesis from a χ^2 of order $O(1)$ to a χ^2 of order $O(1/T)$. For a useful review, see Cribari-Neto and Cordeiro (1996).

The Bartlett correction, where observations are identically and independently distributed, often offers an excellent improvement of fit.[43] This may not hold true in the presence of unit roots. According to Jensen and Wood (1997), the Dickey-Fuller test, in a univariate situation, cannot be corrected by Barlett's method. At this point, it is important to point out that the correction factor suggested above is based upon the model in Equation (4.114) and makes the idealised assumptions that the errors are identically and independently distributed and Gaussian, and that the lag length and cointegrating rank are correctly specified. When applying this correction, it is important that these assumptions are carefully met. In Johansen's own words, 'the calculation is useful as a complement to the asymptotic analysis since it ... demonstrates that an uncritical use of asymptotic tables can be misleading'.[44]

Under the null hypothesis that $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ and $\boldsymbol{\Upsilon} = \boldsymbol{\alpha\rho}'$, the model in Equation (4.114) becomes

$$\Delta\mathbf{X}_t = \boldsymbol{\alpha}(\boldsymbol{\beta}'\mathbf{X}_{t-1} + \boldsymbol{\rho}'t^{n_d}) + \sum_{i=1}^{k-1}\Gamma_i\Delta\mathbf{X}_{t-i} + \sum_{i=0}^{n_d-1}\Phi_i t^i + \varepsilon_t, \tag{4.123}$$

recalling that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $n \times r$ and $\boldsymbol{\rho}$ is $1 \times r$. The expansion of the expectation of the likelihood ratio test of $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ and $\boldsymbol{\Upsilon} = \boldsymbol{\alpha\rho}'$ in Equation (4.114) is derived, under the assumption that $\mathbf{X}_t$ is an $I(1)$ process, with cointegrating rank r, given by Equation (4.123). A corollary of this gives the correction factor for the likelihood ratio test. Assuming that $\mathbf{X}_t$ is an $I(1)$ process with cointegrating rank r given by Equation (4.123), the correction factor for the test of Equation (4.123) in Equation (4.114), i.e., the test for cointegrating rank r in the vector autoregressive model Equation (4.114), is given by Equation (4.122) where

$$\begin{aligned} b(\theta) &= c_1\left(1 + h(n_b, n_d)\right) + \left(n_b c_2 + 2(c_3 + n_d c_1)\right)\frac{g(n_b, n_d)}{n_b^2} \\ &\quad - 2tr\left[\boldsymbol{\psi}'\boldsymbol{\Sigma}^{-1}\sum_{j=0}^{\infty}\psi_j\right] k(n_b, n_d, j). \end{aligned} \tag{4.124}$$

Note that $\boldsymbol{\Sigma}$, $\boldsymbol{\psi}$ and ψ_j are variance and long-run coefficient parameters, where

$$\begin{aligned} c_1 &= tr\{\mathbf{V}_\psi\}, \\ c_2 &= tr\{\mathbf{I}_{n_y} - \mathbf{V}_\theta - \mathbf{V}_\psi\}, \\ c_3 &= tr\{\mathbf{V}\}, \end{aligned} \tag{4.125}$$

and

$$\begin{aligned} \mathbf{V}_\theta &= \boldsymbol{\theta\theta}'\boldsymbol{\Sigma}^{-1}, \\ \mathbf{V}_\psi &= \boldsymbol{\psi\psi}'\boldsymbol{\Sigma}^{-1}, \\ \mathbf{V} &= \boldsymbol{\psi}'\boldsymbol{\Sigma}^{-1}\textstyle\sum_{m=0}^{\infty}\psi_m tr\{\boldsymbol{\Sigma}^{-1}\gamma(m+1)\} + \boldsymbol{\psi}'\boldsymbol{\Sigma}^{-1}\displaystyle\sum_{m=0}^{\infty}\gamma(m+1)'\boldsymbol{\Sigma}^{-1}\psi_m. \end{aligned} \tag{4.126}$$

Note that $\boldsymbol{\Sigma}^{-1}$ is the variance of $\mathbf{y}_t$, which is a function of $\mathbf{X}_t$, and that $\boldsymbol{\theta\theta}'$ and $\boldsymbol{\psi\psi}'$ are

[43] See, for example, Bartlett (1937) and Lawley (1956).
[44] Johansen (2002), p. 1931.

the long-run variances of $\mathbf{y}_{\theta t}$ and $\mathbf{y}_{\psi t}$, where $\mathbf{y}_t = \mathbf{y}_{\theta t} + \mathbf{y}_{\psi t}$. Fuller details can be found in Johansen.[45]

To implement the correction, $f(n_b,\ n_d)$, $a(n_b,\ n_d)$, $g(n_b,\ n_d)$, $h(n_b,\ n_d)$ and $k(n_b,\ n_d,\ j)$ must be calculated. As these are complicated functions, Johansen (2002) tabulated them by simulation, for values of $n_d = 0,\ 1,\ 2$. These estimates may be found in Tables I and II.[46] Given these tabulated simulation values, the proposed correction factor becomes

$$a\left(T, n_b, n_d\right)\left(1 + b\left(\widehat{\theta}\right)\right), \tag{4.127}$$

where

$$b(\theta) = c_1(1 + h(n_b, n_d)) + (n_b c_2 + 2(c_3 + n_d c_1))\frac{g(n_b, n_d)}{n_b^2}. \tag{4.128}$$

4.5.2 Implementing the small sample correction

The previous section details the theory behind the small sample correction. As stated, to implement the correction, $f(n_b,\ n_d)$, $a(n_b,\ n_d)$, $g(n_b,\ n_d)$, $h(n_b,\ n_d)$ and $k(n_b,\ n_d,\ j)$ must be calculated. This would certainly appear to involve considerable effort and is not straightforward. Johansen, Hansen, and Fachin (2002), in an unpublished paper written to accompany Johansen (2002), outlined the calculations required to estimate the correction factor. Source code for the econometric software RATS is included,[47] requiring the user to supply just the following: the number of observations, T, n and r, the dimensions of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, values for the matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\Omega}$ and Γ_i, where $i = 1,\ 2,\ \ldots,\ k-1$, the number of matrices included in total and the type of model to be estimated.[48] The RATS code produces an estimate of the small sample correction given this information, which can then be applied to the Trace test statistic in question.

[45] Johansen (2002), p. 1933-1938.

[46] Johansen (2002), p. 1939.

[47] This code is available from `http://www.math.ku.dk/~sjo/`.

[48] Like any new procedure, implementing the correction may be less than straightforward. Although the RATS code supplied by Johansen, et al. (2002) greatly simplifies the procedure, some issues remain. The small sample correction requires several inputs. These are the matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\Gamma}_i$, $i = 1,\ \ldots,\ k-1$ and $\boldsymbol{\Omega}$. These estimates are obtained after the appropriate number of cointegration vectors have been selected in the usual manner. To obtain the correction factor, the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ must have the same dimensions. While this may not appear to be problematical in many cases, the situation where a restricted constant or trend is included in the cointegrating VAR is noteworthy. In such circumstances, an $\boldsymbol{\alpha}$ vector of dimension $n \times r$ will be accompanied by a $\boldsymbol{\beta}^*$ vector of dimension $(n+1) \times r$. Evaluating the correction factor in this case requires the upper $n \times r$ elements of $\boldsymbol{\beta}^*$ to be taken. This $n \times r$ upper, $\boldsymbol{\beta}$, is used in calculating the correction factor.

Chapter 5

The Demand for Money

Research carried out in the course of writing this chapter appear as 'Demand for Money: A Study in Testing Time Series for Long Memory and Nonlinearity' by Bond, D., M.J. Harrison, and E.J. O'Brien, in *The Economic and Social Review*, 38, 2007. Also published as 'Testing for Long Memory and Nonlinear Time Series: A Demand for Money Study' in Trinity Economic Papers, No. 21 (2005). It was presented at the 20^{th} Irish Economic Association Annual Conference, Bunclody, Co. Wexford, 28^{th}-30^{th} April 2006.

5.1 Introduction

The importance of the concepts of stationarity and regime stability in economic and financial time-series modelling is well established. However, recent concerns about the interrelationship between these two concepts, and the associated problems for applied work, have ensured that they remain a significant focus for research. Early studies, such as those by Bhattacharya, Gupta, and Waymire (1983) and Perron (1989), highlighted the difficulty of distinguishing between time series generated by difference stationary processes and those generated by nonlinear but stationary processes. Since then, an increasing research emphasis has been on the problem of distinguishing between long memory and nonlinearity. The developing interest in long memory models has been stimulated, in particular, by a growing awareness of the limitations of the simple $I(1)/I(0)$ framework. For example, Baillie and Bollerslev (2000) and Maynard and Phillips (2001) showed how the low power of familiar unit root tests, such as those introduced by Dickey and Fuller (1981), could lead to incorrect inference in the Fama (1984) regression model of the relationship between future spot and forward exchange rates, and how the empirical work could be set in a framework of fractional integration using a long memory model. Long memory models and fractional (co)integration are now popular in several other areas of the applied literature; see, for example, Gil-Alana (2003), Liu and Chou (2003), Dittmann (2004) and Masih and Masih (2004). A major problem with such models is that it is not easy to distinguish them empirically from models with regime switching or more general nonlinearities; see, for example, Diebold and Inoue (2001).

In the theoretical literature, two main strands of discussion have developed. The first is that of testing for difference stationarity when the processes are in fact nonstationary; see Perron and Qu (2004) for references. The second concerns testing for structural breaks when long memory is a possibility; see Nunes, Kuan, and Newbold (1995), Krämmer and Sibbertsen (2002) and Hsu (2001). Recent work by Mayoral (2005) and Dolado, Gonzalo, and Mayoral (2005b) has developed specific tests for difference stationarity against the alternative of stationarity with a structural break. All of these studies use conventional parametric techniques for either modelling or testing for nonlinearities. The recent development of random field regression has also provided a suite of tests for structural breaks, nonlinearity and time-varying parameters; for example, chapters 2 and 3 discuss tests suggested by Hamilton (2001) and Dahl and González-Rivera (2003). The strength of this alternative approach is that it does not rely on any functional form being specified prior to estimation.

The purpose of this chapter is to compare the performance of traditional integration analysis, the fractional integration approach and random field regression-based inference, all introduced in previous chapters of this thesis, using a standard economic model and a well-known time-series dataset. The discussion is structured as follows. In Section 5.2, the theoretical background to money demand is briefly explained. The data used in this study is described and the standard $I(1)/I(0)$ analysis is then conducted in sections 5.3 and 5.4. The univariate analysis of the series, using the augmented Dickey-Fuller (ADF) testing strategy proposed by Dolado, Jenkinson, and Sosvilla-Rivero (1990) is implemented to determine whether the individual series are trend stationary or difference stationary. The Engle-Granger (1987) error-correction (ECM), the Johansen (1988, 1991) vector autoregression (VAR) and the common factor approaches are used to investigate the possibility of cointegration, with the augmented Engle-Granger (AEG) test, the cointegrating regression Durbin-Watson (CRDW) test of Sargan and Bhargava (1983) and the ECM test due to Banerjee, Hendry, and Smith (1986) being used in the former case. The p-values from MacKinnon (1996), MacKinnon, Haug, and Michelis (1999), Ericsson and MacKinnon (2002) and standard normal tables are used, as appropriate. The effect of applying Johansen's (2002) small sample correction is also examined. This correction is based on the Bartlett (1937) correction and assumes that the errors are normal, independent and identically distributed, as described in Section 4.5 of the previous chapter.

Having conducted the standard cointegration analysis, the long memory and fractional integration analysis is undertaken in Section 5.5. Inference is problematical here, as none of the usual procedures are appropriate. The classical asymptotics of the $I(0)$ case do not apply when time series are fractionally integrated and neither does the standard cointegration approach. In the $I(1)$ case, conventional tests depend on the statistics converging to known functionals of Brownian motion. When $d \neq 1$, however, these are replaced by functionals of fractional Brownian motion. Taking the approach of testing for $I(1)$ against $I(d)$ is also problematical, since tests such as the ADF test of Dickey and Fuller (1981), while consistent, have very low power; see Diebold and Rudebusch (1991) and Hassler and Wolters (1994). Furthermore, the precision with which the parameters are estimated hinges on the correct specification of the model; see Hauser, Potscher, and Reschenhofer (1999). The situation becomes even more complex when the concept of fractional cointegration is entertained. As Phillips (2003, p. c30) pointed out, 'The problems presented by these models of fractional cointegration seem considerably more complex than the $I(1)/I(0)$ ... case that is now common in applications'. Only univariate analysis is attempted, therefore. In particular, it seems unlikely that the series in either of the two cases considered all have the same level of fractional integration. The 'over differenced' ARFIMA model, using Δy_t rather than y_t, is estimated, as recommended by Smith, Sowell, and Zin (1997), to avoid the problems associated with drift. Four estimates of d are calculated using the Doornik and Ooms (1999) ARFIMA package, namely, Sowell's (1992) exact maximum likelihood (EML) estimator, Beran's (1995) nonlinear least squares (NLS) estimator, the method of Geweke and Porter-Hudak (1983) (GPH) and Robinson's (1994) Gaussian semiparametric (GSP) estimator, all previously introduced in Chapter 4. The fact that the first of these requires $d < 0.5$ is another reason for using the 'over-differenced' model. The estimates of d are then used in the fractional Dickey-Fuller (FDF) and fractional augmented Dickey-Fuller (FADF) tests, with the Schwarz (Bayesian) information criterion (SIC) being used as the basis for the choice of the lag length for the test. In Section 5.6, the random field regression approach is applied to the two cases, using the GAUSS code provided by Hamilton (2001). Finally, in Section 5.7, the results of the analysis are discussed and some practical conclusions drawn.

5.2 The Demand for Money

To investigate the application of both the new long memory tests and the random field approach, a standard applied economics problem, namely, the estimation of a demand for money function, is considered in this section. The well-known datasets for Denmark and Finland, provided by Johansen and Juselius (1990), are used. The well documented instability of the demand for money function in many countries has led to several studies that place the analysis of money demand in the $I(1)/I(0)$ framework; see, for example, Astley and Haldane (1997), Fiess and McDonald (2001), Mark and Sul (2003) and Choi and Saikkonen (2004).

Following Johansen and Juselius (1990), a simple demand for money function can be specified for Denmark and Finland as

$$m_t = \alpha + \beta_1 y_t + \beta_2 p_t + \beta_3 i_t + \beta_4 b_t + \varepsilon_t, \tag{5.1}$$

where m_t is the logarithm of some measure of money demand, y_t is the logarithm of real income, p_t is the logarithm of the inflation rate, i_t is the deposit interest rate and b_t is the bond rate at time t. For Finland, β_4 is assumed to be zero as no data are available. There is a wealth of empirical evidence to support this framework. Regardless of the measure of money supply used, the negative effect of interest rates has been widely confirmed, as has the positive effect of real spending and the positive effect of the real cost of transactions; see, for example, Goldfeld (1973, 1976), Goldfeld and Sichel (1990), Fair (1987) and Mulligan and Sala-i-Martin (1996).

5.3 Data and Preliminary Analysis

The component factors of money demand have long been known to exhibit nonstationary behaviour, and hence modelling the demand for money has commonly been placed in the $I(1)/I(0)$ framework. Two datasets, provided by Johansen and Juselius (1990) are used. The first, for Denmark, samples the period 1974 to 1987, a total of 55 quarterly observations. The variables include a measure of money demand, proxied by $M2$, in logarithmic form, m_t^{Den};[1] national income, as the logarithm of gross domestic product, y_t^{Den}; and the inflation rate, again in logarithmic form, p_t^{Den}. The costs of holding real money balances is proxied by the bond rate, b_t^{Den}, and the deposit rate, i_t^{Den}.[2]

The Finnish data samples the period 1958 to 1984, a total of 106 quarterly observations. It consists of four variables. The demand for money is proxied by the logarithm of $M1$, m_t^{Fin}.[3] The logarithm of real income is included as y_t^{Fin}. To proxy the cost of holding money balances, both the inflation rate, p_t^{Fin}, and the marginal rate of interest of the Bank of Finland, i_t^{Fin}, are used. The data can be found in Appendix C.1; Table C.1 contains the data for Denmark, while the data for Finland can be found in Table C.2.[4]

To begin, each of the data series were plotted. These plots can be found in Appendix C.2: figures C.1 to C.4 for Denmark and figures C.5 to C.7 for Finland. Table C.3 gives the results of the Dolado, et al. (1990) unit root testing strategy for the Danish and Finnish variables, respectively.[5] For Denmark, all of the data series appear to be $I(1)$. In all but

[1] $M2$ is defined as the sum of currency in circulation and in cheque accounts, itself a definition of $M1$, plus consumer time deposits, money-market deposits and some other items.

[2] The bank deposit rate for interest bearing deposits, which are in fact, a large part of $M2$.

[3] See previous definition of $M2$.

[4] These data can also be downloaded from Søren Johansen's website at `http://www.math.ku.dk/~sjo/`.

[5] See Chapter 4 for an introduction to the method of Dolado, et al. (1990).

one case, the constant and trend terms are found to be insignificant; in the remaining case, the standard normal probabilities are used. For the Finnish data, however, only the m_t^{Fin} and y_t^{Fin} variables seem to be $I(1)$, though the inference is marginal for y_t^{Fin}. In the case of Finland's m_t^{Fin} variable, the constant in the ADF test is only marginally insignificant, but if it is treated as significant, the ADF test still supports the null of a unit root, with a test statistic of -0.760 and an associated p-value of 0.826. By contrast, the unit root null is rejected decisively for Finland's p_t^{Fin} and i_t^{Fin} series. It is noteworthy, though, that if, for these last two variables, the Akaike information criterion (AIC) is used instead of the SIC, the choice of lag lengths for the ADF tests, and the test results, are different: the suggestion then is that, like m_t^{Fin} and y_t^{Fin}, the Finnish price and interest rate variables are also $I(1)$. To ensure that those series found to be $I(1)$ were in fact integrated to order one, the first differences of those series were tested for nonstationarity. Each series was found to be stationary after differencing once, indicating that each of those series were indeed $I(1)$. These results generally confirm the earlier findings of Johansen and Juselius (1990) and Johansen (1996).[6]

Although the data are quarterly, the issue of possible seasonal integration has been ignored up to this point. A more detailed examination of the issue of seasonal unit roots was undertaken, using the procedures of Hylleberg, Engle, Granger, Yoo (1990), following the procedures outlined in Ghysels and Osborn (2001).[7] Table C.4 contains the results for Denmark. These results generally confirm the findings of the ADF tests discussed previously and there is little evidence of seasonal integration. For m_t, with and without an intercept and a time trend, there is some evidence that $\pi_2 \neq 0$, suggesting a seasonal unit root, at the 5 per cent significance level. But at the 10 per cent level, where the critical value is -1.59, the null of no seasonal unit root can not be rejected. In cases where seasonal dummies have been included in the test specification, the evidence suggests that only nonseasonal unit roots are present in m_t. For the remaining variables, the HEGY tests finds no seasonal integration; the results of the standard ADF tests are confirmed, with all series being found to be $I(1)$.

The results for Finland can be found in Table C.5. There is some evidence here of seasonal integration in m_t and y_t, regardless of whether seasonal dummy variables are included in the test specification or not. For both of these variables, the null of $\pi \neq 0$ is rejected, and with intercept and trend, F_{234} is only marginally significant in both cases. The remaining hypotheses can not be rejected, suggesting therefore, that if there is seasonal integration in these variables, it is semi-annual in nature. While it appears that p_t and i_t are not seasonally integrated, there is mixed evidence as to whether they are even $I(1)$. This very much confirms the previous findings of the ADF tests.

To investigate further, the KPSS, ERS and NP alternative unit root tests were conducted.[8] While the latter two tests have as their null hypothesis that the series has a unit root, the first has the null that the series is stationary and the alternative hypothesis that it has a unit root. For the Danish data, the additional tests broadly confirm the previous findings. In only a few cases does the KPSS test fail to reject the null hypothesis of stationarity. One case is

[6]The Danish dataset introduced by Johansen and Juselius (1990) was subsequently used by Johansen (1996).

[7]The so-called HEGY test of Hylleberg, et al. (1990) was discussed in Chapter 4. Other tests are available to test for seasonal integration, including Osborn, Chui, Smith, Birchenhall (1988).

[8]These tests were briefly introduced in Chapter 4.

that of the money demand variable, m_t, when Parzen kernel estimation is used and no trend is specified. The other is that of the income variable, y_t, when a trend is allowed for in the specification. In this latter case, the result holds for any of the spectral estimation methods, but not for the moment estimators. For the Finnish data, the results are less clear. For all variables, the NP test, which it has been argued has better power than standard $I(1)/I(0)$ tests, tends to reject the null hypothesis of a unit root. This is often supported by the KPSS and ERS tests. Details of the results have been omitted for compactness.

5.4 Testing for Cointegration

5.4.1 The Engle-Granger 2-Step method

On the assumption that the variables are $I(1)$, which seems to be a far safer assumption to make for Denmark than for Finland, the Engle-Granger two-step approach to cointegration gives the estimated levels models, and associated AEG and CRDW test results for the OLS residuals, presented in Table C.6. Using the 5 per cent significance level, there is little evidence for both countries that a cointegrated money demand relationship might exist. Only in the case of Finland, when p_t and i_t are ignored in view of the fact that they seem to be $I(0)$ using the Dolado, et al. (1990) procedure and the supplementary unit root checks, is cointegration of m_t and y_t suggested by the AEG and CRDW tests, but even then only marginally.

The estimates of parsimonious error-correction models, using the lag of the residuals from the levels regression models as the error-correction terms, are given in Table C.7. The models are statistically acceptable in the sense that they are supported by a range of misspecification diagnostics. Only in the case of the equations for Finland is there a marginal suggestion of heteroscedasticity. However, with R^2 values around 0.5, the fits are quite poor and there is a high incidence of insignificance of the estimated coefficients. In particular, the coefficient on the error-correction term is highly insignificant for Denmark, while the coefficients are perversely signed, albeit significantly, for the three cases relating to Finland. The ECM test decisively rejects cointegration in all cases. Even in the one case for Finland in which the AEG and CRDW tests suggest the possibility of cointegration, the ECM test rejection is unambiguous.

5.4.2 Johansen's maximum likelihood approach

The Danish data have been used extensively by Johansen and it is clear from his various results that the argument that there is a cointegrating money demand relationship depends largely on the VAR specification and the test statistic used; see Johansen (1988), Johansen and Juselius (1990) and Johansen (2002). Table C.8 gives a summary of the results that can be obtained for Denmark using Johansen's approach and a VAR lag length of one, as suggested by the SIC and the adjusted likelihood-ratio test. See tables C.9 and C.10 for further details.[9] As can be seen, a range of specifications concerning intercepts and trends was examined for

[9] The AIC and unadjusted likelihood-ratio test suggested a lag length of two. The choice of lag length one has the advantage of economising on degrees of freedom.

variants of the model with and without centred seasonal dummy variables.[10] Examination of the various VAR estimates suggested that the specification with restricted intercepts and trends was the most appropriate, although for comparison, the case for unrestricted intercept and trend is included also.[11] Moreover, given that the data used were quarterly, the variant with seasonal dummies, sc_i, $i = 1, 2, 3$, was also preferred. There is variability in the suggested number of cointegrating relationships across the range of specifications used, and between the Trace test and the Maximal eigenvalue test used to ascertain this number. The surprise is that despite the results from the static cointegrating regressions and error-correction models, which overwhelmingly point to no cointegration, all of the results in Table C.8, except one, suggest at least one cointegrating vector. In the case of the preferred specification, the suggestion is of one cointegrating relationship, in contrast to the outcome produced by the Engle-Granger (1987) approach.

For the Finnish data, the summary results of the Johansen procedure on the full model are given in Table C.11. There is similar variability in the number of cointegrating relationships suggested for the different specifications and tests to that noted for Denmark, though it is not quite as marked. The preferred specification is again that with restricted intercepts and trends, although the case with unrestricted intercept and unrestricted trends is also considered, and seasonal dummies, for which case the number of cointegrating relationships indicated is two, again in stark contrast to the earlier indications of no cointegration. As Johansen and Juselius (1990) have pointed out, the interpretation of the findings for the Finnish data poses particular problems. Accordingly, two alternative reduced models for Finland were also investigated: one taking p_t to be $I(0)$ in the VAR analysis and the other treating both p_t and i_t as $I(0)$. The summary results for these cases are given in Table C.12 and Table C.13, respectively. Table C.12 contains consistent indications of a single cointegrating vector across all VAR specifications and tests, though once again this finding contradicts the indications from the AEG, CRDW and ECM tests. Slight variability in the results for different specifications and tests is seen in Table C.13, but in this case no cointegration is suggested for the preferred specification. This finding conflicts with the corresponding AEG and CRDW results, which indicate a possibility of cointegration, but it is in agreement with the ECM test result. Tables C.14 and C.15 provide information on the choice of the VAR specification in this case.

The Johansen (2002) bias-correction factor was calculated only for the two variants of the preferred VAR specification in the case of Denmark, and for the preferred specification of the full and the two reduced models in the case of Finland. Table C.16 and C.17 for the alternative specification, presents the Danish results. Although the correction factor relates only to the Trace test, details of the Maximal eigenvalue test are also given. The corresponding results for the full Finnish model and the two reduced versions are given in tables C.18 to C.23,

[10] In this context, an intercept is denoted by c, a trend term by t and centred seasonal dummy variables by sc_i, $i = 1, 2, 3$.

[11] In a cointegrating vector, for unrestricted intercepts and trends, X_t will be trend stationary when the rank of $\mathbf{\Pi}$ is full. But if it is rank deficient, the solution for X_t will contain quadratic trends. For unrestricted intercepts and no trends, a rank deficiency in $\mathbf{\Pi}$ will result in X_t containing linear deterministic trends. To avoid these situations, the use of restricted intercepts and no trends, or unrestricted intercepts and restricted trends is the normal practice. However, this results in the cointegrating vectors containing a deterministic trend in the first case and intercepts in the second case.

inclusive. Interestingly, when the adjusted critical value is used for the Trace test, the result for Denmark changes to one suggesting no cointegrating relationships, in accordance with the AEG, CRDW and ECM test findings. Thus there is conflict between the Trace test and the Maximal eigenvalue test in the case considered, the latter indicating one cointegrating relationship. The correction factors are close to unity for the Finland cases, probably due to the larger sample size. Even so, the outcome for the full Finnish model is similar to that for Denmark; the modified Trace test indicating the reduced number of one cointegrating relationship, while the Maximal eigenvalue test indicates two. However, the correction has no effect in the cases of the two reduced models. In particular, as the correction would increase the critical value of the Trace statistic, and as the test statistic for the second reduced model already lies well below the uncorrected critical value, as can be seen from Table C.23, the correction factor was not even computed for this final case. The conclusion suggested by the modified Johansen procedure remains that the number of cointegrating vectors is one and zero for the first and second reduced Finnish models, respectively.

It can be seen from these various results that the traditional analysis is somewhat confusing. Examination of the Danish data seems to suggest that all variables are $I(1)$ and, using the Engle-Granger (1987) 2-step procedure, that cointegration does not hold and error-correction models are not appropriate. Yet, using the original Johansen VAR approach, there are strong indications of cointegration, which are only challenged if a bias-corrected Trace test is undertaken. The Finnish data give rise to some similar findings, although in contrast to the Danish case, unit root tests suggest that some of the series are possibly not $I(1)$. When allowance is made for this possibility, the Engle-Granger approach marginally supports cointegration. However, when the Johansen technique is applied in this case, it gives contrary results, whether or not a modified Trace test is used, indicating that there is no cointegration.

5.4.3 Common factor analysis

The final test for cointegration to be considered in this chapter is the common factor approach, as outlined in Chapter 4. In considering the money demand relationship, the equation to be estimated is

$$\begin{aligned} m_t^{Den} = c + \beta_1 m_{t-1}^{Den} + \gamma_0 y_t^{Den} + \gamma_1 y_{t-1}^{Den} + \delta_0 p_t^{Den} + \delta_1 p_{t-1}^{Den} + \eta_0 b_t^{Den} + \eta_1 b_{t-1}^{Den} \\ + \mu_0 i_t^{Den} + \mu_1 i_{t-1}^{Den} + \varepsilon_t. \end{aligned} \tag{5.2}$$

To find a cointegrating relationship, the common factor approach must first find evidence of common roots in the specified data, then that those common roots are indeed unity, i.e., unit roots, and finally, that the specification is in terms of an error-correction model. The first restrictions to consider, those that test for common factors are,

$$\gamma_1 + \gamma_0\beta_1 = 0, \tag{5.3}$$

$$\delta_1 + \delta_0\beta_1 = 0, \tag{5.4}$$

$$\eta_1 + \eta_0\beta_1 = 0, \tag{5.5}$$

$$\mu_1 + \mu_0\beta_1 = 0. \tag{5.6}$$

Only when each of these restrictions is found to be significant, i.e., when the restriction is rejected, should the tests of restrictions for unit roots and an error-correction specification be carried out. For simplicity, only the common factor restrictions, as outlined above, will be tested in this case. These tests are sufficient to indicate if a cointegrating relationship is possible, allowing for comparison with the alternative methods. The full OLS estimates of Equation (5.2) can be found in Table C.24. The estimated equation was

$$\begin{aligned} m_t^{Den} = \underset{(0.67)}{2.55} + \underset{(0.08)}{0.70} m_{t-1}^{Den} + \underset{(0.18)}{0.40} y_t^{Den} - \underset{(0.17)}{0.21} y_{t-1}^{Den} - \underset{(0.58)}{0.60} p_t^{Den} - \underset{(0.58)}{0.62} p_{t-1}^{Den} \\ - \underset{(0.38)}{0.87} b_t^{Den} - \underset{(0.45)}{0.23} b_{t-1}^{Den} + \underset{(0.64)}{0.36} i_t^{Den} - \underset{(0.64)}{0.26} i_{t-1}^{Den} + \hat{\varepsilon}_t, \end{aligned} \tag{5.7}$$

where standard errors are given in parenthesis. Table C.25 shows the results for the test of each individual common factor restriction plus the test of joint significance for all restrictions.

As can be seen from these results, three of the individual test results are insignificant. One of the results, however, is significant, as is the test of joint restrictions, rejecting therefore, the restriction. Since all four restrictions must be valid, it can be assumed that Equation (5.2) does not contain common factors for all variables. These findings suggest that cointegration cannot be ruled out in the case of Denmark.

The COMFAC analysis for Finland began by looking at the most general specification for money demand, that is, including all of the variables introduced to date. Recall, however, that the variables, p_t^{Fin} and i_t^{Fin}, are likely to be stationary. As before, several models were estimated here; a full model including all variables and two reduced models, excluding i_t^{Fin} and p_t^{Fin}. To test this most general specification, the following equation was estimated

$$\begin{aligned} m_t^{Fin} = -\underset{(0.10)}{0.23} + \underset{(0.06)}{0.81} m_{t-1}^{Fin} + \underset{(0.09)}{0.60} y_t^{Fin} - \underset{(0.10)}{0.41} y_{t-1}^{Fin} - \underset{(0.43)}{0.91} p_t^{Fin} - \underset{(0.42)}{0.03} p_{t-1}^{Fin} \\ + \underset{(0.13)}{0.35} i_t^{Fin} - \underset{(0.13)}{0.33} i_{t-1}^{Fin} + \hat{\varepsilon}_t. \end{aligned} \tag{5.8}$$

Results in full can be found in Table C.26. The common factor restrictions were tested, producing the results found in Table C.27. As can be seen, two of these restrictions prove to be significant, along with the test of joint restrictions, once again suggesting that a cointegrating relationship cannot be ruled out. Recall that the details of these restrictions have been introduced previously, in Chapter 4, and above. Given the evidence that i_t^{Fin} may be stationary, the relationship without this variable was estimated, giving

$$m_t^{Fin} = -\underset{(0.09)}{0.23} + \underset{(0.06)}{0.81} m_{t-1}^{Fin} + \underset{(0.09)}{0.68} y_t^{Fin} - \underset{(0.10)}{0.48} y_{t-1}^{Fin} - \underset{(0.44)}{0.91} p_t^{Fin} + \underset{(0.43)}{0.09} p_{t-1}^{Fin} + \hat{\varepsilon}_t. \tag{5.9}$$

These results can be found in Table C.28. Once again, one of the common factor restrictions tested here were found to be significant, as was the joint test of restrictions, as can be seen from Table C.29. Again, cointegration cannot be ruled out.

Finally, just the money demand and income variables were included in the estimation, producing

$$m_t^{Fin} = -\underset{(0.08)}{0.18} + \underset{(0.06)}{0.80} m_{t-1}^{Fin} + \underset{(0.09)}{0.73} y_t^{Fin} - \underset{(0.10)}{0.54} y_{t-1}^{Fin} + \hat{\varepsilon}_t. \tag{5.10}$$

Full results for this estimation are available in Table C.30. The common factor restriction proved to be marginally significant, as can be seen from Table C.31, once again rejecting the restriction and failing to reject the possibility of cointegration. Overall, it must be concluded that in the three specifications considered, the COMFAC approach cannot reject the possibility of cointegration in the Finnish data.

5.5 Testing for Fractional Integration

Having raised concerns over the standard $I(1)/I(0)$ analysis, the next step is to consider the possibility of fractional integration.[12] Table C.32 gives the results of the fractional analysis for the Danish data. For each variable, a range of estimates of d is provided, as well as the results of the FDF and FADF tests. The corresponding results for the Finnish data are given in Table C.33.

The previous standard analysis generally found the Danish data to be $I(1)$. From Table C.32, it can be seen that there is mixed evidence in support of the data being $I(1)$, somewhat in contradiction to the previous findings. It is possible, if just the parametric estimators of d are considered, to argue that the Danish b_t variable is fractionally integrated. The nonparametric and semiparametric estimators suggest the series in question may be $I(1)$. For the Finnish data, it would appear that three of the four variables are fractionally integrated, namely, m_t, p_t and i_t. It will be recalled that unit root tests decisively rejected the unit root null for the latter two variables. The results for Finland's y_t variable also give indications that it is fractionally integrated, but the FADF result in this case has the wrong logical sign. Overall, the investigation of fractional integration suggests that the Finnish data series are not generated by $I(1)$ processes but that the Danish data may be.

5.6 Nonlinear Inference

In light of the possibility that the emerging difficulties may be related to parameter instability or some other type of nonlinearity, of what may be stationary data generating processes, Hamilton's (2001) random field approach was used to explore the likely form of the two models, and this leads to some interesting results. Hamilton's LM test statistics for nonlinearity for the Danish and Finnish models were 15.34 and 123.81, respectively, which are significantly greater than the 5 per cent critical χ_1^2 value of 3.84, again suggesting that the models should not be simply linear.

[12] As Parke (1999) pointed out, 'a growing body of empirical evidence supports the notion that important economic data series might be fractionally integrated'. Evidence of such long memory behaviour has been found in business-cycle indicators, price indices, asset prices and exchange rate volatility; see for example, Geweke and Porter-Hudak (1983), Diebold and Rudebusch (1989), Sowell (1992), Ding, Granger, and Engle (1993), Baillie, Bollerslev, and Mikkelsen (1996), Baillie, Chung, and Tieslau (1996), Andersen and Bollerslev (1997), Breidt, Crato, and de Lima (1998), and Andersen, Bollerslev, Diebold, and Labys (1999). Several explanations of long memory in economic data have been put forward. Granger (1980) suggested that fractional integration may arise from data aggregation. Liu (1995) suggested that regime-switching in stock market volatility may result in fractional processes. Apart from Parke (1999), Baillie (1996) offered an excellent survey of fractional integration.

Given the prior belief is that the nonlinear relation

$$m_t^{Den} = f\left(y_t^{Den}, p_t^{Den}, b_t^{Den}, i_t^{Den}\right) + \nu_t, \tag{5.11}$$

may well explain the variations in real money demand, m_t^{Den}, those variables in Equation (5.11) were analysed using Hamilton's (2001) methodology. The results of the flexible nonlinear inference, which can also be found in Table C.34, with the nonlinear optimisation making use of the algorithm switching method STEEPEST DESCENT-NEWTON, $\zeta = 1.0$ and converging after 46 iterations, were

$$\begin{aligned} m_t^{Den} &= \underset{(1.14)}{7.34} + \underset{(0.19)}{0.78}\, y_t^{Den} + \underset{(0.06)}{0.13}\, p_t^{Den} - \underset{(0.04)}{0.11}\, b_t^{Den} - \underset{(0.06)}{0.07}\, i_t^{Den} \\ &+ \underset{(0.01)}{0.01}\left[\underset{(4.01)}{5.38}\, m\left(\underset{(2.34)}{3.41}\, y_t^{Den},\ \underset{(1.39)}{6.49}\, p_t^{Den},\ \underset{(0.57)}{0.000003}\, b_t^{Den},\ \underset{(0.51)}{-0.00002}\, i_t^{Den}\right) + \nu_t\right]. \end{aligned} \tag{5.12}$$

where standard errors are given in parenthesis. Clearly, only p_t^{Den} contributes significantly to the nonlinearity in the estimated relation above. The conditional expectation function, $\widehat{E}\left[\overline{y}_t^{Den}, \overline{b}_t^{Den}, \overline{i}_t^{Den}, p_t^{Den} | \mathbf{m}_T^{Den}\right]$, was plotted, therefore, as a function of p_t^{Den}, in Figure C.8.[13] This represents the demand for money one would expect for any price level, if the income, bond and deposit rates were equal to their average values for the sample period in question. The Hamilton (2001) results from the Danish data are rather disappointing, in so much as both σ and ζ estimates are not statistically significant on the basis of an asymptotic t-test. It could be argued, along the lines of Dahl and González-Rivera (2003), that this is due to nuisance parameter problems, given that under the null of linearity, the g_i parameters are unidentified. If the statistical insignificance of $\widetilde{\sigma}$ and $\widetilde{\zeta}$ is ignored, the significant coefficient of p_t in the linear and the nonlinear components of the Danish model strongly suggests that this inflation variable is the prime source of any parameter instability. This is interesting given that Johansen and Juselius (1990) excluded this variable from their analysis as 'it did not enter significantly into the cointegration relation for money demand' (Johansen and Juselius, 1990, p. 172).

In the case of Finland, the results in Table C.34 are more satisfying. Both $\widetilde{\sigma}$ and $\widetilde{\zeta}$ are statistically significant, in agreement with the implied value of λ in the LM test, and suggesting that there is significant nonlinearity in the money demand relationship. In the Finnish case, it is the income variable, y_t, that proves significant in both the linear and nonlinear parts of the model and, therefore, that needs to be investigated further. As before, it was assumed that a nonlinear relationship of the form

$$m_t^{Fin} = f\left(i_t^{Fin}, p_t^{Fin}, y_t^{Fin}\right) + \nu_t, \tag{5.13}$$

described the Finnish data. The test for nonlinearity overwhelmingly rejected the null of linearity. The results of the nonlinear analysis, for STEEPEST DESCENT-NEWTON, $\zeta = 1.5$

[13] $\mathbf{m}_T^{Den}$ is defined as $\mathbf{m}_T^{Den} = \left(m_T^{Den}, \mathbf{x}_T', m_{T-1}^{Den}, \mathbf{x}_{T-1}', \ldots, m_1^{Den}, \mathbf{x}_1'\right)$, where $\mathbf{x}_T' = \{y_T^{Den}\, p_T^{Den}\, b_T^{Den}\, i_T^{Den}\}$.

and where convergence was achieved after 33 iterations, were

$$m_t^{Fin} = \underset{(0.35)}{-0.55} + \underset{(0.17)}{0.13}\, i_t^{Fin} - \underset{(0.46)}{0.83}\, p_t^{Fin} + \underset{(0.08)}{0.87}\, y_t^{Fin} \tag{5.14}$$
$$+ \underset{(0.005)}{0.05} \left[\underset{(0.31)}{1.29}\, m \left(\underset{(2.17)}{2.24}\, i_t^{Fin}, \underset{(0.36)}{0.01}\, p_t^{Fin}, \underset{(0.75)}{4.79}\, y_t^{Fin} \right) + \nu_t \right].$$

From this equation, it is clear that only y_t^{Fin} is significant, in both the linear and nonlinear elements of the relationship. Therefore, the conditional expectation function $\widehat{E}\left[\bar{i}_t^{Fin}, \bar{p}_t^{Fin}, y_t^{Fin} | \mathbf{m}_T^{Fin}\right]$ was plotted as a function of y_t^{Fin}, which represents the demand for money expected for any income level, if the inflation and interest rate were equal to their average values.[14] This plot is shown in Figure C.9. This, together with the respective plot for Denmark, hint at the possibility of a piecewise linear regression being an adequate model for the money demand relationships. Looking at Figure C.8, it was necessary to attempt to infer the nature of the nonlinearity found in Equation (5.12). The conditional expectation function in Figure C.8 suggests that breaks occur in the data, at approximately -0.44 and 0.26, for m_t^{Den} as a function of p_t^{Den}. In the case of Denmark, such a model is

$$m_t = \alpha + \beta_1 y_t + \beta_2 p_t + \beta_3 (p_t - p_1) D_{1t} + \beta_4 (p_t - p_2) D_{2t} + \beta_5 i_t + \beta_6 b_t + \varepsilon_t, \tag{5.15}$$

$$\begin{aligned} p_1 = -0.44, \quad & D_{1t} = 0, \quad p_t \leq p_1, \\ & D_{1t} = 1, \quad p_t > p_1, \\ p_2 = 0.26, \quad & D_{2t} = 0, \quad p_t \leq p_2, \\ & D_{2t} = 1, \quad p_t > p_2. \end{aligned} \tag{5.16}$$

For Finland, Figure C.9 suggests a break at 4.3 for m_t^{Fin} as a function of y_t^{Fin}. An alternative model is, therefore,

$$m_t = \alpha + \beta_1 i_t + \beta_2 p_t + \beta_3 y_t + \beta_4 (y_t - y_1) D_{1t} + \varepsilon_t, \tag{5.17}$$

$$\begin{aligned} y_1 = 4.3, \quad & D_{1t} = 0, \quad y_t \leq y_1, \\ & D_{1t} = 1, \quad y_t > y_1. \end{aligned} \tag{5.18}$$

In both of the above equations, ε_t is a white noise error. The resulting OLS estimates, with standard errors given in parentheses, are

$$m_t = \underset{(0.67)}{6.66} + \underset{(0.11)}{0.93}\, y_t + \underset{(0.14)}{0.54}\, p_t - \underset{(0.16)}{0.65} (p_t - p_1) D_{1t} + \underset{(0.17)}{1.25} (p_t - p_2) D_{2t} + \underset{(0.58)}{0.61}\, i_t - \underset{(0.31)}{1.48}\, b_t, \tag{5.19}$$

and

$$m_t = \underset{(0.27)}{1.77} + \underset{(0.12)}{0.31}\, i_t - \underset{(0.46)}{0.32}\, p_t + \underset{(0.06)}{0.30}\, y_t + \underset{(0.09)}{0.88} (y_t - y_1) D_{1t}, \tag{5.20}$$

for Denmark and Finland, respectively. In both cases the extra nonlinear terms are highly significant. Furthermore, the R^2 values are about 0.95 for both equations and the misspecification diagnostics for nonnormality, heteroscedasticity and functional form are also

[14] $\mathbf{m}_T^{Fin}$ is defined as $\mathbf{m}_T^{Fin} = \left(m_T^{Fin}, \mathbf{x}'_T, m_{T-1}^{Fin}, \mathbf{x}'_{T-1}, \ldots, m_1^{Fin}, \mathbf{x}'_1\right)$, where $\mathbf{x}'_T = \{i_T^{Fin}\, p_T^{Fin}\, y_T^{Fin}\}$.

satisfactory. However, there are significant indications of first-order autocorrelation from the Durbin-Watson test, as well as fourth-order autocorrelation from the relevant Lagrange multiplier test. Further details can be found in tables C.35 and C.36. Moreover, when the Hamilton (2001) test for nonlinearity is applied to these revised equations, the sample values of the LM statistics for the Danish and Finnish models are 42.99 and 18.35, respectively, which are still higher than the critical χ_1^2 value of 3.84. This finding contradicts the indications provided by the first diagnostic test for nonlinearity (RESET), which suggests that a piecewise linear functional form is appropriate for both countries. Though the substantial fall in the value of the Hamilton test statistic for the Finnish data is encouraging, Hamilton's method suggests that both models are still not adequately specified.

Although fitting piecewise linear regressions to the money demand functions of Denmark and Finland failed to model the nonlinearity effectively, it was encouraging to note the drop in value of the LM statistic, in both cases, but particularly in the case of Finland. With that in mind, and also recalling the nature of the relationship between m_t^{Fin} and y_t^{Fin} from Figure C.9, an attempt was made to fit a STAR model to the Finnish data.[15]

Following Teräsvirta (2004), and using Lütkepohl and Krätzig's (2004) `JMulTi` software, the first step was to discover the transition variable.[16] The results of the STR tests for nonlinearity can be found in Table C.37. They clearly indicate that y_t^{Fin} exhibits the strongest nonlinearity, in agreement with the findings of the Hamilton (2001) approach, and that this, therefore, should be the transition variable in the fitted STR model. Attempts to fit such a model for all variables failed to produce reasonable results. As before, therefore, the relationship was simplified, with just y_t^{Fin} being included as the independent variable. This has little impact on the tests for linearity, although even smaller p-values are reported. An unrestricted LSTR model was fitted, therefore, producing the results

$$m_t^{Fin} = \underset{(0.37)}{0.54} + \underset{(0.09)}{0.61}\, y_t^{Fin} + \tag{5.21}$$
$$\left(\underset{(0.40)}{-2.94} + \underset{(0.10)}{0.65}\, y_t^{Fin}\right)\left(1 + \exp\left\{-\left(\underset{(35.32)}{40.72}/\widehat{\sigma}_y\right)\left(y_t^{Fin} - \underset{(0.01)}{4.28}\right)\right\}\right)^{-1} + \widehat{\varepsilon}_t.$$

A plot of actual and fitted values for this equation can be found in Figure C.10. Interestingly, when the residuals of this regression are tested for nonlinearity with Hamilton's (2001) LM test, the test statistic falls to 1.63 ($\chi^2 = 2.67$), with a p-value of 0.10.[17] This suggests that this model may indeed effectively capture the nonlinearity evident from previous testing.

5.7 Conclusion

This chapter has drawn attention to some of the pitfalls involved in using the conventional $I(1)/I(0)$ framework for economic and financial modelling of time-series data, an approach

[15] See Chapter 2 for a brief overview of STAR modelling.

[16] Given the nonlinear analysis undertaken to this point, it was deemed likely to be y_t^{Fin}.

[17] As with the case of the piecewise linear regressions described above, the Hamilton LM test is applied following Hamilton (2001), where ε_t are the residuals and $\mathbf{M}_T$ is the projection matrix from the estimated equation, in this case, Equation (5.21), while $\mathbf{H}$ is calculated from the original explanatory variables (Hamilton, 2001, p. 561). On a cautionary note, it is uncertain how the Hamilton LM test performs on the residuals of such a nonlinear specification.

involving well-known unit root tests and the cointegration testing and modelling procedures of Engle and Granger (1987), Johansen (1988, 1991) and the COMFAC approach, that has been applied widely during the last decade or so. The practical difficulties of untangling the issues of stationarity, fractional integration, nonlinearity and parameter instability have been highlighted. This chapter has briefly discussed some of the recent research directed at resolving these problems and providing alternative, or at least complementary, approaches to modelling, previously outlined in this thesis.

This chapter has presented a case study intended to illustrate the application of these newer techniques and contrast their findings with those of the standard cointegration modelling approach. The study used the data previously analysed by Johansen and Juselius (1990) in connection with demand for money functions in Denmark and Finland. The results obtained from the various techniques exemplify the problems with the standard approach and the alternative conclusions that might be reached by using different techniques. The findings, using the standard approach, were as follows.

Though ADF tests, implemented using the procedure of Dolado, et al. (1990), appear to suggest unit roots for most variables, they are sensitive to the specification of the test equation and the information criterion used to choose lag length in the case of some variables, especially for Finland. Tests of seasonal integration confirm the $I(1)$ nature of the Danish variables, as no evidence of seasonal integration is found. In the case of the Finnish variables, however, the sensitivity observed in the ADF tests is mirrored in the mixed results. There are indications that some of the variables are $I(1)$, some are $I(0)$, and that some may be seasonally integrated, but these findings very much depend on the specification of the tests. When the matter of unit roots was explored further, using the ERS, KPSS and NP tests, unit roots for the Danish variables tended to be confirmed but not for the Finnish variables.

Proceeding on the assumption that all variables are $I(1)$, the Engle-Granger (1987) 2-step procedure does not support cointegration in general, a result that is confirmed by CRDW tests and ECM tests conducted in an error-correction framework for the money demand relationship for each country. However, the Engle-Granger approach does suggest cointegration for the version of the Finland model that treats two of the variables, p_t and i_t, as $I(0)$.

Using the Johansen (1988, 1991) approach without its small sample bias-correction factor, there is considerably stronger evidence of cointegration in the case of Denmark, though the number of cointegrating vectors suggested varies, depending on the VAR specification chosen. For the preferred VAR specification, one cointegrating vector is suggested for Denmark. The picture that emerges for Finland is similar, although for the version of the model that treats the p_t and i_t variables as $I(0)$, the Johansen method suggests no cointegration, contradicting the finding of the Engle-Granger (1987) procedure in this case.

The Johansen (2002) correction factor has a marked effect on the result in the case of the small sample of data for Denmark, with the modified Trace test agreeing with the conclusion from the Engle-Granger (1987) procedure that there is no cointegrating demand for money relationship. However, it was noted that the modified Trace test provides a different signal from the Maximal eigenvalue test, which indicates cointegration. As might be expected, the Johansen correction has no effect on the findings for Finland, which are based on a much larger sample.

Interestingly, the common factor approach cannot rule out the possibility of cointegration in either the Danish or Finnish samples, including the reduced Finnish models. Given the contradictory evidence provided by the Engle-Granger (1987) and Johansen (1988, 1991) approaches, this adds further weight to the contradictory evidence regarding cointegration in these data.

These results are puzzling, not withstanding the relatively small size of the Danish sample used and the known low power of unit root tests. In particular, the contradictory results from the Engle-Granger (1987), Johansen (1988, 1991) and COMFAC procedures concerning the existence of cointegrating relationships, in the case of both countries, is curious.

Checking for fractional integration by means of a range of estimators of the fractional integration parameter, as well as the new FDF and FADF tests of Dolado, et al. (2002), confirms the $I(1)$ nature of the Danish variables and the lack of a unit root for the variables in the case of Finland. It is difficult to say why the bias-corrected Johansen technique fails to find cointegration in the former case and yet suggests it in the latter.

Assuming that the Finnish data are not $I(1)$, and hence can not be simply cointegrated, what type of model is appropriate? The possibility of stationarity with regime shifts or some other kind of nonlinearity arises. This was explored for both countries by the Hamilton (2001) methodology, which may be appropriate for general, unknown forms of nonlinearity. This method produces strong evidence of structural change/nonlinearity, if underlying stationarity is entertained, although no consideration was given to sample effect, as outlined in Chapter 2, in these cases. An attempt to re-specify the money demand equations as piecewise linear regressions, however, which was suggested by examination of the data, was not very successful. In the case of Finland, where the plot of the conditional expectation function clearly suggests a smooth transition model might be effective, fitting a STAR model produced more satisfactory results. Were this alternative approach to be preferred, further work remains necessary to find an adequate nonlinear functional form, particularly for Denmark.

In conclusion, the messages from this first applied study appear to be that, first, standard $I(1)/I(0)$ modelling strategies for economic and financial time series are fraught with dangers. Secondly, complementary procedures designed to investigate the possibilities of fractional integration and nonlinearity are available and relatively easy to implement. Thirdly, fractional integration analysis may confirm the existence of unit roots, but may also suggest fractional integration of different degrees for different variables. This is a complicated situation that raises challenges for modelling. Fourthly, and recalling that unit root tests may often indicate that a unit root exists when a series is stationary but subject to level shifts, a general analysis of nonlinearity, such as that offered by the Hamilton (2001) procedure, may be an attractive option that can lead to acceptable alternative models. The moral would seem to be that reliance on any one approach may not be a sensible practice in applied work, and that practitioners would be well advised to consider using a range of alternative methods and selecting models according to the balance of the wider body of evidence produced.

Chapter 6

Purchasing Power Parity: The Irish Experience

Research carried out in the course of writing this chapter appears as 'Purchasing Power Parity: The Irish Experience Re-visited' by Bond, D., M.J. Harrison, and E.J. O'Brien, in Trinity Economic Papers, No. 15 (2006). It has been submitted to the Working Paper Series of the European Central Bank and the 5^{th} INFINITI Conference on International Finance, Dublin, Ireland, 11^{th}-12^{th} June 2007.

6.1 Introduction

The theory of purchasing power parity (PPP) has become a major area of research in applied econometrics. In part, this is due to the crucial role of the concept in the theory of both exchange rates and international finance. Recent surveys include Taylor and Taylor (2004), Sarno and Taylor (2002) and Rogoff (1996). The analysis has generally kept pace with developments in econometric time-series analysis. Two major areas of current research are the mean reversion characteristics of the real exchange rate and the nonlinear representation of the real exchange rate.[1] However, the mainstream literature in the area has as yet to fully utilise two developments in econometric theory; long memory models and random field inference. These developments could provide useful tools for investigating both the mean reversion and nonlinearity in PPP analysis.

From the econometric literature it is clear that nonstationarity and nonlinearity may be closely related. It has been well known for many years that it is difficult to statistically distinguish between difference stationary series and nonlinear but stationary series.[2] Recent works in the area include Lee, Kim, and Newbold (2005) and Hong and Phillips (2005). Increasingly, the analysis uses the fractional integration framework rather than the 'knife-edge' $I(1)/I(0)$ approach,[3] to consider the effects nonlinearity has on stationarity tests. This approach has been followed by Diebold and Inoue (2001) and Perron and Qu (2004). Hsu (2001) and Krämmer and Sibbertsen (2002) have used the reverse approach, considering the effects of nonstationarity on tests for nonlinearity. Recent work by Gil-Alana (2004), Mayoral (2005) and Dolado, Gonzalo, and Mayoral (2005b) have tested explicitly for difference stationarity and nonlinearity. However, in most cases the form of the nonlinearity needs to be known.

The aim of this chapter is to use two recent developments in econometric theory discussed in Bond, Harrison, and O'Brien (2007), to explore the time-series characteristics of simple empirical interpretations of PPP using Irish, German and United Kingdom data. The first of these developments is the Dolado, Gonzalo, and Mayoral (2002) fractional augmented Dickey-Fuller test (FADF) and the second is Hamilton's (2001) method of random field estimation to investigate nonlinearity. Both of these developments have been discussed at length in previous chapters of this thesis. The remainder of this chapter is structured as follows. In Section 6.2, the background to the theory of PPP and the notation to be used here will be briefly explained. Section 6.3 contains a description of the data and some preliminary analysis. Sections 6.4 and 6.5 test for cointegration and fractional integration, respectively, employing the techniques, including the Dolado, et al. FADF test, outlined in Chapter 4 and utilised in the previous case study in Chapter 5. Section 6.6 considers the results of using Hamilton's random field methodology and Section 6.7 concludes by considering how the methodology might assist in the development of the general discussion of the theory of PPP.

[1] See, for example, Cashin and McDermott (2004) and Sarno (2005).

[2] See, for example, Perron (1989), Harrison and Bond (1992), Teverosky and Taqqu (1997) and Diebold and Inoue (2001).

[3] Jensen (1999)

6.2 Purchasing Power Parity

The theory of PPP is, according to Wright (1994), 'one of the leading applications of cointegration analysis in applied econometrics'. A simple statement of the purchasing power parity hypothesis is that national price levels should be equal when expressed in a common currency. More formally, if s_t is the logarithm of the nominal exchange rate, p_t and p_t^* are the logarithms of the domestic and foreign price levels, respectively, and q_t is the logarithm of the real exchange rate in period t, then

$$q_t \equiv s_t - p_t + p_t^*, \tag{6.1}$$

for all $t = 1, 2, \ldots, T$. It follows that q_t must be stationary for long run PPP to hold, as nonstationarity in q_t would imply permanent deviations from parity, thereby rejecting the theory.[4] Most of the empirical studies of PPP have either been concerned with testing whether q_t has a mean reversion tendency over time or whether s_t, p_t and p_t^* move together over time.

This latter work has generally been concerned with models whose simplest form is:

$$s_t = \alpha_0 + \beta_1 p_t + \beta_2 p_t^* + \epsilon_t, \tag{6.2}$$

where ϵ_t is white noise. Early works were concerned with whether the estimated values of the parameters of various versions of Equation (6.2) were as predicted; see, for example MacDonald and Taylor (1992). As awareness of time-series dynamics increased, the issue changed to one of whether Equation (6.2) is cointegrated. Papers such as those by Thom (1989), Wright (1994) and Kenny and McGettigan (1999) took such an approach with Irish data. The results of these Irish studies have been somewhat confusing. In some cases the theory of PPP could not be accepted, whereas in others in could not be rejected. Nonrejection seemed most common when other variables were included in the model or where prices related to the traded sector. For example, Wright considered the inclusion of interest rate differentials, and Kenny and McGettigan distinguished between prices in the traded and nontraded sectors. An alternative argument that has been gaining ground in the literature is that of the possibility that the relationship is in fact nonlinear. The argument is that nonlinearities arise because of transaction costs in international arbitrage, as discussed in Sarno (2005).

In recent years the emphasis has generally shifted from considering models of the form of Equation (6.2), to considering directly the behaviour of $\{q_t\}_{t=1}^{\infty}$, the real exchange rate. Within the $I(1)/I(0)$ framework, most of the early studies failed to reject the hypothesis of real exchange rates being $I(1)$ for recent periods of flexible exchange rates. This failure to reject the possibility of unit roots in the real exchange rate series, $\{q_t\}$, implies a lack of mean reversion which undermines the PPP hypothesis. The explanation given for this nonrejection is the recognised low power of traditional unit root tests, such as the standard Dickey-Fuller test. To overcome this problem, two general approaches have been adopted. The first has been the construction and use of long series of exchange rate data and more

[4] When specified as in Equation (6.1), q_t must be stationary for PPP to hold in the long run. In terms of the nominal exchange rate, as in Equation (6.2), the theory of PPP allows for persistent deviations in ϵ_t. See, for example, Enders (1995) and Sarno and Taylor (2003).

'powerful' asymptotic tests.[5] The second, using panel data, attempts to estimate the half life of the mean reversion on the real exchange rate.[6]

Empirically, it is well known that the theory of PPP does not hold in the short run.[7] Naturally, this does not detract from the above theory. It is expected that some deviation from the mean will occur, but that in the long-run this deviation will be stationary. There is, however, conflicting evidence as to whether PPP holds in the long run. Dornbusch (1976) and Ainzenman (1986), using sticky price models, permited sustained deviations from parity but typically maintained PPP as a valid long-run hypothesis. On the other hand, Roll (1979) and Alder and Lehman (1983) suggested PPP is violated in the long run, using models based on efficient international capital markets.

As previously mentioned, for the case of Ireland, several studies have been undertaken. Thom (1989) failed to find cointegrating relationships for Irish and United States, and Irish and German data, but does find it for Irish and United Kingdom data, although only when coefficient restrictions are imposed on the model. Wright (1994), by adding the short-term interest rates for Germany, Ireland and the United Kingdom, found evidence for cointegration in both cases. This study loosely follows that work.

To investigate the usefulness of both the FADF and random field approaches to understanding the issues surrounding PPP, this chapter applies the techniques to data for Ireland and Germany and Ireland and the United Kingdom. The specification for the explanatory model used is that of Wright (1994), namely,

$$s_t = \alpha_0 + \beta_1 p_t + \beta_2 p_t^* + \beta_3 i_t + \beta_4 i_t^* + \epsilon_t, \tag{6.3}$$

where i_t and i_t^* are the domestic and foreign interest rates, respectively.

To place the long memory and random field analysis into context, the standard $I(1)/I(0)$ analysis using the ADF test is conducted. The strategy of Dolado, Jenkinson, and Sosvilla-Rivero (1990), to determine whether the series are trend stationary or difference stationary, is adopted. The lag length for the ADF test is determined using the modified Akaike information criterion (MAIC), which Ng and Perron (2001) showed to be a generally better decision criterion, as it takes account of the persistence found in many series.[8]

Traditional cointegration analysis is then applied to the naïve PPP model of Equation (6.2). Firstly, the Engle-Granger (1987) 2-step approach is applied using the lag of residuals of the levels regression model as the error-correction term. Then the Johansen (1988, 1991) VAR approach and the common factor approach are applied to the data. The effects of applying Johansen's (2002) small sample correction factor is also investigated.

Following on from this traditional analysis, the issue of fractional integration is investigated.[9] Following the approaches outlined in chapters 4 and 5, a consistent parametric

[5] Taylor (2002).

[6] Cashin and McDermott (2004).

[7] Thom (1989).

[8] A time series characterised by 'lagged effects' resulting from, for example, long memory, may be considered to be a persistent time series.

[9] Cheung and Lai (1993b, 2001) explored the fractionally integrated and long memory properties of purchasing power parity and real exchange rate mean reversion. Robinson and Iacone (2005) also considered purchasing power parity in terms of fractional cointegration.

estimate of d, as suggested by Dolado, et al. (2002), is obtained and used for the FADF test. Again, the 'over differenced' ARFIMA model, where Δy_t is used rather than y_t, is estimated, as recommended by Smith, Sowell, and Zin (1997), to avoid the problems associated with drift. Two parametric estimates of d are calculated using the Doornik and Ooms (1999) ARFIMA package, namely, the exact maximum likelihood (EML) estimator, using the algorithm suggested by Sowell (1992) (the algorithm requires that $d < 0.5$, which is another reason for using the 'over-differenced' model), and an approximate maximum likelihood estimator based on the conditional sum of squared naïve residuals, developed by Beran (1995) and termed by Doornik and Ooms (1999) a nonlinear least squares (NLS) estimator. Again, as in Chapter 5, two nonparametric estimators are also employed; the method of Geweke and Porter-Hudak (1983) (GPH) and Robinson's (1994) Gaussian semiparametric (GSP) estimator. The estimates of d are then used in the FADF test, with the MAIC being used as the criterion on which to set the lag length of the test.

Finally, the analysis then turns to an investigation of the possibility of nonlinearity in the models. The parameters from the random field model are estimated, using the GAUSS code provided by Hamilton (2001) and adapted as explained earlier in Chapter 2.

6.3 Data and Preliminary Analysis

For the analysis of the theory relating to Ireland and Germany, five data series are used. The exchange rate $e_t^{DM/IR£}$, is measured in Deutsche Marks per Irish pound. Both the Irish and German price variables, p_t^{Ire} and p_t^{Ger} are the respective producer price indices for manufacturing industries. These three variables are all taken in logarithms. Also included, following Wright (1994), are variables for both the Irish and German short-term (three month) interest rates, i_t^{Ire} and i_t^{Ger}, respectively. These series sample the period 1975 to 2003, a total of 115 quarterly observations. As with the previous case, five data series are used for the Ireland and United Kingdom study. The exchange rate, $e_t^{Stg£/IR£}$, is naturally measured in Pounds sterling per Irish Pound. Euro currency is converted with the relevant exchange rate where necessary. The Irish price variable is as before. The United Kingdom price variable, p_t^{UK}, is the producer price index for manufacturing. These three variables are all taken in logarithms. Also included are short-term Irish and United Kingdom interest rates, i_t^{Ire} and i_t^{UK}. As before, these series are quarterly and sample the period from 1975 to 2003, a total of 115 observations. These data series can be found in Appendix D.1, tables D.1 and D.2. They are plotted in figures D.1, D.2 and D.3, which can be found in the Appendix D.2.[10]

The preliminary results of the basic unit root analysis are given in Table D.3. The Dolado, et al. (1990) testing strategy was adopted and in nearly all cases the existence of a trend or drift could not be rejected, so the probabilities given in Table D.3 are mainly from the standard normal distribution. In the few cases where the existence of a trend or drift could be rejected, the probabilities given are from MacKinnon (1996). These results generally seem to suggest that most series are $I(1)$. The Irish price level may be $I(0)$, as the test statistic is significant at the 5 per cent level, using standard normal critical values. The German price

[10]Only the results of unit root tests are considered for these series, although it should be noted from figures D.1, D.2 and D.3 that there is limited graphical evidence of unit roots.

series is just marginally insignificant at the 5 per cent level, again using standard normal critical values, although in this case the results suggest it is $I(1)$.

Tests for seasonal unit roots are included in tables D.4 and D.5.[11] Interestingly, these results suggest that the Irish German exchange rate is stationary. No evidence of unit roots at any frequency was found. In general, results for the remainder of the series suggest the series are in fact $I(1)$ and that no seasonal integration is present. The exception to this is the Irish price level. As with standard ADF tests, evidence again suggests that the series is $I(0)$ and that it is not seasonally integrated, in almost all specifications examined. Given the strong prior belief that price series are generally $I(1)$, despite some evidence to the contrary, this assumption will be maintained throughout the remainder of the chapter.[12] The same approach is taken with the Irish German exchange rate, as the ADF test suggests that it is $I(1)$.

6.4 Testing for Cointegration

6.4.1 The Engle-Granger 2-Step method

The results of applying the traditional Engle-Granger (1987) analysis to the explanatory models is given in tables D.6 and D.7. Table D.6 reports the findings of the standard levels analysis and in all cases the traditional AEG test fails to reject the hypothesis that the estimated residuals of the model have an unit root. The CRDW test confirms these findings.

Despite these negative findings regarding cointegration, Table D.7 gives the results of trying to estimate a parsimonious error-correction model using the first lag of the residual from the corresponding levels model as the error-correction term. In all cases, the results of the analysis confirm the previous findings. While the coefficients of the error-correction terms have the 'right' sign, the t-ratios are low. The ECM test also rejects cointegration in all cases. Dropping the insignificant constant terms has minimal effect on the results. Interestingly, in all four model specifications, lags of the dependent and independent variables were highly insignificant suggesting a fairly quick adjustment process. These results are broadly in line with Thom (1989).

6.4.2 Johansen's maximum likelihood approach

Tables D.8 to D.15 provide details of the VAR selection and specification for the four models under consideration: Ireland and Germany, and Ireland and the United Kingdom both with and without interest rates. Table D.16 summarises the Johansen (1988, 1991) analysis of the data for all models, while more detailed results are given in tables D.17 to D.20. Table D.17 shows evidence of one cointegrating vector for Ireland and Germany, when interest rates are excluded. Importantly, this result is overturned when the small sample correction is applied; based on the Trace test, no evidence of cointegration is found. One cointegrating vector is found when interest rates are included, as shown in Table D.18. In this case, however,

[11]As in Chapter 5, these tests were carried out using the so-called HEGY test, attributable to Hylleberg, Engle, Granger, and Yoo (1990). See Chapter 4 for further details.

[12]Wright (1994) encountered a similar problem with the Irish price index and the Irish German exchange rate, and made similar assumptions.

the small sample correction has no impact on this result. Both the Trace and Maximal eigenvalue tests find evidence for one cointegrating vector. Tables D.19 and D.20 present the results for Ireland and the United Kingdom. As with the previous case, the finding of one cointegrating vector in the specification without interest rates is overturned by the small sample correction. Two cointegrating vectors are found, however, when the interest rates are included and this result is unaffected by the small sample correction, which strangely is less than 1! These results indicate that there is little evidence of cointegration in a traditional PPP specification, but that the introduction of interest rate differentials appears to be significant. These findings are very similar to those of Wright (1994). Overall, however, as in other studies, this attempt to place the PPP analysis in a cointegrating framework is not entirely satisfactory.

6.4.3 Common factor analysis

To complete this final stage of the conventional analysis, two specifications were examined for the PPP hypothesis, as was the case with both the Engle-Granger (1987) and Johansen (1988, 1991) approaches. For Ireland and Germany, the specification including the short-term interest rates was estimated as follows

$$\begin{aligned} e_t^{DM/IR\pounds} &= \underset{(0.30)}{0.52} + \underset{(0.04)}{0.89}\, e_{t-1}^{DM/IR\pounds} - \underset{(0.22)}{0.44} p_t^{Ire} + \underset{(0.21)}{0.43} p_{t-1}^{Ire} + \underset{(0.55)}{0.73} p_t^{Ger} - \underset{(0.54)}{0.82} p_{t-1}^{Ger} \\ &+ \underset{(0.001)}{0.59E} - 3i_t^{Ire} - \underset{(0.001)}{0.86E} - 3i_{t-1}^{Ire} + \underset{(0.005)}{0.18E} - 3i_t^{Ger} + \underset{(0.005)}{0.38E} - 3i_{t-1}^{Ger} + \hat{\varepsilon}_t, \end{aligned} \tag{6.4}$$

where standard errors are given in parenthesis. Full results can be found in Table D.21. As in Chapter 5, just the common factor restrictions were tested. The results, which can be found in Table D.22, show that the imposed restrictions cannot be rejected, thereby excluding the possibility of cointegration among these variables.

The second specification for Ireland and Germany, without the interest rates, was estimated as follows,

$$e_t^{DM/IR\pounds} = \underset{(0.26)}{0.49} + \underset{(0.04)}{0.89}\, e_{t-1}^{DM/IR\pounds} - \underset{(0.21)}{0.48} p_t^{Ire} + \underset{(0.208)}{0.47}\, p_{t-1}^{Ire} + \underset{(0.44)}{0.84} p_t^{Ger} - \underset{(0.43)}{0.93} p_{t-1}^{Ger} + \hat{\varepsilon}_t. \tag{6.5}$$

The full results may be found in Table D23. Once again, the tests of common factor restrictions prove to be insignificant, as is clear from Table D.24. There is no evidence of cointegration in this case. It can be concluded, therefore, that no evidence of cointegration can be found using the COMFAC approach, for Ireland and Germany, regardless of the inclusion of interest rate variables.

Next, attention was turned to the data for Ireland and the United Kingdom. Two specifications were again estimated and the COMFAC restrictions then tested. As before, no evidence of cointegration was found with this method. The first specification, including prices and short-term interest rates was estimated as

$$\begin{aligned} e_t^{Stg£/IR£} &= \underset{(0.09)}{0.12} + \underset{(0.05)}{0.88} e_{t-1}^{Stg£/IR£} - \underset{(0.30)}{1.06} p_t^{Ire} + \underset{(0.27)}{1.06} p_{t-1}^{Ire} + \underset{(0.48)}{0.54} p_t^{UK} - \underset{(0.48)}{0.57} p_{t-1}^{UK} \\ &+ \underset{(0.002)}{0.005} i_t^{Ire} - \underset{(0.002)}{0.004} i_{t-1}^{Ire} - \underset{(0.003)}{0.002} i_t^{UK} - \underset{(0.003)}{0.58E - 3} i_{t-1}^{UK} + \widehat{\varepsilon}_t. \end{aligned} \tag{6.6}$$

The full OLS results can be found in Table D.25. The tests of the common factor restrictions can be found in Table D.26. Again, the restrictions prove to be insignificant, excluding the possibility of cointegration in this case. The second specification, including just prices as independent variables, was estimated as

$$e_t^{Stg£/IR£} = \underset{(0.09)}{0.15} + \underset{(0.05)}{0.87} e_{t-1}^{Stg£/IR£} - \underset{(0.29)}{1.15} p_t^{Ire} + \underset{(0.27)}{1.11} p_{t-1}^{Ire} + \underset{(0.46)}{0.38} p_t^{UK} - \underset{(0.45)}{0.38} p_{t-1}^{UK} + \widehat{\varepsilon}_t. \tag{6.7}$$

These results are available in Table D.27. The results of the tests of restrictions, in Table D.28 are similar to the previous results. The restrictions are insignificant, although one of the two is marginally significant at the 10 per cent level, which suggests that there is no cointegration in this case. As has been the case for all COMFAC analysis in this chapter, it is again found that there is no evidence of cointegration in the Ireland and United Kingdom PPP data, regardless of the inclusion of interest differentials. This confirms the findings of the Engle-Granger (1987) approach and the Johansen (1988, 1991) results excluding interest rates.

6.5 Testing for Fractional Integration

Table D.33 gives the results of the simple fractional analysis. For each series, four different estimates of d are given together with their standard errors and FADF value. The FADF test is only meaningful if $d \leq 1$ and the probabilities to be applied to the test statistics are the standard normal ones. The results are interesting and would seem to imply that the only series that is likely to be fractionally integrated is Irish interest rates. While all the estimates of d for the nominal exchange rate between Ireland and the United Kingdom are less than one in all cases, the FADF test fails to reject the null hypothesis of a unit root. For all other series, the estimates of d gave conflicting values and the FADF test only gave strong evidence of fractional integration in the case of the Irish German exchange rate. This suggests that the conflicting evidence on cointegration cannot readily be attributable to fractionality.

6.6 Nonlinear Inference

To account for the possibility of parameter instability, or some other type of nonlinearity, Hamilton's (2001) random field approach was used to explore the likely form of the two PPP models examined thus far, and this leads to some interesting results. Hamilton's LM test statistics for nonlinearity for the Irish German case were 575.04 and 180.03, for the models with and without interest rates, respectively, which are significantly greater than the 5 per cent critical χ_1^2 value of 3.84. This clearly suggests the models should not be simply linear.

Nonlinear estimates were obtained for both models, using the STEEPEST DESCENT-NEWTON algorithm switching technique, an initial value of $\zeta = 0.5$, Hamilton's (2001) covariance specification, with convergence being achieved in both cases after 36 iterations. Results can be found in Table D.34. For the model without interest rates, both σ and ζ are significant, and both price variables are found to be nonlinearly significant, i.e., they are both found to contribute to the nonlinearity found previously. For the specification with interest rates, both price variables are again nonlinearly significant, as is the German interest rate. In this case, ζ is marginally insignificant, but as λ and σ are highly significant here, this may be evidence of the 'pile-up' problem associated with numerical optimisation, alluded to in Chapter 2. This may signal that the covariance structure used for the random field, if not the normality assumption itself, may not be entirely appropriate here.

Given the difficulties associated with plotting a conditional likelihood function with three nonlinearly significant variables, the cross plot of the Irish German exchange rate against the Irish price series and the German price and interest series is considered instead. It is clear from Figure D.4 that a number of regime shifts are evident in the data. Whether we consider a model with just prices, or the augmented model with interest rates, the structural shifts remain. The difficulty in this case is to find an appropriate modelling strategy for this. This issue will be mentioned again later in this section.

For the Irish United Kingdom case, the same approach was taken. In this case the nonlinearity test statistics were 650.72 and 205.00, for both the standard and augmented model. Once again, both values significantly exceed the critical value, suggesting a nonlinear model may be appropriate. The same strategy was employed, using the STEEPEST DESCENT-NEWTON algorithm switching technique, an initial value of $\zeta = 0.5$, Hamilton's (2001) covariance specification. Convergence was achieved after 42 and 19 iterations, respectively.

Interestingly, both price variables are found to be nonlinearly significant in the standard model, as are σ and ζ. This mirrors the results of the Irish German case. For the augmented model, both price series are again nonlinearly significant, although in this case, it is the Irish interest that is nonlinearly significant. The United Kingdom interest rate is not. As before, ζ is marginally insignificant, but both λ and σ are highly significant. As in the German case, it is not possible to obtain a graphical representation of the conditional likelihood function here, so once again, simple cross-plots are relied upon. Although these plots, which can be found in figures D.5 to D.7, also show some evidence of distinct regimes, they are certainly less clear than the German case. In fact in the case of the Irish interest rate, no obvious pattern emerges.

Most strikingly, perhaps, is the fact that when nonlinearity is modelled by means of a random field, the coefficients on the domestic and foreign prices in the specifications with and without interest rates, are not statistically significantly different from their -1 and 1 values under purchasing power parity theory. This finding contrasts with the findings in the earlier Irish studies by, for example, Thom (1989) and Wright (1994), both of whom reported cointegrating vectors, corresponding to the vector of variables s_t, p_t and p_t^*, that were markedly different from (1, -1, 1).

Having found strong evidence of nonlinearity in all of these models, and which variables may be instrumental in causing that nonlinearity, the question remains how to model that nonlinearity. In neither the German nor the United Kingdom case does the data suggest an obvious approach, nor is there a theoretical framework within which to work. Given the likely complexity of finding and fitting suitable models to the relationships discussed here, it is felt that such work is beyond the scope of this illustrative exercise. The fact remains, however, that strong evidence of nonlinearity is found in these cases, and to better understand the phenomenon of purchasing power parity, efforts should be directed to exploring this in both economic and econometric terms. This is left for future research.

6.7 Conclusion

This chapter has explored the well-known economic concept of purchasing power parity for Ireland, Germany and the United Kingdom. Very much in keeping with the overall themes explored in this thesis, and the conclusions of Chapter 5, this chapter has shown the potential for difficulties in placing the study of PPP in the $I(1)/I(0)$ framework. It has shown also, that in this case, these difficulties cannot be overcome by moving to a fractional integration framework. There was strong evidence of nonlinearity among the data and perhaps this should be the approach taken to model this type of relationship.

Standard ADF tests, implemented using the procedure of Dolado, et al. (1990), suggested that most of the variables here have unit roots. There was some doubt over this finding for the logarithm of the Irish price level, but following Wright (1994), and given the strong prior that price levels are $I(1)$, this was assumed to be the case. There was no evidence of seasonal integration in the variables; in fact, HEGY tests generally confirmed the findings of the ADF tests.

Under the reasonable assumption that all of the variables were $I(1)$, the Engle-Granger (1987) approach was employed to test for cointegration. In both the standard models, and the models augmented with interest rates, no evidence of cointegration was found for either Germany or the United Kingdom. This result was confirmed by the CRDW test, the estimation of error-correction models and by the error-correction test of cointegration.

The Johansen (1988, 1991) approach found evidence of one cointegrating vector in both the German and United Kingdom standard models. Applying Johansen's (2002) small sample correction overturned this finding, however, confirming the earlier results of the Engle-Granger (1987) approach of no cointegration. For the models augmented with interest rates, one cointegrating vector was again found for each case, but the correction factor had no bearing on either, thus contradicting the results of the Engle-Granger procedure. The results thus far confirmed the findings of Thom (1989) and Wright (1994).

Interestingly, the common factor approach found no evidence whatsoever of cointegration. This was very much in agreement with the findings of the Engle-Granger (1987) methodology and the results of the Johansen (1988, 1991) when just prices were included as independent variables.

Little evidence of fractional integration was found in the series, with only Irish interest rates having fractional estimates significantly different from 1. Unlike the results in Chapter

5, this suggests that the conflicting evidence on cointegration cannot readily be attributable to fractionality.

Testing for nonlinearity with Hamilton's (2001) LM-type test uncovered strong evidence of nonlinearity in both the German and United Kingdom models. Evidence from the random field regressions suggested that both the price and interest variables play a strong role in this nonlinearity. Moreover, if the nonlinearity is modelled using a random field regression, they show, importantly, that the Irish experience vis-à-vis Germany and the United Kingdom accords well with purchasing power parity theory. Examination of simple cross plots revealed evidence of shifts in regime for those variables found to contribute to the nonlinearity. No attempt was made to model such nonlinearity in this case, although this is an interesting avenue for future research.

In conclusion, therefore, it would seem that despite the apparent $I(1)$ nature of the data, the standard and augmented PPP models do not fit well into the cointegration and error-correction framework. As the series do not appear to exhibit fractional integration, entertaining the notion of fractional cointegration would not further the modelling effort. In fact, evidence suggests that these models are highly nonlinear and that such nonlinearity derives from both the prices and interest rates. These findings very much depend on the validity of the Hamilton (2001) methodology for $I(1)$ variables. As mentioned in Chapter 2, Lee, et al. (2005) have found the Hamilton LM-type test to be susceptible to spurious nonlinearity, when applied to nonstationary variables. Despite such reservations, given the evidence presented in this chapter, efforts to model purchasing power parity should perhaps be directed to capturing this nonlinearity.

Chapter 7

Some Empirical Observations on the Forward Exchange Rate Anomaly

Research carried out in the course of writing this chapter has led to the publication of 'Some Empirical Observations on the Forward Exchange Rate Anomaly' by Bond, D., M.J. Harrison, N. Hession, and E.J. O'Brien, in Trinity Economic Papers No. 2 (2006). It was presented at the 26^{th} Conference on Applied Statistics in Ireland, Killarney, Co. Kerry, 17^{th}-19^{th} May 2006 and the 4^{th} INFINITI Conference on International Finance, Dublin, Ireland, 12^{th}-13^{th} June 2006.

7.1 Introduction

This chapter will examine the forward exchange rate anomaly. It is the final substantive chapter of this thesis and the third and last application of the methodologies discussed in chapters 2, 3 and 4. As the focus here will be deliberately narrow, it is somewhat shorter than chapters 5 and 6.

The forward exchange rate anomaly arises from the failure of the forward rate unbiasedness (FRU) property; i.e., the fact that the results of empirical studies suggest that foreign exchange markets are so inefficient at forecasting the future movements of exchange rates that they systematically predict these movements in the wrong direction. A substantial literature exists on both the anomaly and the risk premium, stemming primarily from the seminal paper by Fama (1984), although excellent surveys are also provided by Hodrick (1987) and Engel (1996). Recently, following developments in econometric theory, two approaches have been taken to investigate the forward exchange rate anomaly. The first views the anomaly as the outcome of possible nonlinearity in economic and financial relationships. This nonlinearity can be modelled by using smooth transition regressions, which have been briefly mentioned in chapters 2 and 5. The second approach attempts to explain the anomaly in terms of fractional integration, previously introduced in Chapter 4.

The aim of this chapter is to use two recent developments in econometric theory, introduced in earlier chapters, to explore the forward exchange rate anomaly. The first of these is the Hamilton (2001) method of nonlinear inference, based on random field regression. This was introduced and discussed at considerable length in Chapter 2. The second is the Dolado, Gonzalo, and Mayoral (2002) fractional augmented Dickey-Fuller (FADF) test, that was introduced in Chapter 4. Both methods have also been applied in chapters 5 and 6, where further discussions can be found. These methods may offer further insight into the understanding of the anomaly.

The chapter will proceed as follows. In Section 7.2 the forward exchange rate anomaly will be discussed and the notation used will be explained. A review of the attempts at explaining the anomaly will also be given. In Section 7.3, the data, for the cross-exchange rates for the Australian dollar and sterling, the Canadian dollar and sterling, and the Japanese yen and sterling, will be presented and discussed. A preliminary analysis of the data will also be completed. Sections 7.4 and 7.5, respectively, will consider tests for integration, cointegration and nonlinearity, employing the Dolado, et al. (2002) FADF test and the Hamilton (2001) random field model. Finally, Section 7.6 will offer a brief summary and conclusion, which will consider how the results reported relate to the general discussion of the forward exchange rate anomaly.

7.2 The Forward Exchange Rate Anomaly

The forward exchange rate anomaly has played a central role in the theory of foreign exchange market efficiency. Consider, as a starting point, the covered interest rate parity (CIP) hypothesis of international macroeconomics, which states that

$$f_{t,k} - s_t = i_{t,k} - i^*_{t,k}, \tag{7.1}$$

where s_t and $f_{t,k}$ are the (logarithm) spot and forward exchange rates at time t, k is the length of the forward contract and $i_{t,k}$ and $i^*_{t,k}$ are the k periods to maturity nominal interest rates available on similar domestic and foreign assets, respectively. The validity of the CIP hypothesis is generally accepted.[1]

Closely linked to the CIP hypothesis is the uncovered interest rate parity (UIP) condition, that can be seen as a central parity condition for foreign exchange market efficiency:

$$E_t(s_{t+k} - s_t) = i_{t,k} - i^*_{t,k}, \tag{7.2}$$

or

$$E_t(\Delta_k s_{t+k}) = i_{t,k} - i^*_{t,k}, \tag{7.3}$$

where $E_t(\cdot)$ denotes the expectation based on information available at time t and $\Delta_k = 1 - L^k$, with L being the usual lag operator. Making use of Equation (7.1) in Equation (7.3) gives

$$E_t(\Delta_k s_{t+k}) = f_{t,k} - s_t, \tag{7.4}$$

and therefore

$$E_t(s_{t+k}) = f_{t,k}. \tag{7.5}$$

Equation (7.5) is also known as the forward rate unbiasedness hypothesis. Simple tests of the CIP and FRU hypotheses consist of inference on the coefficients of the following regressions:

$$\Delta_k s_{t+k} = \alpha_1 + \beta_1(f_{t,k} - s_t) + \epsilon_{1,t+k}, \tag{7.6}$$

and

$$s_{t+k} = \alpha_2 + \beta_2 f_{t,k} + \epsilon_{2,t+k}, \tag{7.7}$$

where $\epsilon_{1,t+k}$ and $\epsilon_{2,t+k}$ are hypothesised white noise error terms.

Under the UIP and FRU hypotheses, $\alpha_i = 0$ and $\beta_i = 1$, $i = 1, 2$. Early analysis, such as Frenkel (1976), used Equation (7.7) and the results appeared encouraging, with estimates of β_2 being found to be close to 1. However, the results also had most of the hallmarks of the spurious regression problem alluded to in Chapter 4.[2] Therefore, most of the next round of empirical work used Equation (7.6) as its basis, with the seminal work being that of Fama (1984). Findings, based on a large variety of currencies and time periods, generally failed to accept the UIP hypothesis and the efficient market hypothesis; see, for example, Hodrick (1987), Lewis (1995) and Engel (1996). The estimates of β_1 obtained were usually negative and insignificantly different from zero. This negative estimate for β_1 is the main feature of the forward exchange rate anomaly; it implies that the more the forward currency is at a premium in the forward market the less the home currency is predicted to depreciate.

[1] See Sarno and Taylor (2003), Chapter 2 for a survey of evidence.

[2] See also Phillips (1986).

In recent years, two interrelated topics, which may have considerable relevance to the investigation of the forward rate anomaly, have attracted much attention. The first, deriving mainly from economic theory, is the possibility of nonlinearity in economic and financial relationships and its investigation using variations of the smoothed transition regression model of Granger and Teräsvirta (1993).[3] The second, based mainly on econometric theory, is the role of time-series dynamics and, in particular, the possibility of fractional integration in explaining the anomaly.[4]

These developments mirror in several ways the developments in econometric theory dealing with nonstationarity and nonlinearity of time-series processes. As mentioned in Chapter 6, it has been well known for many years that it is difficult to distinguish statistically between difference stationary series and nonlinear but stationary series. Recent research has not only considered the effects of nonlinearity on unit root tests such as the augmented Dickey-Fuller test, but also the reverse scenario; the effect of nonstationarity on tests for nonlinearity.

The empirical investigation of the forward exchange rate anomaly has been handicapped by a lack of appropriate econometric procedures. As discussed in Chapter 5, inference can be problematical in the fractionally integrated environment, as none of the normal procedures are appropriate. The classical asymptotics of the $I(0)$ case do not apply when the series are fractional and neither does the conventional $I(1)$ approach. Early tests of fractional integration were based on the frequency domain approach of Robinson (1994).[5] In this approach, a semiparametric test statistic is calculated for various values of d and inference is made on the tabulated results. For further details and an application of the methodology to the forward rate anomaly, using the Canadian US dollar exchange rate.[6]

Testing for nonlinearity in a context of nonstationarity is also problematical. In Gil-Alana (2004), an attempt was made to extend the semiparametric approach of Robinson (1994) but this requires knowledge of the form of the nonlinearity. Other recent papers by Dolado, et al. (2005b) and Mayoral (2005), considered testing for fractional integration against the alternative of stationarity and nonlinearity in the form of structural breaks.

In this chapter, the usefulness of two recent tests in helping to explain the forward exchange rate anomaly are investigated. The first is the fractional augmented Dickey-Fuller (FADF) test introduced by Dolado, et al. (2002), which is a simple-to-implement parametric test;[7] and the second is the random field regression-based approach to testing for nonlinearities introduced by Hamilton (2001), discussed at length in Chapter 2 and used throughout this thesis. The strength of Hamilton's approach, unlike others, is that it does not rely on any functional form being specified prior to estimation.

7.3 Data

To investigate the usefulness of both the Dolado, et al. (2002) FADF test and the Hamilton (2001) random field regression approach in helping explain the forward exchange rate anom-

[3]See, for example, Sarno, Valente and Hyginus (2006), Baillie and Kiliç (2006) and Sarno (2005).
[4]See Baillie and Bollerslev (2000), Maynard and Phillips (2001) and Gil-Alana (2002).
[5]See Gil-Alana and Robinson (1997, 2001).
[6]See Gil-Alana (2002).
[7]See Chapter 4 for a full discussion of this test procedure.

aly, this chapter applies the two techniques to the cross-exchange rates for the Australian dollar and sterling, the Canadian dollar and sterling, and the Japanese yen and sterling.

These data were taken from Thomson Financial's Datastream service and are provided in Appendix E.1, tables E.1, E.2 and E.3 for Australia, Canada and Japan, respectively. Just the relevant data are provided here: the exchange rate, the logarithm of the spot rate and the logarithm of the spot rate 3-month future. The remaining series examined here can all be derived from these. Specifically, the logarithm of the forward rate (3 months) is simply the logarithm of the relevant exchange rate, the forward premium is the difference between the logarithm of the forward rate and the logarithm of the spot rate, and the spot premium is the difference between the forward premium and the logarithm of the spot rate. In each case, the data used are daily series for the following periods:[8] 31$^{\text{th}}$ December 1996 to 16$^{\text{th}}$ June 2005, inclusive, for the Australian exchange rate and logarithm of the spot rate, and for all of the Japanese series; 31$^{\text{th}}$ December 1996 to 18$^{\text{th}}$ March 2005, inclusive, for the Australian logarithm of the spot rate 3-months future; 30$^{\text{th}}$ December 1994 to 16$^{\text{th}}$ June 2005, inclusive, for the Canadian exchange rate and the logarithm of the spot rate and 30$^{\text{th}}$ December 1994 to 18$^{\text{th}}$ March 2005, inclusive, for the Canadian logarithm of the spot rate 3-months future. For illustration, the relevant data series are plotted in Appendix E.2, figures E.1 to E.6.

7.4 Testing for (Co)integration

For purposes of comparison, the study starts with the standard $I(1)/I(0)$ analysis, using the ADF test and using the methodology of Dolado, et al. (1990), to determine whether the series are trend stationary or difference stationary.[9] The lag lengths for the ADF test were determined by means of the modified Akaike information criterion (MAIC), which Ng and Perron (2001) have shown to be a generally better decision criterion than the standard AIC, as it takes account of the persistence found in many series. As well as testing the individual exchange rate and exchange premium series, unit root tests were also carried out on the ordinary least squares residuals from a number of static regressions to assess the possibility of cointegration.

The results of the preliminary unit root tests are given in Table E.4. The Dolado, et al. (1990) testing strategy failed to support the existence of a trend or drift in all cases, so the p-values given in Table E.4 are those provided by MacKinnon (1996). A clear picture emerges from Table E.4. In all three cases considered, the ADF test does not suggest rejection of the null hypothesis of a unit root in the forward exchange rate, the spot rate and the forward premium, but it does point to clear rejection of the unit root null for all three spot premiums.

Following this analysis, the issue of fractional integration was investigated with the Dolado, et al. (2002) FADF test, which itself requires a consistent parametric estimate of the parameter d. Following previous chapters, two parametric estimates of d are calculated, using the ARFIMA package of Doornik and Ooms (1999). The first is Sowell's (1992) exact maximum likelihood (EML) estimator, which requires that $d < 0.5$. As discussed in Chapter

[8]Although daily series are available, the data in Appendix E.1 show only every fifth observation. The daily series were analysed throughout this chapter, with the exception of the random field regressions, where weekly data were used to reduce the computational burden.

[9]See Chapter 4 for further details of Dolado, et al. (1990).

4, estimates of d are obtained from the first differences of series, for this reason. The other estimator is Beran's (1995) nonlinear least squares estimator (NLS), which is based on the conditional sum of squared naïve residuals. The nonparametric estimators of Geweke and Porter-Hudak (1983) (GPH), and the semiparametric estimator discussed in Robinson (1994), are also computed. Both of these complementary approaches are also available in the Doornik and Ooms ARFIMA package. The parametric EML estimates of d are then used in the FADF test, with the MAIC again being used to set the lag length of the test. To investigate the results of Maynard and Phillips (2001) and Zivot (2000), equations (7.6) and (7.7) were estimated and the order of integration of the respective estimated error terms explored.[10] For Equation (7.7), both s_t and s_{t+k} are regressed on $f_{t,k}$, following Maynard and Phillips.

Table E.5 contains the results of the fractional integration analysis. The FADF test fails to reject the null hypothesis that the spot and forward rates are $I(1)$ against the alternative of fractional integration, in agreement with the findings of Heravi and Patterson (2005). However, from this table it can be seen that in all cases, it is unlikely that the forward premium is either $I(1)$ or $I(0)$. Whereas the ADF test could not reject the $I(1)$ hypothesis, the FADF test clearly rejects it, if the EML estimate of d is used for the alternative hypothesis. It therefore seems very likely that the forward premium is $I(d)$, where $0 < d < 1$. The weight of evidence from the parametric estimators is that d is around 0.5.

Estimation of the value of d for the spot premium proved interesting. Given the ADF results in Table E.4, it would seem reasonable to assume that d is close to zero. However, both parametric and semiparametric estimates of d are close to unity. The converse seems to be true for the forward premium, for which Table E.4 suggests a value of d close to one, while the corresponding results in Table E.5 suggest a much lower value of d. The values of d obtained for the forward premium are more in line with those reported by Baillie and Bollerslev (1994), than those found by Maynard and Phillips (2001).

The results for the standard regression models of equations (7.6) and (7.7) are presented in Table E.6. The estimated coefficients are generally in line with the corresponding results from previous studies and it can be concluded from the standard Engle-Granger (1987) approach that most of the levels regressions may constitute a cointegrating regression. There are clear contradictions, however, between the Engle-Granger procedure and the results of the CRDW test; for the regressions s_t on $f_{t,k}$, the CRDW test rejects the null of no cointegration for all three currencies, in direct contradiction to the Engle-Granger approach. This situation is reversed for the remainder of regressions, with the CRDW test failing to reject the null of no cointegration, while the AEG test finds evidence of cointegration in all cases. Table E.7 presents estimates of d obtained from the four alternative methods of estimation applied to

[10] Maynard and Phillips (2001), using both parametric and nonparametric estimation, found evidence of nonstationary long-memory behaviour of the forward premium. They went on to show how the implied imbalance in the traditional Fama-type regressions leads to nonstandard limiting distributions for the estimators and the test statistics. The slope and R^2 coefficients converge to zero, the t statistic diverges and its left-tailed limiting distribution appears consistent with the forward rate anomaly. They also showed that regression in the levels would be fractionally cointegrated, with nonstationary residuals and a slope coefficient estimate that is consistent, but with a t statistic that diverges. If the forward premium is indeed a long memory process, then the simple FRU hypothesis must be rejected. Zivot (2000) showed that if s_t and $f_{t,1}$ are cointegrated, the cointegrating model for s_{t+1} and $f_{t,1}$ is not a simple finite-order ECM and that estimating a first-order ECM for s_{t+1} and $f_{t,1}$ can lead to mistaken inferences concerning the exogeneity of the spot rate and the unbiasedness of the forward rate.

the residuals from the regressions reported in Table E.6. The results obtained are broadly in line with the theory given in Maynard and Phillips (2001), as previously mentioned. While estimates of d for the residuals of all three s_t on $f_{t,k}$ regressions are generally small, those from the other two regressions are considerably larger in the three cases. Correspondingly, the FDF and FADF tests clearly reject the $I(1)$ null in favour of the alternative of fractional integration for the residuals of all regressions of s_t on $f_{t,k}$ and for the residuals of the Canadian dollar-sterling regression of s_{t+k} on $f_{t,k}$. This latter rejection of the null appears to agree with the corresponding result in Table E.6. However, in the case of Table E.7, the indication is clearly that the alternative is $0.5 < d < 1$, rather than $0 \leq d \leq 0.5$, though $I(0)$ is the conclusion from the Engle-Granger type analysis. There is one case of clear disagreement between the findings of the unit root test in Table E.6 and the corresponding result in Table E.7. This is the Japanese yen-sterling regression of $\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$, for which the standard unit root test strongly suggest rejection of the unit root null, while the fractional tests indicate nonrejection. Given that the Dickey-Fuller test is well known to have low power in distinguishing between series that are $I(1)$ and $I(d)$, where d is less than but close to unity, this is somewhat puzzling, as the estimates of d in this case are all very close to unity. This case further underlines the difficulties in relying on one approach or procedure.

7.5 Nonlinear Inference

Finally, the random field regression approach is applied to the data. To do this, the GAUSS program code provided by Hamilton (2001) is used. Given the large size of the dataset, the approach of Hansen and Hodrick (1980) is adopted to ease the considerable computational burden involved in the random field analysis. Weekly data points are chosen, using every fifth observation.

The results from the Hamilton (2001) analysis are given in Table E.8. In producing these results, two variants of equations (7.6) and (7.7) were used. After some exploratory checking of cross-plots, a time trend was included in both equations, and they were estimated with and without a constant. For computational reasons, some re-scaling of the data was undertaken and an algorithm-switching strategy was used in the numerical optimisation. Specifically, the observations on the explanatory variable $f_{t,k}$ were scaled up by a factor of ten in the case of the Canadian dollar-sterling data, while switching between the GAUSS algorithms STEEPEST DESCENT and NEWTON was used, along with selected initial values of ζ, ranging from 0.1 to 1.9, and the default value of the GAUSS parameter `_oprteps`; see Chapter 2 for further details on the computational aspects of the Hamilton (2001) methodology. Furthermore, both the original Hamilton covariance matrix and the Dahl and González-Rivera (2003) forms of the covariance matrix for the random field were utilised.[11] Despite the extensive experimentation in the approach to the calculations, it did not prove possible to obtain nonlinear estimates for the premium equations in the Australian dollar-sterling and the Japanese yen-sterling cases. In the cases in which the nonlinear estimation was successful, the number of iterations required to determine the maximum likelihood estimates ranged from 6 to 28.

[11] See chapters 2 and 3 for further details.

There is overwhelming evidence of nonlinearity in these models, with the Hamilton (2001) Lagrange multiplier test statistics ranging from 381.46 to 5925.76. As can be seen from Table E.8, the nonlinearity in the equations is consistently associated with the time variable, which has a statistically significant coefficient in the nonlinear component of all of the models successfully estimated. It should be noted however, that $f_{t,k}$ also has a statistically significant coefficient in the nonlinear component in the case of Japan, which suggests that unlike the other cases, the nonlinearity here may be associated with both $f_{t,k}$ and t.

A feature of the Hamilton (2001) results is the high significance of the σ and ζ estimates in the equations for exchange rates, and the contrasting lack of significance of these estimates in the Canadian dollar-sterling premium equations, even though the latter estimates are much bigger numerically. It seems reasonable to assume that these particular insignificant results are related to what, in the time-series literature, is known as the 'pile-up' phenomenon associated with numerical optimisation, and that this may signal that the covariance structure used for the random field, if not the normality assumption itself, may not be entirely appropriate.[12] Recall that this problem was also encountered in Chapter 6, the study of purchasing power parity.

However, the most significant aspect of the Hamilton (2001) analysis is that it shows that when nonlinearity is allowed for by means of a random field in the exchange rate equation, the intercept is not significantly different from zero and the slope coefficient is estimated, with great precision, to be unity in each of the cases considered, in accordance with exchange rate theory. Similarly, in the Canadian dollar-sterling exchange premium equation, the intercept and slope are not significantly different from zero and unity, respectively, though as the standard errors are larger in this case, the result is not quite as striking as it is for the rate equations. Modelling nonlinearity using the Hamilton method seems to remove the forward anomaly.

7.6 Conclusion

This chapter has focused on the well-known foreign exchange rate anomaly, brought to prominence by Fama (1984). It has given brief descriptions of the anomaly and the main early approaches that were used in trying to explain it. In particular, it has drawn attention to the theoretical work by Dolado, et al. (2002) on testing for fractional integration, and that of Hamilton (2001) on random field regression and nonlinear inference, as developments that offer relevant new approaches to the study of the anomaly. Finally, to assess the usefulness of these two new methods, the chapter reports on an investigation of their application to three sets of exchange rate and exchange premium data. The main findings are as follows.

In all three cases considered, the standard $I(1)/I(0)$ approach to testing for unit roots and cointegration suggests that spot and forward exchange rates, as well as the forward exchange premium, behave as nonstationary $I(1)$ series, and that the spot premium is $I(0)$. There are mixed findings on the possibility of cointegration; this possibility is clearest in the Canadian dollar-sterling case.

[12]See Chapter 2, DeJong and Whiteman (1993), and Hamilton (2005) for further details.

While the results of the fractional integration analysis accord with the finding that spot and forward rates are $I(1)$, they contradict those of the standard analysis with regard to the properties of the exchange premiums. Whereas the ADF test suggests that $d = 1$ for the forward premium, EML and other estimates indicate a value closer to $d = 0.5$, and the FDF and FADF tests give a strong rejection of the unit root null hypothesis. Similarly, rejection of the unit root null in the standard analysis suggests that the spot premium may be treated as $I(0)$, while fractional parameter estimation indicates that d is fairly close to unity. This latter conflict is very puzzling and deserves attention in any future research, perhaps using the new test of Dolado, et al. (2005b), which would permit testing of the null hypothesis that a series is $I(0)$ against the alternative that it is fractionally integrated.

Similar discrepancies emerge between the outcomes of standard unit root test and the fractional analysis when the ordinary least squares residuals from a variety of regressions are examined. The FDF and FADF tests tend to support the standard tests with regard to their finding that the unit root null should be rejected for the residuals, but the fractional analysis suggests that $0 < d < 1$, calling into question the standard conclusion that the residuals may be deemed to be $I(0)$.

There are strong indications of time-dependent nonlinearity when the data are subjected to examination using the Hamilton (2001) nonlinearity test and random field regression procedure, though in two of our cases nonlinear estimates for the premium equation could not be obtained. This matter also deserves further investigation in future research.[13] It is of considerable interest that in all cases when the nonlinearity is successfully modelled by means of a random field, exchange rate theory is confirmed and the forward rate anomaly removed. This key finding adds weight to the earlier work on the relevance of nonlinearity or parameter instability to the forward anomaly debate, referred to in Section 7.2. It points clearly to the possibility that the kind of relationships that have been estimated are nonlinear.

[13]See Bond, Harrison, and O'Brien (2005a) for a detailed discussion of Hamilton algorithm failures.

Chapter 8

Conclusion

This thesis has discussed a variety of issues in applied time series econometrics. Specifically, it has considered some of the problems associated with modelling stationary, non-stationary and nonlinear data, and the difficulties in distinguishing between them. The Hamilton (2001) framework for exploring nonlinearity has been examined, as has the LM-type test for nonlinearity, proposed therein. The concept of fractional integration has been highlighted and discussed, and its potential in further developing modelling strategies has been explored. Johansen's (2002) small-sample Trace test correction has also been employed. These techniques have been applied to three case studies: the demand for money in Denmark and Finland, purchasing power parity for Ireland, Germany, and the United Kingdom, and the forward exchange rate anomaly for the currencies of Australia, Canada, Japan, and the United Kingdom. What follows is a more detailed summary of each chapter.

After a short introductory chapter which discussed the motivation behind this thesis, Chapter 2 began by giving a brief review of nonlinear economic modelling. The motivation for, and several methods of, modelling nonlinear economic relationships were introduced and discussed. Most importantly, a new approach to nonlinear econometric modelling, proposed by Hamilton (2001), was introduced. An account of this new approach was given, as was a brief description of some of the methods of nonlinear optimisation that may be used in the GAUSS computer program provided by Hamilton for the implementation of his methodology. The performance of this program was investigated using data relating to Hamilton's three examples, using not only randomly generated data, but also data concerning the US Phillips curve, two versions of the GAUSS software and a range of alternative numerical optimisation options, including the GAUSS parameter `_oprteps`, and parameter starting values. The performance of algorithm switching procedures was also examined. Finally, the effects of changes in the sample data on the results produced by Hamilton's procedure were explored. The focus was designedly on the GAUSS implementation of the procedure and, while several changes in the GAUSS procedures were investigated, no attempt was made at modification of Hamilton's methodology. The results presented suggested some clear conclusions, which will hopefully be of value to those contemplating working with Hamilton's method. Different algorithms used for the numerical optimisation have different chances of success. Hamilton's choice of the Broyden, Fletcher, Goldfarb and Shanno algorithm fails in over 60 per cent of the cases examined in the study of his Example 3, while the less computationally efficient STEEPEST DESCENT method succeeds in all cases. When different algorithms work, they may produce significantly different numerical results, including different signs for parameter estimates. The use of procedures that employ two algorithms and a switching criterion appears to result in less variation in the parameter estimates, when compared to any one algorithm used on its own, and to be less sensitive to the choice of initial parameter estimates. Minor changes in data can have significant effects, both in terms of whether an algorithm operates or not and, in the case of it operating, the numerical results it produces. Despite the sensitivity of results to choice of algorithm, initial values and data changes, the statistical significance of the nonlinear parameter estimates, hence the inference about the form of nonlinearity, generally seems to be little affected according to the findings that have been reported.

For the simulated datasets used in Examples 1 and 2, the sensitivity of results to choice of algorithm and size of sample was less pronounced than has been found using the real data of Example 3. Hamilton (2005) offered some insight into the difficulties which may have caused the numerical instability of the algorithms revealed in this study. His findings seem to confirm that 'difficult' likelihood surfaces make convergence challenging. Another possibility is that difficulties exist within Hamilton's methodology, and that the use of the alternative covariance functions for specifying the random field regression model may help to reduce numerical instability, as may different procedures for estimating the parameters of Hamilton's random field regression model. These matters constitute an interesting agenda for future research. Finally, Hamilton (2005) drew attention to the phenomenon known as 'pile-up'. While issues relating to this were not discussed in Chapter 2, they are of considerable importance, and were discussed in Chapter 5. With regard to the implementation of the Hamilton random field regression methodology, the main recommendation of Chapter 2 is to employ an algorithm switching approach supplemented as necessary by changes of the GAUSS `_oprteps` parameter, the starting value of ζ, and using both the Hamilton and Dahl and the González-Rivera's (2003) covariance specifications. On the basis of the evidence provided here and supplemented in later chapters, such an approach is more efficient than the use of single algorithms and appears to be less susceptible to numerical instability and failure.

Chapter 3 assessed the power of several tests of nonlinearity, some of which are well known, across a range of specifications often encountered in economics and econometrics. The results provide some clear evidence regarding the comparative power of these tests. The well-known tests of Durbin and Watson (1950), Harvey and Collier (1977) and Ramsey (1969) are powerful against misspecification and nonlinearity, particularly when the former tests are applied to ordered data. Given the relative simplicity of these tests and their wide availability, with the exception of the Harvey-Collier test, these results certainly endorse their use in applied research. The LM-type test, proposed by Hamilton (2001), does offer a more powerful solution, but the increase in power is small, and given its more complex nature and lack of widespread availability, it may remain under-utilised. Of the three tests proposed by Dahl and González-Rivera (2003), the λ_{OP}^{A} test appeared to be most powerful across the range of specifications examined here, particularly when the test was implemented with bootstrapped, as opposed to asymptotic, p-values. Interestingly, there does not appear to be a large difference between the powers obtained from the bootstrapped p-values and the asymptotic p-values, despite the relatively small sample sizes used. The exception to this was the g_{OP} test, where asymptotic and bootstrapped p-values differed considerably. Also, the λ_{OP}^{A} test appeared to be somewhat more powerful than Hamilton's λ_{H}^{E} test, although the powers for this test are based on the asymptotic p-values. Once again, however, differences in power are small, suggesting perhaps that Ramsey's RESET should be favoured over the random field methods. Avenues for further research would be to compare the performance of the bootstrapped λ_{H}^{E} test with its asymptotic equivalent; to consider the performance of the Keenan (1985), Tsay (1986) and Luukkonen, Saikkonen, and Teräsvirta (1988) tests, which are adaptations of RESET, against the random field methods discussed here, as these are potentially more powerful than the RESET procedure and perhaps, therefore, more powerful than the random field approaches. It would also be interesting to consider a wider range of model specifications and data, including nonnormal distributions, to gain a greater understanding of the properties of these tests.

Chapter 4 introduced the necessary theoretical material for the remaining chapters. It contained standard treatments of the concepts of stationarity and nonstationarity, testing for unit roots, cointegration and three approaches to testing for it, namely, the Engle-Granger (1987) 2-step method, Johansen's (1988, 1991) VAR tests and the COMFAC approach. Consideration was also given to the fractional augmented Dickey-Fuller test proposed by Dolado, Gonzalo, and Mayoral (2002) and Johansen's (2002) small sample correction to the Trace test. Given the relatively recent nature of these topics, this chapter aimed to give just an account of these procedures.

Chapter 5 drew attention to some of the pitfalls involved in using the conventional $I(1)/I(0)$ framework for economic and financial modelling of time-series data, an approach involving well-known unit root tests and the cointegration testing and modelling procedures of Engle and Granger (1987), Johansen (1988, 1991) and COMFAC analysis that have been applied widely in applied economics. The practical difficulties of untangling the issues of stationarity, fractional integration, nonlinearity and parameter instability were highlighted.

These issues were highlighted by presenting a case study intended to illustrate the application of these newer techniques and to contrast their findings with those of the standard cointegration modelling approach. The study used the data previously analysed by Johansen and Juselius (1990) and Johansen (1996) in connection with demand for money functions in Denmark and Finland. The results obtained from the various techniques exemplify the problems with the standard approach and the alternative conclusions that might be reached by using different techniques. The findings were as follows.

Though ADF tests, implemented using the procedure of Dolado, Jenkinson, and Sosvilla-Rivero (1990), appear to suggest unit roots for most variables, they are sensitive to the specification of the test equation and the information criterion used to choose lag length in the case of some variables, especially for Finland. When the matter of unit roots was explored further, using the Elliot, Rothenberg, and Stock (1996), Kwiatkowski, Phillips, Schmidt, and Shin (1992), and Ng and Perron (2001) tests, unit roots for the Danish variables tended to be confirmed but not for the Finnish variables.

Proceeding on the assumption that all variables are $I(1)$, the Engle-Granger (1987) 2-step procedure does not support cointegration in general, a result that is confirmed by ECM tests conducted in an error-correction framework for the money demand relationship for each country. However, the Engle-Granger approach does suggest cointegration for the Finland model that treats both the inflation and deposit rate variables as $I(0)$.

Using the Johansen (1988, 1991) approach without small sample bias-correction for the Trace test, there is considerably stronger evidence of cointegration in the case of Denmark, though the number of cointegrating vectors suggested varies, depending on the VAR specification chosen. The picture that emerges for Finland is similar, although for the model that treats the inflation and deposit rate variables as $I(0)$, the Johansen (1988, 1991) method suggests no cointegration, contradicting the finding of the Engle-Granger (1987) procedure in this case.

The Johansen (2002) correction factor has a marked effect on the result in the case of the small sample of data for Denmark; the modified trace test agreeing with the conclusion from the Engle-Granger (1987) procedure that there is no cointegrating demand for money relationship. The Johansen correction has no effect on the findings for Finland, which are based on a much larger sample. These results are puzzling, and in particular, the contradictory results from the Engle-Granger and Johansen (1988, 1991, 2002) procedures concerning the existence of cointegrating relationships, in the case of both countries, is curious.

The common factor approach failed to reject the possibility of cointegration in either the Danish or Finnish samples, including the reduced Finnish models. This adds further weight to the contradictory evidence regarding cointegration in these data.

Checking for fractional integration by means of a range of estimators of the fractional integration parameter, as well as the new fractional Dickey-Fuller and fractional augmented Dickey-Fuller tests of Dolado, et al. (2002), confirms the $I(1)$ nature of the Danish variables and the lack of a unit root for the variables in the case of Finland.

Assuming that the Finnish data are not $I(1)$, and hence can not be simply cointegrated, the possibility of stationarity with regime shifts or some other kind of nonlinearity arises. This was explored, for both countries, by the Hamilton (2001) procedure, which is more appropriate for general, unknown forms of nonlinearity. This method produced strong evidence of structural change/nonlinearity, if underlying stationarity is entertained. However, an attempt to re-specify the money demand equations as piecewise linear regressions, which was suggested by examination of the data, was not very successful. Clearly, further work would be necessary to find an adequate nonlinear functional form, were this alternative approach to be preferred.

The messages from this study appear to be as follows. The standard $I(1)/I(0)$ modelling strategies for economic and financial time series appear to be fraught with dangers. Fractional integration analysis, which is relatively easy to implement, may confirm the existence of unit roots, but may also suggest fractional integration to different degrees for different variables. This is a complicated situation that raises challenges for modelling. Recalling that unit root tests may often indicate that a unit root exists when a series is stationary but subject to level shifts, a general analysis of nonlinearity, such as that offered by the Hamilton (2001) procedure, may be an attractive option that can lead to acceptable alternative models. It would seem that the reliance on any one approach may not be a sensible practice in applied work, and that practitioners would be well advised to consider using a range of alternative methods and selecting models according to the balance of the wider body of evidence produced.

Chapter 6 examined the testing of purchasing power parity using Irish data. Tests for fractional integration and nonlinearity were used to investigate the behaviour of the Irish exchange rate for Germany and the United Kingdom. This chapter, much like Chapter 5, showed the potential for difficulties in placing the study of purchasing power parity in the $I(1)/I(0)$ framework. It also showed that in this case these difficulties cannot be overcome by moving to a fractional integration framework. In fact, there was strong evidence of nonlinearity in the data. This suggests that perhaps a nonlinear approach should be taken to model this type of relationship.

Standard augmented Dickey-Fuller tests, implemented using the procedure of Dolado, et al. (1990), suggested that most of the variables here have unit roots. Although there was some doubt over this finding for the Irish price level, this was assumed to be the case. There was no evidence of seasonal integration in the variables. Under the assumption that all of the variables were $I(1)$, the Engle-Granger (1987) approach was employed to test for cointegration, but no evidence of cointegration was found for either Germany or the United Kingdom. This result was confirmed by the estimation of error-correction models and ECM tests of cointegration.

The Johansen (1988, 1991) approach found evidence of one cointegrating vector in both the German and United Kingdom standard models. Applying Johansen's (2002) small sample correction overturned this finding, however. For the models augmented with interest rates, one cointegrating vector was again found, but the correction factor had no bearing in this case, thus contradicting the results of the Engle-Granger (1987) approach. Interestingly, the common factor approach found no evidence whatsoever of cointegration.

Little evidence of fractional integration was found in the series, with only Irish interest rates having an estimated parameter of fractional integration significantly different from 1. Unlike the results in Chapter 5, this suggests that the conflicting evidence on cointegration cannot readily be attributable to fractionality. Testing for nonlinearity with Hamilton's (2001) LM-type test uncovered strong evidence of nonlinearity in both the German and United Kingdom models. Evidence from the random field regressions suggested that the price and interest variables play a strong role in this nonlinearity. Examination of simple cross plots revealed evidence of shifts in regime for those variables found to contribute to the nonlinearity. No attempt was made to model such nonlinearity in this case, although this is an interesting avenue for future research.

It would seem, therefore, that despite the apparent $I(1)$ nature of the data, the standard and augmented PPP models do not fit well into the cointegration and error-correction framework. As the series do not appear to exhibit fractional integration, entertaining the notion of fractional cointegration would not further the modelling effort. In fact, evidence suggests that these PPP relationships are highly nonlinear, and that such nonlinearity derives from both the prices and interest rates. Efforts to model purchasing power parity should be directed to capturing this nonlinearity.

Chapter 7 focused on the well-known foreign exchange rate anomaly, brought to prominence by Fama (1984). It gave a brief description of the anomaly, the main early approaches that were used in trying to explain it, and how fractional integration and nonlinearity underlie some recent attempts at explanation, using both long memory time-series models and tests for nonlinearity. In particular, it has drawn attention to the theoretical work by Dolado, et al. (2002) on testing for fractional integration, and that of Hamilton (2001) on random field regression and nonlinear inference, as developments that offer relevant new approaches to the study of the anomaly. Finally, to illustrate and assess the usefulness of these two new methods, the chapter reported on an investigation of their application to three sets of exchange rate and exchange premium data. The main findings were as follows.

In all three cases considered, the standard $I(1)/I(0)$ approach to testing for unit roots and cointegration suggested that spot and forward exchange rates, as well as the forward exchange premium, behave as nonstationary $I(1)$ series, and that the spot premium is $I(0)$. Furthermore, there were mixed findings on the possibility of cointegration; the possibility was clearest in the Canadian dollar-sterling case. While the results of the fractional integration analysis accorded with the finding that spot and forward rates were $I(1)$, they contradicted those of the standard analysis with regard to the properties of the exchange premiums. Whereas the ADF tests suggested that $d = 1$ for the forward premium, EML and other estimates indicated a value closer to $d = 0.5$, and the FDF and FADF tests gave a strong rejection of the unit root null hypothesis. Similarly, rejection of the unit root null in the standard analysis suggested that the spot premium may be treated as $I(0)$, while fractional parameter estimation indicated that d is fairly close to unity. This latter conflict is very puzzling and deserves attention in any future research, perhaps using the new test of Dolado, Gonzalo, and Mayoral (2005b), which would permit testing of the null hypothesis that a series is $I(0)$ against the alternative that it is fractionally integrated. Similar discrepancies emerged between the outcomes of standard unit root tests and the fractional analysis when the ordinary least squares residuals from a variety of regressions were examined. The FDF and FADF tests tended to support the standard tests with regard to their finding that the unit root null should be rejected for the residuals, but the fractional analysis suggested that $0 < d < 1$, calling into question the standard conclusion that the residuals may be deemed to be $I(0)$. There were strong indications of time-dependent nonlinearity when the data are subjected to examination using the Hamilton (2001) nonlinearity test and random field regression procedure, though in two of our cases, nonlinear estimates for the premium equation could not be obtained. This matter also deserves further investigation in future research to develop the findings of Bond, Harrison, and O'Brien (2005a), concerning the failure of the Hamilton algorithm. It is of considerable interest that in all cases when the nonlinearity was successfully modelled by means of a random field, exchange rate theory was confirmed and the forward rate anomaly removed. This key finding adds weight to the earlier work on the relevance of nonlinearity or parameter instability to the forward anomaly debate. It points clearly to the possibility that the kind of relationships that have been estimated are nonlinear.

In conclusion, this thesis has outlined methods for modelling nonlinearity, and in particular, the new approach attributable to Hamilton (2001). It has discussed the implementation of this procedure and explored the power of its integral test for nonlinearity. It has highlighted the potential difficulties associated with the standard $I(1)/I(0)$ framework, and has used the Hamilton approach, in tandem with other recent developments in testing for fractional integration and cointegration, to compare modelling strategies, using as illustrative case studies the demand for money, purchasing power parity, and the forward exchange rate anomaly. This work has offered clear insights into the difficulties that may be encountered in numerical optimisation, regardless of the procedure or program in which it is used, and has contributed further to the understanding of the Hamilton methodology. It has ascertained the power of a range of random field-based tests for nonlinearity against several well-known tests. The problems of relying on any one approach to modelling, a central theme of this thesis, have been clearly outlined in three case studies. These studies have also contributed to the understanding of the areas of money demand, purchasing power parity, and the forward exchange rate anomaly.

Bibliography

AHN, S. K., AND C. G. REINSEL (1990): "Estimation for Partially Nonstationary Multivariate Autoregressive Models," *Journal of the American Statistical Association*, 85, 813–823.

AINZENMAN, J. (1986): "Monopolistic Competition, Relative Prices and Output Adjustment in the Open Economy," Working Paper No. 1787, National Bureau of Economic Research.

AKAIKE, H. (1973): "Information Theory and an Extension of the Maximum Likelihood Principle," in *Second International Symposium on Information Theory*, ed. by B. Petrov, and C. Csake. Akademiai Kiado, Budapest.

ALDER, M., AND B. LEHMAN (1983): "Deviations from Purchasing Power Parity in the Long Run," *Journal of Finance*, 39, 1471–1487.

ANDERSEN, T. G., AND T. BOLLERSLEV (1997): "Heterogeneous Information Arrivals and Returns Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975–1005.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

ANDERSON, T. W. (1951): "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions," *Annals of Mathematical Statistics*, 22, 327–351.

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

ANDREWS, D. W. K., AND W. PLOBERGER (1994): "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383–1414.

ANSLEY, C. F., AND P. NEWBOLD (1980): "Finite Sample Properties of Estimators for Autoregressive Moving Average Models," *Journal of Econometrics*, 13, 159–183.

ASTLEY, M. S., AND A. HALDANE (1997): "The Information in Money," Bank of England Quarterly Bulletin, May.

BACON, D. W., AND D. G. WATTS (1971): "Estimating the Transition between Two Intersecting Straight Lines," *Biometrika*, 58, 525–534.

BAILLIE, R. T. (1996): "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5–59.

BAILLIE, R. T., AND T. BOLLERSLEV (1994): "Cointegration, Fractional Cointegration and Exchange Rate Dynamics," *Journal of Finance*, 49, 737–745.

——— (2000): "The Forward Premium Anomaly is Not as Bad as You Think," *Journal of International Money and Finance*, 19, 471–488.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996): "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 3–30.

Baillie, R. T., C.-F. Chung, and M. A. Tieslau (1996): "Analysing Inflation by the Fractionally Integrated Arfima-Garch Model," *Journal of Applied Econometrics*, 11, 23–40.

Baillie, R. T., and R. Kiliç (2006): "Do Asymmetric and Nonlinear Adjustments Explain the Forward Premium Anomaly?," *Journal of International Money and Finance*, 25, 22–47.

Banerjee, A., J. J. Dolado, J. W. Galbraith, and D. F. Hendry (1993): *Co-integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford University Press, Oxford.

Banerjee, A., D. F. Hendry, and G. W. Smith (1986): "Exploring Equilibrium Relations in Economics through Static Models," *Oxford Bulletin of Economics and Statistics*, 48, 253–277.

Bartlett, M. S. (1937): "Properties of Sufficiency and Statistical Tests," *Proceedings of the Royal Society of London, Series A*, 160, 268–282.

——— (1950): "Periodogram Analysis and Continuous Spectra," *Biometrika*, 37, 1–16.

Beran, J. (1995): "Maximum Likelihood Estimation of the Differencing Parameter for Invertible and Short and Long Memory Autoregressive Integrated Moving Average Models," *Journal of the Royal Statistical Society, Series B*, 57, 659–672.

Bhargava, A. (1986): "On the Theory of Testing for Unit Roots in Observed Time Series," *The Review of Economic Studies*, 53, 369–384.

Bhattacharya, R. N., V. K. Gupta, and E. Waymire (1983): "The Hurst Effect Under Trends," *Journal of Applied Probability*, 20, 649–662.

Bierens, H. J. (2001): "Unit Roots," in *A Companion to Econometric Theory*, ed. by B. Baltagi. Blackwell, Oxford.

——— (2004): *Introduction to the Mathematical and Statistical Foundations of Econometrics.* Cambridge University Press, Cambridge.

Bond, D., M. J. Harrison, N. Hession, and E. J. O'Brien (2006): "Some Empirical Observations on the Forward Exchange Rate Anomaly," Trinity Economic Papers, No. 2.

Bond, D., M. J. Harrison, and E. J. O'Brien (2003): "Investigating Nonlinearity: A Note on the Implementation of Hamilton's Methodology," Trinity Economic Papers, No. 12.

——— (2005a): "Investigating Nonlinearity: A Note on the Estimation of Hamilton's Random Field Regression Model," *Studies in Nonlinear Dynamics and Econometrics*, 9, Article 2.

——— (2005b): "Investigating Nonlinearity: A Note on the Estimation of Hamilton's Random Field Regression Model," Trinity Economic Papers, No. 4.

——— (2005c): "Testing for Long Memory and Nonlinear Time Series: A Demand for Money Study," Trinity Economic Papers, No. 21.

——— (2006): "Purchasing Power Parity: The Irish Experience Re-visited," Trinity Economic Papers, No. 15.

——— (2007): "Demand for Money: A Study in Testing Time Series for Long Memory and Nonlinearity," *The Economic and Social Review*, forthcoming.

Box, G. E. P., and G. M. Jenkins (1976): *Time Series Analysis, Forecasting and Control.* Holden Day, San Francisco, revised edn.

Breidt, F. J., N. Crato, and P. de Lima (1998): "The Detection and Estimation of Long Memory in Stochastic Volatility," *Journal of Econometrics*, 83, 325–348.

Brent, R. P. (1972): "A New Algorithm for Minimizing a Function of Several Variables without Calculating Derivatives," in *Optimization*, ed. by R. Anderssen, L. Jennings, and D. Ryan. University of Queensland Press, Brisbane.

——— (1973): *Algorithms for Minimization without Derivatives.* Prentice-Hall, Englewood Cliffs, New Jersey.

Brock, W., W. D. Dechert, and J. Scheinkman (1987): "A Test for Independence Based on the Correlation Dimension," Working Paper, Department of Economics, University of Wisconsin, Madison.

Broyden, C. G. (1967): "Quasi-Newton Methods and their Application to Function Minimization," *Mathematics of Computation*, 21, 368–381.

Caner, M., and B. E. Hansen (1998): "Threshold Autoregressions with a Unit Root," Working Paper in Economics No. 381, Boston College.

Caporale, G. M., and L. A. Gil-Alana (2005): "Fractional Cointegration and Aggregate Money Demand Functions," *The Manchester School*, 73, 737–753.

Cashin, P., and C. J. McDermott (2004): "Parity Reversion in Real Exchange Rates: Fast, Slow or Not at All?," Working Paper No. 128, International Monetary Fund.

Chen, W. W., and C. M. Hurvich (2003a): "Estimating Fractional Cointegration in the Presence of Polynomial Trends," *Journal of Econometrics*, 117, 95–121.

——— (2003b): "Semiparametric Estimation of Multivariate Fractional Cointegration," *Journal of the American Statistical Association*, 98, 629–642.

Cheung, Y. W., and K. S. Lai (1993a): "Finite Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration," *Oxford Bulletin of Economics and Statistics*, 55, 313–332.

——— (1993b): "A Fractional Cointegration Analysis of Purchasing Power Parity," *Journal of Business and Economic Statistics*, 11, 103–12.

——— (2001): "Long Memory and Nonlinear Mean Reversion in Japanese Yen-based Real Exchange Rates," *Journal of International Money and Finance*, 20, 115–132.

Choi, I., and P. Saikkonen (2004): "Testing Linearity in Cointegrating Smooth Transition Regression," *The Econometric Journal*, 7, 341–365.

Cooper, D. M., and R. Thompson (1977): "A Note on the Estimation of the Parameters of the Autoregressive-moving Average Process," *Biometrika*, 64, 625–628.

Cox, D., and E. Koh (1989): "A Smoothing Spline Based Test of Model Adequacy in Polynomial Regression," *Annals of the Institute for Statistical Mathematics*, 41, 383–400.

Cox, D., E. Koh, G. Wahba, and B. Yandell (1988): "Testing the (Parametric) Null Model Hypothesis in (Semiparametric) Partial and Generalized Spline Models," *The Annals of Statistics*, 16, 113–119.

Cribari-Neto, F., and G. M. Cordeiro (1996): "On Bartlett and Bartlett-type Corrections," *Econometric Reviews*, 15, 339–367.

Cryer, J. D., and J. Ledolter (1981): "Small-sample Properties of the Maximum Likelihood Estimator in the First-order Moving Average Model," *Biometrika*, 68, 691–694.

da Silva Lopes, A. C. B., and A. Montanes (1999): "The Behaviour of Seasonal Unit Roots Tests with Seasonal Mean Shifts," Discussion Paper No. 10-99, Universidade Técnica de Lisboa.

Dahl, C. M. (2002): "An Investigation of Tests for Linearity and the Accuracy of Likelihood Based Inference Using Random Fields," *Econometrics Journal*, 5, 263–284.

Dahl, C. M., and G. González-Rivera (2003): "Testing for Neglected Nonlinearity in Regression Models Based on the Theory of Random Fields," *Journal of Econometrics*, 114, 141–164.

Dahl, C. M., G. González-Rivera, and Y. Qin (2005): "Statistical Inference and Prediction in Nonlinear Models using Additive Random Fields," Working Paper, Department of Economics, Purdue University.

Dahl, C. M., and S. Hylleberg (2004): "Flexible Regression Models and Relative Forecast Performance," *International Journal of Forecasting*, 20, 201–217.

Dahl, C. M., and Y. Qin (2004): "The Limiting Behaviour of the Estimated Parameters in a Misspecified Random Field Regression Model," Working Paper, Department of Economics, Purdue University.

Davidon, W. C. (1968): "Variance Algorithms for Minimization," *The Computer Journal*, 10, 406–410.

Davidson, J. E. H. (2002): "A Model of Fractional Cointegration and Tests For Cointegration Using the Bootstrap," *Journal of Econometrics*, 110, 187–212.

Davidson, J. E. H., D. F. Hendry, F. Srba, and S. Yeo (1978): "Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers Expenditure and Income in the United Kingdom," *Economic Journal*, 88, 661–692.

Davidson, R., and J. G. MacKinnon (1993): *Estimation and Inference in Econometrics*. Oxford University Press, Oxford.

——— (2004): *Econometric Theory and Methods*. Oxford University Press, Oxford.

Davies, R. B. (1977): "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64, 247–254.

——— (1987): "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74, 33–43.

DeGroot, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York.

DeJong, D. N., and C. H. Whiteman (1993): "Estimating Moving Average Parameters: Classical Pileups and Bayesian Posteriors," *Journal of Business and Economic Statistics*, 11, 311–317.

Dennis, Jr., J. E., and S. B. Schnabel (1983): *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewood Cliffs, New Jersey.

Dickey, D. A., and W. A. Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

——— (1981): "Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

Dickey, D. A., D. P. Hasza, and W. A. Fuller (1984): "Testing for Unit Roots in Seasonal Time Series," *Journal of the American Statistical Association*, 79, 355–367.

Diebold, F. X., and A. Inoue (2001): "Long Memory and Regime Switching," *Journal of Econometrics*, 105, 131–159.

Diebold, F. X., and G. D. Rudebusch (1989): "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics*, 24, 189–209.

Diebold, F. X., and G. D. Rudebusch (1991): "On the Power of the Dickey-Fuller Test Against Fractional Alternatives," *Economics Letters*, 35, 155–160.

Ding, Z., C. W. J. Granger, and R. F. Engle (1993): "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, 83–106.

Dittmann, I. (2004): "Error Correction Models for Fractionally Cointegrated Time Series," *Journal of Time Series Analysis*, 25, 27–32.

Dolado, J. J., J. Gonzalo, and L. Mayoral (2002): "A Fractional Dickey-Fuller Test for Unit Roots," *Econometrica*, 70, 1963–2006.

——— (2005a): "Testing $I(1)$ Against $I(d)$ Alternatives in the Presence of Deterministic Components," Working Paper, Universitat Pompeu Fabra.

——— (2005b): "What is What? A Simple Test of Long-Memory versus Structural Breaks in the Time Domain," Working Paper, Universitat Pompeu Fabra.

Dolado, J. J., T. Jenkinson, and S. Sosvilla-Rivero (1990): "Cointegration and Unit Roots," *Journal of Economic Surveys*, 4, 249–273.

Doornik, J., and M. Ooms (1999): "A Package for Estimating, Forecasting and Simulating Arfima Models: Arfima Package 1.0 for Ox," Discussion Paper, Nuffield College, Oxford.

Dornbusch, R. (1976): "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84, 1161–1174.

Durbin, J., and G. S. Watson (1950): "Testing for Serial Correlation in Least Squares Regression I," *Biometrika*, 37, 409–428.

Elliot, G., T. Rothenberg, and J. H. Stock (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Enders, W. (1995): *Applied Econometric Time Series.* Wiley, New York.

Enders, W., and C. W. J. Granger (1998): "Unit Root Tests and Asymmetric Adjustment," *Journal of Business and Economic Statistics*, 16, 304–311.

Engel, C. (1996): "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance*, 3, 123–192.

Engle, R. F., and C. W. J. Granger (1987): "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251–276.

Engle, R. F., and B. S. Yoo (1987): "Forecasting and Testing in Cointegrated Systems," *Journal of Econometrics*, 35, 143–159.

Ericsson, N. R., and J. G. MacKinnon (2002): "Distributions of Error Correction Tests for Cointegration," *Econometrics Journal*, 5, 285–318.

Everitt, B. S. (2002): *The Cambridge Dictionary of Statistics.* Cambridge University Press, Cambridge.

Fair, R. C. (1987): "International Evidence on the Demand for Money," *Review of Economics and Statistics*, 69, 473–480.

Fama, E. (1984): "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319–338.

Fiess, N., and R. McDonald (2001): "The Instability of the Money Demand Function: an $I(2)$ Interpretation," *Oxford Bulletin of Economics and Statistics*, 63, 475–495.

Fletcher, R. (1970): "A New Approach to Variable Metric Algorithms," *The Computer Journal*, 13, 317–322.

Fletcher, R., and M. J. D. Powell (1963): "A Rapidly Convergent Descent Method for Minimization," *The Computer Journal*, 6, 163–168.

Fox, R., and M. S. Taqqu (1986): "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series," *The Annals of Statistics*, 14, 517–532.

Franses, P. H. (1998): *Time Series Models for Business and Economic Forecasting.* Cambridge University Press, Cambridge.

Franses, P. H., and B. Hobijn (1997): "Critical Values for Unit Root Tests in Seasonal Time Series," *Journal of Applied Statistics*, 24, 25–46.

Frenkel, J. A. (1976): "A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence," *Scandinavian Journal of Economics*, 78, 200–224.

Fuller, W. A. (1976): *Introduction to Statistical Time Series.* Wiley, New York.

Gallant, A. R., and D. W. Nychka (1987): "Semi-Parametric Maximum Likelihood Estimation," *Econometrica*, 55, 363–390.

Geweke, J., and S. Porter-Hudak (1983): "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

Ghysels, E., and D. R. Osborn (2001): *The Econometric Analysis of Seasonal Time Series.* Cambridge University Press, Cambridge.

Gil-Alana, L. A. (2002): "Empirical Evidence of the Spot and the Forward Exchange Rates in Canada," *Economics Letters*, 77, 405–409.

——— (2003): "Testing of Fractional Cointegration in Macroeconomic Time Series," *Oxford Bulletin of Economics and Statistics*, 65, 517–528.

——— (2004): "A Joint Test of Fractional Integration and Structural Breaks at a Known Period of Time," *Journal of Time Series Analysis*, 25, 691–700.

Gil-Alana, L. A., and P. M. Robinson (1997): "Testing of Unit Root and Other Nonstationary Hypotheses in Macroeconomic Time Series," *Journal of Econometrics*, 80, 241–268.

——— (2001): "Testing of Seasonal Fractional Integration in UK and Japanese Consumption and Income," *Journal of Applied Econometrics*, 16, 95–114.

Gilbert, C. (1986): "Professor Hendry's Econometric Methodology," *Oxford Bulletin of Economics and Statistics*, 48, 283–307.

Goffe, W. L. (1996): "Simann: A Global Optimization Algorithm Using Simulated Annealing," *Studies in Nonlinear Dynamics and Econometrics*, 1, Algorithm 1.

Goldfarb, D. (1970): "A Family of Variable-Metric Methods Derived by Variational Means," *Mathematics of Computation*, 24, 23–26.

Goldfeld, S. M. (1973): "The Demand for Money Revisited," *Brookings Papers on Economic Activity*, 3, 577–638.

——— (1976): "The Case of the Missing Money," *Brookings Papers on Economic Activity*, 3, 683–730.

Goldfeld, S. M., and D. E. Sichel (1990): "The Demand for Money," in *Handbook for Monetary Economics*, ed. by B. M. Friedman, and F. H. Hahn, vol. 1. North Holland, Amsterdam.

Gonzalo, J., and J. Y. Pitarakis (1999): "Dimensionality Effect in Cointegration Analysis," in *Cointegration, Causality and Forecasting: Festschrift in Honour of Clive Granger*, ed. by R. F. Engle, and H. White. Oxford University Press, Oxford.

Granger, C. W. J. (1980): "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227–238.

Granger, C. W. J., and R. Joyeux (1980): "An Introduction to Long-Memory Time Series," *Journal of Time Series Analysis*, 1, 15–30.

Granger, C. W. J., and P. Newbold (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 111–120.

Granger, C. W. J., and T. Teräsvirta (1993): *Modelling Nonlinear Economic Relationships.* Oxford University Press, Oxford.

Granger, C. W. J., T. Teräsvirta, and H. M. Anderson (1993): "Modelling Nonlinearity over the Business Cycle," in *Business Cycles, Indicators, and Forecasting*, ed. by J. H. Stock, and M. W. Watson. The University of Chicago Press, Chicago.

Greene, W. (2003): *Econometric Analysis.* Prentice Hall International, New Jersey, 5^{th} edn.

Guckenheimer, J., and P. Holmes (1997): *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag, New York.

Hall, A. (1989): "Testing for a Unit Root in the Presence of Moving Average Errors," *Biometrika*, 76, 49–56.

Hamilton, J. D. (1994): *Time Series Analysis.* Princeton University Press, Princeton, New Jersey.

——— (2001): "A Parametric Approach to Flexible Nonlinear Inference," *Econometrica*, 69, 537–573.

——— (2003): "What is an Oil Shock?," *Journal of Econometrics*, 113, 363–398.

——— (2005): "Comment on 'Investigating Nonlinearity'," *Studies in Nonlinear Dynamics and Econometrics*, 9, Article 3.

HANNAN, E. J., AND B. G. QUINN (1979): "The Determination of the Order of an Autoregression," *Journal of the Royal Statistical Society, Series B*, 41, 190–195.

HANSEN, B. E. (1996): "Inference When a Nuisance Parameter is not Identified under the Null Hypothesis," *Econometrica*, 64, 413–430.

HANSEN, G., J.-R. KIM, AND S. MITTNIK (1998): "Testing Cointegrating Coefficients in Vector Autoregressive Error Correction Models," *Economics Letters*, 58, 1–5.

HANSEN, L. P., AND R. J. HODRICK (1980): "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829–853.

HARRISON, M. J., AND D. BOND (1992): "Testing and Estimation in Unstable Dynamic Models: A Case Study," *The Economic and Social Review*, 24, 25–49.

HARRISON, M. J., AND G. KEOGH (1985): "A Lagrange Multiplier Interpretation of Disturbance Estimators with an Application to Testing for Nonlinearity," *Journal of Econometrics*, 27, 259–269.

HARRISON, M. J., AND E. J. O'BRIEN (2007): "Testing for Nonlinearity: A Note on the Power of Random Field LM-Type Approaches," Trinity Economic Papers, forthcoming.

HARVEY, A. C. (1981): *Time Series Models*. John Wiley, New York.

——— (1990): *The Econometric Analysis of Time Series*. Philip Allan, New York, 2^{nd} edn.

HARVEY, A. C., AND P. COLLIER (1977): "Testing for Functional Misspecification in Regression Analysis," *Journal of Econometrics*, 6, 103–119.

HASSLER, U., AND J. WOLTERS (1994): "On the Power of Unit Root Tests Against Fractional Alternatives," *Economics Letters*, 45, 1–5.

HAUG, A. A. (1996): "Tests for Cointegration: A Monte Carlo Comparison," *Journal of Econometrics*, 71, 89–115.

HAUSER, M. A., B. M. POTSCHER, AND E. RESCHENHOFER (1999): "Measuring Persistence in Aggregate Output: ARMA Models, Fractionally Integrated ARMA Models and Nonparametric Procedures," *Empirical Economics*, 24, 243–269.

HENDRY, D. F., AND G. J. ANDERSON (1977): "Testing Dynamic Specification in Small Simultaneous Systems: An Application to a Model of Building Society Behaviour in the United Kingdom," in *Frontiers in Quantitative Economics*, ed. by M. D. Intriligator. North Holland, Amsterdam.

HENDRY, D. F., AND G. E. MIZON (1978): "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Study of the Demand for Money by the Bank of England," *The Economic Journal*, 88, 549–563.

HERAVI, S., AND K. PATTERSON (2005): "Optimal and Adaptive Semi-Parametric Narrowband and Broadband and Maximum Likelihood Estimation of the Long-Memory Parameter for Real Exchange Rates," *The Manchester School*, 73, 165–213.

HODRICK, R. J. (1987): *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets.* Harwood, London.

HONG, S. H., AND P. C. B. PHILLIPS (2005): "Testing Linearity in Cointegrating Relations with an Application to Purchasing Power Parity," Discussion Paper No. 1541, Cowles Foundation, Yale University.

HOSKING, J. R. M. (1981): "Fractional Differencing," *Biometrika*, 68, 165–176.

HSU, C. C. (2001): "Change Point Estimation in Regressions with $I(d)$ Variables," *Economics Letters*, 70, 147–155.

HUALDE, J., AND P. M. ROBINSON (2002): "Root-n-Consistent Estimation of Weak Fractional Cointegration," Working Paper No. 08/02, Facultad de Ciencias Econmicas y Empresariales Universidad de Navarra.

HYLLEBERG, S., R. F. ENGLE, C. W. J. GRANGER, AND B. S. YOO (1990): "Seasonal Integration and Cointegration," *Journal of Econometrics*, 99, 215–238.

JEFFREYS, H., AND B. S. JEFFREYS (1988): *Methods of Mathematical Physics.* Cambridge University Press, Cambridge.

JENSEN, J. L., AND A. T. A. WOOD (1997): "On the Non-existence of a Bartlett Correction for Unit Root Tests," *Statistics and Probability Letters*, 35, 181–187.

JENSEN, M. J. (1999): "Using Wavelets to Obtain a Consistent Ordinary Least Squares Estimator of the Long-Memory Parameter," *Journal of Forecasting*, 18, 17–32.

JERRELL, M. E., AND W. A. CAMPIONE (2001): "Global Optimization of Econometric Functions," *Journal of Global Optimization*, 20, 273–295.

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 Autoregressive Models," *Econometrica*, 59, 1551–1580.

——— (1996): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, Oxford, 2^{nd} edn.

——— (2002): "A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model," *Econometrica*, 70, 1929–1961.

JOHANSEN, S., H. HANSEN, AND S. FACHIN (2002): "A Simulation Study of Some Functionals of Random Walk," Department of Statistics and Operations Research, University of Copenhagen.

JOHANSEN, S., AND K. JUSELIUS (1990): "Maximum Likelihood Estimation and Inference on Cointegration - With Applications to Demand for Money," *Oxford Bulletin of Economics and Statistics*, 52, 169–210.

KAC, M. (1969): "Some Mathematical Models in Science," *Science*, 166, 695–699.

KEENAN, D. M. (1985): "A Tukey Non-Additivity-Type Test for Time Series Nonlinearity," *Biometrika*, 72, 39–44.

KENNEDY, P. (2003): *A Guide to Econometrics.* Blackwell, Oxford, 5^{th} edn.

Kenny, G., and D. McGettigan (1999): "Modelling Traded, Non-traded and Aggregate Inflation in a Small Open Economy: The Case of Ireland," *The Manchester School*, 67, 60–88.

Kim, C. S., and P. C. B. Phillips (2000): "Fully Modified Estimation of Fractional Cointegration Models," preprint, Yale University, New Haven.

Kim, D. H., D. R. Osborn, and M. Sensier (2005): "Nonlinearity in the Fed's Monetary Policy Rule," *Journal of Applied Econometrics*, 20, 621–639.

Koopman, S. J., and M. Ooms (2003): "Times Series Modelling of Daily Tax Revenues," *Statistica Neerlandica*, 57, 439–469.

Kopka, H., and P. W. Daly (2004): *Guide to LaTeX*. Addison-Wesley, Boston, 4^{th} edn.

Krämmer, W., and P. Sibbertsen (2002): "Testing for Structural Changes in the Presence of Long Memory," *International Journal of Business and Economics*, 1, 235–242.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178.

Lamport, L. (1994): *LaTeX: A Document Preparation System. User's Guide and Reference Manual.* Addison-Wesley, Boston, 2^{nd} edn.

Lawley, D. N. (1956): "A General Method for Approximating the Distribution of Likelihood Ratio Criteria," *Biometrika*, 43, 296–303.

Lee, T.-H., H. White, and C. W. J. Granger (1993): "Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests," *Journal of Econometrics*, 56, 269–290.

Lee, Y. S., T.-H. Kim, and P. Newbold (2005): "Spurious Nonlinear Regression in Econometrics," *Economics Letters*, 87, 301–306.

Lewis, K. K. (1995): "Puzzles in International Financial Markets," in *The Handbook of International Economics*, ed. by G. M. Grossman, and K. Rogoff, vol. 3. Elsevier North Holland, Amsterdam.

Leybourne, S. J., and B. P. M. McCabe (1994): "A Consistent Test for a Unit Root," *Journal of Business and Economic Statistics*, 12, 157–166.

——— (1999): "Modified Stationarity Tests with Data-Dependent Model-Selection Rules," *Journal of Business and Economic Statistics*, 17, 261–286.

Lien, D., and Y. K. Tse (1999): "Forecasting the Nikkei Spot Index with Fractional Cointegration," *Journal of Forecasting*, 18, 259–273.

Litterman, R. B. (1986): "Forecasting with Bayesian Vector Autoregressions - Five Years of Experience," *Journal of Business and Economic Statistics*, 4, 25–38.

Liu, M. (2000): "Modelling Long Memory in Stock Market Volatility," *Journal of Econometrics*, 99, 139–171.

Liu, S., and C. Chou (2003): "Parities and Spread Trading in Gold and Silver Markets: A Fractional Cointegration Analysis," *Applied Financial Economics*, 13, 899–911.

Lütkepohl, H., and M. Krätzig (eds.) (2004): *Applied Time Series Econometrics*. Cambridge University Press, Cambridge.

Lütkepohl, H., T. Teräsvirta, and J. Wolters (1999): "Investigating Stability and Linearity of a German $M1$ Money Demand Function," *Journal of Applied Econometrics*, 14, 511–525.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

MacDonald, R., and M. P. Taylor (1992): "Exchange Rate Economics: A Survey," *International Monetary Fund Staff Papers*, 39, 1–57.

MacKinnon, J. G. (1996): "Numerical Distribution Functions for Unit Root and Cointegration Tests," *Journal of Applied Econometrics*, 11, 601–618.

MacKinnon, J. G., A. A. Haug, and L. Michelis (1999): "Numerical Distribution Functions of Likelihood Ratio Tests For Cointegration," *Journal of Applied Econometrics*, 14, 563–577.

Maddala, G. S., and F. D. Nelson (1974): "Maximum Likelihood Methods for Models of Markets in Disequillibrium," *Econometrica*, 42, 1013–1030.

Mardia, K. V., and A. J. Watkins (1989): "On Multimodality of the Likelihood in the Spatial Linear Model," *Biometrika*, 76, 289–295.

Mark, N. C., and D. Sul (2003): "Cointegration Vector Estimation by Panel DOLS and Long-Run Money Demand," *Oxford Bulletin of Economics and Statistics*, 65, 655–680.

Martin, G. M. (2001): "Bayesian Analysis of a Fractional Cointegration Model," *Econometric Reviews*, 20, 217–234.

Masih, A. M. M., and R. Masih (2004): "Fractional Cointegration, Low Frequency Dynamics and Long-Run Purchasing Power Parity: an Analysis of the Australian Dollar Over its Recent Float," *Applied Economics*, 36, 593–605.

Maynard, A., and P. C. B. Phillips (2001): "Rethinking an Old Empirical Puzzle: Econometric Evidence on the Forward Discount Anomaly," *Journal of Applied Econometrics*, 16, 671–708.

Mayoral, L. (2003): "A New Minimum Distance Estimator for Arfima Processes," Working Paper, Universitat Pompeu Fabra.

——— (2005): "Is the Observed Persistence Spurious or Real? A Test for Fractional Integration versus Short Memory and Structural Breaks," Working Paper, Universitat Pompeu Fabra.

McLeod, A. I., and W. K. Li (1983): "Diagnostic Checking Arma Time Series Models Using Squared-Residual Autocorrelations," *Journal of Time Series Analysis*, 4, 269–273.

Mittelbach, F., M. Goossens, J. Braams, D. Carlisle, and C. Rowley (2004): *The LaTeX Companion*. Addison-Wesley, Boston, 2^{nd} edn.

Mooney, C. Z. (1997): *Monte Carlo Simulation*. Sage University Paper Series on Quantitative Applications in the Social Sciences, Thousand Oak, California.

Mulligan, C. B., and X. Sala-i Martin (1996): "Financial Technologies: Implications for Money Demand and Monetary Policy," Working Paper No. 5504, National Bureau of Economic Research.

Murray, W. (ed.) (1972): *Numerical Methods for Unconstrained Optimization.* Academic Press, London.

Nankervis, J. C., and N. E. Savin (1985): "Testing the Autoregressive Parameter with the t Statistic," *Journal of Econometrics*, 27, 143–161.

Nelson, C. R., and C. I. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139–162.

Newey, W., and K. West (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

Ng, S., and P. Perron (2001): "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69, 1519–1554.

Nunes, L. C., C. M. Kuan, and P. Newbold (1995): "Spurious Breaks," *Econometric Theory*, 11, 736–749.

Osborn, D. R., A. P. L. Chui, J. P. Smith, and C. R. Birchenhall (1988): "Seasonality and the Order of Integration for Consumption," *Oxford Bulletin of Economics and Statistics*, 50, 361–377.

Osterwald-Lenum, M. (1992): "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Economics and Statistics*, 54, 461–472.

Pantula, S. G. (1989): "Testing for Unit Roots in Time Series Data," *Econometric Theory*, 5, 256–271.

Parke, W. R. (1999): "What is Fractional Integration?," *The Review of Economics and Statistics*, 81, 632–638.

Parzen, E. (1961): "Mathematical Considerations in the Estimation of Spectra," *Technometrics*, 3, 167–190.

Perron, P. (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica*, 57, 1361–1401.

Perron, P., and Z. Qu (2004): "An Analytical Evaluation of the Log-Periodogram Estimate in the Presence of Level Shifts and its Implication for Stock Market Volatility," Working Paper, Department of Economics, Boston University.

Pesaran, M. H., and S. Potter (1997): "A Floor and Ceiling Model of US Output," *Journal of Economic Dynamics and Control*, 21, 661–696.

Pesaran, M. H., Y. Shin, and R. J. Smith (1996): "Structural Analysis of Vector Error Correction Models with Exogenous $I(1)$ Variables," Working Paper No. 9706, Department of Applied Economics, University of Cambridge.

Phillips, A. W. (1958): "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957," *Economica*, 25, 283–299.

Phillips, P. C. B. (1983): "ERA's: A New Approach to Small Sample Theory," *Econometrica*, 51, 1505–1527.

——— (1986): "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, 311–340.

——— (1987): "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

——— (2003): "Law and Limits of Econometrics," *The Economic Journal*, 113, c26–c52.

PHILLIPS, P. C. B., AND S. OULIARIS (1990): "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58, 165–193.

PHILLIPS, P. C. B., AND P. PERRON (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335–346.

PHILLIPS, P. C. B., AND Z. XIAO (1998): "A Primer on Unit Root Testing," in *Practical Issues in Cointegration Analysis*, ed. by M. McAleer, and L. Oxley. Blackwell, Oxford.

POIRIER, D. (1976): *The Econometrics of Structural Change*. North Holland, Amsterdam.

POLAK, E., AND G. RIBIERE (1969): "Note sur la Convergence de Method des Direction Conjugees," Working Paper, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley.

PRIESTLEY, M. B. (1981): *Spectral Analysis of Time Series*. Academic Press, London.

——— (1988): *Nonlinear and Nonstationary Time Series Analysis*. Academic Press, London.

RAMSEY, J. B. (1969): "Tests of Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of the Royal Statistical Society, Series B*, 31, 350–371.

REIMERS, H. E. (1992): "Comparison of Tests for Multivariate Cointegration," *Statistical Papers*, 33, 335–359.

ROBINSON, P. M. (1994): "Semiparametric Analysis of Long-Memory Time Series," *The Annals of Statistics*, 22, 515–539.

ROBINSON, P. M., AND J. HUALDE (2003): "Cointegration in Fractional Systems with Unknown Integration Orders," *Econometrica*, 71, 1727–1766.

ROBINSON, P. M., AND F. IACONE (2005): "Cointegration in Fractional Systems with Deterministic Trends," *Journal of Econometrics*, 127, 263–298.

RODRIGUES, P. M. M. (2001): "Near Seasonal Integration," *Econometric Theory*, 17, 70–86.

RODRIGUES, P. M. M., AND D. R. OSBORN (1999): "Asymptotic Confidence Intervals and Seasonal Unit Root Test Statistics," Discussion Paper No. 9916, School of Economic Studies, University of Manchester.

ROGOFF, K. (1996): "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 34, 647–668.

ROLL, R. (1979): "Violations of Purchasing Power Parity and Their Implications for Efficient Commodity Markets," in *International Finance and Trade*, ed. by M. Sarnat, and G. Szego. Ballinger, Cambridge, Massachusetts.

SARGAN, J. D. (1964): "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in *Econometric Analysis for National Economic Planning*, ed. by P. Hart, G. Mills, and J. Whitaker. Butterworth, London.

——— (1975): "A Suggested Technique for Computing Approximations to Wald Criteria with Applications to Testing Dynamic Specifications," Discussion Paper A2, London School of Economics.

Sargan, J. D., and A. S. Bhargava (1983): "Testing Residuals for Least Squares Regression for Being Generated by the Gaussian Random Walk," *Econometrica*, 52, 213–248.

Sarno, L. (2005): "Towards a Solution to the Puzzles in Exchange Rate Economics: Where Do We Stand?," *Canadian Journal of Economics*, 35, 673–708.

Sarno, L., and M. P. Taylor (2002): "Purchasing Power Parity and the Real Exchange Rate," *International Monetary Fund Staff Papers*, 49, 65–105.

——— (2003): *The Economics of Exchange Rates*. Cambridge University Press, Cambridge.

Sarno, L., G. Valente, and L. Hyginus (2006): "Nonlinearity in Deviations from Uncovered Interest Parity: An Explanation of the Forward Bias Puzzle," *Review of Finance*, forthcoming.

Schoenberg, R. (2001): *Optimization:* Gauss *Reference Manual*. Aptech Systems, Inc., Maple Valley, Washington.

Schwarz, G. (1978): "Estimating the Dimensions of a Model," *The Annals of Statistics*, 6, 461–464.

Shanno, D. F. (1970): "Conditioning of Quasi-Newton Methods for Function Minimization," *Mathematics of Computation*, 24, 647–656.

Sims, C. A. (1993): "A Nine-Variable Probabilistic Macroeconomic Forecasting Model," in *Business Cycles, Indicators, and Forecasting*, ed. by J. H. Stock, and M. W. Watson. Chicago University Press, Chicago.

Smith, A. A., F. Sowell, and S. E. Zin (1997): "Fractional Integration with Drift: Estimation in Small Samples," *Empirical Economics*, 22, 103–116.

Sowell, F. (1992): "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics*, 53, 165–188.

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

Stewart, J., and L. Gill (1998): *Econometrics*. Prentice Hall Europe, London, 2^{nd} edn.

Stock, J. H. (1987): "Asymptotic Properties of Least-Squares Estimators of Cointegrating Vectors," *Econometrics*, 55, 1035–1056.

Tanaka, K. (1999): "The Nonstationary Fractional Unit Root," *Econometric Theory*, 15, 549–582.

Taylor, A. M. (2002): "A Century of Purchasing Power Parity," *The Review of Economics and Statistics*, 84, 139–150.

Taylor, A. M., and M. P. Taylor (2004): "The Purchasing Power Parity Debate," *Journal of Economic Perspectives*, 18, 135–158.

Taylor, M. P., D. A. Peel, and L. Sarno (2002): "Nonlinear Mean-Reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzle," *International Economic Review*, 42, 1015–1042.

Teräsvirta, T. (1994): "Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89, 208–218.

——— (2004): "Smooth Transition Regression Modelling," in *Applied Time Series Econometrics*, ed. by H. Lütkepohl, and M. Krätzig. Cambridge University Press, Cambridge.

TERÄSVIRTA, T., C. F. LIN, AND C. W. J. GRANGER (1993): "Power of the Neural Network Linearity Test," *Journal of Time Series Analysis*, 14, 209–220.

TEVEROSKY, V., AND M. TAQQU (1997): "Testing for Long-Range Dependence in the Presence of Shifting Means or a Slowly Declining Trend, Using a Variance-Type Estimator," *Journal of Time Series Analysis*, 18, 279–304.

THOM, R. (1989): "Real Exchange Rates, Cointegration and Purchasing Power Parity: Irish Experience in the EMS," *The Economic and Social Review*, 20, 147–163.

TODA, H. Y. (1995): "Finite Sample Performance of Likelihood Ratio Tests for Cointegrating Rank in Vector Autoregressions," *Econometric Theory*, 11, 1015–1032.

TONG, H. (1978): "On a Threshold Model," in *Pattern Recognition and Signal Processing*, ed. by C. H. Chen. Sijhoff and Noordhoff, Amsterdam.

——— (1990): *Nonlinear Time Series*. Oxford University Press, Oxford.

TONG, H., AND K. S. LIM (1980): "Threshold Autoregression, Limit Cycles and Cyclical Data," *Journal of the Royal Statistical Society, Series B*, 42, 245–292.

TSAY, R. S. (1986): "Nonlinearity Tests for Time Series," *Biometrika*, 73, 461–466.

——— (1989): "Testing and Modelling Threshold Autoregressive Processes," *Journal of the American Statistical Association*, 84, 231–240.

——— (2002): *Analysis of Financial Time Series*. John Wiley and Sons, New York.

TSAY, W.-J., AND C.-F. CHUNG (2000): "The Spurious Regression of Fractionally Integrated Processes," *Journal of Econometrics*, 96, 155–182.

TUCCI, M. P. (2002): "A Note on Global Optimization in Adaptive Control, Econometrics and Macroeconomics," *Journal of Economic Dynamics and Control*, 26, 1739–1764.

VELASCO, C. (2003): "Gaussian Semi-parametric Estimation of Fractional Cointegration," *Journal of Time Series Analysis*, 24, 345–378.

WAHBA, G. (1978): "Improper Priors, Spline Smoothing and the Problem of Guarding Against Model Errors in Regression," *Journal of the Royal Statistical Society, Series B*, 40, 364–372.

WARNES, J. J., AND B. D. RIPLEY (1987): "Problems with Likelihood Estimation of Covariance Functions of Spatial Gaussian Processes," *Biometrika*, 74, 640–642.

WECKER, W. E., AND C. F. ANSLEY (1983): "The Signal Extraction Approach to Nonlinear Regression and Spline Smoothing," *Journal of the American Statistical Association*, 78, 81–89.

WHITE, H. (1989): "An Additional Hidden Unit Test for Neglected Nonlinearity in Multilayer Feedforward Networks," *Proceedings of the International Joint Conference on Neural Networks, Washington DC*, 2, 451–455.

WRIGHT, J. H. (1994): "A Cointegration Based Analysis of Irish Purchasing Power Parity Relationships using the Johansen Procedure," *The Economic and Social Review*, 25, 261–78.

ZHOU, S. (2000): "Testing Structural Hypotheses on Cointegration Relations with Small Samples," *Economic Inquiry*, 38, 629–640.

ZIVOT, E. (2000): "Cointegration and Forward and Spot Exchange Rate Regressions," *Journal of International Money and Finance*, 19, 785–812.

Appendix A

Appendix to Chapter 2

A.1 Data for Hamilton's Examples

Table A.1: Data for Hamilton's (2001) Example 3.

	u_t	π_t		u_t	π_t		u_t	π_t
1948	3.8	3.0	1966	3.8	3.5	1984	7.5	3.9
1949	5.9	-2.1	1967	3.8	3.0	1985	7.2	3.8
1950	5.3	5.9	1968	3.6	4.7	1986	7.0	1.1
1951	3.3	6.0	1969	3.5	6.2	1987	6.2	4.4
1952	3.0	0.8	1970	4.9	5.6	1988	5.5	4.4
1953	2.9	0.7	1971	5.9	3.3	1989	5.3	4.6
1954	5.5	-0.7	1972	5.6	3.4	1990	5.5	6.1
1955	4.4	0.4	1973	4.9	8.7	1991	6.7	3.1
1956	4.1	3.0	1974	5.6	12.3	1992	7.4	2.9
1957	4.3	2.9	1975	8.5	6.9	1993	6.8	2.7
1958	6.8	1.8	1976	7.7	4.9	1994	6.1	2.7
1959	5.5	1.7	1977	7.1	6.7	1995	5.6	2.5
1960	5.5	1.4	1978	6.1	9.0	1996	5.4	3.3
1961	6.7	0.7	1979	5.8	13.3	1997	4.9	1.7
1962	5.5	1.3	1980	7.1	12.5	1998	4.5	1.6
1963	5.7	1.6	1981	7.6	8.9	1999	4.2	2.7
1964	5.2	1.0	1982	9.7	3.8	2000	4.0	3.4
1965	4.5	1.9	1983	9.6	3.8	2001	4.8	1.6

Source: `http://www.bls.gov/`.

Table A.2: Data for Hamilton's (2001) Example 1.

	x_1	x_2		x_1	x_2		x_1	x_2
1	-1.496	6.446	14	2.330	-30.162	27	21.392	-13.275
2	10.780	-3.349	15	7.583	15.786	28	-10.061	-1.464
3	1.292	-13.326	16	-6.042	2.789	29	-4.808	3.554
4	-13.721	1.057	17	6.664	4.697	30	-2.069	-7.897
5	-1.710	-5.035	18	7.061	-12.153	31	-1.749	-8.883
6	21.174	3.378	19	4.436	20.871	32	12.357	-17.726
7	9.814	-1.193	20	2.684	-11.124	33	-0.082	4.900
8	0.062	-6.746	21	13.161	-6.302	34	3.163	-1.763
9	0.997	-5.987	22	3.042	14.675	35	-0.055	1.623
10	3.393	11.574	23	-4.022	26.122	36	0.798	-8.917
11	11.553	7.599	24	10.679	5.702	37	4.418	11.649
12	18.400	12.391	25	11.823	14.381	38	11.070	-24.143
13	-16.851	9.096	26	-0.234	5.186	39	12.252	8.078

Continued on next page.

	x_1	x_2		x_1	x_2		x_1	x_2
40	0.733	2.749	61	6.452	6.521	82	6.195	-13.680
41	-5.691	-20.963	62	4.305	17.726	83	-7.320	20.653
42	-9.451	-20.542	63	8.908	-0.187	84	-0.712	5.740
43	-1.506	-5.739	64	10.301	-4.572	85	6.259	6.612
44	-10.855	8.069	65	16.945	0.629	86	-6.585	-4.628
45	-9.493	-11.404	66	1.942	5.272	87	-20.608	17.380
46	20.376	10.213	67	0.580	16.792	88	-8.588	1.660
47	0.179	6.106	68	0.901	10.418	89	-0.247	4.028
48	-2.044	-4.997	69	-8.067	10.618	90	3.399	-11.128
49	2.039	-5.677	70	3.531	-21.965	91	-2.327	4.502
50	-4.391	-16.671	71	1.484	-9.801	92	8.782	6.354
51	-14.555	-3.992	72	16.703	-6.006	93	-0.311	23.368
52	-8.492	-14.156	73	-3.149	-9.161	94	3.547	-25.837
53	3.047	9.015	74	2.736	6.730	95	0.683	-1.138
54	-7.127	-3.087	75	13.615	-12.554	96	-18.421	-16.145
55	-9.599	10.729	76	-18.257	3.661	97	-7.658	-0.321
56	-23.986	3.794	77	1.114	-0.135	98	-6.609	-5.254
57	7.778	-2.436	78	16.268	4.419	99	0.666	-9.137
58	0.897	7.265	79	-16.130	4.878	100	1.479	5.199
59	-2.747	15.277	80	23.938	10.156			
60	0.239	-3.764	81	-16.481	7.123			

Table A.3: Data for Hamilton's (2001) Example 2.

	x_1	x_2	x_3		x_1	x_2	x_3
1	-6.696	1.659	0.101	17	3.189	3.624	-0.008
2	-0.772	3.193	0.516	18	-0.352	2.480	2.949
3	-0.502	1.476	-0.162	19	-3.397	-1.941	-2.709
4	0.472	-4.107	1.235	20	-2.175	-1.574	-1.328
5	-0.441	-2.305	1.069	21	1.055	-1.457	-1.217
6	0.293	1.869	-0.058	22	3.109	-0.051	-3.784
7	1.235	-0.483	3.499	23	5.398	-1.166	2.970
8	-0.039	1.704	0.206	24	-2.545	1.781	-4.032
9	3.982	1.524	3.212	25	0.647	-0.744	0.625
10	0.302	0.524	-1.084	26	0.877	2.354	0.478
11	0.208	0.492	-0.551	27	4.210	-2.416	2.432
12	2.830	2.108	2.106	28	-2.153	1.648	0.681
13	-1.707	-1.318	0.464	29	-0.727	0.228	-0.390
14	3.277	-0.857	-1.732	30	0.682	0.064	0.475
15	-0.565	-0.691	-1.189	31	0.258	3.323	-1.187
16	0.323	3.961	0.201	32	2.608	-0.383	-0.706

Continued on next page.

	x_1	x_2	x_3		x_1	x_2	x_3
33	-1.582	-2.581	0.486	67	2.217	-0.125	-0.377
34	-0.505	-0.849	2.601	68	1.987	0.432	-0.021
35	0.354	0.958	0.533	69	-0.296	-2.564	-0.741
36	1.120	1.070	-1.661	70	2.574	-1.470	-0.235
37	-0.664	-1.062	0.874	71	-2.204	0.531	-0.075
38	-0.982	1.261	1.898	72	0.168	-0.513	0.358
39	-4.550	0.707	1.346	73	-1.085	-3.269	-0.625
40	1.039	-3.511	-1.461	74	1.128	-3.206	-0.973
41	0.709	-2.427	-0.421	75	-1.764	-1.627	-0.503
42	1.431	2.226	2.464	76	-3.498	-1.441	-3.605
43	-0.229	-0.708	2.218	77	0.891	0.285	0.465
44	-3.533	-1.932	-1.955	78	0.355	1.905	-1.792
45	-0.175	-0.176	-1.629	79	-1.481	-0.088	-1.099
46	-1.605	2.082	-0.407	80	-3.269	-0.052	0.047
47	-4.787	0.699	-3.491	81	0.364	1.693	0.325
48	-0.309	-3.005	2.496	82	0.073	3.624	0.350
49	1.564	-4.833	0.582	83	1.476	-0.108	-2.133
50	0.057	-2.830	1.530	84	1.773	0.470	-1.425
51	0.553	-3.132	0.535	85	-0.119	1.059	-1.009
52	-1.117	-0.559	3.703	86	0.029	0.564	-1.821
53	5.332	-0.924	-3.975	87	1.385	2.147	3.289
54	-0.537	0.628	-1.922	88	-0.523	-1.541	0.793
55	0.693	-2.306	-2.304	89	0.985	-3.997	4.155
56	3.949	-2.756	-1.629	90	-5.002	-0.965	0.884
57	-0.580	-2.684	6.955	91	-2.190	-1.004	-2.068
58	1.761	-1.452	1.270	92	0.887	2.413	-2.624
59	3.703	-0.805	0.341	93	-0.309	0.669	-0.615
60	0.391	2.937	-2.207	94	-0.903	-1.911	0.359
61	-0.455	-1.230	5.471	95	-0.407	-0.595	0.936
62	-1.866	2.371	-2.227	96	1.792	-1.011	0.182
63	-2.833	2.302	-1.052	97	-1.222	-3.166	1.078
64	1.568	1.451	-3.363	98	-1.450	3.823	-1.402
65	-0.354	-0.631	-2.324	99	1.026	-1.537	-1.014
66	-1.122	1.511	1.854	100	-0.611	-3.249	-2.840

A.2 Results

Table A.4: **Dataset 1 (Hamilton's US Phillips curve data), GAUSS 3 results.**

ALGORITHM	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: STEEPEST DESCENT	150	Coefficient	0.089	0.127	0.072	-51.903	-0.873	0.406	0.030	1.827	1.431
		Standard error	0.140	0.071	0.016	153.654	0.431	0.253	0.078	1.019	0.355
2: BFGS	28	Coefficient	0.142	-0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.436	0.455	0.232	0.065	1.294	0.441
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: DFP	150	Coefficient	0.143	-0.155	0.136	-88.728	-0.919	0.437	0.049	2.036	1.240
		Standard error	0.168	0.077	0.031	126.952	0.457	0.231	0.064	1.283	0.440
5: NEWTON	17	Coefficient	0.142	0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.214	0.455	0.232	0.065	1.294	0.441
6: PRCG	11	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.5: Dataset 1 (Hamilton's US Phillips curve data), GAUSS 5 results, _oprteps = 0.00001.

ALGORITHM	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: STEEPEST DESCENT	150	Coefficient	0.144	0.157	-0.135	-89.083	-0.912	0.440	0.049	1.988	1.256
		Standard error	0.170	0.078	0.033	126.027	0.456	0.230	0.064	1.228	0.432
2: BFGS	28	Coefficient	0.142	-0.155	0.136	-88.483	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.113	0.455	0.232	0.065	1.294	0.441
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: DFP	150	Coefficient	-0.051	0.131	0.541	-123.141	-0.654	0.418	0.066	7.010	0.365
		Standard error	0.382	0.127	0.202	108.257	0.335	0.336	0.056	10.033	0.502
5: NEWTON	15	Coefficient	0.142	-0.155	-0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.058	0.455	0.232	0.065	1.294	0.441
6: PRCG	150	Coefficient	-2.260	0.930	-0.078	-71.093	-0.429	0.668	0.038	4.562	0.491
		Standard error	1.198	0.473	0.029	63.154	0.288	0.129	0.032	4.782	0.490

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.6: Dataset 1 (Hamilton's US Phillips curve data), GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	_oprteps	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.144	0.157	-0.135	-89.083	-0.912	0.440	0.049	1.988	1.256
			Standard error	0.170	0.078	0.033	126.027	0.456	0.230	0.064	1.228	0.432
2: Bfgs	0.00001	28	Coefficient	0.142	-0.155	0.136	-88.483	-0.922	0.436	0.049	2.047	1.237
			Standard Error	0.167	0.078	0.032	127.113	0.455	0.232	0.065	1.294	0.441
	0.001	29	Coefficient	0.142	-0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	0.441
			Standard error	0.167	0.078	0.032	127.193	0.455	0.232	0.065	1.294	0.441
	0.1	47	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	1.0	20	Coefficient	0.142	-0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
			Standard error	0.167	0.078	0.032	127.353	0.455	0.232	0.065	1.294	0.441
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp	0.00001	150	Coefficient	-0.051	0.131	0.541	-123.141	-0.654	0.418	0.066	7.010	0.365
			Standard error	0.382	0.127	0.202	108.257	0.335	0.336	0.056	10.033	0.502
	0.001	150	Coefficient	-0.009	0.175	0.540	-111.722	-0.671	0.489	0.060	8.410	0.302
			Standard error	0.360	0.081	0.134	87.365	0.321	0.199	0.045	15.513	0.545
	0.1	150	Coefficient	-0.005	0.150	0.559	-114.967	-0.650	0.462	0.061	5.559	0.450
			Standard error	0.359	0.086	0.178	90.850	0.334	0.227	0.047	6.615	0.508
	1.0	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
5: Newton*	0.00001	15	Coefficient	0.142	-0.155	-0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
			Standard error	0.167	0.078	0.032	127.058	0.455	0.232	0.065	1.294	0.441
6: Prcg*	0.00001	150	Coefficient	-2.260	0.930	-0.078	-71.093	-0.429	0.668	0.038	4.562	0.491
			Standard error	1.198	0.473	0.029	63.154	0.288	0.129	0.032	4.782	0.490

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.7: **Effect of initial value of ζ, Dataset 1 (Hamilton's US Phillips curve data) : single algorithm (Steepest Descent), GAUSS 5 results.**

Initial Value of ζ		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.4	Coefficient	0.141	0.155	0.136	-88.126	-0.927	0.434	0.049	2.082	1.225
	Standard error	0.165	0.077	0.032	128.072	0.455	0.233	0.065	1.336	0.447
0.5	Coefficient	0.144	0.157	-0.135	-89.083	-0.912	0.440	0.049	1.988	1.256
	Standard error	0.170	0.078	0.033	126.027	0.456	0.230	0.064	1.228	0.432
0.6	Coefficient	1.681	-1.078	0.081	-72.623	-0.431	0.670	0.039	1.986	1.026
	Standard error	0.886	1.211	0.037	63.130	0.284	0.129	0.032	1.593	0.644
0.7	Coefficient	0.140	0.155	0.136	-87.877	-0.930	0.433	0.049	2.107	1.218
	Standard error	0.164	0.088	0.033	128.815	0.457	0.243	0.065	1.375	0.452
0.8	Coefficient	0.079	0.123	0.071	-45.273	-0.921	0.385	0.027	2.093	1.354
	Standard error	0.125	0.072	0.019	167.746	0.426	0.268	0.085	1.231	0.372
0.9	Coefficient	0.094	0.129	-0.072	-54.224	-0.849	0.416	0.031	1.715	1.467
	Standard error	0.147	0.073	0.017	147.030	0.433	0.248	0.075	0.944	0.350
1.0	Coefficient	0.082	-0.123	0.071	-46.407	-0.907	0.390	0.028	2.026	1.373
	Standard error	0.128	0.071	0.015	164.028	0.429	0.264	0.083	1.169	0.367
1.1	Coefficient	0.095	0.129	-0.072	-54.393	-0.847	0.417	0.031	1.706	1.470
	Standard error	0.149	0.073	0.017	149.786	0.434	0.248	0.076	0.940	0.349
1.2	Coefficient	0.088	0.127	0.072	-51.034	-0.877	0.404	0.030	1.850	1.424
	Standard error	0.136	0.071	0.016	153.303	0.428	0.254	0.078	1.030	0.356
1.3	Coefficient	0.154	0.160	-0.135	-92.380	-0.861	0.460	0.051	1.691	1.361
	Standard error	0.191	0.101	0.034	117.212	0.463	0.227	0.060	0.962	0.399
1.4	Coefficient	0.158	0.162	-0.134	-93.275	-0.843	0.466	0.051	1.609	1.393
	Standard error	0.197	0.102	0.032	114.707	0.465	0.223	0.058	0.902	0.393
1.5	Coefficient	0.156	0.161	-0.135	-92.776	-0.854	0.463	0.051	1.656	1.374
	Standard error	0.193	0.101	0.034	115.900	0.463	0.225	0.059	0.935	0.396

Note: in all cases, `_oprteps` = 0.00001 and number of iterations = 150. Algorithm failed for initial values of ζ = 0.1, 0.2 and 0.3.

Table A.8: Effect of initial value of ζ, Dataset 1 (Hamilton's US Phillips curve data): algorithm switching (Steepest Descent/BFGS), GAUSS 5 results.

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.5	28	Coefficient	0.142	-0.155	0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.298	0.455	0.232	0.065	1.294	0.441
0.7	29	Coefficient	0.142	-0.155	-0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.220	0.455	0.232	0.065	1.294	0.441
0.9	24	Coefficient	0.142	-0.155	0.136	-88.484	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.885	0.455	0.232	0.065	1.294	0.441
1.1	18	Coefficient	0.142	-0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.281	0.455	0.232	0.065	1.294	0.441
1.3	28	Coefficient	0.142	0.155	-0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.283	0.455	0.232	0.065	1.294	0.441
1.5	23	Coefficient	0.142	0.155	-0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.238	0.455	0.232	0.065	1.294	0.441

Note: algorithm failed for initial values of $\zeta = 0.1$ and 0.3. In all cases, `_oprteps` = 0.00001.

Table A.9: **Effect of initial value of ζ, Dataset 1 (Hamilton's US Phillips curve data): algorithm switching (Steepest Descent/Newton), GAUSS 5 results.**

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.5	28	Coefficient	0.142	-0.155	0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.371	0.455	0.232	0.065	1.294	0.441
0.7	25	Coefficient	0.142	-0.155	-0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.198	0.455	0.232	0.065	1.294	0.441
0.9	21	Coefficient	0.142	-0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	126.874	0.455	0.232	0.064	1.294	0.441
1.1	11	Coefficient	0.142	-0.155	0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.673	0.455	0.232	0.065	1.294	0.441
1.3	24	Coefficient	0.142	0.155	-0.136	-88.482	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.275	0.455	0.232	0.065	1.294	0.441
1.5	19	Coefficient	0.142	0.155	-0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.381	0.455	0.232	0.065	1.294	0.441

Note: algorithm failed for initial values of ζ = 0.1 and 0.3. In all cases, `_oprteps` = 0.00001.

Table A.10: Dataset 2, GAUSS 3 results.

ALGORITHM	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: STEEPEST DESCENT	150	Coefficient	-0.082	$1.1E-9$	0.220	-25.236	-0.833	0.239	0.017	3.497	0.803
		Standard error	0.131	0.040	0.030	132.164	0.331	0.176	0.037	1.905	0.332
2: BFGS	54	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: DFP	150	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
5: NEWTON	7	Coefficient	0.500	-1.283	-0.082	-38.335	-0.335	0.660	0.021	1.154	1.468
		Standard error	0.365	0.805	0.037	0.678	0.280	0.135	0.035	0.678	0.451
6: PRCG	150	Coefficient	0.047	$4.0E-4$	0.221	-28.182	-0.868	0.207	0.019	3.744	0.771
		Standard error	0.162	0.033	0.045	140.213	0.311	0.172	0.071	1.914	0.311

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.11: Dataset 5, GAUSS 3 results.

Coefficient Estimates			$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	Iterations										
1: Steepest Descent	150	Coefficient	0.153	-0.156	-0.140	-136.197	-0.984	0.439	0.073	1.978	1.236
		Standard error	0.172	0.078	0.033	126.625	0.439	0.221	0.064	1.259	0.443
2: Bfgs	62	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: Dfp	150	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
5: Newton	11	Coefficient	-0.152	0.155	0.140	-136.081	-0.991	0.436	0.073	2.025	1.221
		Standard error	0.169	0.078	0.032	127.704	0.439	0.223	0.065	1.315	0.451
6: Prcg	150	Coefficient	-0.146	-0.151	0.141	-135.445	-1.029	0.421	0.073	2.329	1.129
		Standard error	0.154	0.076	0.032	134.591	0.444	0.233	0.068	1.794	0.525

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.12: Dataset 7, GAUSS 3 results.

COEFFICIENT ESTIMATES			$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	Iterations										
1: STEEPEST DESCENT	150	Coefficient	1.538	-1.175	-0.083	-103.345	-0.474	0.750	0.054	4.402	0.438
		Standard error	0.583	0.441	0.022	62.183	0.257	0.118	0.032	3.830	0.361
2: BFGS	26	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: DFP	150	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
5: NEWTON	10	Coefficient	0.137	-0.039	-0.128	-201.127	-0.991	0.391	0.106	5.063	0.575
		Standard error	0.122	0.044	0.017	166.789	0.471	0.163	0.085	6.120	0.576
6: PRCG	12	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.13: Dataset 8, GAUSS 3 results.

COEFFICIENT ESTIMATES			$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	Iterations										
1: STEEPEST DESCENT	150	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
2: BFGS	42	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: DFP	150	Coefficient	1.966	0.447	0.115	-60.980	-0.363	0.648	0.033	14.822	0.158
		Standard error	0.708	0.172	0.057	64.908	0.305	0.137	0.033	45.736	0.484
5: NEWTON	10	Coefficient	0.154	0.148	-0.138	-60.245	-0.875	0.453	0.034	2.135	1.202
		Standard error	0.177	0.082	0.039	125.218	0.495	0.236	0.063	1.387	0.452
6: PRCG	150	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.14: Dataset 11, GAUSS 3 results.

Coefficient Estimates			$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	Iterations										
1: Steepest Descent	150	Coefficient	-0.091	0.120	-0.070	-10.413	-0.871	0.411	0.009	1.977	1.348
		Standard error	0.126	0.057	0.013	149.804	0.430	0.244	0.076	0.978	0.316
2: Bfgs	27	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: Dfp	150	Coefficient	1.895	0.386	0.127	-46.959	-0.307	0.639	0.025	7.190	0.316
		Standard error	0.750	0.188	0.070	59.678	0.289	0.137	0.031	10.033	0.432
5: Newton	14	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
6: Prcg	150	Coefficient	-0.144	-0.157	-0.069	-22.715	-0.776	0.439	0.015	2.331	1.203
		Standard error	0.257	0.067	0.015	150.104	0.738	0.260	0.075	1.625	0.399

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.15: Summary (Hamilton's US Phillips curve data), GAUSS 5 results.

Dataset		1		2		3		4		5		6		7		8		9		10		11	
Algorithm	_oprteps																						
1: Steepest Descent	0.00001	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S
	0.001	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S
	0.1	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S
	1.0	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S	150	S
2: Bfgs	0.00001	28	S	45	F	31	S	26	S	118	F	60	F	52	F	57	F	65	F	29	F	27	S
	0.001	29	S	45	F	31	S	26	S	118	F	60	F	50	F	57	F	65	F	29	F	27	S
	0.1	47	F	45	F	31	S	26	S	118	F	60	F	46	F	57	F	65	F	29	F	27	S
	1.0	20	S	45	F	31	S	26	S	118	F	60	F	32	S	57	F	65	F	29	F	27	S
4: Dfp	0.00001	150	S	150	S	74	S	150	F	150	S	150	F	150	F	150	S	150	S	150	S	150	S
	0.001	150	S	150	S	74	S	150	F	150	S	150	F	150	S	150	S	150	S	150	S	150	S
	0.1	150	S	150	S	74	S	150	F	150	S	150	F	150	S	150	S	150	S	150	S	150	S
	1.0	150	S	150	S	74	S	150	F	150	S	150	F	139	S	150	S	150	S	150	S	150	S
5: Newton	0.00001	15	S	23	F	8	S	18	S	36	F	11	S	11	S	150	F	22	S	150	F	10	S
	0.001	15	S	23	F	8	S	18	S	36	F	11	S	9	S	134	S	22	S	39	S	10	S
	0.1	15	S	23	F	8	S	18	S	36	F	11	S	9	S	16	S	22	S	15	S	10	S
	1.0	15	S	23	F	8	S	18	S	36	F	11	S	9	S	13	S	22	S	15	S	10	S
6: Prcg	0.00001	150	S	150	S	150	S	88	F	150	S	150	F	150	F	150	S	150	F	150	S	150	S
	0.001	150	S	150	S	150	S	150	S	150	S	150	F	150	F	150	S	150	F	150	S	150	S
	0.1	150	S	150	S	150	S	150	S	150	S	150	S	150	F	150	S	150	F	150	S	150	S
	1.0	150	S	150	S	150	S	150	F	150	S	150	S	150	F	150	S	150	S	150	S	150	S

Note: algorithm 3 (Bfgs-sc), omitted from the table, failed in all cases.
Key: S - success, F - failure.

Table A.16: Dataset 2, GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.099	$2.0E-10$	0.221	-26.144	-0.806	0.265	0.017	3.011	0.892
			Standard error	0.154	0.044	0.032	130.494	0.357	0.194	0.066	1.654	0.356
2: Bfgs*	0.00001	45	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp*	0.00001	150	Coefficient	0.111	0.001	0.447	-67.915	-0.603	0.257	0.038	113.968	0.022
			Standard error	0.244	0.060	0.093	97.525	0.372	0.180	0.050	2662.126	0.520
5: Newton*	0.00001	23	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
6: Prcg	0.00001	150	Coefficient	-0.043	$5.4E-5$	0.222	-24.321	-0.883	0.174	0.017	6.401	0.491
			Standard error	0.107	0.031	0.027	149.158	0.289	0.143	0.076	4.692	0.325
	0.001	150	Coefficient	-0.042	$7.7E-6$	0.222	-24.257	-0.884	0.173	0.017	6.539	0.481
			Standard error	0.106	0.031	0.027	146.673	0.288	0.142	0.074	4.881	0.326
	0.1	150	Coefficient	-0.041	$6.0E-5$	0.221	-24.113	-0.885	0.172	0.017	6.700	0.471
			Standard error	0.109	0.030	0.044	150.050	0.289	0.145	0.076	5.106	0.327
	1.0	150	Coefficient	-0.032	$2.0E-5$	0.221	-23.576	-0.893	0.153	0.016	11.967	0.274
			Standard error	0.111	0.028	0.036	151.022	0.281	0.137	0.077	15.682	0.347

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.17: Dataset 3, GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.171	-1.1$E-9$	0.143	-119.44	-0.857	0.524	0.064	1.632	1.214
			Standard error	0.187	0.039	0.032	114.294	0.452	0.142	0.058	0.778	0.255
2: Bfgs*	0.00001	31	Coefficient	-0.298	-2.7$E-6$	0.471	-100.439	-0.642	0.543	0.054	2.708	0.752
			Standard error	0.242	0.104	0.079	80.059	0.331	0.231	0.041	3.676	0.791
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp*	0.00001	74	Coefficient	0.298	7.6$E-6$	0.471	-100.441	-0.642	0.543	0.054	2.710	0.751
			Standard error	0.242	0.111	0.079	80.090	0.331	0.231	0.041	3.686	0.792
5: Newton*	0.00001	8	Coefficient	0.299	4.4$E-6$	0.471	-100.438	-0.642	0.543	0.054	2.705	0.752
			Standard error	0.242	0.108	0.079	80.034	0.331	0.230	0.041	3.664	0.789
6: Prcg	0.00001	150	Coefficient	0.119	4.0$E-4$	0.136	-109.930	-0.971	0.470	0.059	2.272	1.070
			Standard error	0.146	0.034	0.035	140.275	0.451	0.177	0.071	1.756	0.385
	0.001	150	Coefficient	0.107	-3.0$E-4$	-0.137	-106.029	-1.000	0.455	0.058	2.481	1.029
			Standard error	0.154	0.033	0.035	149.642	0.471	0.215	0.076	2.607	0.520
	0.1	150	Coefficient	0.107	-6.0$E-4$	0.137	-106.168	-0.100	0.455	0.058	2.473	1.031
			Standard error	0.153	0.033	0.035	148.540	0.468	0.213	0.075	2.560	0.512
	1.0	150	Coefficient	0.115	-2.0$E-4$	0.137	-109.364	-0.979	0.467	0.059	2.302	1.065
			Standard error	0.143	0.034	0.035	141.438	0.446	0.180	0.072	1.825	0.396

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.18: Dataset 4, GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.106	$2.1E{-}8$	-0.132	-191.748	-1.003	0.408	0.101	2.997	0.870
			Standard error	0.120	0.127	0.021	153.395	0.424	0.145	0.078	1.684	0.289
2: Bfgs*	0.00001	26	Coefficient	0.281	$-3.1E{-}6$	0.458	-142.422	-0.614	0.445	0.075	92.029	0.024
			Standard error	0.211	0.103	0.064	89.201	0.349	0.152	0.046	3360.837	0.894
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp*	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
5: Newton*	0.00001	18	Coefficient	0.137	0.039	0.128	-201.126	-0.991	0.391	0.106	5.063	0.575
			Standard error	0.122	0.044	0.017	166.279	0.471	0.163	0.085	6.120	0.576
6: Prcg	0.00001	88	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.001	150	Coefficient	1.544	1.173	-0.087	-102.800	-0.474	0.751	0.054	4.273	0.450
			Standard error	0.622	0.518	0.057	62.306	0.258	0.118	0.032	3.649	0.363
	0.1	150	Coefficient	1.465	1.190	0.084	-103.752	-0.473	0.751	0.054	3.725	0.513
			Standard error	0.598	0.539	0.061	63.237	0.259	0.118	0.032	2.901	0.371
	1.0	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.19: Dataset 5, GAUSS 5 results.

COEFFICIENT ESTIMATES				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: STEEPEST DESCENT*	0.00001	150	Coefficient	0.157	-0.158	-0.140	-136.518	-0.965	0.446	0.073	1.852	1.280
			Standard error	0.180	0.080	0.033	123.255	0.440	0.217	0.063	1.123	0.423
2: BFGS*	0.00001	118	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
3: BFGS-SC*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: DFP*	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
5: NEWTON*	0.00001	36	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
6: PRCG	0.00001	150	Coefficient	-0.146	0.151	-0.142	-135.371	-1.035	0.419	0.073	2.376	1.115
			Standard error	0.153	0.076	0.032	135.720	0.446	0.235	0.069	1.890	0.540
	0.001	150	Coefficient	-0.147	0.152	-0.141	-135.525	-1.023	0.424	0.073	2.271	1.145
			Standard error	0.157	0.076	0.034	133.185	0.443	0.232	0.068	1.684	0.508
	0.1	150	Coefficient	-0.143	0.149	-0.142	-135.172	-1.052	0.412	0.073	2.546	1.070
			Standard error	0.146	0.075	0.026	139.223	0.451	0.242	0.071	2.303	0.605
	1.0	150	Coefficient	-0.152	0.155	-0.140	-136.044	-0.993	0.436	0.073	2.039	1.216
			Standard error	0.168	0.077	0.032	127.990	0.439	0.223	0.065	1.332	0.453

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.20: Dataset 6, GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	_oprteps	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.151	-0.155	0.141	-143.357	-0.987	0.438	0.077	1.874	1.286
			Standard error	0.179	0.081	0.034	129.535	0.441	0.222	0.066	1.203	0.449
2: Bfgs*	0.00001	60	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp*	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
5: Newton*	0.00001	11	Coefficient	0.151	0.154	0.141	-143.253	-0.990	0.437	0.077	1.895	1.279
			Standard error	0.177	0.081	0.034	130.389	0.441	0.223	0.066	1.228	0.453
6: Prcg	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.1	150	Coefficient	-0.151	0.154	-0.141	-143.245	-0.990	0.437	0.077	1.897	1.278
			Standard error	0.177	0.081	0.034	130.254	0.441	0.223	0.066	1.229	0.453
	1.0	150	Coefficient	-0.151	0.154	-0.141	-143.251	-0.990	0.437	0.077	1.895	1.279
			Standard error	0.177	0.081	0.034	130.267	0.441	0.222	0.066	1.228	0.453

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.21: Dataset 7, GAUSS 5 results.

COEFFICIENT ESTIMATES				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: STEEPEST DESCENT*	0.00001	150	Coefficient	0.155	0.150	0.143	-174.889	-1.032	0.418	0.093	1.889	1.283
			Standard error	0.189	0.083	0.034	136.196	0.450	0.225	0.069	1.258	0.465
2: BFGS	0.00001	52	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.001	50	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.1	46	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	1.0	32	Coefficient	0.159	-0.152	0.143	-174.902	-1.013	0.425	0.093	1.775	1.324
			Standard error	0.198	0.084	0.035	132.935	0.450	0.221	0.069	1.123	0.444
3: BFGS-SC*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: DFP	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.001	150	Coefficient	-0.193	-0.094	0.682	-187.177	-0.681	0.413	0.098	8.305	0.291
			Standard error	0.371	0.109	0.238	104.643	0.380	0.233	0.054	15.097	0.518
	0.1	150	Coefficient	-0.116	-0.137	0.565	-192.728	-0.780	0.419	0.101	7.017	0.356
			Standard error	0.366	0.098	0.280	125.823	0.372	0.275	0.065	10.953	0.537
	1.0	139	Coefficient	0.159	-0.152	0.143	-174.902	-1.013	0.425	0.093	1.775	1.324
			Standard error	0.198	0.084	0.035	132.867	0.450	0.221	0.068	1.123	0.443
5: NEWTON	0.00001	11	Coefficient	0.159	0.152	0.143	-174.902	-1.013	0.425	0.093	1.775	1.324
			Standard error	0.198	0.084	0.035	133.131	0.450	0.221	0.068	1.122	0.443
	0.001	9	Coefficient	0.159	0.152	0.143	-174.902	-1.013	0.425	0.093	1.775	1.324
			Standard error	0.198	0.084	0.035	133.131	0.450	0.221	0.068	1.122	0.443
6: PRCG*	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.22: Dataset 8, GAUSS 5 results.

COEFFICIENT ESTIMATES				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: STEEPEST DESCENT*	0.00001	150	Coefficient	0.101	-0.120	0.072	-27.747	-0.814	0.431	0.018	1.743	1.452
			Standard error	0.147	0.069	0.015	146.613	0.457	0.243	0.074	0.948	0.345
2: BFGS*	0.00001	57	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
3: BFGS-SC*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: DFP*	0.00001	150	Coefficient	-2.013	0.453	-0.115	-60.304	-0.363	0.651	0.032	11.627	0.200
			Standard error	1.071	0.206	0.068	65.050	0.303	0.140	0.033	27.879	0.476
5: NEWTON	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.001	134	Coefficient	0.154	0.148	0.138	-60.245	-0.875	0.453	0.034	2.135	1.202
			Standard error	0.177	0.082	0.039	125.621	0.495	0.236	0.064	1.387	0.452
	0.1	16	Coefficient	-0.086	0.114	0.071	-14.095	-0.896	0.396	0.011	2.211	1.316
			Standard error	0.118	0.059	0.015	164.111	0.448	0.261	0.083	1.281	0.372
	1.0	13	Coefficient	0.086	0.114	-0.071	-14.095	-0.896	0.396	0.011	2.211	1.316
			Standard error	0.119	0.059	0.015	194.625	0.449	0.261	0.099	1.288	0.373
6: PRCG*	0.00001	150	Coefficient	0.117	0.087	-0.134	-57.145	-1.029	0.322	0.033	3.229	0.980
			Standard error	0.132	0.110	0.023	160.908	0.609	0.383	0.082	5.990	1.151

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.23: Dataset 9, GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.105	-0.123	-0.072	-31.254	-0.809	0.431	0.019	1.736	1.437
			Standard error	0.146	0.067	0.015	137.396	0.449	0.239	0.070	0.915	0.334
2: Bfgs*	0.00001	65	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp*	0.00001	150	Coefficient	1.819	0.446	-0.122	-61.274	-0.363	0.646	0.033	18.626	0.125
			Standard error	0.870	0.162	0.064	64.172	0.301	0.138	0.033	74.657	0.500
5: Newton*	0.00001	22	Coefficient	0.086	-0.116	-0.071	-19.980	-0.904	0.391	0.014	2.253	1.290
			Standard error	0.114	0.057	0.015	147.587	0.436	0.259	0.075	1.274	0.362
6: Prcg	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.1	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	1.0	150	Coefficient	0.081	-0.113	0.071	-16.504	-0.935	0.378	0.013	2.472	1.236
			Standard error	0.109	0.057	0.014	186.950	0.439	0.268	0.095	1.496	0.384

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.24: Dataset 10, GAUSS 5 results.

COEFFICIENTS ESTIMATES				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	_oprteps	Iterations										
1: STEEPEST DESCENT*	0.00001	150	Coefficient	0.170	-0.152	0.137	-62.455	-0.828	0.464	0.035	1.986	1.220
			Standard error	0.191	0.074	0.039	113.734	0.494	0.217	0.058	1.131	0.396
2: BFGS*	0.00001	29	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
3: BFGS-SC*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: DFP*	0.00001	150	Coefficient	1.742	0.439	-0.121	-56.969	-0.342	0.642	0.031	7.573	0.303
			Standard error	1.555	0.198	0.061	68.276	0.296	0.144	0.035	14.237	0.567
5: NEWTON	0.00001	150	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
	0.001	39	Coefficient	-0.089	-0.115	-0.072	-28.476	-0.904	0.391	0.018	2.243	1.278
			Standard error	0.114	0.059	0.015	154.898	0.431	0.255	0.078	1.249	0.354
	0.1	15	Coefficient	0.161	-0.148	-0.138	-60.681	-0.872	0.447	0.035	2.267	1.135
			Standard error	0.175	0.078	0.038	118.902	0.494	0.232	0.060	1.453	0.436
	1.0	15	Coefficient	-0.089	0.115	-0.072	-28.476	-0.904	0.391	0.018	2.243	1.278
			Standard error	0.114	0.059	0.015	155.825	0.431	0.255	0.079	1.248	0.354
6: PRCG*	0.00001	150	Coefficient	0.088	0.114	-0.072	-27.608	-0.914	0.387	0.018	2.312	1.261
			Standard Error	0.114	0.060	0.021	153.456	0.431	0.259	0.078	1.311	0.361

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.25: Dataset 11, GAUSS 5 results.

Coefficient Estimates				$\widetilde{g}_1$	$\widetilde{g}_2$	$\widetilde{g}_3$	$\widetilde{\alpha}_0$	$\widetilde{\alpha}_1$	$\widetilde{\alpha}_2$	$\widetilde{\alpha}_3$	$\widetilde{\zeta}$	$\widetilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: Steepest Descent*	0.00001	150	Coefficient	0.108	0.127	0.071	-18.092	-0.780	0.450	0.013	1.555	1.480
			Standard error	0.154	0.065	0.014	120.389	0.437	0.227	0.061	0.744	0.303
2: Bfgs*	0.00001	27	Coefficient	1.935	0.450	0.110	-46.994	-0.316	0.645	0.025	99.894	0.023
			Standard error	0.661	0.158	0.046	58.592	0.287	0.135	0.030	3595.458	0.829
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: Dfp*	0.00001	150	Coefficient	1.935	0.450	0.110	-46.994	-0.316	0.645	0.025	92.147	0.025
			Standard error	0.661	0.158	0.046	58.622	0.287	0.135	0.030	3319.922	0.899
5: Newton*	0.00001	10	Coefficient	0.136	0.127	0.136	-35.965	-0.957	0.404	0.022	2.561	1.073
			Standard error	0.140	0.160	0.024	124.981	0.508	0.354	0.063	2.076	0.495
6: Prcg*	0.00001	150	Coefficient	0.085	-0.118	0.070	-6.998	-0.905	0.397	0.008	2.167	1.297
			Standard error	0.123	0.061	0.027	132.466	0.427	0.254	0.067	1.107	0.325

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.26: Evaluating algorithm switching (Steepest Descent/BFGS) with eleven datasets, GAUSS 5 results.

DATASET	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
1	28	Coefficient	0.142	-0.155	0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.298	0.455	0.232	0.065	1.294	0.441
2	51	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3	29	Coefficient	-0.161	$6.6E-7$	0.143	-118.441	-0.880	0.515	0.063	1.730	1.188
		Standard error	0.178	0.038	0.031	117.924	0.449	0.145	0.060	0.865	0.264
4	56	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
5	20	Coefficient	0.152	-0.156	0.140	-136.081	-0.991	0.436	0.073	2.025	1.221
		Standard error	0.169	0.078	0.032	127.940	0.439	0.223	0.065	1.315	0.451
6	28	Coefficient	0.151	-0.154	-0.141	-143.252	-0.990	0.437	0.077	1.895	1.279
		Standard error	0.177	0.081	0.034	130.185	0.441	0.222	0.066	1.227	0.453
7	51	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
8	55	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
9	28	Coefficient	0.156	0.149	0.138	-61.034	-0.881	0.450	0.035	2.225	1.161
		Standard error	0.173	0.080	0.038	122.477	0.490	0.234	0.062	1.437	0.444
10	29	Coefficient	0.090	-0.115	-0.072	-28.476	-0.904	0.392	0.019	2.243	1.278
		Standard error	0.114	0.059	0.015	155.081	0.430	0.255	0.079	1.248	0.354
11	31	Coefficient	0.079	0.116	0.069	-2.441	-0.939	0.385	0.005	2.368	1.247
		Standard error	0.112	0.061	0.026	76.025	0.424	0.265	0.039	1.280	0.338

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` $= 0.00001$ and $\zeta = 0.5$.

Table A.27: Evaluating algorithm switching (Steepest Descent/Newton) with eleven datasets, GAUSS 5 results.

DATASET	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
1	28	Coefficient	0.142	-0.155	0.136	-88.481	-0.922	0.436	0.049	2.047	1.237
		Standard error	0.167	0.078	0.032	127.371	0.456	0.232	0.065	1.294	0.441
2	37	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
3	25	Coefficient	-0.161	$-9.7E-9$	0.143	-118.442	-0.880	0.515	0.063	1.730	1.188
		Standard error	0.178	0.038	0.031	117.894	0.450	0.145	0.060	0.865	0.264
4	32	Coefficient	1.543	-0.456	-0.124	-121.801	-0.475	0.728	0.064	172.707	0.012
		Standard error	0.395	0.195	0.052	71.611	0.286	0.130	0.037	5065.979	0.345
5	15	Coefficient	0.152	-0.156	0.140	-136.081	-0.991	0.436	0.073	2.025	1.221
		Standard error	0.169	0.078	0.032	127.768	0.439	0.223	0.065	1.315	0.451
6	24	Coefficient	0.151	-0.154	-0.141	-143.253	-0.990	0.437	0.077	1.895	1.279
		Standard error	0.177	0.081	0.034	130.413	0.441	0.223	0.0664	1.228	0.453
7	44	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
8	37	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
9	23	Coefficient	0.156	0.149	0.138	-61.032	-0.881	0.450	0.035	2.225	1.161
		Standard error	0.173	0.080	0.038	122.837	0.490	0.234	0.062	1.437	0.444
10	22	Coefficient	0.090	-0.115	-0.072	-28.476	-0.904	0.392	0.019	2.243	1.278
		Standard error	0.114	0.059	0.015	151.061	0.431	0.255	0.076	1.248	0.354
11	25	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` = 0.00001 and $\zeta = 0.5$.

Table A.28: Example 1, GAUSS 3 results.

Coefficient Estimates			$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	Iterations								
1: Steepest Descent	150	Coefficient	0.075	$4.1E-6$	4.667	0.308	0.196	1.554	1.047
		Standard error	0.010	0.002	1.000	0.053	0.010	0.331	0.083
2: Bfgs	40	Coefficient	0.074	$-7.8E-11$	4.747	0.307	0.196	1.758	1.027
		Standard error	0.006	0.002	1.096	0.058	0.010	0.398	0.082
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: Dfp	37	Coefficient	0.074	$2.7E-9$	4.747	0.307	0.196	1.758	1.027
		Standard error	0.006	0.002	1.096	0.058	0.010	0.398	0.082
5: Newton	19	Coefficient	-0.074	$2.1E-12$	4.747	0.307	0.196	1.758	1.027
		Standard error	0.006	0.002	1.096	0.058	0.010	0.398	0.082
6: Prcg	150	Coefficient	0.093	$-2.0E-4$	4.262	0.309	0.197	0.987	1.131
		Standard error	0.007	0.004	0.647	0.038	0.010	0.162	0.090

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.29: Example 2, GAUSS 3 results.

Coefficient Estimates			$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	Iterations										
1: Steepest Descent	150	Coefficient	0.329	0.329	3.0$E-4$	7.861	0.896	1.200	0.794	2.991	0.909
		Standard error	0.041	0.040	0.015	1.088	0.281	0.356	0.070	0.563	0.124
2: Bfgs	45	Coefficient	0.325	0.321	3.3E-8	7.929	0.924	1.232	0.782	4.340	0.716
		Standard error	0.045	0.034	0.013	1.242	0.315	0.399	0.066	1.176	0.143
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-	-	-
4: Dfp	150	Coefficient	0.327	-0.318	2.0$E-4$	7.935	0.925	1.233	0.781	4.503	0.698
		Standard error	0.047	0.043	0.013	1.257	0.320	0.403	0.066	1.294	0.148
5: Newton	15	Coefficient	0.325	-0.321	1.1$E-8$	7.929	0.924	1.232	0.782	4.340	0.716
		Standard error	0.045	0.034	0.013	1.242	0.315	0.399	0.066	1.176	0.143
6: Prcg	150	Coefficient	0.326	0.323	-5.0$E-4$	7.912	0.916	1.223	0.786	3.827	0.778
		Standard error	0.040	0.035	0.013	1.192	0.303	0.385	0.067	0.884	0.131

Note: a dash (-) denotes no estimate due to algorithm failure.

Table A.30: Example 1, GAUSS 5 results.

Coefficient Estimates				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	_oprteps	Iterations								
1: Steepest Descent*	0.00001	150	Coefficient	0.103	$2.0E-4$	3.966	0.286	0.214	0.774	1.177
			Standard error	0.012	0.004	0.523	0.032	0.011	0.119	0.094
2: Bfgs*	0.00001	20	Coefficient	0.080	$-3.1E-9$	4.387	0.276	0.213	1.549	1.042
			Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
3: Bfgs-sc*	0.00001	1	Coefficient	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-
4: Dfp*	0.00001	41	Coefficient	0.080	$-1.0E-9$	4.387	0.276	0.213	1.549	1.042
			Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
5: Newton*	0.00001	18	Coefficient	0.080	$1.1E-9$	4.387	0.276	0.213	1.549	1.042
			Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
6: Prcg*	0.00001	150	Coefficient	-0.081	$-7.3E-5$	4.330	0.278	0.213	1.323	1.068
			Standard error	0.021	0.002	0.882	0.047	0.010	0.295	0.083

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of _oprteps.

Table A.31: Example 2, GAUSS 5 results.

COEFFICIENT ESTIMATES				$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
Algorithm	`_oprteps`	Iterations										
1: STEEPEST DESCENT*	0.00001	150	Coefficient	0.398	0.334	-0.002	8.020	0.796	1.368	0.825	2.783	0.936
			Standard error	0.038	0.069	0.021	0.995	0.261	0.328	0.072	0.512	0.124
2: BFGS	0.00001	24	Coefficient	0.394	0.321	$-4.1E-7$	8.155	0.806	1.391	0.834	4.073	0.735
			Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.040	0.138
	0.001	22	Coefficient	0.394	0.321	$1.1E-7$	8.155	0.806	1.391	0.834	4.073	0.735
			Standard error	0.026	0.033	0.034	1.131	0.296	0.367	0.067	1.039	0.138
	0.1	19	Coefficient	0.394	0.321	$1.7E-6$	8.155	0.806	1.391	0.834	4.073	0.735
			Standard error	0.026	0.033	0.037	1.131	0.296	0.367	0.067	1.040	0.138
	1.0	53	Coefficient	0.394	0.321	$2.0E-7$	8.155	0.806	1.391	0.834	4.073	0.735
			Standard error	0.026	0.033	0.037	1.131	0.296	0.367	0.067	1.040	0.138
3: BFGS-SC*	0.00001	1	Coefficient	-	-	-	-	-	-	-	-	-
			Standard error	-	-	-	-	-	-	-	-	-
4: DFP*	0.00001	22	Coefficient	0.394	0.321	$2.4E-6$	8.155	0.806	1.391	0.834	4.073	0.735
			Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.040	0.138
5: NEWTON*	0.00001	10	Coefficient	-0.394	0.321	$-4.2E-8$	8.155	0.806	1.391	0.834	4.073	0.735
			Standard error	0.026	0.033	0.038	1.131	0.296	0.367	0.067	1.040	0.138
6: PRCG*	0.00001	150	Coefficient	0.394	-0.321	$7.3E-7$	8.147	0.806	1.389	0.833	3.928	0.753
			Standard error	0.026	0.034	0.030	1.118	0.293	0.364	0.068	0.956	0.134

Note: a dash (-) denotes no estimate due to algorithm failure.
* indicates same results for all values of `_oprteps`.

Table A.32: **Effect of initial value of ζ: single algorithm (Steepest Descent), Hamilton's example 1, GAUSS 5 results.**

Initial Value of ζ		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.3	Coefficient	0.337	-0.010	3.669	0.289	0.214	1.042	1.118
	Standard error	0.031	0.013	0.403	0.028	0.013	0.177	0.098
0.4	Coefficient	0.104	0.001	3.953	0.287	0.215	0.755	1.184
	Standard error	0.010	0.006	0.514	0.032	0.011	0.115	0.095
0.5	Coefficient	0.103	0.000	3.966	0.286	0.214	0.774	1.177
	Standard error	0.012	0.004	0.523	0.032	0.010	0.119	0.094
0.6	Coefficient	0.103	0.000	3.987	0.286	0.214	0.803	1.166
	Standard error	0.015	0.004	0.538	0.033	0.011	0.125	0.093
0.7	Coefficient	0.103	0.000	4.012	0.285	0.214	0.842	1.154
	Standard error	0.010	0.004	0.552	0.034	0.011	0.133	0.092
0.8	Coefficient	0.102	0.000	4.053	0.284	0.214	0.911	1.134
	Standard error	0.011	0.003	0.583	0.035	0.010	0.148	0.089
0.9	Coefficient	0.102	0.000	4.080	0.283	0.214	0.963	1.121
	Standard error	0.010	0.003	0.606	0.037	0.010	0.161	0.088
1.0	Coefficient	0.101	0.000	4.115	0.281	0.214	1.037	1.105
	Standard error	0.012	0.003	0.640	0.038	0.010	0.180	0.087
1.1	Coefficient	0.101	0.000	4.156	0.280	0.214	1.138	1.086
	Standard error	0.013	0.002	0.685	0.041	0.010	0.208	0.085
1.2	Coefficient	0.101	0.000	4.182	0.279	0.214	1.211	1.074
	Standard error	0.012	0.002	0.717	0.042	0.010	0.231	0.084
1.3	Coefficient	0.101	0.000	4.212	0.278	0.213	1.306	1.060
	Standard error	0.007	0.002	0.758	0.045	0.010	0.263	0.083
1.4	Coefficient	0.101	0.000	4.239	0.277	0.213	1.404	1.048
	Standard error	0.008	0.002	0.801	0.047	0.010	0.299	0.082
1.5	Coefficient	0.101	0.000	4.263	0.276	0.213	1.499	1.037
	Standard error	0.006	0.002	0.843	0.049	0.010	0.338	0.082

Note: in all cases, `_oprteps` = 0.00001 and number of iterations was 150. Algorithm failed for initial values of ζ = 0.1 and 0.2.

Table A.33: **Effect of initial value of ζ: single algorithm (Steepest Descent), Hamilton's example 2, GAUSS 5 results.**

Initial Value of ζ		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.3	Coefficient	0.397	0.330	0.002	8.046	0.798	1.373	0.826	2.934	0.907
	Standard error	0.036	0.052	0.021	1.010	0.266	0.334	0.071	0.550	0.123
0.4	Coefficient	-0.395	-0.323	0.001	8.106	0.803	1.382	0.829	3.406	0.827
	Standard error	0.028	0.033	0.023	1.065	0.280	0.349	0.069	0.710	0.126
0.5	Coefficient	0.398	0.334	-0.002	8.020	0.796	1.368	0.825	2.783	0.936
	Standard error	0.038	0.069	0.021	0.995	0.261	0.328	0.072	0.512	0.124
0.6	Coefficient	0.399	0.335	0.003	8.009	0.795	1.367	0.825	2.732	0.947
	Standard error	0.041	0.066	0.021	0.986	0.259	0.326	0.072	0.497	0.124
0.7	Coefficient	-0.396	-0.327	-0.002	8.070	0.800	1.376	0.827	3.088	0.880
	Standard error	0.034	0.052	0.021	1.033	0.271	0.339	0.070	0.601	0.124
0.8	Coefficient	0.399	0.335	-0.003	8.009	0.795	1.367	0.825	2.730	0.947
	Standard error	0.041	0.066	0.021	0.985	0.259	0.326	0.072	0.496	0.124
0.9	Coefficient	0.399	0.335	0.003	8.010	0.795	1.367	0.825	2.738	0.945
	Standard error	0.041	0.075	0.021	0.990	0.259	0.326	0.072	0.502	0.124
1.0	Coefficient	0.399	0.335	-0.001	8.009	0.795	1.367	0.825	2.733	0.946
	Standard error	0.041	0.066	0.020	0.986	0.259	0.326	0.072	0.497	0.124
1.1	Coefficient	0.398	0.334	-0.001	8.015	0.795	1.368	0.825	2.762	0.941
	Standard error	0.040	0.069	0.020	0.993	0.260	0.327	0.072	0.507	0.123
1.2	Coefficient	0.399	0.335	-0.001	8.009	0.795	1.367	0.825	2.733	0.946
	Standard error	0.041	0.066	0.020	0.986	0.259	0.326	0.072	0.497	0.124
1.3	Coefficient	0.399	0.335	-0.003	8.010	0.795	1.367	0.825	2.738	0.945
	Standard error	0.041	0.075	0.021	0.990	0.259	0.326	0.072	0.502	0.124
1.4	Coefficient	0.399	0.335	-0.003	8.011	0.795	1.367	0.825	2.744	0.944
	Standard error	0.041	0.077	0.021	0.993	0.260	0.326	0.072	0.505	0.124
1.5	Coefficient	0.399	0.335	-0.003	8.011	0.795	1.367	0.825	2.744	0.944
	Standard error	0.041	0.077	0.021	0.993	0.260	0.326	0.072	0.505	0.124

Note: in all cases, `_oprteps` = 0.00001 and number of iterations was 150. Algorithm failed for initial values of $\zeta = 0.1$ and 0.2.

Table A.34: **Effect of initial value of ζ: Example 1, algorithm switching (Steepest Descent/BFGS), GAUSS 5 results.**

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.1	42	Coefficient	0.080	$6.7E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.3	39	Coefficient	0.080	$1.5E-9$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.5	37	Coefficient	-0.080	$-2.0E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.7	39	Coefficient	-0.080	$1.7E-9$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.9	38	Coefficient	-0.080	$1.0E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.0	38	Coefficient	-0.080	$-4.5E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.1	44	Coefficient	0.080	$-5.8E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.3	44	Coefficient	-0.080	$1.7E-9$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.5	40	Coefficient	0.080	$-5.2E-11$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081

Note: in all cases, `_oprteps` = 0.00001.

Table A.35: Effect of initial value of ζ: Example 1, algorithm switching (Steepest Descent/Newton), GAUSS 5 results.

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.1	35	Coefficient	-0.080	$-1.9E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.3	32	Coefficient	-0.080	$9.7E-13$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.5	32	Coefficient	0.080	$3.0E-13$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.7	41	Coefficient	-0.080	$9.0E-13$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
0.9	31	Coefficient	0.080	$2.0E-13$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.0	36	Coefficient	0.080	$3.1E-12$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.1	50	Coefficient	-0.080	$-2.7E-10$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.3	42	Coefficient	0.080	$3.7E-12$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081
1.5	36	Coefficient	-0.080	$7.2E-11$	4.387	0.276	0.213	1.549	1.042
		Standard error	0.005	0.001	0.950	0.052	0.010	0.335	0.081

Note: in all cases, `_oprteps` = 0.00001 and ζ.

Table A.36: Effect of initial value of ζ: Example 2, algorithm switching (Steepest Descent/BFGS), GAUSS 5 results.

INITIAL VALUE OF ζ	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.1	55	Coefficient	0.394	0.321	-1.7$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.039	0.138
0.3	38	Coefficient	0.394	0.321	7.6$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.039	0.138
0.5	33	Coefficient	0.394	0.321	3.6$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.035	1.131	0.296	0.367	0.067	1.039	0.138
0.7	33	Coefficient	-0.394	-0.321	5.4$E-8$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.037	1.131	0.296	0.367	0.067	1.040	0.138
0.9	34	Coefficient	0.394	0.321	5.1$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.040	0.1378
1.0	36	Coefficient	0.394	0.321	-2.0$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.040	0.138
1.1	34	Coefficient	0.394	0.321	-4.3$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.034	1.131	0.296	0.367	0.067	1.039	0.138
1.3	43	Coefficient	0.394	0.321	7.2$E-8$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.036	1.131	0.296	0.367	0.067	1.039	0.138
1.5	34	Coefficient	0.394	0.321	-1.5$E-6$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	0.035	1.131	0.296	0.367	0.067	1.040	0.138

Note: in all cases, `_oprteps` = 0.00001 and ζ.

Table A.37: **Effect of initial value of ζ: Example 2, algorithm switching (Steepest Descent/Newton), GAUSS 5 results.**

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\zeta}$	$\tilde{\sigma}$
0.1	31	Coefficient	0.394	0.321	-1.3$E-6$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.7$E-2$	1.131	0.296	0.367	0.067	1.040	0.138
0.3	28	Coefficient	0.394	0.321	9.7$E-9$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.6$E-2$	1.131	0.296	0.367	0.067	1.039	0.138
0.5	28	Coefficient	0.394	0.321	2.4$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.6$E-2$	1.131	0.296	0.367	0.067	1.040	0.138
0.7	22	Coefficient	-0.394	-0.321	-4.6$E-8$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.7$E-2$	1.131	0.296	0.367	0.067	1.040	0.138
0.9	28	Coefficient	0.394	0.321	-4.5$E-7$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.56$E-2$	1.131	0.296	0.367	0.067	1.039	0.138
1.0	29	Coefficient	0.394	0.321	-1.2$E-8$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.54$E-2$	1.131	0.296	0.367	0.067	1.039	0.138
1.1	30	Coefficient	0.394	0.321	6.9$E-8$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.65$E-2$	1.131	0.296	0.367	0.067	1.039	0.138
1.3	27	Coefficient	0.394	0.321	-1.4$E-6$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.52$E-2$	1.131	0.296	0.368	0.067	1.039	0.138
1.5	32	Coefficient	0.394	0.321	1.8$E-9$	8.155	0.806	1.391	0.834	4.073	0.735
		Standard error	0.026	0.033	3.7$E-2$	1.131	0.296	0.367	0.067	1.039	0.138

Note: in all cases, `_oprteps` = 0.00001 and ζ.

Table A.38: Results for all algorithms, dataset 2, Hamilton's Example 1, GAUSS 5.

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: Steepest Descent	150	Coefficient	-0.115	-0.008	4.173	0.281	0.215	1.336	0.997
		Standard error	0.008	0.008	0.690	0.041	0.014	0.274	0.091
2: Bfgs	22	Coefficient	0.125	-0.009	4.282	0.278	0.216	1.941	0.894
		Standard error	0.008	0.007	0.846	0.051	0.017	0.540	0.103
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: Dfp	31	Coefficient	0.125	-0.009	4.282	0.278	0.216	1.941	0.894
		Standard error	0.008	0.007	0.846	0.051	0.017	0.540	0.103
5: Newton	11	Coefficient	0.078	0.005	4.537	0.272	0.214	2.037	0.929
		Standard error	0.005	0.005	1.109	0.059	0.015	0.551	0.094
6: Prcg	150	Coefficient	0.084	0.006	4.313	0.279	0.213	1.261	1.032
		Standard error	0.020	0.007	0.790	0.043	0.013	0.257	0.088
7: Steepest Descent/Bfgs	36	Coefficient	-0.078	-0.005	4.537	0.272	0.214	2.037	0.929
		Standard error	0.005	0.005	1.109	0.059	0.015	0.551	0.094
8: Steepest Descent/Newton	25	Coefficient	-0.105	-0.007	4.323	0.276	0.215	1.905	0.920
		Standard error	0.006	0.006	0.918	0.053	0.015	0.520	0.097

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` = 0.00001 and $\zeta = 0.5$.

Table A.39: Results for all algorithms, dataset 3, Hamilton's Example 1, GAUSS 5.

INITIAL VALUE OF ζ	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: STEEPEST DESCENT	150	Coefficient	0.104	$3.0E-4$	3.952	0.287	0.215	0.761	1.190
		Standard error	0.010	0.004	0.520	0.032	0.011	0.117	0.097
2: BFGS	25	Coefficient	0.080	$4.5E-10$	4.384	0.276	0.214	1.572	1.043
		Standard error	0.005	0.001	0.964	0.053	0.010	0.343	0.083
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: DFP	60	Coefficient	0.080	$2.7E-9$	4.384	0.276	0.214	1.572	1.043
		Standard error	0.005	0.001	0.964	0.053	0.010	0.343	0.083
5: NEWTON	9	Coefficient	0.097	$-6.5E-10$	4.266	0.276	0.214	1.468	1.047
		Standard error	0.007	0.002	0.847	0.049	0.010	0.322	0.083
6: PRCG	150	Coefficient	-0.081	$-1.9E-5$	4.330	0.278	0.214	1.343	1.070
		Standard error	0.007	0.002	0.855	0.048	0.010	0.263	0.084
7: STEEPEST DESCENT/BFGS	40	Coefficient	-0.080	$8.1E-11$	4.384	0.276	0.214	1.572	1.043
		Standard error	0.005	0.001	0.964	0.053	0.010	0.343	0.083
8: STEEPEST DESCENT/NEWTON	38	Coefficient	0.080	$-8.6E-12$	4.384	0.276	0.214	1.572	1.043
		Standard error	0.005	0.001	0.964	0.053	0.010	0.343	0.083

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` $= 0.00001$ and $\zeta = 0.5$.

Table A.40: Results for all algorithms, dataset 4, Hamilton's example 1, GAUSS 5.

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: Steepest Descent	150	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
2: Bfgs	34	Coefficient	0.067	0.002	4.968	0.297	0.205	1.821	1.050
		Standard error	0.002	0.004	1.190	0.061	0.012	0.426	0.092
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: Dfp	103	Coefficient	0.067	0.002	4.968	0.297	0.205	1.821	1.050
		Standard error	0.002	0.004	1.190	0.061	0.012	0.428	0.092
5: Newton	15	Coefficient	0.075	0.002	4.791	0.298	0.205	1.726	1.052
		Standard error	0.005	0.004	1.091	0.058	0.012	0.406	0.092
6: Prcg	42	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
7: Steepest Descent/Bfgs	46	Coefficient	0.067	-0.002	4.968	0.297	0.205	1.821	1.050
		Standard error	0.002	0.004	1.190	0.061	0.012	0.428	0.092
8: Steepest Descent/Newton	33	Coefficient	0.075	-0.002	4.791	0.298	0.205	1.726	1.052
		Standard error	0.005	0.004	1.091	0.058	0.012	0.406	0.092

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` = 0.00001 and $\zeta = 0.5$.

Table A.41: **Results for all algorithms, dataset 5, Hamilton's example 1, GAUSS 5.**

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: Steepest Descent	150	Coefficient	0.105	$3.0E-4$	3.896	0.290	0.217	0.750	1.200
		Standard error	0.022	0.004	0.526	0.033	0.011	0.120	0.099
2: Bfgs	32	Coefficient	-0.080	$-3.5E-10$	4.356	0.277	0.215	1.569	1.048
		Standard error	0.006	0.001	0.969	0.053	0.010	0.347	0.084
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: Dfp	83	Coefficient	0.079	$-2.1E-10$	4.378	0.277	0.215	1.581	1.048
		Standard error	0.008	0.001	0.989	0.054	0.010	0.355	0.084
5: Newton	13	Coefficient	0.080	$-7.4E-12$	4.356	0.277	0.215	1.569	1.048
		Standard error	0.006	0.001	0.969	0.053	0.010	0.347	0.084
6: Prcg	150	Coefficient	0.102	$4.2E-5$	4.166	0.280	0.215	1.282	1.073
		Standard error	0.010	0.002	0.754	0.045	0.010	0.259	0.087
7: Steepest Descent/Bfgs	32	Coefficient	-0.080	$-3.9E-10$	4.356	0.277	0.215	1.569	1.048
		Standard error	0.006	0.001	0.969	0.053	0.010	0.347	0.084
8: Steepest Descent/Newton	48	Coefficient	-0.080	$6.4E-10$	4.356	0.277	0.215	1.569	1.048
		Standard error	0.006	0.001	0.969	0.053	0.010	0.347	0.084

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, _oprteps = 0.00001 and $\zeta = 0.5$.

Table A.42: Results for all algorithms, dataset 6, Hamilton's example 1, GAUSS 5.

Initial Value of ζ	Iterations		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: Steepest Descent	150	Coefficient	0.087	-0.003	4.191	0.308	0.192	1.147	1.093
		Standard error	0.008	0.007	0.734	0.042	0.012	0.222	0.093
2: Bfgs	17	Coefficient	0.087	-0.003	4.288	0.308	0.192	1.498	1.042
		Standard error	0.008	0.006	0.896	0.050	0.012	0.357	0.092
3: Bfgs-sc	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: Dfp	24	Coefficient	-0.087	-0.003	4.288	0.308	0.192	1.498	1.042
		Standard error	0.008	0.006	0.896	0.050	0.012	0.357	0.092
5: Newton	9	Coefficient	-0.087	0.003	4.288	0.308	0.192	1.498	1.042
		Standard error	0.008	0.006	0.896	0.050	0.012	0.357	0.092
6: Prcg	34	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
7: Steepest Descent/Bfgs	34	Coefficient	0.087	0.003	4.288	0.308	0.192	1.498	1.042
		Standard error	0.008	0.006	0.896	0.050	0.012	0.357	0.092
8: Steepest Descent/Newton	27	Coefficient	0.078	0.002	4.383	0.308	0.191	1.549	1.044
		Standard error	0.012	0.005	0.973	0.053	0.011	0.383	0.091

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` $= 0.00001$ and $\zeta = 0.5$.

Table A.43: Results for all algorithms, dataset 7, Hamilton's example 1, GAUSS 5.

INITIAL VALUE OF ζ	ITERATIONS		$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{\alpha}_0$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\zeta}$	$\tilde{\sigma}$
1: STEEPEST DESCENT	150	Coefficient	0.104	$7.0E-4$	3.908	0.289	0.216	0.765	1.202
		Standard error	0.012	0.006	0.533	0.033	0.012	0.123	0.102
2: BFGS	28	Coefficient	0.077	$-5.4E-11$	4.393	0.277	0.214	1.580	1.053
		Standard error	0.009	0.001	1.000	0.054	0.011	0.363	0.087
3: BFGS-SC	1	Coefficient	-	-	-	-	-	-	-
		Standard error	-	-	-	-	-	-	-
4: DFP	35	Coefficient	-0.080	$6.3E-11$	4.350	0.277	0.214	1.554	1.053
		Standard error	0.006	0.001	0.966	0.053	0.011	0.348	0.087
5: NEWTON	10	Coefficient	0.101	$-5.3E-12$	4.211	0.278	0.214	1.443	1.056
		Standard error	0.008	0.002	0.830	0.048	0.011	0.325	0.089
6: PRCG	150	Coefficient	0.081	$2.0E-4$	4.296	0.279	0.214	1.339	1.080
		Standard error	0.006	0.002	0.860	0.048	0.011	0.268	0.089
7: STEEPEST DESCENT/BFGS	24	Coefficient	0.080	$-1.5E-9$	4.350	0.277	0.214	1.554	1.053
		Standard error	0.006	0.001	0.966	0.053	0.011	0.348	0.087
8: STEEPEST DESCENT/NEWTON	27	Coefficient	0.080	$1.2E-11$	4.350	0.277	0.214	1.554	1.053
		Standard error	0.006	0.001	0.966	0.053	0.011	0.348	0.087

Note: a dash (-) denotes no estimate due to algorithm failure. In all cases, `_oprteps` = 0.00001 and $\zeta = 0.5$.

Appendix B

Appendix to Chapter 3

B.1 Simulation Results

Table B.1: Key to simulation specifications.

Number	Specification	Parameters	
		Null	
1	$y_t = 1 + 0.25x_{1t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$	
2	$y_t = 1 + 0.5x_{1t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 4)$	
3	$y_t = 1 + 0.75x_{1t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 8)$	
4	$y_t = 1 + 0.25x_{2t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.4)$	
5	$y_t = 1 + 0.5x_{2t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.7)$	
6	$y_t = 1 + 0.75x_{2t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.6)$	
7	$y_t = 1 + 0.25x_{3t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.3)$	
8	$y_t = 1 + 0.5x_{3t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.8)$	
9	$y_t = 1 + 0.75x_{3t} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$	
		Quadratic	
1	$y_t = 1 + 0.25x_{1t} + 0.25x_{1t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$	
2	$y_t = 1 + 0.5x_{1t} + 0.5x_{1t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 4)$	
3	$y_t = 1 + 0.75x_{1t} + 0.75x_{1t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 8)$	
4	$y_t = 1 + 0.25x_{2t} + 0.25x_{2t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.4)$	
5	$y_t = 1 + 0.5x_{2t} + 0.5x_{2t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.7)$	
6	$y_t = 1 + 0.75x_{2t} + 0.75x_{2t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.6)$	
7	$y_t = 1 + 0.25x_{3t} + 0.25x_{3t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.3)$	
8	$y_t = 1 + 0.5x_{3t} + 0.5x_{3t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 0.8)$	
9	$y_t = 1 + 0.75x_{3t} + 0.75x_{3t}^2 + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$	
		Square	
1	$y = x_{4t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 1)$	$\varepsilon_t \sim N(0, 1)$
2	$y = x_{4t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 25)$	$\varepsilon_t \sim N(0, 1)$
3	$y = x_{4t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 400)$	$\varepsilon_t \sim N(0, 1)$
4	$y = x_{1t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 1)$	
5	$y = x_{1t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 25)$	
6	$y = x_{1t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 400)$	
7	$y = x_{2t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 1)$	
8	$y = x_{2t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 25)$	
9	$y = x_{2t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 400)$	
10	$y = x_{3t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 1)$	
11	$y = x_{3t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 25)$	
12	$y = x_{3t}^2 + \epsilon_t$	$\epsilon_t \sim N(0, 400)$	
		Autoregressive	
1	$y_t = 1 + 0.25x_{1t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$
2	$y_t = 1 + 0.25x_{1t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$
3	$y_t = 1 + 0.25x_{1t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0, 1.5)$

Continued on next page.

Number	Specification	Parameters	
4	$y_t = 1 + 0.5x_{1t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,4)$
5	$y_t = 1 + 0.5x_{1t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,4)$
6	$y_t = 1 + 0.5x_{1t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,4)$
7	$y_t = 1 + 0.75x_{1t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,8)$
8	$y_t = 1 + 0.75x_{1t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,8)$
9	$y_t = 1 + 0.75x_{1t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,8)$
10	$y_t = 1 + 0.25x_{2t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.4)$
11	$y_t = 1 + 0.25x_{2t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.4)$
12	$y_t = 1 + 0.25x_{2t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.4)$
13	$y_t = 1 + 0.5x_{2t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.7)$
14	$y_t = 1 + 0.5x_{2t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.7)$
15	$y_t = 1 + 0.5x_{2t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.7)$
16	$y_t = 1 + 0.75x_{2t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1.6)$
17	$y_t = 1 + 0.75x_{2t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1.6)$
18	$y_t = 1 + 0.75x_{2t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1.6)$
19	$y_t = 1 + 0.25x_{3t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.3)$
20	$y_t = 1 + 0.25x_{3t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.3)$
21	$y_t = 1 + 0.25x_{3t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.3)$
22	$y_t = 1 + 0.5x_{3t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.8)$
23	$y_t = 1 + 0.5x_{3t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.8)$
24	$y_t = 1 + 0.5x_{3t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,0.8)$
25	$y_t = 1 + 0.75x_{3t} + \epsilon_t$	$\epsilon_t = 0.1\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1.5)$
26	$y_t = 1 + 0.75x_{3t} + \epsilon_t$	$\epsilon_t = 0.5\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1.5)$
27	$y_t = 1 + 0.75x_{3t} + \epsilon_t$	$\epsilon_t = 0.9\epsilon_{t-1} + \varepsilon_t$	$\varepsilon_t \sim N(0,1.5)$
	Hamilton		
1	$y_t = 0.6x_{2t}1_{[x_{2t}>5.5]} + 0.2x_{2t} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$	
2	$y_t = 0.6x_{2t}1_{[x_{2t}>5.5]} + 0.2x_{3t} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$	
3	$y_t = 0.6x_{2t}1_{[x_{2t}>5.5]} + 0.2x_{1t} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$	
4	$y_t = 0.6x_{3t}1_{[x_{3t}>5.5]} + 0.2x_{3t} + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$	
5	$y_t = 0.6x_{3t}1_{[x_{3t}1>5.5]} + 0.2x_{1t}2 + \varepsilon_t$	$\varepsilon_t \sim N(0,1)$	

Note: $x_{1t} = 20t/T$.
$x_{2t} = U(1,10)$.
$x_{3t} = N(5.5, 2.25^2)$.
$x_{4t} = 0.6x_{1t-1} + \varepsilon_t$.

Table B.2: The Durbin-Watson test - unordered null case.

SPECIFICATION	VALUE OF d	T =	25	75	125	175	225
1	$d < d_L$		1.685	3.080	3.410	3.845	3.600
	$d_L < d < d_U$		3.035	1.765	1.685	1.485	1.300
2	$d < d_L$		1.870	3.020	3.525	3.490	3.550
	$d_L < d < d_U$		3.140	2.045	1.655	1.375	1.205
3	$d < d_L$		1.910	2.815	3.295	3.630	3.795
	$d_L < d < d_U$		3.190	2.150	1.660	1.355	1.265
4	$d < d_L$		3.130	3.750	4.055	4.265	4.330
	$d_L < d < d_U$		4.650	2.660	1.850	1.610	1.340
5	$d < d_L$		2.975	3.495	4.015	3.920	4.215
	$d_L < d < d_U$		4.835	2.365	1.990	1.490	1.455
6	$d < d_L$		3.270	3.890	4.265	4.100	4.070
	$d_L < d < d_U$		4.550	2.440	2.035	1.625	1.325
7	$d < d_L$		3.035	4.095	4.050	4.080	3.945
	$d_L < d < d_U$		4.845	2.460	1.835	1.590	1.345
8	$d < d_L$		2.395	4.070	3.905	4.230	4.235
	$d_L < d < d_U$		4.355	2.490	1.875	1.500	1.385
9	$d < d_L$		2.980	3.845	4.255	4.480	4.160
	$d_L < d < d_U$		5.160	2.420	1.990	1.580	1.400

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.3: The Durbin-Watson test - unordered quadratic case.

Specification	Value of d	T = 25	75	125	175	225
1	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
2	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
3	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
4	$d < d_L$	0.925	0.690	0.115	3.360	0.000
	$d_L < d < d_U$	5.865	1.400	0.140	2.765	0.000
5	$d < d_L$	0.000	13.455	9.660	43.445	0.080
	$d_L < d < d_U$	0.000	18.265	11.200	16.280	0.180
6	$d < d_L$	0.235	0.000	80.420	0.000	0.165
	$d_L < d < d_U$	2.000	0.000	10.735	0.000	0.380
7	$d < d_L$	0.000	24.355	0.000	0.680	0.000
	$d_L < d < d_U$	0.000	25.745	0.000	1.055	0.000
8	$d < d_L$	1.030	0.000	0.000	0.000	0.000
	$d_L < d < d_U$	2.870	0.000	0.000	0.000	0.000
9	$d < d_L$	89.090	0.000	0.000	0.000	0.000
	$d_L < d < d_U$	10.840	0.000	0.000	0.000	0.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.4: The Durbin-Watson test - unordered square case.

Specification	Value of d	T = 25	75	125	175	225
1	$d < d_L$	1.480	40.885	91.440	99.620	100.000
	$d_L < d < d_U$	3.675	15.485	4.320	0.185	0.000
2	$d < d_L$	3.355	9.875	9.245	22.595	12.400
	$d_L < d < d_U$	5.640	4.670	3.635	5.080	3.140
3	$d < d_L$	2.130	3.580	3.780	3.990	4.795
	$d_L < d < d_U$	4.060	2.415	1.870	1.640	1.555
4	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
5	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
6	$d < d_L$	99.175	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.630	0.000	0.000	0.000	0.000
7	$d < d_L$	0.000	0.000	0.000	0.000	0.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
8	$d < d_L$	0.210	1.875	0.730	0.290	9.565
	$d_L < d < d_U$	0.790	1.935	0.530	0.190	3.045
9	$d < d_L$	2.880	3.475	4.545	4.525	3.980
	$d_L < d < d_U$	4.500	2.380	2.195	1.735	1.460
10	$d < d_L$	0.000	0.000	0.000	0.000	0.000
	$d_L < d < d_U$	0.110	0.000	0.000	0.000	0.000
11	$d < d_L$	3.805	0.705	0.525	1.725	1.480
	$d_L < d < d_U$	5.740	0.875	0.370	0.875	0.935
12	$d < d_L$	2.160	2.955	4.225	4.980	3.940
	$d_L < d < d_U$	3.700	1.975	1.660	1.800	1.415

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.5: **The Durbin-Watson test - unordered autoregressive case.**

Specification	Value of d T =	25	75	125	175	225	T =	25	75	125	175	225
1	$d < d_L$	4.660	14.585	23.290	30.875	37.260	15	95.410	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	6.500	6.130	5.900	5.565	5.250		2.405	0.000	0.000	0.000	0.000
2	$d < d_L$	51.220	98.600	99.970	100.000	100.000	16	8.005	16.850	26.370	33.355	39.550
	$d_L < d < d_U$	16.220	0.555	0.010	0.000	0.000		9.490	6.860	6.330	5.725	5.430
3	$d < d_L$	91.610	100.000	100.000	100.000	100.000	17	64.620	99.050	99.985	100.000	100.000
	$d_L < d < d_U$	4.040	0.000	0.000	0.000	0.000		13.795	0.345	0.005	0.000	0.000
4	$d < d_L$	4.635	14.330	23.195	30.935	37.760	18	96.930	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	6.510	6.340	5.935	5.510	5.435		1.580	0.000	0.000	0.000	0.000
5	$d < d_L$	51.365	98.550	99.975	100.000	100.000	19	7.350	17.385	26.470	33.060	39.995
	$d_L < d < d_U$	16.135	0.640	0.020	0.000	0.000		9.070	6.870	6.430	5.735	5.605
6	$d < d_L$	91.550	100.000	100.000	100.000	100.000	20	60.035	98.950	99.995	100.000	100.000
	$d_L < d < d_U$	4.235	0.000	0.000	0.000	0.000		15.300	0.430	0.000	0.000	0.000
7	$d < d_L$	4.400	14.220	23.280	30.070	37.790	21	95.745	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	6.725	6.120	5.880	5.965	5.415		2.280	0.000	0.000	0.000	0.000
8	$d < d_L$	51.525	98.620	99.975	100.000	100.000	22	8.300	17.290	25.325	33.200	39.450
	$d_L < d < d_U$	15.900	0.510	0.005	0.000	0.000		9.535	6.850	6.155	5.950	5.665
9	$d < d_L$	91.350	100.000	100.000	100.000	100.000	23	65.405	98.915	99.980	100.000	100.000
	$d_L < d < d_U$	4.080	0.000	0.000	0.000	0.000		14.000	0.440	0.005	0.000	0.000
10	$d < d_L$	6.840	17.400	25.815	33.375	40.580	24	94.790	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	8.435	6.970	6.110	5.825	5.680		2.810	0.000	0.000	0.000	0.000
11	$d < d_L$	61.610	99.140	99.980	100.000	100.000	25	8.140	17.325	25.590	33.975	39.575
	$d_L < d < d_U$	14.515	0.405	0.005	0.000	0.000		8.780	6.825	6.645	5.825	5.440
12	$d < d_L$	97.080	100.000	100.000	100.000	100.000	26	64.080	98.955	99.995	100.000	100.000
	$d_L < d < d_U$	1.650	0.000	0.000	0.000	0.000		13.810	0.465	0.000	0.000	0.000
13	$d < d_L$	8.970	17.175	25.730	33.335	39.425	27	97.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	9.775	6.985	6.015	5.935	5.745		1.660	0.000	0.000	0.000	0.000
14	$d < d_L$	62.470	99.055	99.990	100.000	100.000						
	$d_L < d < d_U$	13.870	0.415	0.000	0.000	0.000						

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.6: The Durbin-Watson test - unordered Hamilton case.

SPECIFICATION	VALUE OF d	T =	25	75	125	175	225
1	$d < d_L$		1.840	2.905	3.085	3.535	3.875
	$d_L < d < d_U$		10.395	4.960	4.090	3.465	2.985
2	$d < d_L$		1.655	3.000	3.340	3.490	3.695
	$d_L < d < d_U$		10.380	4.955	3.725	3.140	2.800
3	$d < d_L$		1.160	2.150	2.755	2.970	3.215
	$d_L < d < d_U$		7.250	4.395	3.245	2.725	2.530
4	$d < d_L$		1.880	2.785	3.305	3.690	3.725
	$d_L < d < d_U$		10.715	5.185	3.895	3.050	3.005
5	$d < d_L$		0.920	2.270	2.895	2.860	2.750
	$d_L < d < d_U$		7.285	4.230	3.365	2.610	2.590

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.7: The Durbin-Watson test - ordered null case.

Specification	Value of d T =	25	75	125	175	225
1	$d < d_L$	1.835	2.845	3.480	3.470	3.530
	$d_L < d < d_U$	3.485	2.125	1.710	1.390	1.215
2	$d < d_L$	1.765	2.965	3.420	3.340	3.575
	$d_L < d < d_U$	3.220	1.775	1.625	1.335	1.185
3	$d < d_L$	1.725	2.935	3.500	3.635	3.665
	$d_L < d < d_U$	3.340	2.015	1.760	1.375	1.225
4	$d < d_L$	1.925	3.070	3.395	3.455	3.855
	$d_L < d < d_U$	3.360	2.030	1.705	1.310	1.225
5	$d < d_L$	1.805	2.920	3.455	3.725	3.535
	$d_L < d < d_U$	3.190	1.975	1.395	1.425	1.295
6	$d < d_L$	1.870	2.810	3.475	3.645	3.485
	$d_L < d < d_U$	3.505	2.075	1.685	1.425	1.275
7	$d < d_L$	1.735	3.025	3.380	3.555	3.470
	$d_L < d < d_U$	3.195	2.215	1.715	1.365	1.310
8	$d < d_L$	1.815	2.910	3.405	3.575	3.605
	$d_L < d < d_U$	3.375	2.075	1.695	1.385	1.270
9	$d < d_L$	1.700	3.075	3.415	3.450	3.565
	$d_L < d < d_U$	3.390	1.905	1.790	1.255	1.115

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.8: The Durbin-Watson test - ordered quadratic case.

Specification	Value of d T =	25	75	125	175	225
1	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
2	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
3	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
4	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
5	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
6	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
7	$d < d_L$	99.120	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.670	0.000	0.000	0.000	0.000
8	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
9	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.9: The Durbin-Watson test - ordered square case.

Specification	Value of d	T = 25	75	125	175	225
1	$d < d_L$	99.900	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.095	0.000	0.000	0.000	0.000
2	$d < d_L$	41.180	78.930	83.195	61.925	62.590
	$d_L < d < d_U$	16.465	5.700	3.880	5.145	4.915
3	$d < d_L$	2.335	3.505	4.475	4.415	4.310
	$d_L < d < d_U$	3.725	2.150	2.120	1.615	1.450
4	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
5	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
6	$d < d_L$	99.225	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.590	0.000	0.000	0.000	0.000
7	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
8	$d < d_L$	94.160	99.995	100.000	100.000	100.000
	$d_L < d < d_U$	3.820	0.005	0.000	0.000	0.000
9	$d < d_L$	5.405	16.040	21.275	23.085	33.340
	$d_L < d < d_U$	6.140	6.120	5.240	4.765	4.925
10	$d < d_L$	100.000	100.000	100.000	100.000	100.000
	$d_L < d < d_U$	0.000	0.000	0.000	0.000	0.000
11	$d < d_L$	95.045	99.975	100.000	100.000	100.000
	$d_L < d < d_U$	3.450	0.010	0.000	0.000	0.000
12	$d < d_L$	4.405	12.615	26.910	22.975	64.080
	$d_L < d < d_U$	5.930	5.540	5.990	4.600	4.770

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.10: The Durbin-Watson test - ordered autoregressive case.

Specification	Value of d	T = 25	75	125	175	225		T = 25	75	125	175	225
1	$d < d_L$	4.780	14.920	22.680	30.160	37.495	15	91.495	99.995	99.995	100.000	99.995
	$d_L < d < d_U$	7.005	5.880	6.070	5.605	5.365		3.985	0.000	0.000	0.000	0.000
2	$d < d_L$	51.540	98.605	99.980	100.000	100.000	16	4.745	14.460	22.725	31.160	37.395
	$d_L < d < d_U$	16.090	0.540	0.020	0.000	0.000		6.415	6.175	6.130	5.645	5.075
3	$d < d_L$	91.305	100.000	100.000	100.000	100.000	17	52.175	98.530	99.965	99.995	99.995
	$d_L < d < d_U$	4.295	0.000	0.000	0.000	0.000		16.080	0.585	0.015	0.000	0.000
4	$d < d_L$	4.695	14.580	22.865	30.755	37.505	18	92.380	99.995	99.995	99.995	99.995
	$d_L < d < d_U$	6.730	6.550	5.705	5.595	5.570		3.770	0.000	0.000	0.000	0.000
5	$d < d_L$	51.325	98.540	99.970	100.000	100.000	19	4.780	14.660	23.700	30.855	36.765
	$d_L < d < d_U$	16.530	0.610	0.015	0.000	0.000		6.595	6.310	5.735	5.385	5.520
6	$d < d_L$	91.675	100.000	100.000	100.000	100.000	20	51.385	98.540	99.980	99.995	99.995
	$d_L < d < d_U$	3.940	0.000	0.000	0.000	0.000		16.715	0.615	0.005	0.000	0.000
7	$d < d_L$	4.880	14.705	23.110	30.770	37.615	21	92.840	99.995	99.995	99.995	99.995
	$d_L < d < d_U$	6.760	5.915	6.030	5.595	5.310		3.500	0.000	0.000	0.000	0.000
8	$d < d_L$	51.195	98.465	99.965	100.000	100.000	22	4.485	14.550	23.090	30.400	37.240
	$d_L < d < d_U$	16.065	0.575	0.015	0.000	0.000		6.760	6.390	6.170	5.635	5.375
9	$d < d_L$	91.270	100.000	100.000	100.000	100.000	23	53.275	98.675	99.965	99.995	99.995
	$d_L < d < d_U$	4.265	0.000	0.000	0.000	0.000		15.910	0.580	0.020	0.000	0.000
10	$d < d_L$	4.640	14.430	22.555	30.930	36.940	24	92.100	99.995	99.995	99.995	99.995
	$d_L < d < d_U$	6.305	6.250	5.960	5.445	5.310		4.075	0.000	0.000	0.000	0.000
11	$d < d_L$	51.545	98.640	99.985	99.995	99.995	25	4.690	14.350	23.135	31.275	37.240
	$d_L < d < d_U$	16.070	0.560	0.010	0.000	0.000		6.845	5.740	5.855	5.490	5.310
12	$d < d_L$	91.890	99.995	99.995	99.995	99.995	26	52.950	98.530	99.970	99.995	99.995
	$d_L < d < d_U$	3.915	0.000	0.000	0.000	0.000		15.685	0.615	0.020	0.000	0.000
13	$d < d_L$	4.765	14.835	22.800	31.095	37.395	27	91.790	99.995	99.995	99.995	99.995
	$d_L < d < d_U$	6.670	6.105	5.650	5.665	5.390		4.035	0.000	0.000	0.000	0.000
14	$d < d_L$	51.430	98.500	99.965	99.995	99.995						
	$d_L < d < d_U$	16.055	0.610	0.030	0.000	0.000						

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.11: The Durbin-Watson test - ordered Hamilton case.

Specification	Value of d	T =	25	75	125	175	225
1	$d < d_L$		28.740	94.130	99.725	99.980	100.000
	$d_L < d < d_U$		33.850	3.505	0.155	0.015	0.000
2	$d < d_L$		29.360	94.275	99.665	99.980	99.995
	$d_L < d < d_U$		33.725	3.500	0.190	0.015	0.005
3	$d < d_L$		28.130	94.115	99.645	99.985	100.000
	$d_L < d < d_U$		33.830	3.395	0.210	0.010	0.000
4	$d < d_L$		42.745	99.045	99.995	100.000	100.000
	$d_L < d < d_U$		33.845	0.635	0.000	0.000	0.000
5	$d < d_L$		42.990	98.990	100.000	100.000	100.000
	$d_L < d < d_U$		33.490	0.720	0.000	0.000	0.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of $d < d_L$ and $d_L < d < d_U$.

Table B.12: The Harvey-Collier test - unordered.

Specification	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	5.135	4.985	4.960	4.815	5.090	Auto	1	5.850	6.205	6.670	6.330	6.570
	2	5.205	4.815	5.005	5.080	4.810		2	10.690	12.605	13.270	14.470	14.425
	3	4.915	4.925	4.875	4.935	5.135		3	22.980	27.300	28.205	28.650	29.510
	4	4.645	4.860	4.875	5.175	5.100		4	5.685	6.110	6.360	6.460	6.300
	5	5.025	5.150	5.140	5.010	5.170		5	10.340	12.260	13.070	13.930	14.820
	6	4.700	4.950	4.865	5.155	5.025		6	22.840	27.765	27.985	29.085	29.340
	7	5.115	4.870	4.995	5.170	4.925		7	5.825	6.035	6.605	6.430	6.535
	8	4.920	5.040	5.015	4.950	5.025		8	10.520	12.515	13.225	14.420	13.920
	9	5.385	5.005	4.840	5.070	4.995		9	23.230	27.245	28.590	29.215	28.910
								10	6.700	6.710	6.605	6.600	6.935
								11	13.495	15.070	15.755	16.005	15.750
Quad	1	100.000	100.000	100.000	100.000	100.000		12	26.750	30.895	32.310	31.895	32.515
	2	100.000	100.000	100.000	100.000	100.000		13	6.655	7.030	6.180	6.460	6.505
	3	100.000	100.000	100.000	100.000	100.000		14	10.905	15.640	15.350	16.030	15.835
	4	5.075	88.760	0.000	47.360	0.080		15	27.900	30.970	31.125	32.170	32.235
	5	0.000	99.965	0.000	0.000	0.000		16	6.095	6.485	6.440	6.895	6.630
	6	0.340	0.000	0.000	0.000	43.185		17	13.930	15.175	16.165	15.885	16.035
	7	11.845	7.535	33.030	0.000	0.000		18	27.415	30.835	31.900	32.450	32.940
	8	0.665	0.000	0.000	0.000	0.000		19	6.315	6.675	6.525	6.750	6.705
	9	0.010	0.000	0.000	0.000	97.255		20	13.230	14.555	15.460	15.625	16.345
								21	27.655	30.125	31.715	32.315	32.530
								22	5.980	6.405	6.595	6.820	6.980
Square	1	16.780	33.680	57.760	16.195	0.000		23	12.975	15.410	15.190	16.205	16.000
	2	18.440	17.760	0.480	1.915	1.505		24	27.900	30.680	31.100	31.810	32.440
	3	5.305	6.120	5.220	3.605	5.630		25	6.300	6.655	6.690	6.945	6.745
	4	100.000	100.000	100.000	100.000	100.000		26	13.370	14.910	15.810	15.390	16.185
	5	100.000	100.000	100.000	100.000	100.000		27	27.625	31.245	31.840	32.100	33.180
	6	99.940	100.000	100.000	100.000	100.000							
	7	0.000	0.000	0.000	0.000	0.000							
	8	0.275	2.160	0.030	1.835	0.020							
	9	3.520	6.420	4.470	3.360	2.825							
	10	0.000	0.000	0.380	0.000	8.820							
	11	0.000	2.165	1.010	11.980	0.950							
	12	5.235	2.505	2.935	1.955	2.995							

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.13: The Harvey-Collier test - ordered.

SPECIFICATION	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	4.935	4.870	4.865	4.940	4.825	Auto	1	5.700	6.345	6.560	6.055	6.140
	2	4.890	4.945	4.970	4.915	5.045		2	10.840	12.535	13.470	14.165	14.595
	3	4.565	5.075	4.815	4.885	5.180		3	22.730	27.830	28.240	28.885	29.080
	4	5.155	5.055	4.870	4.750	4.890		4	5.820	6.200	6.235	6.680	6.460
	5	4.910	4.985	4.990	5.070	4.765		5	10.565	12.460	13.460	13.890	14.130
	6	4.670	5.005	4.905	4.985	4.975		6	22.925	27.575	28.035	28.505	29.270
	7	5.065	5.170	5.200	5.105	4.680		7	6.050	6.590	6.355	6.570	6.285
	8	5.220	4.880	5.080	5.030	5.205		8	10.490	12.835	13.240	13.865	14.150
	9	4.940	4.980	5.060	5.115	5.160		9	23.125	27.555	28.385	28.670	29.715
								10	4.635	4.725	4.990	4.965	5.280
								11	6.165	6.190	5.390	4.980	5.360
Quad	1	99.540	100.000	100.000	100.000	100.000		12	3.555	5.780	7.175	1.730	5.280
	2	100.000	100.000	100.000	100.000	100.000		13	5.520	5.525	4.835	5.280	5.070
	3	99.980	100.000	100.000	100.000	100.000		14	6.605	7.915	4.220	4.235	5.230
	4	100.000	100.000	100.000	100.000	100.000		15	11.090	4.005	3.830	4.235	4.680
	5	100.000	100.000	100.000	100.000	100.000		16	5.155	5.300	4.715	5.160	5.290
	6	100.000	100.000	100.000	100.000	100.000		17	5.370	5.830	5.305	4.425	5.890
	7	100.000	100.000	100.000	100.000	100.000		18	3.775	1.495	10.260	4.260	6.795
	8	100.000	100.000	100.000	100.000	100.000		19	5.075	4.700	5.015	5.175	4.850
	9	99.995	100.000	100.000	100.000	100.000		20	6.430	4.960	4.570	4.645	4.885
								21	8.505	3.720	2.905	2.265	3.805
								22	5.095	4.785	4.995	4.670	4.780
Square	1	97.885	100.000	100.000	100.000	100.000		23	5.440	5.420	5.055	7.255	4.635
	2	15.545	79.775	86.865	99.855	99.370		24	11.105	2.000	3.590	4.265	2.490
	3	8.765	15.965	33.865	20.685	24.620		25	5.065	5.360	5.155	4.955	5.185
	4	100.000	100.000	100.000	100.000	100.000		26	4.890	4.040	5.195	4.525	4.645
	5	100.000	100.000	100.000	100.000	100.000		27	8.215	3.555	6.290	4.345	5.770
	6	99.915	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	87.560	100.000	100.000	100.000	100.000							
	9	38.160	51.755	76.450	88.425	96.785							
	10	100.000	100.000	100.000	100.000	100.000							
	11	72.585	100.000	100.000	100.000	100.000							
	12	25.250	49.305	86.505	90.205	95.990							

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.14: The Ramsey Reset test.

SPECIFICATION	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	4.990	5.085	5.315	5.050	5.175	Auto	1	6.765	7.320	7.410	7.570	7.555
	2	4.940	4.885	5.045	4.810	5.060		2	22.840	25.200	25.740	25.680	26.035
	3	5.010	5.410	5.095	4.915	4.910		3	52.140	62.680	63.895	64.190	65.010
	4	5.140	5.140	5.205	5.000	5.025		4	7.230	7.400	7.310	7.160	7.775
	5	4.815	4.870	5.070	5.135	5.290		5	22.550	24.975	25.490	25.020	25.065
	6	5.200	5.000	5.145	4.925	4.905		6	52.110	62.505	64.495	64.655	64.885
	7	4.855	5.010	4.730	4.830	4.740		7	7.040	7.645	7.395	7.710	7.020
	8	4.915	5.095	4.955	5.260	4.880		8	22.815	25.300	24.875	25.695	25.590
	9	4.960	4.745	5.075	5.040	5.150		9	52.340	63.040	63.850	64.580	65.170
								10	4.290	5.080	5.260	5.025	4.685
								11	6.415	5.915	7.735	3.760	5.550
Quad	1	100.000	100.000	100.000	100.000	100.000		12	2.575	2.610	3.265	6.925	9.235
	2	100.000	100.000	100.000	100.000	100.000		13	4.515	4.980	4.860	4.970	5.045
	3	100.000	100.000	100.000	100.000	100.000		14	6.260	7.575	2.955	5.790	3.800
	4	100.000	100.000	100.000	100.000	100.000		15	1.955	10.645	4.475	5.290	4.305
	5	100.000	100.000	100.000	100.000	100.000		16	4.930	4.615	5.085	5.180	4.605
	6	100.000	100.000	100.000	100.000	100.000		17	7.390	4.280	4.705	5.705	3.650
	7	100.000	100.000	100.000	100.000	100.000		18	2.900	5.245	0.400	6.595	2.970
	8	100.000	100.000	100.000	100.000	100.000		19	4.965	4.755	4.790	5.195	4.650
	9	100.000	100.000	100.000	100.000	100.000		20	2.825	7.465	5.040	3.220	4.415
								21	1.345	1.505	12.435	9.950	5.810
								22	4.615	4.760	5.130	5.135	4.865
Square	1	97.885	100.000	100.000	100.000	100.000		23	4.635	3.980	4.290	5.245	5.625
	2	8.475	98.210	98.140	99.980	100.000		24	13.160	1.825	0.220	5.310	8.310
	3	6.905	12.455	16.565	68.895	45.500		25	4.560	5.105	4.630	4.880	4.825
	4	100.000	100.000	100.000	100.000	100.000		26	2.450	8.240	5.440	6.580	5.270
	5	100.000	100.000	100.000	100.000	100.000		27	2.895	10.040	0.920	1.140	10.600
	6	100.000	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	100.000	100.000	100.000	100.000	100.000	Ham	1	21.435	56.135	79.200	90.840	96.205
	9	13.735	67.475	83.745	97.690	99.555		2	22.655	58.685	80.305	91.740	96.920
	10	100.000	100.000	100.000	100.000	100.000		3	12.265	31.150	50.960	66.315	77.445
	11	99.995	100.000	100.000	100.000	100.000		4	25.950	54.605	73.605	84.935	91.140
	12	16.795	84.120	96.030	99.590	99.980		5	16.400	39.675	56.750	69.985	79.790

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.15: The $\lambda_H^E(g)$ test.

Specification	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	4.570	5.075	4.555	4.895	5.105	Auto	1	8.125	9.270	9.420	9.500	9.710
	2	4.440	5.085	4.880	4.925	4.775		2	41.780	51.080	54.055	54.470	54.820
	3	4.640	4.955	4.980	4.810	4.875		3	80.950	96.985	98.640	99.240	99.580
	4	4.545	4.630	4.780	4.905	4.980		4	7.975	9.620	9.600	9.580	9.870
	5	4.585	4.875	5.080	5.055	4.815		5	41.965	50.755	53.600	54.405	54.625
	6	4.690	4.740	5.035	5.080	4.885		6	81.150	97.005	98.805	99.270	99.510
	7	4.480	4.920	4.905	4.805	4.535		7	8.140	9.230	9.850	9.705	9.220
	8	4.545	4.795	5.050	4.865	4.960		8	41.945	51.245	53.385	53.880	54.925
	9	4.115	4.915	5.005	4.995	4.820		9	81.185	96.945	98.705	99.250	99.565
								10	4.560	4.600	4.970	4.880	5.045
								11	3.195	5.250	2.975	3.645	5.840
Quad	1	100.000	100.000	100.000	100.000	100.000		12	4.210	4.250	1.260	10.035	6.975
	2	100.000	100.000	100.000	100.000	100.000		13	4.600	5.445	4.630	4.885	5.125
	3	100.000	100.000	100.000	100.000	100.000		14	3.010	4.115	4.465	3.995	5.630
	4	100.000	100.000	100.000	100.000	100.000		15	1.260	3.760	2.210	1.555	4.465
	5	100.000	100.000	100.000	100.000	100.000		16	4.475	4.820	4.925	5.180	5.000
	6	100.000	100.000	100.000	100.000	100.000		17	3.555	5.170	4.240	5.300	5.575
	7	100.000	100.000	100.000	100.000	100.000		18	2.040	4.150	6.705	1.865	3.040
	8	100.000	100.000	100.000	100.000	100.000		19	4.260	4.570	4.930	4.980	4.825
	9	100.000	100.000	100.000	100.000	100.000		20	5.285	5.315	4.435	4.715	5.090
								21	2.940	7.985	5.455	3.885	2.470
								22	4.510	5.000	4.825	5.105	5.095
Square	1	100.000	100.000	100.000	100.000	100.000		23	5.820	5.165	5.115	3.565	6.120
	2	20.480	44.365	95.745	97.665	97.175		24	1.340	4.955	2.105	1.975	3.980
	3	4.740	8.315	8.215	16.325	22.330		25	4.150	4.945	4.410	5.195	4.965
	4	100.000	100.000	100.000	100.000	100.000		26	1.655	6.040	3.940	4.290	4.885
	5	100.000	100.000	100.000	100.000	100.000		27	7.845	1.995	14.585	4.975	8.800
	6	99.995	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	99.530	100.000	100.000	100.000	100.000	Ham	1	28.555	92.895	99.765	99.990	100.000
	9	13.100	52.190	79.050	92.250	94.905		2	30.715	93.040	99.855	99.995	100.000
	10	100.000	100.000	100.000	100.000	100.000		3	28.735	91.975	99.840	100.000	100.000
	11	94.760	99.710	100.000	100.000	100.000		4	39.160	98.330	99.975	100.000	100.000
	12	12.960	34.385	70.200	95.050	96.905		5	36.830	98.395	99.995	100.000	100.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.16: The λ_{OP}^{A} test - bootstrapped p-values.

SPECIFICATION	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	4.500	5.400	5.100	3.900	6.300	Auto	1	8.100	9.800	8.300	10.200	8.600
	2	3.800	4.300	4.600	4.200	4.900		2	37.000	49.100	51.400	52.500	53.500
	3	5.400	5.100	4.500	4.900	4.900		3	74.800	96.800	99.300	99.900	99.600
	4	4.100	4.700	5.000	4.100	4.900		4	6.400	8.900	9.600	8.900	8.700
	5	3.800	3.800	3.500	4.900	4.400		5	39.800	50.600	53.200	51.700	54.500
	6	5.000	5.100	4.300	5.100	4.900		6	75.800	96.900	99.200	99.600	99.700
	7	4.700	4.500	5.800	4.700	5.000		7	7.400	6.700	9.500	9.900	10.700
	8	5.000	5.400	4.200	5.400	3.500		8	37.800	48.900	49.800	54.200	52.700
	9	4.700	5.900	3.300	4.100	4.900		9	75.600	98.300	99.200	99.800	99.900
								10	5.600	5.500	4.900	4.500	5.800
								11	3.500	4.900	6.100	4.600	5.200
Quad	1	100.000	100.000	100.000	100.000	100.000		12	6.700	2.500	4.500	4.400	5.000
	2	100.000	100.000	100.000	100.000	100.000		13	4.600	4.100	5.500	4.900	5.900
	3	100.000	100.000	100.000	100.000	100.000		14	5.600	9.400	3.400	4.000	5.000
	4	100.000	100.000	100.000	100.000	100.000		15	0.700	13.300	1.300	2.900	6.100
	5	100.000	100.000	100.000	100.000	100.000		16	4.000	4.200	4.000	5.800	4.800
	6	100.000	100.000	100.000	100.000	100.000		17	5.800	3.900	3.700	4.500	4.600
	7	100.000	100.000	100.000	100.000	100.000		18	3.400	3.100	9.300	2.700	9.700
	8	100.000	100.000	100.000	100.000	100.000		19	4.700	4.900	5.000	3.700	4.400
	9	100.000	100.000	100.000	100.000	100.000		20	12.600	5.100	3.400	2.500	4.600
								21	4.400	6.500	5.000	1.000	3.900
								22	7.100	4.500	5.600	6.200	4.500
Square	1	99.500	100.000	100.000	100.000	100.000		23	3.300	4.000	5.400	4.000	4.700
	2	25.500	61.300	79.000	100.000	99.900		24	8.300	7.500	4.900	1.900	7.100
	3	9.700	13.400	10.200	31.600	24.500		25	5.800	5.000	4.800	3.900	4.900
	4	100.000	100.000	100.000	100.000	100.000		26	4.200	3.300	4.000	3.900	8.000
	5	100.000	100.000	100.000	100.000	100.000		27	0.900	12.000	1.700	2.500	10.100
	6	100.000	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	100.000	100.000	100.000	100.000	100.000	Ham	1	36.900	97.800	100.000	100.000	100.000
	9	27.200	57.900	91.800	94.200	97.900		2	39.800	98.500	100.000	100.000	100.000
	10	100.000	100.000	100.000	100.000	100.000		3	38.200	97.300	100.000	100.000	100.000
	11	100.000	100.000	100.000	100.000	100.000		4	52.400	99.700	100.000	100.000	100.000
	12	37.000	51.900	98.400	100.000	99.800		5	52.700	99.700	100.000	100.000	100.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.17: The λ_{OP}^{A} test - asymptotic p-values.

SPECIFICATION	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	1.600	1.635	1.610	1.495	1.345	Auto	1	3.180	3.265	3.435	3.480	3.550
	2	1.785	1.545	1.550	1.415	1.425		2	22.785	32.365	34.125	35.805	35.815
	3	1.555	1.510	1.520	1.785	1.600		3	65.660	94.275	98.085	98.955	99.410
	4	1.730	1.185	1.165	1.760	1.345		4	3.005	3.455	3.615	3.260	3.350
	5	1.165	1.290	1.550	1.370	1.375		5	23.005	32.290	34.700	35.220	35.380
	6	0.810	1.395	1.665	1.235	1.355		6	65.940	94.305	97.990	99.070	99.370
	7	1.030	1.470	0.955	1.585	1.285		7	3.165	3.240	3.300	3.385	3.405
	8	2.150	1.260	1.375	1.085	1.375		8	23.295	32.185	34.320	35.375	35.785
	9	1.155	1.305	0.820	1.185	1.105		9	65.590	94.150	97.890	98.995	99.290
								10	1.735	1.745	1.775	1.510	1.595
								11	2.680	1.230	1.460	1.820	0.770
Quad	1	100.000	100.000	100.000	100.000	100.000		12	0.290	1.060	0.890	1.720	0.455
	2	100.000	100.000	100.000	100.000	100.000		13	1.730	1.270	1.345	1.435	1.505
	3	100.000	100.000	100.000	100.000	100.000		14	3.920	0.965	1.570	1.515	1.540
	4	100.000	100.000	100.000	100.000	100.000		15	0.810	0.205	3.555	0.915	0.340
	5	100.000	100.000	100.000	100.000	100.000		16	1.100	1.585	1.745	1.485	1.350
	6	100.000	100.000	100.000	100.000	100.000		17	0.660	1.805	1.655	2.260	1.040
	7	100.000	100.000	100.000	100.000	100.000		18	0.155	2.105	1.505	0.665	0.125
	8	100.000	100.000	100.000	100.000	100.000		19	1.645	1.265	0.975	1.005	1.080
	9	100.000	100.000	100.000	100.000	100.000		20	1.570	1.265	2.455	1.485	1.440
								21	0.350	3.965	2.405	0.215	1.435
								22	1.135	0.920	1.925	1.230	0.940
Square	1	99.860	100.000	100.000	100.000	100.000		23	2.010	2.060	2.080	0.645	0.690
	2	19.005	97.175	80.705	86.015	99.940		24	1.275	0.020	0.175	0.395	0.905
	3	4.225	2.050	6.040	6.980	13.155		25	0.975	1.200	1.160	1.180	1.155
	4	100.000	100.000	100.000	100.000	100.000		26	1.365	1.000	0.610	1.035	1.290
	5	100.000	100.000	100.000	100.000	100.000		27	3.045	0.345	3.075	0.270	0.805
	6	100.000	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	99.830	100.000	100.000	100.000	100.000	Ham	1	28.005	95.290	99.980	100.000	100.000
	9	12.610	45.925	81.605	91.550	92.590		2	27.620	95.230	99.945	100.000	100.000
	10	100.000	100.000	100.000	100.000	100.000		3	27.955	95.550	99.970	100.000	100.000
	11	99.990	100.000	100.000	100.000	100.000		4	37.500	99.115	100.000	100.000	100.000
	12	5.950	28.305	92.340	94.890	98.700		5	36.975	98.865	100.000	100.000	100.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.18: The $\lambda^{E}_{OP}(g)$ test - bootstrapped p-values.

SPECIFICATION	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	4.100	5.500	4.800	4.500	4.800	Auto	1	4.000	4.700	4.300	4.200	4.700
	2	5.100	5.600	4.900	5.400	5.000		2	28.200	52.700	55.800	54.800	58.300
	3	3.900	4.600	4.500	3.400	6.000		3	76.200	99.500	100.000	99.900	99.900
	4	5.300	5.100	6.600	5.800	3.700		4	2.900	3.900	4.200	3.700	5.000
	5	5.600	3.800	5.000	4.700	4.900		5	28.400	50.200	53.600	54.700	58.300
	6	5.200	5.000	4.900	5.000	5.300		6	74.600	99.000	99.700	99.900	100.000
	7	7.000	3.600	5.300	5.700	5.300		7	4.400	4.800	4.600	5.700	4.000
	8	5.100	5.100	5.300	4.300	5.900		8	30.200	46.800	55.600	56.600	58.600
	9	4.900	5.300	4.900	5.500	5.400		9	78.700	98.900	99.800	100.000	100.000
								10	5.200	5.600	4.300	5.100	5.300
								11	3.400	5.400	5.900	3.500	4.300
Quad	1	100.000	100.000	100.000	100.000	100.000		12	2.900	3.000	5.400	5.100	4.400
	2	100.000	100.000	100.000	100.000	100.000		13	4.000	4.800	4.900	5.900	5.300
	3	100.000	100.000	100.000	100.000	100.000		14	4.000	4.100	5.300	3.900	6.500
	4	100.000	100.000	100.000	100.000	100.000		15	5.300	5.500	4.400	4.400	5.100
	5	100.000	100.000	100.000	100.000	100.000		16	5.400	4.700	4.200	4.400	4.100
	6	100.000	100.000	100.000	100.000	100.000		17	5.900	5.200	4.000	5.400	5.500
	7	99.800	100.000	100.000	100.000	100.000		18	14.100	5.800	6.200	3.900	6.600
	8	100.000	100.000	100.000	100.000	100.000		19	4.600	4.300	4.900	4.800	4.400
	9	100.000	100.000	100.000	100.000	100.000		20	4.200	5.500	4.600	4.100	5.600
								21	4.400	3.000	3.800	8.000	4.100
								22	4.100	4.200	6.100	6.000	4.900
Square	1	69.100	100.000	100.000	100.000	100.000		23	3.900	5.300	6.200	4.800	5.100
	2	4.800	23.500	53.900	79.200	92.500		24	6.500	3.600	5.000	4.500	4.300
	3	4.000	5.200	4.400	5.800	5.300		25	4.000	5.500	4.800	5.400	5.000
	4	100.000	100.000	100.000	100.000	100.000		26	5.500	4.600	4.300	4.700	4.300
	5	100.000	100.000	100.000	100.000	100.000		27	6.200	9.000	3.900	5.200	4.400
	6	99.700	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	62.400	100.000	100.000	100.000	100.000	Ham	1	8.600	76.700	98.400	100.000	100.000
	9	3.900	38.900	59.200	77.300	80.800		2	11.200	79.900	98.400	100.000	100.000
	10	100.000	100.000	100.000	100.000	100.000		3	7.900	78.400	99.200	100.000	100.000
	11	70.600	100.000	100.000	100.000	100.000		4	17.200	92.900	99.900	100.000	100.000
	12	3.700	13.200	61.600	63.100	71.000		5	15.200	93.100	99.900	100.000	100.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.19: The $\lambda_{OP}^{E}(g)$ test - asymptotic p-values.

Specification	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	1.285	0.650	0.690	0.615	0.560	Auto	1	1.465	1.935	2.315	2.130	2.245
	2	1.280	0.765	0.575	0.655	0.525		2	23.960	43.445	47.420	49.695	50.400
	3	1.230	0.610	0.565	0.545	0.640		3	72.620	98.615	99.755	99.910	99.980
	4	1.355	0.250	0.355	0.590	0.630		4	1.625	1.985	2.140	2.265	2.160
	5	0.360	0.465	0.355	0.610	0.640		5	23.970	43.625	47.765	49.555	50.750
	6	1.205	0.725	0.680	0.640	0.530		6	72.080	98.660	99.790	99.925	99.985
	7	1.155	0.870	0.980	0.825	0.880		7	1.705	1.825	2.070	2.080	2.195
	8	0.655	1.010	0.840	0.685	0.835		8	24.015	44.095	46.980	49.320	49.905
	9	0.350	0.755	0.895	0.800	1.035		9	72.770	98.720	99.770	99.930	99.975
								10	1.080	0.830	0.520	0.500	0.715
								11	0.255	0.635	1.110	0.510	0.325
Quad	1	100.000	100.000	100.000	100.000	100.000		12	0.020	0.010	1.005	0.065	0.940
	2	100.000	100.000	100.000	100.000	100.000		13	0.645	0.765	0.795	0.760	0.540
	3	100.000	100.000	100.000	100.000	100.000		14	0.435	0.345	1.320	0.510	0.705
	4	100.000	100.000	100.000	100.000	100.000		15	0.340	1.090	0.160	0.115	2.430
	5	100.000	100.000	100.000	100.000	100.000		16	0.855	0.700	0.415	0.660	0.640
	6	99.990	100.000	100.000	100.000	100.000		17	0.580	0.820	0.875	0.480	0.575
	7	72.040	100.000	100.000	100.000	100.000		18	0.175	0.095	0.075	0.295	0.710
	8	76.575	100.000	100.000	100.000	100.000		19	0.830	0.790	0.640	0.945	0.935
	9	59.415	100.000	100.000	100.000	100.000		20	0.575	1.900	2.215	1.405	0.975
								21	0.470	0.055	0.105	0.965	0.825
								22	1.700	0.760	0.970	0.890	0.960
Square	1	86.745	100.000	100.000	100.000	100.000		23	0.510	0.705	1.000	0.930	1.195
	2	6.050	19.300	49.470	96.045	75.595		24	0.160	0.555	1.675	0.530	0.295
	3	0.910	1.295	1.240	2.415	3.435		25	0.845	1.085	0.955	0.770	0.865
	4	100.000	100.000	100.000	100.000	100.000		26	0.570	0.865	0.975	1.875	1.105
	5	100.000	100.000	100.000	100.000	100.000		27	0.405	2.655	1.270	1.815	0.865
	6	99.015	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	74.635	100.000	100.000	100.000	100.000	Ham	1	10.745	78.515	98.845	99.970	100.000
	9	2.190	25.475	51.870	78.375	92.570		2	12.390	77.355	98.680	99.960	100.000
	10	100.000	100.000	100.000	100.000	100.000		3	10.890	78.690	98.940	99.960	100.000
	11	95.475	99.935	100.000	100.000	100.000		4	17.970	92.705	99.945	100.000	100.000
	12	1.910	13.080	61.425	20.395	49.750		5	16.275	93.460	99.955	100.000	100.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.20: The g_{OP} test - bootstrapped p-values.

Specification	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	4.700	5.500	5.200	5.000	5.800	Auto	1	3.900	3.300	3.000	2.900	4.000
	2	5.400	5.700	4.700	5.800	5.600		2	10.600	17.700	15.800	16.900	19.500
	3	3.400	4.100	6.200	4.400	5.800		3	38.000	73.600	80.900	84.100	87.800
	4	6.000	6.300	5.400	5.500	5.400		4	4.600	4.100	4.700	3.800	2.500
	5	4.800	5.400	5.300	5.700	6.600		5	8.400	17.000	21.600	19.500	22.000
	6	5.000	5.700	4.200	3.500	5.700		6	36.100	71.700	79.800	85.800	86.900
	7	5.600	5.400	4.800	5.100	4.600		7	3.700	3.800	2.400	3.400	2.700
	8	5.200	4.600	5.000	5.000	4.800		8	10.600	17.400	17.800	18.500	17.900
	9	4.600	4.700	4.300	5.000	5.200		9	35.800	73.800	81.200	84.600	86.000
								10	4.700	5.500	3.900	4.500	5.600
								11	4.400	4.200	5.000	5.200	4.400
Quad	1	100.000	100.000	100.000	100.000	100.000		12	2.000	2.200	4.700	6.300	6.800
	2	100.000	100.000	100.000	100.000	100.000		13	4.300	4.300	5.700	5.900	4.600
	3	100.000	100.000	100.000	100.000	100.000		14	7.000	4.600	5.300	4.800	6.200
	4	69.000	100.000	100.000	100.000	100.000		15	1.400	7.800	6.700	7.100	7.000
	5	100.000	100.000	100.000	100.000	100.000		16	5.900	5.700	3.800	4.900	5.500
	6	100.000	100.000	100.000	100.000	100.000		17	5.900	5.300	4.900	5.500	5.700
	7	100.000	100.000	100.000	100.000	100.000		18	4.300	5.000	2.700	4.200	5.300
	8	67.400	100.000	100.000	100.000	100.000		19	5.800	4.600	4.300	4.500	5.500
	9	100.000	100.000	100.000	100.000	100.000		20	5.000	4.100	4.400	4.900	4.800
								21	2.700	4.400	3.600	6.500	6.100
								22	4.300	6.000	5.000	5.200	5.300
Square	1	100.000	100.000	100.000	100.000	100.000		23	5.400	5.000	5.300	5.400	4.300
	2	15.400	95.500	98.000	99.200	100.000		24	4.500	6.500	9.500	5.000	3.500
	3	4.600	22.200	11.300	23.700	27.300		25	5.200	3.800	6.000	5.100	4.900
	4	100.000	100.000	100.000	100.000	100.000		26	6.700	4.400	4.500	5.000	4.500
	5	100.000	100.000	100.000	100.000	100.000		27	2.700	3.700	5.700	3.100	3.300
	6	92.900	100.000	100.000	100.000	100.000							
	7	100.000	100.000	100.000	100.000	100.000							
	8	93.800	100.000	100.000	100.000	100.000	Ham	1	8.400	54.500	93.900	99.800	100.000
	9	3.600	20.700	64.500	71.000	90.100		2	7.500	58.000	94.700	99.900	100.000
	10	100.000	100.000	100.000	100.000	100.000		3	8.000	45.400	89.000	98.500	100.000
	11	82.700	100.000	100.000	100.000	100.000		4	14.400	85.100	99.700	100.000	100.000
	12	3.300	21.200	68.500	85.000	95.300		5	13.100	85.600	99.800	100.000	100.000

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

Table B.21: The g_{OP} test - asymptotic p-values.

Specification	T =	25	75	125	175	225		T =	25	75	125	175	225
Null	1	0.000	0.000	0.000	0.000	0.000	Auto	1	0.000	0.000	0.000	0.000	0.000
	2	0.000	0.000	0.000	0.000	0.000		2	0.000	0.250	0.475	0.635	0.725
	3	0.000	0.000	0.000	0.000	0.000		3	0.000	27.870	42.055	48.305	51.960
	4	0.000	0.000	0.000	0.000	0.000		4	0.000	0.000	0.000	0.000	0.005
	5	0.000	0.000	0.000	0.000	0.000		5	0.000	0.265	0.500	0.685	0.760
	6	0.000	0.000	0.000	0.000	0.000		6	0.000	28.540	41.490	47.885	52.125
	7	0.000	0.000	0.000	0.000	0.000		7	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000		8	0.000	0.315	0.495	0.680	0.775
	9	0.000	0.000	0.000	0.000	0.000		9	0.000	28.515	42.180	48.550	51.625
								10	0.000	0.000	0.000	0.000	0.000
								11	0.000	0.000	0.000	0.005	0.000
Quad	1	0.000	100.000	100.000	100.000	100.000		12	0.000	0.000	0.000	0.000	0.000
	2	0.000	100.000	100.000	100.000	100.000		13	0.000	0.000	0.000	0.000	0.000
	3	0.000	100.000	100.000	100.000	100.000		14	0.000	0.000	0.000	0.000	0.000
	4	0.000	100.000	100.000	100.000	100.000		15	0.000	0.000	0.000	0.000	0.000
	5	0.000	100.000	100.000	100.000	100.000		16	0.000	0.000	0.000	0.000	0.000
	6	0.000	100.000	100.000	100.000	100.000		17	0.000	0.000	0.000	0.000	0.000
	7	0.025	100.000	100.000	100.000	100.000		18	0.000	0.000	0.000	0.000	0.000
	8	0.000	100.000	100.000	100.000	100.000		19	0.000	0.000	0.000	0.000	0.005
	9	4.130	100.000	100.000	100.000	100.000		20	0.000	0.000	0.000	0.000	0.000
								21	0.000	0.000	0.000	0.000	0.000
								22	0.000	0.000	0.000	0.005	0.000
Square	1	44.315	100.000	100.000	100.000	100.000		23	0.000	0.000	0.000	0.000	0.000
	2	11.400	73.900	96.250	99.985	99.485		24	0.000	0.000	0.000	0.000	0.000
	3	0.535	1.960	4.240	4.995	6.010		25	0.000	0.000	0.000	0.000	0.000
	4	0.000	100.000	100.000	100.000	100.000		26	0.000	0.000	0.000	0.000	0.000
	5	0.000	100.000	100.000	100.000	100.000		27	0.000	0.000	0.000	0.000	0.000
	6	0.000	99.985	100.000	100.000	100.000							
	7	0.000	100.000	100.000	100.000	100.000							
	8	0.000	99.305	100.000	100.000	100.000	Ham	1	0.000	0.035	0.905	8.940	36.155
	9	0.000	0.055	0.355	2.300	9.030		2	0.000	0.010	0.905	10.560	39.015
	10	0.000	100.000	100.000	100.000	100.000		3	0.000	0.030	0.675	6.460	27.215
	11	0.140	99.705	100.000	100.000	100.000		4	0.005	1.060	18.920	67.540	95.515
	12	0.000	4.600	11.010	5.185	59.220		5	0.005	0.905	18.205	65.980	94.985

Note: this table contains the simulated sizes of this test, i.e., the percentage frequencies of rejection of the null hypotheses.

B.2 Simulation Code

This code, along with the other files, produces the Monte Carlo simulations used in Chapter 5. In its current set-up, it simulates datasets with five different sample sizes and uses 20,000 Monte Carlo trials, saving the results in a $20,000 \times 5$ matrix as an `*.fmt` file in the working directory.

Monte Carlo Code

```
new;
max_n = 225; min_n = 25;
```
The minimum number of observations to be used in the model is min_n and the maximum is max_n
```
step_n = 50;
```
The increment in number of observations between experiments is step_n
```
n = min_n;
mc_trials = 20000;
```
The number of Monte Carlo Trials per experiment
```
exp_n = (((max_n - min_n) / step_n) + 1);
```
exp_n denotes the total number of experiments
```
resultsmatrix = zeros(mc_trials,exp_n);
```
The matrix where results from each simulation will be stored
```
do while exp_n >= 1;
```
This loop command controls and repeats the Monte Carlo simulation until all experiments are complete

Enter the FIXED parameters HERE

```
a = 0;
b1 = ones(n, 1);
b2 = 0.25;
/* b2 = 0.50; */
/* b2 = 0.75; */
x = (seqa(1, 1, n) / n) * 20; b0 = 1.5; /* 4.0, 8.0, 1, 25, 400 */
/* x = (rndu(n, 1) * 9) + 1; b0 = 0.4; 0.7, 1.6, 1 25, 400 */
/* x = 5.5 + ((sqrt(5.0625)) * rndn(n, 1); b0 = 0.3; 0.8, 1.5, 1 ,25, 400 */

index = 1;
```
The index number for the Monte Carlo simulation
```
resultsvector = zeros(mc_trials, 1);
```
This is the vector of results for each Monte Carlo experiment
```
do until index > mc_trials;
```
A loop to carry out the test until mc_trails number of simulations are complete

Enter the VARIABLE parameters HERE

```
e = a + (sqrt(b0) * rndn(n, 1));
```

Enter the Model Specification HERE
```
y = b1 + b2 * x + e; /* Null */
y = b1 + b2 * x + b2 * x^2 + e; /* Quadratic */
y = x^2 + e; x0 = rndn(1, 1); e0 = rndn(n, 1); x = recserar(e0, x0, 0.6); /*
Square */
```

```
y = b1 + b2 * x + u; e = a + (sqrt(b0) * rndn(n, 1); e0 = a + (sqrt(b0) *
rndn(1, 1); u = recserar(e,e0,rho); rho = 0.1, 0.5, 0.9; /* Autoregressive
*/
y = 0.6 * x[., 1] .* (x1[., 1] .> 0) + 0.2 * x2[., 2] + rndn(n, 1); /*
Hamilton */
```

Enter Testing Procedure Code HERE

```
resultsvector[index, 1] = result;
```

Collect the result from each test into the results vector for that experiment

```
index = index + 1;
```

Increments the Monte Carlo Simulation index number by 1

```
endo;
```

Completes the Monte Carlo simulation for the given sample size, n

```
resultsmatrix[1:mc_trials, 6-exp_n] = resultsvector;
```

Collects the results of the experiment into the results matrix

```
if n < max_n;
n = n + step_n;
```

If n is less than the maximum, the simulation moves to the next experiment where n equals n plus its increment, step_n

```
else;
n = min_n;
endif;
```

This stops the simulation when n reaches its preset minimum value

```
exp_n = exp_n - 1;
```

Increments the experiment number while n $\leq$ max_n

```
endo;
save resultsmatrix;
clear all;
```

The Tests

The Durbin-Watson Test

```
_olsres = 1;
screen off; output off;
{vnam,m,slopes,stb,vc,stderr,sigma,cx,r2,resid,dwstat} = ols(0, y, x);
```

OLS command, without printed results, to estimate the DW Statistic

```
screen on; output on;
```

The Harvey-Collier Test

```
#include hc;
{t} = recrsidb(zz);
hcn_1[index, 1] = t;
```

The *proc* hc is defined below: [1]

```
proc (1) = recrsidb(zz);
local
```

[1] Thanks to Curt Wells, Lund University for making his recursive residual code available.

```
t,f1,f2,l1,l2,phitest,nl,r,m,n,j,p,p1,b,bt,bt1,z,yt,xt,xxt,xxt1,vv,vt,ft;
p = cols(zz);
```
The number of independent variables, including constant
```
p1 = p + 1;
```
The number of the dependent variable
```
m = p;
n = rows(zz);
```
This reads length internally
```
b = zeros(n, p);
vv = zeros(n, 1);
z = ones(n, 1) ~ zz;
r = (n - m);
nl = (n - r);
```
The initial OLS estimate
```
xxt1 = inv(z[1:m, 1:p]' * z[1:m, 1:p]);
bt1 = xxt1 * z[1:m, 1:p]' * z[1:m, p1];
b[m,.]  = bt1';
j = m + 1;
```
The loop to calculate the b's and the recursive residuals
```
do until j > n;
xt = z[j, 1:p]; yt = z[j, p1];
ft = 1 + xt * xxt1 * xt';
xxt = xxt1 - xxt1 * xt' * xt * xxt1 / ft;
vt = yt - xt * bt1;
bt = bt1 + xxt1 * xt' * vt / ft;
```
Update and save
```
vv[j, 1] = vt / sqrt(ft);
b[j,.]  = bt';
xxt1 = xxt;
bt1 = bt;
j = j + 1;
endo;
f1 = vv[nl + 1:n, 1];
f2 = meanc(f1);
l2 = sumc(f1);
l1 = sumc((f1 - f2)^2);
phitest = (((r - 1)^-1 * l1)^-0.5) * ((r)^-0.5) * l2;
t = cdftc(phitest, r - 1);
```

The Ramsey Reset Test

```
j = 1; /* or 2 or 3, depending on the model specification */
```
The number of columns of x
```
r_old = zeros(1, 1);
r_new = zeros(1, 1);
```
Initialises the two vectors for holding the R-squared values
```
screen off; output off;
{vnam,m,b,stb,vc,stderr,sigma,cx,rsq,resid,dwstat} = ols(0, y, x);
```
Run OLS on x and y
```
screen on; output on;
r_old = rsq;
```
Record the R-Squared value
```
yhat = (b1 ~ x) * b;
```

```
yhat2 = yhat^2;
```
Square those values
```
x1 = x ~ yhat2;
screen off; output off;
{vnam,m,b,stb,vc,stderr,sigma,cx,r2,resid,dwstat } = ols(0, y, x1);
```
Recalculate R-squared for the new x vector, including the estimates of Yhat
```
screen on; output on;
r_new = r2;
```
Assign the new R-squared value
```
F = (((r_new - r_old) / 1)) / ((1 - r_new) / (n- (j + 1)));
```
Calculate the F-Statistic
```
p = cdffc(F, 1, (n - (j + 1)));
```
Calculate the probability associated with the F-Statistic
```
rrn_1[index, 1] = p;
```
Places the results in the vector res_1

The $\lambda_H^E(g)$ test

This code is adapted from Hamilton (2001). The two included *procs* are unchanged.

```
kqopt = 2;
#include dist2;
#include covary;
index = 1;
```
Index is the index for each Monte Carlo simulation
```
lmn_1 = zeros(mc_trials, 1);
```
This is the vector of results for each Monte Carlo simulation
```
do until index > mc_trials;
```
A loop to carry out the test until mc_trails number of simulations are complete
```
n = rows (x1);
k = cols (x1);
screen off; output off;
_olsres=1;
{vnam,m,slopes,stb,vc,stderr,sigma,cx,r2,resid,dwstat} = ols(0, y, x1);
screen on;
output on;
```
OLS procedure to return the residuals required for the Hamilton (2001) LM test
```
e = resid;
xwhole = ones(n, 1) ~ x1;
xsig = meanc(x1);
xsig = sqrt(meanc((x1 - xsig')^2));
xsig = sqrt(k) * xsig / 2;
sige = e' * e / (n - k - 1);
sig0 = sige;
xq = dist2(x1, 1 ./ xsig );
ht = covary (k, xq);
m0 = xwhole * invpd(xwhole' * xwhole) * xwhole';
m0 = eye(n) - m0;
a0 = m0 * ht * m0;
zeta = a0 - m0 * sumc(diag(a0)) / (n - k - 1);
zeta = sqrt(2 * sumc(diag(zeta * zeta)));
zeta = (e' * ht * e - sige * sumc(diag(a0))) / (sig0 * zeta);
p = cdfchic(zeta^2, 1);
```
Computes the p-value of the Hamilton (2001) LM test

```
lmn_1[index, 1] = p;
```
Returns the results for each of the five experiments

The *proc* dist2:

```
proc dist2(xmat, gam);
local iter, xc_tmp, nx, nk, xc;
nk = rows(gam);
nx = rows(xmat);
iter = 1;
xc = 0;
do until iter > nk;
xc_tmp = (xmat[.,iter] * gam[iter, 1]) .* ones(nx, nx);
xc = xc + (xc_tmp - xc_tmp')^2;
iter = iter + 1;
endo;
retp(sqrt(xc));
endp;
```

The *proc* covary:

```
proc covary(kv, xmat);
local kodd, k0, G0, Gh, hoo, kb, xc, xrow, xcol, xall;
xmat = xmat / 2;
```
Now the elements of xmat correspond to h in Theorem 2.2 and Table 1
```
xrow = rows(xmat);
xcol = cols(xmat);
xall = xrow * xcol;
xmat = reshape(xmat, 1, xall);
xmat = minc(xmat | ones(1, xall));
xmat = reshape(xmat, xrow, xcol);
```
Now the elements of xmat are no larger than unity
```
if kqopt == 1;
kb = kv - 1;
kodd = (kb % 2);
```
kodd is 1 if kb is odd and 0 if kb is even

First calculate G0 = G < kb > (0, 1)
```
if kodd == 0;
G0 = 1;
```
This is G < 0 > (0, 1)
```
elseif kodd == 1;
G0 = (pi / 4);
```
This is G < 1 > (0, 1)
```
else;
"integer error; this line should not be executed";
"kodd is";; kodd;
end;
endif;
k0 = kodd;
do until k0 >= kb;
k0 = k0 + 2;
G0 = (k0 / (1 + k0)) * G0;
```

```
endo;
```

Next calculate Gh = G < kb > (h, 1)

```
if kodd == 0;
Gh = 1 - xmat;
```

This is G < 0 < (h, 1)

```
elseif kodd == 1;
Gh = (pi / 4) - 0.5 * xmat .* sqrt(1 - xmat^2) - 0.5 * arcsin(xmat);
```

This is G < 1 > (h, 1)

```
else;
"integer error; this line should not be executed";
endif;
k0 = kodd;
do until k0 >= kb;
k0 = k0 + 2;
Gh = (k0 / (k0 + 1)) * Gh - (xmat / (1 + k0)) .* (1 - xmat^2)^(k0 / 2);
endo;
hoo = Gh / G0;
elseif kqopt == 2;
if kv == 1;
hoo = 1 - xmat;
elseif kv == 2;
hoo = 1 - (2 / pi) * (xmat .* sqrt(1 - xmat^2) + arcsin(xmat));
elseif kv == 3;
hoo = 1 - (3 / 2) * xmat + (1 / 2) * xmat^3;
elseif kv == 4;
hoo = 1 - (2 / pi) * ((2 * xmat / 3) .* ((1 - xmat^2)^1.5));
hoo = hoo - (2 / pi) * (xmat .* sqrt(1 - xmat^2) + arcsin(xmat));
elseif kv == 5;
hoo = 1 - 1.5 * xmat + (1 / 2) * xmat^3 - (3 / 8) * xmat .* (1 - xmat^2)^2;
else;
"Table 1 only calculates the case for 5 variables or less";
"Change global parameter kqopt to 1";
end;
endif;
endif;
retp(hoo);
endp;
```

The λ^A_{OP} test

```
#include RFtestsNew;
{test, pa, pb} = Alambdatestbootc(y, x1, x1, 100);
lan_1[index, 1] = pa;
```

The *proc* λ^A_{OP}:

```
proc(3) = Alambdatestbootc(y, x, xnon, boot);
local nobs, xwhole, eps, SSR0, SSR1, sigmasq, uhat, vhat, xc, gam, Rsq, I,
g;
local stat,iter,jter,Z,ttemp,indx,nx,auxwhole,yf,numb;
local beta, xtest, torg, tstar, epshat, epshatb, betab, sigmasqb, rejection,
```

```
capT;
local p_asymp, p_boot;
capT = rows(x);
xtest = xnon;
tstar = zeros(boot, 1);
nobs = rows(y);
xwhole = ones(nobs, 1) ~ x;
beta = invpd(xwhole' xwhole) * (xwhole' y);
eps = y - xwhole * beta;
sigmasq = meanc(eps^2);
uhat = vec(eye(nobs) - (eps * eps') ./ sigmasq);
SSR0 = uhat' uhat;
iter = 1;
nx = cols(xtest);
do until iter > nx;
xc = xtest[., iter] .* ones(nobs, nobs);
if iter == 1;
Z = vec(abs(xc - xc'));
else;
Z = Z ~ vec(abs(xc - xc'));
endif;
iter = iter + 1;
endo;
```
Change here to include higher order terms for better approximation
```
if nx > 1;
auxwhole = Z ~ remidel(Z * ~ Z);
else;
auxwhole = Z ~ Z^2;
endif;
numb = cols(auxwhole) + 1;
auxwhole = ones(nobs^2, 1) ~ auxwhole ~ vec(eye(nobs));
vhat = uhat - auxwhole * invpd(auxwhole' auxwhole) * (auxwhole' uhat);
SSR1 = vhat' vhat;
Rsq = (SSR0 - SSR1) / SSR0;
torg = 0.5 * rows(Z) * Rsq;
iter = 1;
rejection = 0;
do until iter > boot;
locate 12, 1;
"LambdaA test boot# ";; iter;; " out of ";; boot;
epshatb = sqrt(sigmasq) * rndn(capT, 1);
y = xwhole * beta + epshatb;
betab = invpd(xwhole'xwhole)*xwhole'y;
epshat = y - xwhole * betab;
sigmasqb = meanc(epshat^2);
uhat = vec(eye(nobs) - (epshat * epshat' ) ./ sigmasqb);
SSR0 = uhat' uhat;
vhat = uhat - auxwhole * invpd(auxwhole' auxwhole) * (auxwhole' uhat);
SSR1 = vhat' vhat;
Rsq = (SSR0 - SSR1) / SSR0;
tstar[iter] = 0.5 * rows(Z) * Rsq;
rejection = rejection + (tstar[iter] > torg);
```

```
iter = iter + 1;
endo;
p_asymp = cdfchic(torg, numb);
p_boot = ((1 + rejection) / (1+boot));
retp(torg, p_asymp, p_boot);
endp;
```

The *proc* remidel:

```
proc(1) = remidel(x);
local colsx, rowsx, indx, c, jter, indx1, cs, precis;
precis = 10^(-5);
indx1 = 0;
colsx = cols(x);
rowsx = rows(x);
cs = sumc(x);
indx = seqa(1, 1, cols(x));
c = indx ~ cs;
c = sortc(c, 2);
c = c';
jter = 1;
indx1 = c[1,1];
do until jter == cols(x);
if c[2, jter + 1] > c[2, jter] + precis;
indx1 = indx1|c[1,jter+1];
endif;
jter = jter + 1;
endo;
indx1 = sortc(indx1, 1);
retp(x[., indx1']);
endp;
```

The $\lambda^E_{OP}(\boldsymbol{g})$ test

```
#include RFtestsNew;
{test, pa, pb} = Elambdatestbootc(y, x1, x1, 100);
len_1[index, 1] = pa;
```

The *proc* $\lambda^E_{OP}(\boldsymbol{g})$:

```
proc(3)=Elambdatestbootc(y, x, xnon, boot);
local nobs, xwhole, eps, SSR0, SSR1, sigmasq, uhat, vhat, xc, gam, Rsq, I,
g, stat;
local iter, jter, Z, ttemp, indx, nx, auxwhole, yf, hplus, hoo;
local beta, xtest, torg, tstar, epshat, epshatb, betab, sigmasqb, rejection;
local p_asymp, p_boot;
nobs = rows(y);
tstar = zeros(boot, 1);
xtest = xnon;
nx = cols(xtest);
xwhole = ones(nobs, 1) ~ x;
```

```
beta = invpd(xwhole' xwhole) * (xwhole' y);
eps = y - xwhole * beta;
sigmasq = meanc(eps^2);
uhat = vec(eye(nobs) - (eps * eps' ) ./ sigmasq);
SSR0 = uhat' uhat;
iter = 1;
do until iter > nx;
xc = xtest[., iter] .* ones(nobs, nobs);
if iter == 1;
Z = vec(abs(xc - xc' ));
else;
Z = Z ~ vec(abs(xc - xc' ));
endif;
iter = iter + 1;
endo;
gam = meanc(abs(xtest - meanc(xtest)' ));
gam = 2 ./ (nx * gam);
hplus = 0.5 * (Z * gam);
hoo = (1 - hplus)^(2 * cols(xnon));
I = (hplus .<= 1);
auxwhole = (hoo .* I) ~ vec(eye(nobs));
vhat = uhat - auxwhole * pinv(auxwhole' auxwhole) * (auxwhole' uhat);
SSR1 = vhat' vhat;
Rsq = (SSR0 - SSR1) / SSR0;
torg = 0.5 * rows(Z) * Rsq;
iter = 1;
rejection = 0;
do until iter > boot;
locate 12,1;
"LambdaE test boot# ";; iter;; " out of ";; boot;
epshatb = sqrt(sigmasq) * rndn(nobs, 1);
y = xwhole * beta + epshatb;
betab = invpd(xwhole' xwhole) * xwhole' y;
epshat = y - xwhole * betab;
sigmasqb = meanc(epshat^2);
uhat = vec(eye(nobs) - (epshat * epshat' ) ./ sigmasqb);
SSR0 = uhat' uhat;
vhat = uhat - auxwhole * pinv(auxwhole' auxwhole) * (auxwhole' uhat);
SSR1 = vhat' vhat;
Rsq = (SSR0 - SSR1) / SSR0;
tstar[iter] = 0.5 * rows(Z) * Rsq;
rejection = rejection + (tstar[iter] > torg);
iter = iter + 1;
endo;
p_asymp = cdfchic(torg, 1);
p_boot = ((1 + rejection) / (1 + boot));
retp(torg, p_asymp, p_boot);
endp;
```

The g_{OP} test

```
#include RFtestsNew;
{test, pa, pb} = gtestbootc(y, x1, x1, 100);
gn_1[index, 1] = pa;
```

The *proc* g_{OP}:

```
proc(3) = gtestbootc(y, x, xnon, boot);
local nobs, xwhole, eps, SSR0, SSR1, sigmasq, uhat, vhat, xc, gam, Rsq, I,
g;
local stat, iter, jter, Z, ttemp, indx, nx, auxwhole, yf, numb;
local beta, xtest, torg, tstar, epshat, epshatb, betab, sigmasqb, rejection,
capT;
local p_asymp, p_boot;
xtest = xnon;
tstar = zeros(boot, 1);
nobs = rows(y);
xwhole = ones(nobs, 1) ~ x;
beta = invpd(xwhole' xwhole) * (xwhole' y);
eps = y - xwhole * beta;
sigmasq = meanc(eps^2);
uhat = vec(eye(nobs) - (eps * ep'') ./ sigmasq);
SSR0 = uhat' uhat;
iter = 1;
nx = cols(xtest);
do until iter > nx;
xc = xtest[., iter] .* ones(nobs, nobs);
if iter == 1;
Z = vec(abs(xc - xc' )) - vec(abs(xc)) - vec(abs(xc' ));
else;
Z = Z ~ (vec(abs(xc - xc' )) - vec(abs(xc)) - vec(abs(xc' )));
endif;
iter = iter + 1;
endo;
```

Change here to include higher order terms for better approximation

```
if nx > 1;
auxwhole = Z ~ remidel(Z * ~ Z);
else;
auxwhole = Z ~ Z^2;
endif;
numb = cols(auxwhole);
auxwhole = auxwhole ~ vec(eye(nobs));
vhat = uhat - auxwhole * invpd(auxwhole' 'auxwhole) * (auxwhole' uhat);
SSR1 = vhat' vhat;
Rsq = (SSR0 - SSR1) / SSR0;
torg = 0.5 * rows(Z) * Rsq;
iter = 1;
rejection = 0;
do until iter > boot;
locate 12, 1;
```

```
"g-test boot# ";; iter;;" out of ";; boot;
epshatb = sqrt(sigmasq) * rndn(nobs, 1);
y = xwhole * beta + epshatb;
betab = invpd(xwhole' xwhole) * xwhole' y;
epshat = y - xwhole * betab;
sigmasqb = meanc(epshat^2);
uhat = vec(eye(nobs) - (epshat * epshat' ) ./ sigmasqb);
SSR0 = uhat' uhat;
vhat = uhat - auxwhole * invpd(auxwhole' auxwhole) * (auxwhole' uhat);
SSR1 = vhat' vhat;
Rsq = (SSR0 - SSR1) / SSR0;
tstar[iter] = 0.5 * rows(Z) * Rsq;
rejection = rejection + (tstar[iter] > torg);
iter = iter + 1;
endo;
p_asymp = cdfchic(torg, numb);
p_boot = ((1 + rejection) / (1 + boot));
retp(torg, p_asymp, p_boot);
endp;
```

Appendix C

Appendix to Chapter 5

C.1 Data used in Chapter 5

Table C.1: Money demand data, Denmark.

	m_t	y_t	p_t	b_t	i_t
1974:1	11.633	5.904	-0.619	0.155	0.094
1974:2	11.604	5.874	-0.581	0.178	0.096
1974:3	11.582	5.838	-0.543	0.171	0.096
1974:4	11.602	5.812	-0.505	0.152	0.096
1975:1	11.586	5.804	-0.487	0.134	0.089
1975:2	11.605	5.787	-0.454	0.134	0.079
1975:3	11.653	5.833	-0.441	0.128	0.076
1975:4	11.764	5.930	-0.439	0.129	0.074
1976:1	11.753	5.938	-0.404	0.141	0.072
1976:2	11.766	5.935	-0.373	0.153	0.078
1976:3	11.781	5.932	-0.358	0.161	0.080
1976:4	11.770	5.941	-0.325	0.162	0.103
1977:1	11.746	5.928	-0.310	0.167	0.097
1977:2	11.769	5.937	-0.288	0.163	0.088
1977:3	11.750	5.957	-0.264	0.169	0.095
1977:4	11.749	5.940	-0.229	0.173	0.097
1978:1	11.705	5.932	-0.212	0.172	0.099
1978:2	11.703	5.931	-0.197	0.176	0.088
1978:3	11.703	5.972	-0.185	0.171	0.081
1978:4	11.704	5.969	-0.163	0.182	0.077
1979:1	11.679	5.963	-0.139	0.170	0.075
1979:2	11.708	5.987	-0.121	0.169	0.077
1979:3	11.676	5.985	-0.087	0.178	0.086
1979:4	11.686	5.977	-0.059	0.180	0.101
1980:1	11.642	5.985	-0.036	0.191	0.109
1980:2	11.635	5.945	-0.009	0.197	0.121
1980:3	11.609	5.902	0.014	0.192	0.121
1980:4	11.658	5.902	0.033	0.183	0.107
1981:1	11.628	5.897	0.062	0.185	0.105
1981:2	11.629	5.894	0.104	0.193	0.109
1981:3	11.604	5.884	0.126	0.203	0.111
1981:4	11.629	5.895	0.151	0.192	0.109
1982:1	11.602	5.902	0.174	0.203	0.107
1982:2	11.604	5.924	0.198	0.211	0.111
1982:3	11.595	5.950	0.220	0.209	0.111
1982:4	11.604	5.929	0.246	0.197	0.110
1983:1	11.617	5.916	0.259	0.161	0.106
1983:2	11.688	5.939	0.270	0.138	0.087
1983:3	11.727	5.932	0.285	0.142	0.083
1983:4	11.781	5.974	0.301	0.134	0.085

Continued on next page.

	m_t	y_t	p_t	b_t	i_t
1984:1	11.798	5.964	0.316	0.134	0.085
1984:2	11.842	6.001	0.331	0.142	0.083
1984:3	11.840	5.991	0.343	0.145	0.085
1984:4	11.924	5.969	0.356	0.141	0.092
1985:1	11.903	6.014	0.371	0.132	0.090
1985:2	11.927	6.003	0.383	0.120	0.088
1985:3	11.971	6.045	0.387	0.107	0.080
1985:4	12.026	6.081	0.393	0.104	0.076
1986:1	12.051	6.096	0.398	0.098	0.072
1986:2	12.076	6.090	0.413	0.098	0.068
1986:3	12.056	6.099	0.418	0.112	0.068
1986:4	12.072	6.081	0.428	0.114	0.075
1987:1	12.028	6.061	0.447	0.119	0.077
1987:2	12.040	6.064	0.458	0.117	0.076
1987:3	12.015	6.051	0.468	0.119	0.075

Source: `http://www.math.ku.dk/~sjo/`.

Table C.2: Money demand data, Finland.

	m_t	y_t	p_t	i_t		m_t	y_t	p_t	i_t
1958:02	2.849	3.837	0.015	0.173	1969:01	3.135	4.420	0.000	0.068
1958:03	2.828	3.903	0.000	0.101	1969:02	3.194	4.444	0.009	0.068
1958:04	2.908	3.897	0.000	0.132	1969:03	3.196	4.490	0.000	0.068
1959:01	2.934	3.847	0.000	0.070	1969:04	3.333	4.519	0.000	0.072
1959:02	2.986	3.916	0.000	0.065	1970:01	3.411	4.491	0.018	0.082
1959:03	3.007	3.986	0.015	0.065	1970:02	3.460	4.537	0.000	0.151
1959:04	3.016	3.984	0.014	0.065	1970:03	3.452	4.562	0.009	0.231
1960:01	2.954	3.956	0.014	0.065	1970:04	3.530	4.619	0.009	0.185
1960:02	3.003	3.990	0.000	0.161	1971:01	3.297	4.510	0.026	0.079
1960:03	3.042	4.049	0.014	0.243	1971:02	3.280	4.575	0.025	0.118
1960:04	3.034	4.053	0.000	0.163	1971:03	3.308	4.570	0.024	0.093
1961:01	3.059	4.066	0.000	0.110	1971:04	3.451	4.632	0.008	0.082
1961:02	3.035	4.089	0.000	0.122	1972:01	3.437	4.603	0.008	0.075
1961:03	3.087	4.146	0.000	0.117	1972:02	3.485	4.623	0.031	0.075
1961:04	3.101	4.125	0.027	0.148	1972:03	3.501	4.644	0.015	0.075
1962:01	3.059	4.108	0.000	0.175	1972:04	3.607	4.716	0.015	0.075
1962:02	3.039	4.096	0.027	0.229	1973:01	3.509	4.689	0.029	0.075
1962:03	3.075	4.157	0.000	0.242	1973:02	3.581	4.692	0.035	0.075
1962:04	3.107	4.159	0.013	0.240	1973:03	3.494	4.704	0.054	0.096
1963:01	3.138	4.101	0.013	0.211	1973:04	3.668	4.810	0.026	0.231
1963:02	3.128	4.170	0.013	0.171	1974:01	3.496	4.745	0.044	0.168
1963:03	3.128	4.200	0.013	0.189	1974:02	3.553	4.809	0.036	0.148
1963:04	3.200	4.197	0.012	0.171	1974:03	3.536	4.752	0.057	0.090
1964:01	3.068	4.179	0.060	0.137	1974:04	3.689	4.881	0.017	0.138
1964:02	3.073	4.178	0.023	0.169	1975:01	3.647	4.762	0.059	0.130
1964:03	3.085	4.193	0.011	0.178	1975:02	3.664	4.816	0.036	0.175
1964:04	3.135	4.260	0.000	0.197	1975:03	3.566	4.758	0.044	0.162
1965:01	3.099	4.222	0.011	0.183	1975:04	3.783	4.808	0.028	0.246
1965:02	3.091	4.220	0.022	0.135	1976:01	3.641	4.695	0.045	0.179
1965:03	3.079	4.277	0.000	0.068	1976:02	3.636	4.773	0.013	0.178
1965:04	3.111	4.293	0.011	0.068	1976:03	3.580	4.767	0.039	0.159
1966:01	3.025	4.200	0.011	0.068	1976:04	3.647	4.843	0.021	0.167
1966:02	3.083	4.276	0.011	0.115	1977:01	3.629	4.695	0.040	0.141
1966:03	3.041	4.299	0.021	0.220	1977:02	3.645	4.733	0.035	0.152
1966:04	3.117	4.342	0.010	0.294	1977:03	3.586	4.730	0.026	0.166
1967:01	3.027	4.294	0.010	0.068	1977:04	3.595	4.802	0.011	0.181
1967:02	3.047	4.315	0.010	0.122	1978:01	3.599	4.699	0.015	0.171
1967:03	2.994	4.322	0.010	0.092	1978:02	3.693	4.763	0.018	0.091
1967:04	3.007	4.313	0.029	0.121	1978:03	3.663	4.739	0.014	0.087
1968:01	2.992	4.306	0.038	0.092	1978:04	3.694	4.843	0.007	0.093
1968:02	3.047	4.333	0.018	0.068	1979:01	3.661	4.772	0.028	0.081
1968:03	3.074	4.381	0.000	0.068	1979:02	3.789	4.842	0.020	0.080
1968:04	3.181	4.392	0.009	0.068	1979:03	3.753	4.817	0.017	0.081

Continued on next page.

	m_t	y_t	p_t	i_t		m_t	y_t	p_t	i_t
1979:04	3.813	4.890	0.019	0.112	1982:02	3.813	4.891	0.025	0.129
1980:01	3.725	4.785	0.035	0.116	1982:03	3.822	4.891	0.007	0.128
1980:02	3.771	4.837	0.043	0.139	1982:04	3.850	4.968	0.021	0.141
1980:03	3.713	4.867	0.026	0.148	1983:01	3.804	4.835	0.019	0.134
1980:04	3.745	4.922	0.026	0.153	1983:02	3.863	4.923	0.037	0.140
1981:01	3.722	4.800	0.031	0.163	1983:03	3.852	4.936	0.013	0.145
1981:02	3.733	4.875	0.032	0.122	1983:04	3.826	4.981	0.027	0.155
1981:03	3.733	4.869	0.018	0.124	1984:01	3.783	4.866	0.004	0.161
1981:04	3.788	4.937	0.013	0.140	1984:02	3.830	4.960	0.019	0.154
1982:01	3.706	4.803	0.033	0.131	1984:03	3.825	4.952	0.015	0.152

Source: `http://www.math.ku.dk/~sjo/`.

C.2 Figures from Chapter 5

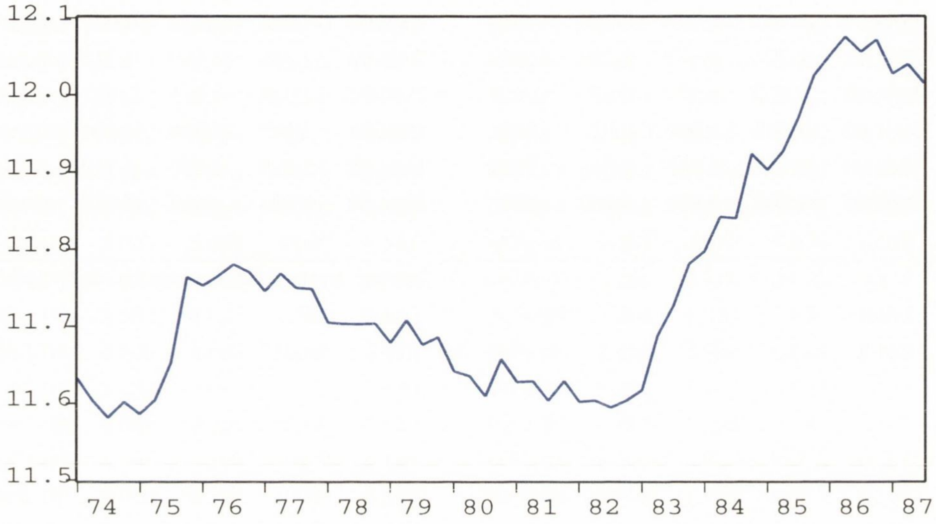

Figure C.1: The log of real money balances, Denmark.

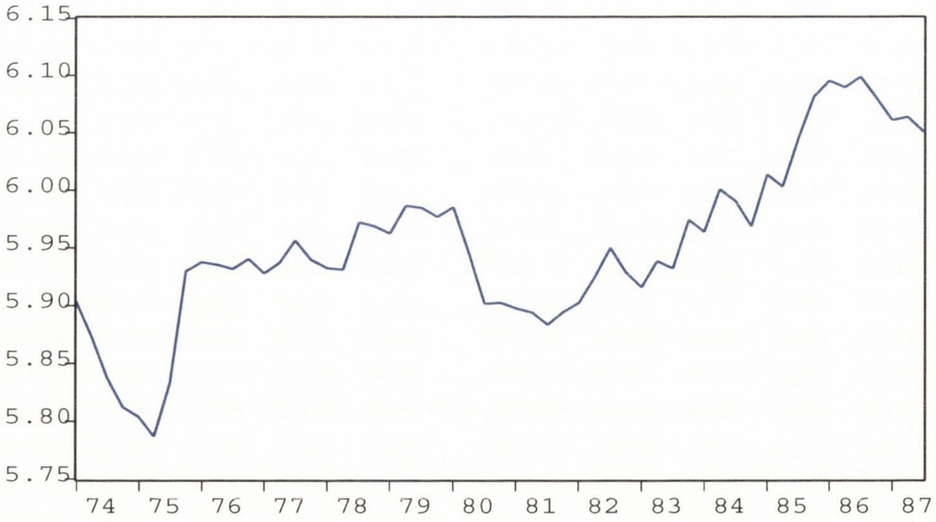

Figure C.2: The log of real income, Denmark.

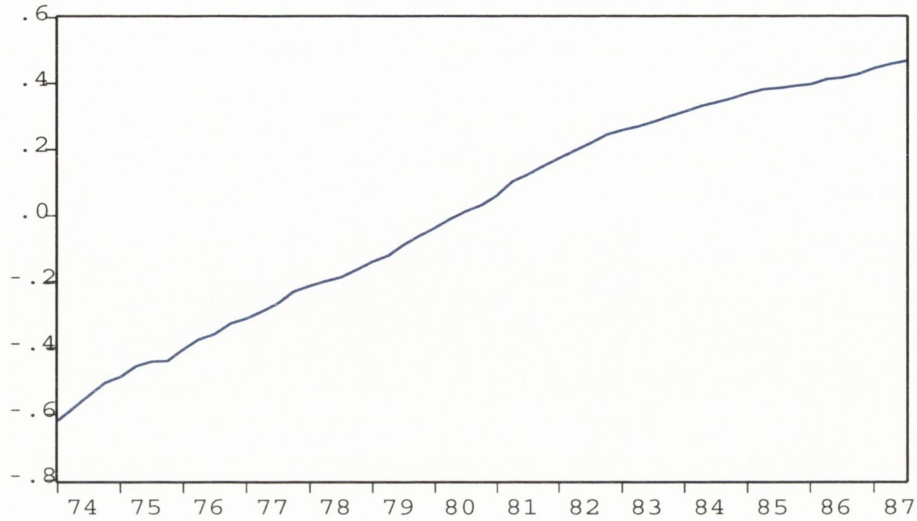

Figure C.3: The log of prices, Denmark.

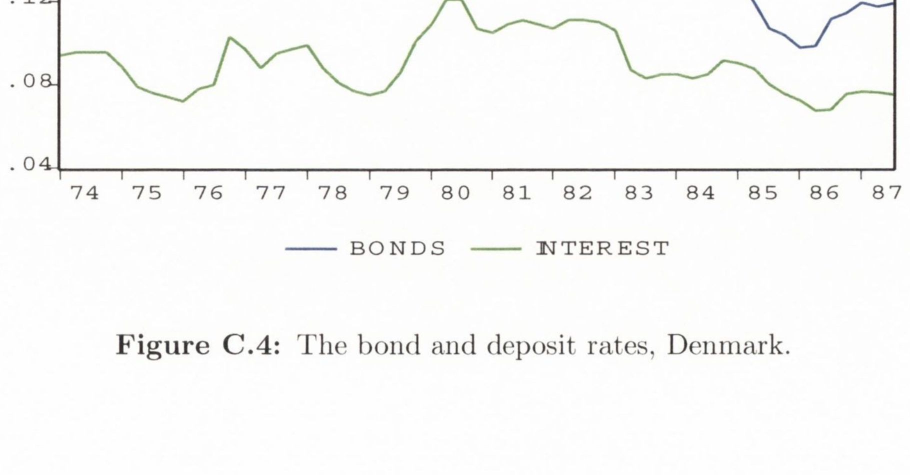

Figure C.4: The bond and deposit rates, Denmark.

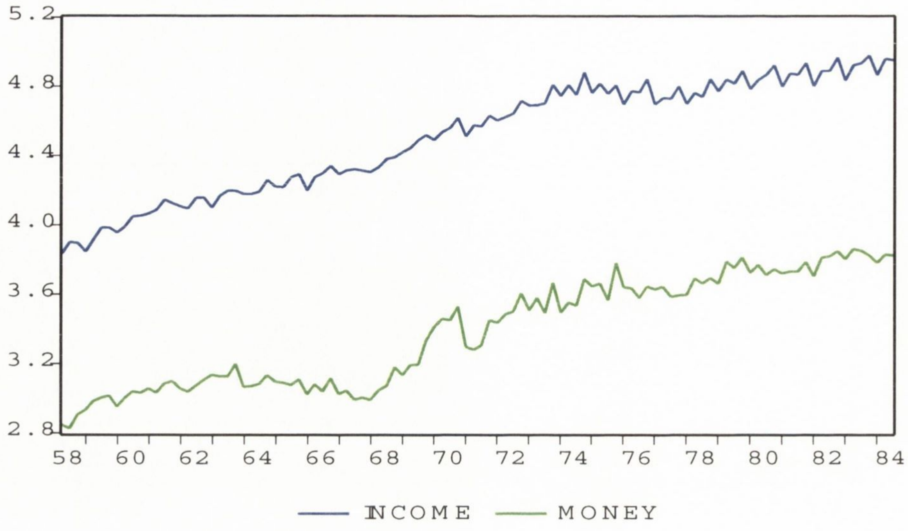

Figure C.5: The logs of real money balances and real income, Finland.

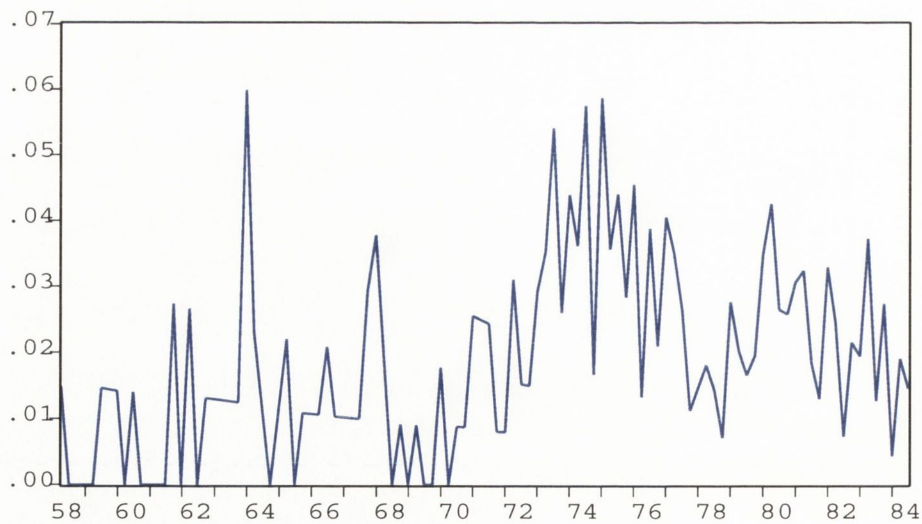

Figure C.6: The inflation rate, Finland.

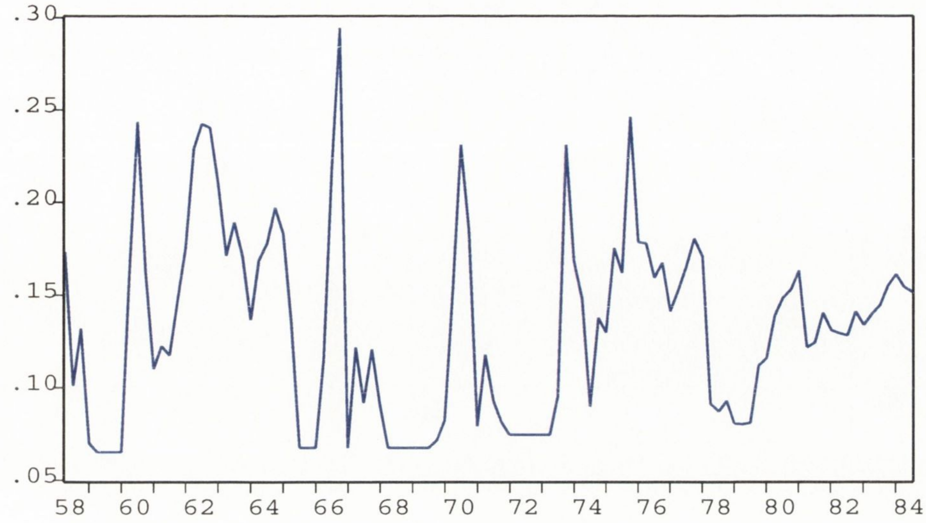

Figure C.7: The marginal rate of interest, Finland.

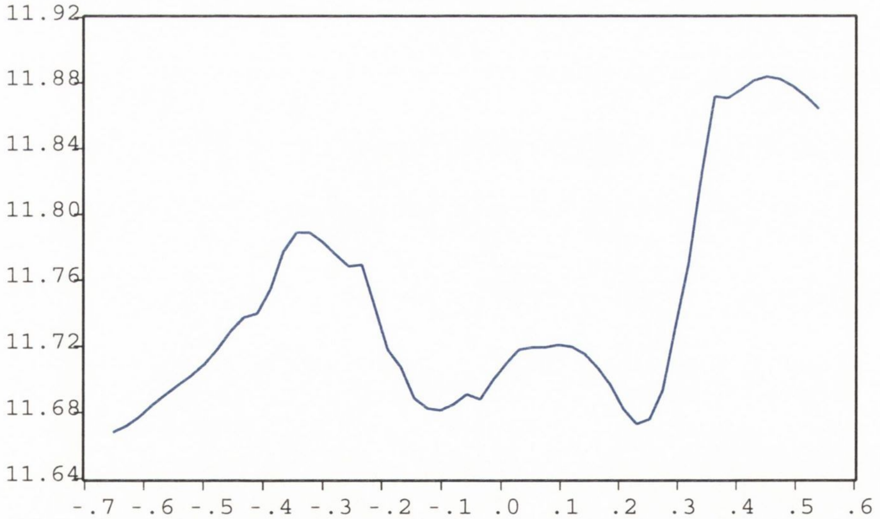

Figure C.8: Plot of conditional expectation function, Denmark.

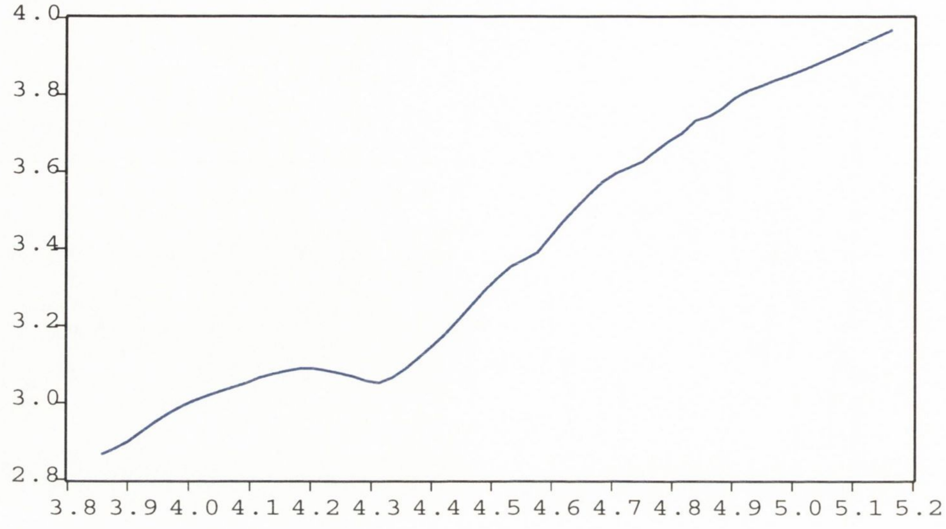

Figure C.9: Plot of conditional expectation function, Finland.

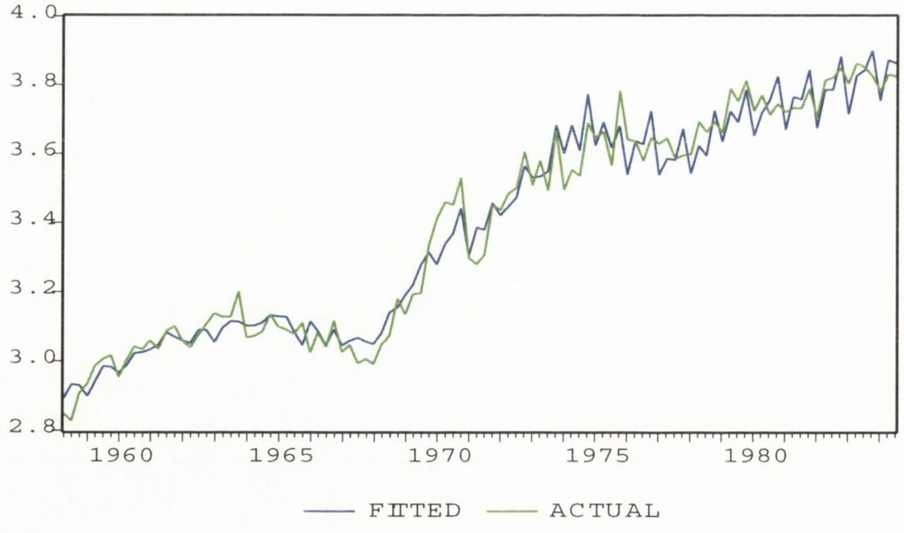

Figure C.10: Plot of actual versus fitted, smooth transition model, Finland.

C.3 Preliminary Analysis

Table C.3: Unit root tests, Denmark & Finland.

Variable	Test Statistic (probability)†	Constant	Trend	Lag Length using Sic
		Denmark ($T = 55$)		
m_t	1.123 (0.930)	Not significant	Not significant	2
y_t	0.776 (0.878)	Not significant	Not significant	0
p_t	0.437 (0.664)‡	significant	significant	0
i_t	−0.616 (0.446)	Not significant	Not significant	0
b_t	−0.982 (0.288)	Not significant	Not significant	1
		Finland ($T = 106$)		
m_t	1.720 (0.979)	Not significant	Not significant	4
y_t	−1.951 (0.054)‡	Significant	Not significant	4
p_t	−4.200 (0.006)	n/a	n/a	1
i_t	−4.874 (0.001)	n/a	n/a	0

†MacKinnon (1996) one-sided p-values.
‡normal probability.
Note: n/a denotes not applicable.

Table C.4: **Seasonal unit root tests, Denmark ($T = 55$).**

Variable	Test Statistics [critical values]					Lag Length
	t_{π_1}	t_{π_2}	F_{34}	F_{234}	F_{1234}	
	HEGY test without deterministic terms					
m_t	0.858 [−1.91]	−1.734 [−1.93]	11.871 [3.11]	9.138 [2.78]	7.034 [2.55]	1
y_t	1.124 [−1.91]	−4.245 [−1.93]	22.988 [3.11]	65.781 [2.78]	54.536 [2.55]	0
p_t	0.204 [−1.91]	−5.893 [−1.93]	33.861 [3.11]	1579.675 [2.78]	1236.130 [2.55]	0
i_t	−0.622 [−1.91]	−4.492 [−1.93]	18.312 [3.11]	80.487 [2.78]	61.219 [2.55]	0
b_t	−0.472 [−1.91]	−4.141 [−1.93]	26.199 [3.11]	118.904 [2.78]	91.335 [2.55]	0
	HEGY test with intercept and time trend					
m_t	−2.091 [−3.40]	−1.609 [−1.93]	8.529 [3.05]	6.576 [2.74]	8.185 [4.19]	1
y_t	−2.670 [−3.40]	−4.537 [−1.93]	25.627 [3.05]	73.599 [2.74]	56.780 [4.19]	0
p_t	0.520 [−3.40]	−4.949 [−1.93]	21.503 [3.05]	66.969 [2.74]	59.789 [4.19]	0
i_t	−2.193 [−3.40]	−4.637 [−1.93]	18.264 [3.05]	84.322 [2.74]	64.529 [4.19]	0
b_t	−1.764 [−3.40]	−4.046 [−1.93]	26.614 [3.05]	119.292 [2.74]	90.462 [4.19]	0
	HEGY test with intercept and seasonal dummy variables					
m_t	−1.526 [−2.84]	−4.941 [−2.83]	20.533 [6.57]	158.129 [5.95]	121.698 [5.56]	0
y_t	−2.059 [−2.84]	−4.254 [−2.83]	24.068 [6.57]	68.422 [5.95]	51.575 [5.56]	0
p_t	−1.471 [−2.84]	−2.963 [−2.83]	8.121 [6.57]	10.644 [5.95]	9.709 [5.56]	12
i_t	−2.131 [−2.84]	−4.280 [−2.83]	19.913 [6.57]	88.446 [5.95]	67.476 [5.56]	0
b_t	−0.878 [−2.84]	−4.062 [−2.83]	18.668 [6.57]	25.239 [5.95]	22.767 [5.56]	1
	HEGY test with intercept, time trend and seasonal dummy variables					
m_t	−1.899 [−3.39]	−4.881 [−2.82]	20.416 [6.55]	152.703 [5.93]	114.532 [6.31]	0
y_t	−2.596 [−3.39]	−4.314 [−2.82]	24.731 [6.55]	70.139 [5.93]	54.095 [6.31]	0
p_t	−0.046 [−3.39]	−2.877 [−2.82]	7.664 [6.55]	10.062 [5.93]	7.586 [6.31]	12
i_t	−2.153 [−3.39]	−4.221 [−2.82]	19.550 [6.55]	86.066 [5.93]	65.846 [6.31]	0
b_t	−1.585 [−3.39]	−4.174 [−2.82]	19.909 [6.55]	26.622 [5.93]	25.432 [6.31]	1

Note: critical values for the 5 per cent level of significance, taken from Franses and Hobijn (1997).

Table C.5: Seasonal unit root tests, Finland (T = 106).

VARIABLE	TEST STATISTICS [critical values]					LAG LENGTH
	t_{π_1}	t_{π_2}	F_{34}	F_{234}	F_{1234}	
	HEGY test without deterministic terms					
m_t	1.700 [−1.91]	−1.000 [−1.93]	4.093 [3.11]	3.084 [2.78]	2.965 [2.55]	5
y_t	2.852 [−1.91]	−0.659 [−1.93]	4.265 [3.11]	2.949 [2.78]	4.472 [2.55]	1
p_t	−1.026 [−1.91]	−3.039 [−1.93]	20.725 [3.11]	18.094 [2.78]	13.762 [2.55]	0
i_t	−0.810 [−1.91]	−4.848 [−1.93]	39.543 [3.11]	60.993 [2.78]	45.886 [2.55]	0
	HEGY test with intercept and time trend					
m_t	−2.321 [−3.40]	−0.958 [−1.93]	3.652 [3.05]	2.755 [2.74]	3.750 [4.19]	5
y_t	−1.565 [−3.40]	−0.661 [−1.93]	4.295 [3.05]	2.971 [2.74]	2.824 [4.19]	1
p_t	−2.755 [−3.40]	−3.103 [−1.93]	22.263 [3.05]	19.295 [2.74]	15.910 [4.19]	0
i_t	−3.813 [−3.40]	−5.124 [−1.93]	43.777 [3.05]	68.029 [2.74]	54.867 [4.19]	0
	HEGY test with intercept and seasonal dummy variables					
m_t	−0.691 [−2.84]	−3.942 [−2.83]	52.354 [6.57]	65.165 [5.95]	49.482 [5.56]	0
y_t	−1.831 [−2.84]	−1.715 [−2.83]	14.251 [6.57]	10.436 [5.95]	9.045 [5.56]	1
p_t	−2.747 [−2.84]	−3.209 [−2.83]	26.482 [6.57]	22.466 [5.95]	18.551 [5.56]	0
i_t	−3.973 [−2.84]	−5.425 [−2.83]	43.500 [6.57]	78.001 [5.95]	62.712 [5.56]	0
	HEGY test with intercept, time trend and seasonal dummy variables					
m_t	−2.106 [−3.39]	−4.020 [−2.82]	54.344 [6.55]	67.671 [5.93]	51.739 [6.31]	0
y_t	−1.611 [−3.39]	−1.681 [−2.82]	14.382 [6.55]	10.493 [5.93]	8.591 [6.31]	1
p_t	−2.928 [−3.39]	−3.215 [−2.82]	27.126 [6.55]	22.926 [5.93]	18.710 [6.31]	0
i_t	−3.960 [−3.39]	−5.401 [−2.82]	43.077 [6.55]	77.244 [5.93]	62.095 [6.31]	0

Note: critical values for the 5 per cent level of significance, taken from Franses and Hobijn (1997).

C.4 Testing for Cointegration

C.4.1 The Engle-Granger 2-Step method

Table C.6: Engle-Granger levels models, Denmark & Finland.

	Coefficients (t-statistics)			
Variable	Denmark ($T = 55$)	Finland ($T = 106$)		
	Full	Full	Reduced 1	Reduced 2
c	4.472 (5.178)	−0.841 (−6.552)	−0.784 (−6.104)	−0.766 (−6.448)
y_t	1.283 (9.135)	0.928 (31.988)	0.926 (31.300)	0.921 (35.021)
p_t	0.004 (0.122)	−0.366 (−0.552)	−0.254 (−0.377)	-
i_t	0.569 (0.705)	0.375 (2.285)	-	-
b_t	−2.601 (−7.370)	-	-	-
R^2	0.926	0.926	0.922	0.922
AEG test on residuals	−3.301 [−4.694]	−3.541 [−4.204]	−3.473 [−3.824]	−3.461 [−3.395]
CRDW test	0.737 [1.19]	0.418 [0.58]	0.399 [0.48]	0.398 [0.38]

Note: 5 per cent AEG and CRDW critical values in square brackets.

Table C.7: Error-correction models, Denmark & Finland.

	Coefficients (t-statistics)			
Variable	Denmark ($T = 55$)	Finland ($T = 106$)		
	Full	Full	Reduced 1	Reduced 2
c	0.008 (0.657)	0.004 (0.852)	0.004 (0.795)	0.004 (0.807)
Δm_{t-1}	−0.088 (−0.710)	−0.166 (−1.873)	−0.158 (−1.779)	−0.161 (−1.806)
Δm_{t-4}	0.387 (2.861)	0.093 (0.978)	0.099 (1.028)	0.098 (1.021)
Δy_t	0.497 (2.815)	0.502 (4.412)	0.502 (4.367)	0.501 (4.371)
Δp_t	−0.233 (−0.449)	−0.416 (−1.218)	−0.392 (−1.140)	−0.372 (−1.084)
Δi_t	−1.137 (−1.700)	0.247 (2.058)	0.222 (1.817)	0.221 (1.814)
Δb_t	−0.860 (−1.744)	-	-	-
Ecm	−0.009 (−0.082)	0.175 (2.852)	0.154 (2.571)	0.157 (2.636)
R^2	0.498	0.532	0.525	0.526
Ecm test critical values	−4.231 [0.9894]	−4.011 [1.000]	−4.011 [1.000]	−4.011 [1.000]
Serial correlation $\chi^2(4)$	5.119 [0.275]	6.460 [0.167]	5.187 [0.269]	5.553 [0.235]
Functional form $\chi^2(1)$	1.443 [0.230]	0.322 [0.570]	0.372 [0.542]	0.371 [0.542]
Normality $\chi^2(2)$	3.457 [0.178]	1.114 [0.573]	1.202 [0.548]	1.264 [0.532]
Heteroscedasticity $\chi^2(1)$	0.099 [0.753]	3.804 [0.051]	3.929 [0.047]	4.072 [0.044]

Note: for diagnostics and Ecm test, p-values in square brackets.

C.4.2 Johansen's maximum likelihood approach

Table C.8: Number of cointegrating relations by model, Denmark.

Test Type	No Inpts No Trends		Rest'd Inpts No Trends		Unrest'd Inpts No Trends		Unrest'd Inpts Rest'd Trends		Unrest'd Inpts Unrest'd Trends	
	0.05 and 0.10 significance levels, excluding seasonal dummies									
Trace	3	4	3	3	3	4	3	3	1	2
Max. Eig.	1	2	3	3	3	4	3	3	0	1
	0.05 and 0.10 significance levels, including seasonal dummies									
Trace	4	4	3	3	3	3	2	3	1	2
Max. Eig.	1	2	2	3	2	3	2	3	1	1

Note: critical values based on Osterwald-Lenum (1992).

Table C.9: Testing the order of the VAR, Denmark.

Based on 52 observations from 1974Q4 to 1987Q3. Order of VAR = 3. List of variables included in the unrestricted VAR: m_t^{Den}, y_t^{Den}, p_t^{Den}, i_t^{Den}, b_t^{Den}. List of deterministic variables: c, t, sc_1, sc_2, sc_3.

Order	LL	AIC	SBC	χ^2	LR Test	Prob.	Adjusted LR Test	Prob.
3	916.641	816.641	719.079	-	-	-	-	-
2	892.844	817.844	744.672	χ^2 (25) =	47.594	0.004	29.289	0.252
1	863.031	813.031	764.250	χ^2 (50) =	107.221	0.000	65.982	0.064
0	605.546	580.546	556.155	χ^2 (75) =	622.190	0.000	382.886	0.000

Table C.10: Test of restrictions for deterministic variables, Denmark.

Based on 53 observations from 1974 Q3 to 1987 Q3. Order of VAR = 2.

Variable / Statistic	Max. LL without Variable	Max. LL with Variable	LR Test χ^2		Prob.
c	895.691	883.467	χ^2 (5) =	24.449	0.000
t	895.691	887.774	χ^2 (5) =	15.834	0.007
sc_1	895.691	866.470	χ^2 (5) =	58.443	0.000
sc_2	895.691	888.423	χ^2 (5) =	14.537	0.013
sc_3	895.691	875.185	χ^2 (5) =	41.013	0.000
$sc_1 + sc_2 + sc_3$	895.691	857.987	χ^2 (15) =	75.408	0.000

Note: Max. LL - Maximised value of loglikelihood.

Table C.11: Number of cointegrating relations by model, Finland model 1.

Test Type	No Inpts No Trends		Rest'd Inpts No Trends		Unrest'd Inpts No Trends		Unrest'd Inpts Rest'd Trends		Unrest'd Inpts Unrest'd Trends	
	0.05 and 0.10 significance levels, excluding seasonal dummies									
Trace	3	4	2	4	2	2	2	2	2	2
Max. Eig.	2	4	2	2	2	2	2	2	2	2
	0.05 and 0.10 significance levels, including seasonal dummies									
Trace	3	4	3	4	2	2	2	2	2	2
Max. Eig.	3	4	2	4	2	2	2	2	2	2

Note: critical values based on Osterwald-Lenum (1992).

Table C.12: Number of cointegrating relations by model, Finland model 2.

Test Type	No Inpts No Trends		Rest'd Inpts No Trends		Unrest'd Inpts No Trends		Unrest'd Inpts Rest'd Trends		Unrest'd Inpts Unrest'd Trends	
	0.05 and 0.10 significance levels, excluding seasonal dummies									
Trace	1	1	1	1	1	1	1	1	1	1
Max. Eig.	1	1	1	1	1	1	1	1	1	1
	0.05 and 0.10 significance levels, including seasonal dummies									
Trace	1	1	1	1	1	1	1	1	1	1
Max. Eig.	1	1	1	1	1	1	1	1	1	1

Note: critical values based on Osterwald-Lenum (1992).

Table C.13: Number of cointegrating relations by model, Finland model 3.

Test Type	No Inpts No Trends		Rest'd Inpts No Trends		Unrest'd Inpts No Trends		Unrest'd Inpts Rest'd Trends		Unrest'd Inpts Unrest'd Trends	
	0.05 and 0.10 significance levels, excluding seasonal dummies									
Trace	0	1	0	1	0	0	0	0	0	0
Max. Eig.	0	1	0	0	0	0	0	0	0	0
	0.05 and 0.10 significance levels, including seasonal dummies									
Trace	1	1	1	2	0	0	0	0	0	0
Max. Eig.	1	1	1	1	0	0	0	0	0	0

Note: critical values based on Osterwald-Lenum (1992).

Table C.14: Testing the order of the VAR, Finland.

Based on 103 observations from 1959 Q1 to 1984 Q3. Order of VAR = 3. List of variables included in the unrestricted VAR: m_t^{Fin}, y_t^{Fin}, p_t^{Fin}, i_t^{Fin}. List of deterministic variables: c, t, sc_1, sc_2, sc_3.

Order	Ll	Aic	Sbc	χ^2	Lr Test	Prob.	Adjusted Lr Test	Prob.
3	942.320	874.320	784.739	-	-	-	-	-
2	926.526	874.526	806.023	χ^2 (16) =	31.588	0.011	26.374	0.049
1	888.360	852.360	804.935	χ^2 (32) =	107.920	0.000	90.108	0.000
0	709.097	689.097	662.749	χ^2 (48) =	466.447	0.000	389.460	0.000

Table C.15: Test of restrictions for deterministic variables, Finland.

Based on 104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2.

Variable / Statistic	Max. Ll without Variable	Max. Ll with Variable	Lr Test χ^2		Prob.
c	936.354	929.002	χ^2 (4) =	14.704	0.005
t	936.354	929.992	χ^2 (4) =	12.724	0.013
sc_1	936.354	895.711	χ^2 (4) =	81.285	0.000
sc_2	936.354	922.702	χ^2 (4) =	27.303	0.000
sc_3	936.354	927.038	χ^2 (4) =	18.631	0.001
$sc_1 + sc_2 + sc_3$	936.354	882.676	χ^2(12) =	107.356	0.000

Note: Max. Ll - Maximised value of loglikelihood.

Table C.16: Johansen's cointegration with restricted trends, Denmark.

54 observations from 1974 Q2 to 1987 Q3. Order of VAR = 1. List of variables included in the cointegrating vector: m_t^{Den}, y_t^{Den}, p_t^{Den}, b_t^{Den}, i_t^{Den}. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

COINTEGRATION RANK TEST (TRACE)

HYPOTHESES		TRACE STATISTIC	0.05 CRITICAL VALUE	0.10 CRITICAL VALUE	MODIFIED 0.05 CRITICAL VALUE
r = 0	r $\geq$ 1	142.297	87.170	82.880	138.391
r $\leq$ 1	r $\geq$ 2	86.394	63.000	59.160	100.019
r $\leq$ 2	r $\geq$ 3	40.821	42.340	39.340	67.219
r $\leq$ 3	r = 4	16.447	25.770	23.080	40.913
r $\leq$ 4	r = 5	3.644	12.390	10.550	19.670

COINTEGRATION RANK TEST (MAXIMAL EIGENVALUE)

HYPOTHESES		MAXIMAL EIGENVALUE STATISTIC	0.05 CRITICAL VALUE	0.10 CRITICAL VALUE
r = 0	r = 1	55.903	37.860	35.040
r $\leq$ 1	r = 2	45.573	31.790	29.130
r $\leq$ 2	r = 3	24.374	25.420	23.100
r $\leq$ 3	r = 4	12.802	19.220	17.180
r $\leq$ 4	r = 5	3.644	12.390	10.550

Note: the correction factor is 1.588.

Table C.17: Johansen's cointegration with unrestricted trends, Denmark.

54 observations from 1974 Q2 to 1987 Q3. Order of VAR = 1. List of variables included in the cointegrating vector: m_t^{Den}, y_t^{Den}, p_t^{Den}, b_t^{Den}, i_t^{Den}. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
r = 0	r $\geq$ 1	111.641	82.230	77.550	112.820
r $\leq$ 1	r $\geq$ 2	55.809	58.930	55.010	80.852
r $\leq$ 2	r $\geq$ 3	29.843	39.330	36.280	53.961
r $\leq$ 3	r = 4	7.640	23.830	21.230	32.695
r $\leq$ 4	r = 5	0.094	11.540	9.750	15.833

Cointegration Rank Test (Maximal Eigenvalue)

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r = 1	55.831	37.070	34.160
r $\leq$ 1	r = 2	25.966	31.000	28.320
r $\leq$ 2	r = 3	22.203	24.350	22.260
r $\leq$ 3	r = 4	7.546	18.330	16.280
r $\leq$ 4	r = 5	0.094	11.540	9.750

Note: the correction factor is 1.372.

Table C.18: Johansen's cointegration with unrestricted trends, Finland model 1.

104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: m_t^{Fin}, y_t^{Fin}, p_t^{Fin}, i_t^{Fin}. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
$r = 0$	$r \geq 1$	85.152	58.930	55.010	63.998
$r \leq 1$	$r \geq 2$	42.629	39.330	36.280	42.712
$r \leq 2$	$r \geq 3$	12.209	23.830	21.230	25.879
$r \leq 3$	$r = 4$	2.247	11.540	9.750	12.532

Cointegration Rank Test (Maximal Eigenvalue)

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
$r = 0$	$r = 1$	42.523	31.000	28.320
$r \leq 1$	$r = 2$	30.419	24.350	22.260
$r \leq 2$	$r = 3$	9.963	18.330	16.280
$r \leq 3$	$r = 4$	2.247	11.540	9.750

Note: the correction Factor is 1.086.

Table C.19: Johansen's cointegration with restricted trends, Finland model 1.

104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: m_t^{Den}, y_t^{Den}, p_t^{Den}, i_t^{Den}, t. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
r = 0	r $\geq$ 1	88.886	63.000	59.160	66.320
r $\leq$ 1	r $\geq$ 2	45.261	42.340	39.340	44.571
r $\leq$ 2	r $\geq$ 3	14.771	25.770	23.080	27.128
r $\leq$ 3	r = 4	4.722	12.390	10.550	13.043

Cointegration Rank Test (Maximal Eigenvalue)

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r = 1	43.625	31.790	29.130
r $\leq$ 1	r = 2	30.490	25.420	23.100
r $\leq$ 2	r = 3	10.049	19.220	17.180
r $\leq$ 3	r = 4	4.722	12.390	10.550

Note: the correction factor is 1.053.

Table C.20: Johansen's cointegration with unrestricted trends, Finland model 2.

104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: m_t^{Fin}, y_t^{Fin}, p_t^{Fin}. List of $I(0)$ variables included in the VAR: i_t^{Fin}, sc_1, sc_2, sc_3

Cointegration Rank Test (Trace)

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
$r = 0$	$r \geq 1$	43.818	39.330	36.280	40.309
$r \leq 1$	$r \geq 2$	10.798	23.830	21.230	24.423
$r \leq 2$	$r = 3$	2.052	11.540	9.750	11.827

Cointegration Rank Test (Maximal Eigenvalue)

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
$r = 0$	$r = 1$	33.020	24.350	22.260
$r \leq 1$	$r = 2$	8.746	18.330	16.280
$r \leq 2$	$r = 3$	2.052	11.540	9.750

Note: the correction Factor is 1.025.

Table C.21: Johansen's cointegration with restricted trends, Finland model 2.

104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: m_t^{Fin}, y_t^{Fin}, p_t^{Fin}, t. List of $I(0)$ variables included in the VAR: i_t^{Fin}, sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
r = 0	r $\geq$ 1	47.635	42.340	39.340	43.712
r $\leq$ 1	r $\geq$ 2	13.784	25.770	23.080	26.605
r $\leq$ 2	r = 3	5.025	12.390	10.550	12.791

Cointegration Rank Test (Maximal Eigenvalue)

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r = 1	33.851	25.420	23.100
r $\leq$ 1	r = 2	8.758	19.220	17.180
r $\leq$ 2	r = 3	5.025	12.390	10.550

Note: the correction factor is 1.032.

Table C.22: Johansen's cointegration with unrestricted trends, Finland model 3.

104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: m_t^{Fin}, y_t^{Fin}. List of $I(0)$ variables included in the VAR: p_t^{Fin}, i_t^{Fin}, sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)				
Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r $\geq$ 1	8.824	23.830	21.230
r $\leq$ 1	r = 2	0.300	11.540	9.750

Cointegration Rank Test (Maximal Eigenvalue)				
Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r = 1	8.524	18.330	16.280
r $\leq$ 1	r = 2	0.300	11.540	9.750

Table C.23: Johansen's cointegration with restricted trends, Finland model 3.

104 observations from 1958 Q4 to 1984 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: m_t^{Fin}, y_t^{Fin}, t. List of $I(0)$ variables included in the VAR: p_t^{Den}, i_t^{Den}, sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)				
Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r $\geq$ 1	9.732	25.770	23.080
r $\leq$ 1	r = 2	0.328	12.390	10.550

Cointegration Rank Test (Maximal Eigenvalue)				
Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
r = 0	r = 1	9.404	19.220	17.180
r$\leq$ 1	r = 2	0.3279	12.390	10.550

C.4.3 Common factor analysis

Table C.24: COMFAC model, Denmark.

Dependent variable is m_t^{Den}. 54 observations used for estimation from 1974 Q2 to 1987 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	2.550	0.668	3.820	0.000
m_{t-1}^{Den}	0.700	0.080	8.712	0.000
y_t^{Den}	0.401	0.183	2.194	0.034
y_{t-1}^{Den}	-0.206	0.173	-1.195	0.239
p_t^{Den}	-0.597	0.581	-1.027	0.310
p_{t-1}^{Den}	0.616	0.578	1.067	0.292
b_t^{den}	-0.874	0.381	-2.293	0.027
b_{t-1}^{den}	-0.227	0.453	-0.502	0.618
i_t^{Den}	0.365	0.641	0.570	0.572
i_{t-1}^{Den}	-0.260	0.643	-0.405	0.688

R-squared	0.980	R-bar-squared	0.976
S.E. of regression	0.024	F-stat. F (9, 44)	244.793 [0.000]
Mean of dependent variable	11.757	S.D. of dependent variable	0.153
Residual sum of squares	0.024	Equation loglikelihood	131.505
Akaike info. criterion	121.505	Schwarz Bayesian criterion	111.560
Durbin's h-statistic	-4.228 [0.000]		

Test Statistics	LM Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	20.977	0.000	F (4, 40) =	6.352	0.000
B:Functional form	χ^2 (1) =	0.725	0.394	F (1, 43) =	0.585	0.448
C:Normality	χ^2 (2) =	11.446	0.003	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	0.022	0.883	F (1, 52) =	0.021	0.886

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's RESET test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table C.25: Wald tests of restrictions, Denmark.

54 observations used for estimation from 1974 Q2 to 1987 Q3.

Restriction(s) for Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	0.485	0.486
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	1.231	0.267
$\eta_1 + \eta_0\beta_1 = 0$	$\chi^2(1) =$	9.149	0.002
$\mu_1 + \mu_0\beta_1 = 0$	$\chi^2(1) =$	$0.9491E-4$	0.992
$\gamma_1 + \gamma_0\beta_1 = 0$, $\delta_1 + \delta_0\beta_1 = 0$, $\eta_1 + \eta_0\beta_1 = 0$, $\mu_1 + \mu_0\beta_1 = 0$	$\chi^2(4) =$	14.984	0.005

Table C.26: COMFAC model 1, Finland.

Dependent variable is m_t^{Fin}. 105 observations used for estimation from 1958 Q3 to 1984 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	-0.227	0.096	-2.374	0.020
m_{t-1}^{Fin}	0.813	0.057	14.172	0.000
y_t^{Fin}	0.603	0.094	6.409	0.000
y_{t-1}^{Fin}	-0.408	0.103	-3.964	0.000
p_t^{Fin}	-0.907	0.427	-2.125	0.036
p_{t-1}^{Fin}	-0.034	0.422	-0.080	0.937
i_t^{Fin}	0.348	0.127	2.738	0.007
i_{t-1}^{Fin}	-0.328	0.128	-2.560	0.012

R-squared	0.976	R-bar-squared	0.974
S.E. of regression	0.050	F-stat. F (7, 97)	564.952 [0.000]
Mean of dependent variable	3.391	S.D. of dependent variable	0.312
Residual sum of squares	0.243	Equation loglikelihood	169.712
Akaike info. criterion	161.712	Schwarz Bayesian criterion	151.096
Durbin's h-statistic	-1.160 [0.246]		

Test Statistics	Lm Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	8.487	0.075	F (4, 93) =	2.045	0.094
B:Functional form	χ^2 (1) =	0.774	0.379	F (1, 96) =	0.713	0.400
C:Normality	χ^2 (2) =	0.366	0.833	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	$0.29E-4$	0.996	F (1, 103) =	$0.28E-4$	0.996

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table C.27: Wald test of restrictions, model 1, Finland.

105 observations used for estimation from 1958 Q3 to 1984 Q3.

Restriction(s) for the Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	7.783	0.005
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	2.768	0.096
$\eta_1 + \eta_0\beta_1 = 0$	$\chi^2(1) =$	0.209	0.647
$\gamma_1 + \gamma_0\beta_1 = 0$, $\delta_1 + \delta_0\beta_1 = 0$, $\eta_1 + \eta_0\beta_1 = 0$	$\chi^2(2) =$	8.910	0.031

Table C.28: COMFAC model 2, Finland.

Dependent variable is m_t^{Fin}. 105 observations used for estimation from 1958 Q3 to 1984 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	-0.231	0.093	-2.494	0.014
m_{t-1}^{Fin}	0.809	0.058	14.067	0.000
y_t^{Fin}	0.676	0.093	7.274	0.000
y_{t-1}^{Fin}	-0.477	0.104	-4.610	0.000
p_t^{Fin}	-0.910	0.441	-2.063	0.042
p_{t-1}^{Fin}	0.090	0.427	0.211	0.833

R-squared	0.974	R-bar-squared	0.973
S.E. of regression	0.052	F-stat. F (5, 99)	739.501 [0.000]
Mean of dependent variable	3.391	S.D. of dependent variable	0.312
Residual sum of squares	0.264	Equation loglikelihood	165.225
Akaike info. criterion	159.225	Schwarz Bayesian criterion	151.263
Durbin's h-statistic	-1.689 [0.091]		

Test Statistics	Lm Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	8.359	0.079	F (4, 95) =	2.054	0.093
B:Functional form	χ^2 (1) =	0.606	0.436	F (1, 98) =	0.569	0.452
C:Normality	χ^2 (2) =	2.611	0.271	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	$0.13E-3$	0.991	F (1, 103) =	$0.13E-3$	0.991

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table C.29: Wald test of restrictions, model 2, Finland.

105 observations used for estimation from 1958 Q3 to 1984 Q3.

Restriction(s) for the Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	6.266	0.012
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	1.890	0.169
$\gamma_1 + \gamma_0\beta_1 = 0$, $\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(2) =$	6.642	0.036

Table C.30: COMFAC model 3, Finland.

Dependent variable is m_t^{Fin}. 105 observations used for estimation from 1958 Q3 to 1984 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	-0.175	0.085	-2.072	0.041
m_{t-1}^{Fin}	0.803	0.058	13.837	0.000
y_t^{Fin}	0.731	0.086	8.525	0.000
y_{t-1}^{Fin}	-0.544	0.098	-5.535	0.000

R-squared	0.973	R-bar-squared	0.972
S.E. of regression	0.052	F-stat. F (3, 101)	1202.000 [0.000]
Mean of dependent variable	3.391	S.D. of dependent variable	0.312
Residual sum of squares	0.276	Equation loglikelihood	162.923
Akaike info. criterion	158.923	Schwarz Bayesian criterion	153.615
Durbin's h-statistic	-1.638 [0.101]		

Test Statistics		Lm Version	Prob.		F Version	Prob.
A:Serial correlation	χ^2 (4) =	8.598	0.072	F (4, 97) =	2.163	0.079
B:Functional form	χ^2 (1) =	1.195	0.274	F (1, 100) =	1.151	0.286
C:Normality	χ^2 (2) =	1.545	0.462	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	$0.37E-4$	0.995	F (1, 103) =	$0.37E-4$	0.995

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table C.31: Wald test of restrictions, model 3, Finland.

105 observations used for estimation from 1958 Q3 to 1984 Q3.

Restriction(s) for the Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	3.309	0.069

C.5 Testing for Fractional Integration

Table C.32: Fractional integration analysis, Denmark.

	EML	NLS	GPH	GSP	FDF†	FADF†
m_t	1.159 (0.123)	1.176 (0.133)	1.168 (0.171)	0.993 (0.096)	-0.850	-1.490
y_t	0.360 (0.281)	0.577 (0.276)	1.23 (0.171)	1.08 (0.096)	-0.076	2.267
p_t	0.741 (0.301)	0.674 (0.231)	0.994 (0.171)	0.870 (0.096)	0.077	-1.569
i_t	0.574 (0.275)	0.521 (0.260)	1.171 (0.171)	1.084 (0.962)	-0.223	-1.125
b_t	0.738 (0.278)	0.727 (0.218)	1.377 (0.171)	1.275 (0.096)	2.339	-2.191

Note: standard errors in parentheses.
†Based on the EML estimate of d.

Table C.33: Fractional integration analysis, Finland.

	EML	NLS	GPH	GSP	FDF†	FADF†
m_t	0.778 (0.090)	0.762 (0.090)	0.830 (0.112)	0.590 (0.069)	-2.94	-2.046
y_t	0.559 (0.084)	0.570 (0.086)	0.745 (0.113)	0.523 (0.693)	1.250	8.502
p_t	0.236 (0.099)	0.210 (0.096)	0.410 (0.114)	0.394 (0.113)	-6.45	-2.63
i_t	0.621 (0.108)	0.622 (0.103)	0.759 (0.112)	0.796 (0.069)	-2.73	-3.54

Note: standard errors in parentheses.
†Based on the EML estimate of d.

C.6 Nonlinear Inference

Table C.34: Hamilton Analysis, Denmark & Finland.

	DENMARK	FINLAND
	ESTIMATES (standard errors)	
Linear		
c	7.338 (1.142)	−0.554 (0.348)
y_t	0.781 (0.190)	0.877 (0.079)
p_t	0.129 (0.061)	−0.826 (0.456)
i_t	−0.066 (0.063)	0.133 (0.171)
b_t	−0.111 (0.039)	
Non-linear		
σ	0.009 (0.006)	0.050 (0.005)
ζ	5.376 (4.014)	1.289 (0.311)
y_t	3.412 (2.335)	4.791 (0.748)
p_t	6.490 (1.394)	0.009 (0.360)
i_t	−0.00002 (0.510)	2.238 (2.167)
b_t	0.000003 (0.569)	

Table C.35: Post-Hamilton piecewise linear regression, Denmark.

Dependent variable is m_t^{Den}. 55 observations used for estimation from 1974 Q1 to 1987 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	6.664	0.6709	9.933	0.000
y_t^{Den}	0.926	0.110	8.424	0.000
p_t^{Den}	0.539	0.144	3.742	0.000
$\left(p_t^{Den} - p_1^{Den^*}\right) D_{1t}$	-0.652	0.157	-4.159	0.000
$\left(p_t^{Den} - p_2^{Den^*}\right) D_{2t}$	1.252	0.173	7.240	0.000
b_t^{Den}	-1.475	0.311	-4.742	0.000
i_t^{Den}	0.612	0.583	1.049	0.299

R-squared	0.966	R-bar-squared	0.961
S.E. of regression	0.030	F-stat. F (6, 48)	224.489 [0.000]
Mean of dependent variable	11.754	S.D. of dependent variable	0.152
Residual sum of squares	0.043	Equation loglikelihood	118.605
Akaike info. criterion	111.605	Schwarz Bayesian criterion	104.580
DW-statistic	0.938		

Test Statistics	Lm Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	23.521	0.000	F (4, 44) =	8.219	0.000
B:Functional form	χ^2 (1) =	1.660	0.198	F (1, 47) =	1.462	0.233
C:Normality	χ^2 (2) =	2.806	0.246	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	0.021	0.886	F (1, 53) =	0.020	0.888

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table C.36: Post-Hamilton piecewise linear regression, Finland.

Dependent variable is m_t^{Fin}. 106 observations used for estimation from 1958 Q2 to 1984 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	1.771	0.268	6.611	0.000
i_t^{Fin}	0.310	0.115	2.691	0.008
p_t^{Fin}	-0.324	0.464	-0.698	0.487
y_t^{Fin}	0.297	0.064	4.630	0.000
$\left(y_t^{Fin} - y_t^{Fin^*}\right) D_t$	0.879	0.085	10.350	0.000

R-squared	0.964	R-bar-squared	0.963
S.E. of regression	0.061	F-stat. F (4, 101)	675.355 [0.000]
Mean of dependent variable	3.386	S.D. of dependent variable	0.315
Residual sum of squares	0.376	Equation loglikelihood	148.649
Akaike info. criterion	143.649	Schwarz Bayesian criterion	136.990
DW-statistic	0.971		

Test Statistics	LM Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	29.723	0.000	F (4, 97) =	9.450	0.000
B:Functional form	χ^2 (1) =	1.937	0.164	F (1, 100) =	1.861	0.176
C:Normality	χ^2 (2) =	3.123	0.210	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	$0.25E-3$	0.987	F (1, 104) =	$0.24E-3$	0.988

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's RESET test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

C.7 Smooth Transition Regression

Table C.37: STR tests of linearity

Model	Transition Variable	P Values of F-test for Hypotheses			
		H_0	H_{04}	H_{03}	H_{02}
1	i_t^{Fin}	$1.02E-3$	$7.52E-4$	$1.19E-1$	$1.92E-1$
	p_t^{Fin}	$6.68E-2$	$9.73E-1$	$2.31E-2$	$9.16E-2$
	y_t^{Fin}	$2.47E-15$	$1.23E-6$	$7.76E-3$	$2.46E-10$
2	y_t^{Fin}	$3.05E-15$	$3.52E-6$	$1.64E-1$	$5.46E-12$

Appendix D

Appendix to Chapter 6

D.1 Data used in Chapter 6

Table D.1: Purchasing power parity data, Ireland.

	e_t^{DM}	$e_t^{£}$	p_t^{Ire}	i_t^{Ire}		e_t^{DM}	$e_t^{£}$	p_t^{Ire}	i_t^{Ire}
1975:1	1.061	0.000	3.318	10.377	1985:1	0.472	-0.204	4.432	14.413
1975:2	0.973	0.000	3.374	9.577	1985:2	0.472	-0.235	4.436	12.313
1975:3	1.021	0.000	3.388	10.270	1985:3	0.454	-0.194	4.438	10.183
1975:4	0.998	0.000	3.418	11.417	1985:4	0.448	-0.150	4.431	10.790
1976:1	0.911	0.000	3.472	9.420	1986:1	0.435	-0.130	4.430	15.227
1976:2	0.852	0.000	3.523	10.050	1986:2	0.437	-0.106	4.425	10.230
1976:3	0.737	0.000	3.564	11.137	1986:3	0.332	-0.072	4.419	10.580
1976:4	0.721	0.000	3.611	14.493	1986:4	0.328	-0.052	4.413	14.047
1977:1	0.742	0.000	3.666	11.967	1987:1	0.311	-0.082	4.425	13.810
1977:2	0.721	0.000	3.704	8.047	1987:2	0.315	-0.095	4.433	10.917
1977:3	0.723	0.000	3.723	7.100	1987:3	0.316	-0.110	4.444	9.543
1977:4	0.719	0.000	3.731	5.493	1987:4	0.304	-0.111	4.450	9.063
1978:1	0.652	0.000	3.757	6.180	1988:1	0.314	-0.153	4.459	8.667
1978:2	0.680	0.000	3.784	7.913	1988:2	0.319	-0.146	4.474	7.687
1978:3	0.670	0.000	3.809	9.363	1988:3	0.315	-0.167	4.488	7.810
1978:4	0.643	0.000	3.820	10.817	1988:4	0.316	-0.183	4.494	8.020
1979:1	0.673	-0.008	3.867	12.273	1989:1	0.311	-0.180	4.513	8.230
1979:2	0.658	-0.059	3.902	11.457	1989:2	0.307	-0.129	4.526	9.710
1979:3	0.644	-0.028	3.918	13.560	1989:3	0.310	-0.130	4.534	10.283
1979:4	0.641	-0.036	3.932	14.547	1989:4	0.301	-0.031	4.527	11.917
1980:1	0.655	-0.112	3.972	16.430	1990:1	0.313	-0.040	4.515	12.207
1980:2	0.649	-0.104	4.002	16.713	1990:2	0.316	-0.082	4.511	11.123
1980:3	0.653	-0.141	4.018	14.743	1990:3	0.316	-0.089	4.511	11.000
1980:4	0.642	-0.229	4.041	12.780	1990:4	0.305	-0.082	4.499	10.897
1981:1	0.625	-0.255	4.096	12.790	1991:1	0.313	-0.108	4.504	11.000
1981:2	0.622	-0.243	4.149	14.407	1991:2	0.313	-0.094	4.519	10.167
1981:3	0.620	-0.139	4.179	16.363	1991:3	0.313	-0.086	4.522	10.073
1981:4	0.600	-0.189	4.200	17.253	1991:4	0.305	-0.067	4.525	10.470
1982:1	0.573	-0.215	4.233	17.897	1992:1	0.309	-0.070	4.532	10.523
1982:2	0.566	-0.216	4.261	18.260	1992:2	0.310	-0.084	4.541	10.167
1982:3	0.555	-0.228	4.284	15.727	1992:3	0.295	0.044	4.535	12.583
1982:4	0.529	-0.145	4.295	13.423	1992:4	0.296	0.075	4.527	24.000
1983:1	0.479	-0.128	4.307	14.510	1993:1	0.218	0.002	4.555	15.897
1983:2	0.476	-0.212	4.321	13.590	1993:2	0.220	-0.044	4.576	7.473
1983:3	0.469	-0.233	4.346	12.740	1993:3	0.177	-0.047	4.593	6.607
1983:4	0.458	-0.245	4.362	11.760	1993:4	0.219	-0.049	4.590	6.500
1984:1	0.447	-0.200	4.381	12.690	1994:1	0.204	-0.033	4.591	6.120
1984:2	0.448	-0.207	4.396	12.523	1994:2	0.213	-0.015	4.592	5.767
1984:3	0.461	-0.197	4.407	13.293	1994:3	0.213	-0.011	4.587	5.917
1984:4	0.467	-0.154	4.415	14.417	1994:4	0.203	-0.010	4.589	5.917

Continued on next page.

	e_t^{DM}	$e_t^{£}$	p_t^{Ire}	i_t^{Ire}		e_t^{DM}	$e_t^{£}$	p_t^{Ire}	i_t^{Ire}
1995:1	0.135	0.003	4.601	6.523	1999:3	0.239	-0.196	4.658	2.703
1995:2	0.149	0.029	4.602	6.730	1999:4	0.239	-0.237	4.667	3.430
1995:3	0.159	0.023	4.605	6.153	2000:1	0.239	-0.274	4.690	3.543
1995:4	0.163	0.035	4.611	5.583	2000:2	0.239	-0.220	4.709	4.260
1996:1	0.172	0.031	4.608	5.107	2000:3	0.239	-0.278	4.722	4.737
1996:2	0.215	0.028	4.605	5.110	2000:4	0.239	-0.233	4.735	5.020
1996:3	0.222	0.024	4.595	5.700	2001:1	0.239	-0.240	4.725	4.747
1996:4	0.290	-0.010	4.591	5.757	2001:2	0.239	-0.266	4.740	4.593
1997:1	0.297	-0.038	4.586	5.777	2001:3	0.239	-0.237	4.733	4.267
1997:2	0.297	-0.097	4.601	6.233	2001:4	0.239	-0.259	4.727	3.443
1997:3	0.273	-0.103	4.608	6.220	2002:1	0.239	-0.251	4.738	3.363
1997:4	0.271	-0.145	4.617	6.137	2002:2	0.239	-0.193	4.733	3.447
1998:1	0.252	-0.210	4.644	5.920	2002:3	0.239	-0.222	4.707	3.357
1998:2	0.255	-0.175	4.638	6.253	2002:4	0.239	-0.191	4.701	3.107
1998:3	0.246	-0.128	4.628	5.897	2003:1	0.239	-0.133	4.671	2.683
1998:4	0.241	-0.112	4.608	3.633	2003:2	0.239	-0.129	4.635	2.367
1999:1	0.239	-0.167	4.622	3.093	2003:3	0.239	-0.123	4.624	2.140
1999:2	0.239	-0.183	4.645	2.637					

Table D.2: Purchasing power parity data, Germany & UK.

	p_t^{Ger}	p_t^{UK}	i_t^{Ger}	i_t^{UK}		p_t^{Ger}	p_t^{UK}	i_t^{Ger}	i_t^{UK}
1975:1	4.146	3.127	6.627	9.800	1979:4	4.292	3.768	9.330	15.137
1975:2	4.149	3.182	4.920	9.390	1980:1	4.317	3.820	9.157	16.197
1975:3	4.149	3.211	4.157	10.430	1980:2	4.336	3.861	10.197	15.933
1975:4	4.153	3.246	4.133	11.057	1980:3	4.337	3.877	9.217	14.663
1976:1	4.165	3.288	3.797	8.780	1980:4	4.345	3.896	9.577	13.443
1976:2	4.179	3.318	3.843	10.643	1981:1	4.367	3.926	11.247	11.910
1976:3	4.191	3.360	4.530	11.393	1981:2	4.387	3.959	13.160	11.540
1976:4	4.191	3.408	4.823	13.990	1981:3	4.405	3.978	12.787	14.037
1977:1	4.205	3.466	4.740	10.423	1981:4	4.413	4.000	11.227	14.697
1977:2	4.211	3.512	4.447	7.460	1982:1	4.426	4.024	10.200	12.847
1977:3	4.211	3.541	4.187	6.340	1982:2	4.438	4.038	9.263	12.357
1977:4	4.211	3.555	4.087	5.733	1982:3	4.446	4.055	8.880	9.973
1978:1	4.211	3.578	3.517	5.917	1982:4	4.450	4.072	7.170	9.293
1978:2	4.217	3.595	3.613	8.250	1983:1	4.445	4.089	5.700	10.667
1978:3	4.217	3.614	3.717	9.110	1983:2	4.449	4.108	5.367	9.543
1978:4	4.221	3.630	3.953	11.137	1983:3	4.458	4.121	5.720	9.237
1979:1	4.238	3.656	4.170	11.920	1983:4	4.466	4.124	6.320	8.843
1979:2	4.260	3.694	5.973	12.013	1984:1	4.475	4.142	5.977	8.697
1979:3	4.278	3.742	7.273	13.350	1984:2	4.481	4.164	6.023	8.713

Continued on next page.

	p_t^{Ger}	p_t^{UK}	i_t^{Ger}	i_t^{UK}		p_t^{Ger}	p_t^{UK}	i_t^{Ger}	i_t^{UK}
1984:3	4.483	4.174	5.990	10.423	1994:2	4.582	4.563	5.287	4.867
1984:4	4.491	4.193	5.953	9.353	1994:3	4.585	4.569	5.013	5.473
1985:1	4.501	4.220	6.140	12.527	1994:4	4.590	4.574	5.277	5.690
1985:2	4.506	4.228	5.847	11.890	1995:1	4.601	4.593	5.110	6.137
1985:3	4.504	4.233	4.947	11.000	1995:2	4.606	4.603	4.600	6.437
1985:4	4.500	4.234	4.860	11.110	1995:3	4.607	4.609	4.403	6.573
1986:1	4.492	4.238	4.607	11.563	1995:4	4.606	4.614	4.013	6.330
1986:2	4.481	4.240	4.600	9.497	1996:1	4.607	4.626	3.440	5.893
1986:3	4.474	4.240	4.597	9.510	1996:2	4.606	4.631	3.337	5.717
1986:4	4.470	4.248	4.737	10.630	1996:3	4.605	4.631	3.263	5.523
1987:1	4.472	4.264	4.203	10.027	1996:4	4.607	4.636	3.180	6.050
1987:2	4.474	4.272	3.817	8.650	1997:1	4.609	4.639	3.197	5.910
1987:3	4.476	4.278	3.967	9.527	1997:2	4.610	4.638	3.180	6.253
1987:4	4.478	4.286	4.143	8.520	1997:3	4.615	4.642	3.237	6.850
1988:1	4.480	4.296	3.423	8.373	1997:4	4.616	4.642	3.687	7.047
1988:2	4.488	4.308	3.647	8.230	1998:1	4.614	4.642	3.533	6.930
1988:3	4.494	4.317	5.107	11.047	1998:2	4.613	4.642	3.607	7.117
1988:4	4.503	4.324	5.147	12.197	1998:3	4.610	4.640	3.510	7.050
1989:1	4.516	4.340	6.267	12.403	1998:4	4.605	4.636	3.527	6.120
1989:2	4.526	4.354	6.840	13.040	1999:1	4.600	4.637	3.093	5.080
1989:3	4.526	4.363	7.190	13.367	1999:2	4.604	4.645	2.637	4.850
1989:4	4.529	4.376	8.177	14.497	1999:3	4.611	4.645	2.703	5.003
1990:1	4.530	4.392	8.363	14.460	1999:4	4.617	4.647	3.430	5.360
1990:2	4.534	4.414	8.300	14.473	2000:1	4.627	4.651	3.543	5.813
1990:3	4.541	4.426	8.393	14.283	2000:2	4.634	4.660	4.260	5.920
1990:4	4.550	4.443	8.897	12.960	2000:3	4.644	4.662	4.737	5.823
1991:1	4.554	4.454	9.173	12.093	2000:4	4.649	4.663	5.020	5.660
1991:2	4.557	4.473	9.107	10.887	2001:1	4.649	4.657	4.747	5.367
1991:3	4.565	4.476	9.243	10.073	2001:2	4.656	4.660	4.593	5.017
1991:4	4.567	4.481	9.463	10.070	2001:3	4.653	4.658	4.267	4.640
1992:1	4.571	4.490	9.613	9.943	2001:4	4.647	4.652	3.443	3.907
1992:2	4.579	4.503	9.763	9.483	2002:1	4.650	4.651	3.363	3.917
1992:3	4.580	4.504	9.720	9.177	2002:2	4.655	4.658	3.447	3.957
1992:4	4.578	4.511	8.973	6.527	2002:3	4.656	4.659	3.357	3.770
1993:1	4.579	4.525	8.323	5.340	2002:4	4.656	4.660	3.107	3.787
1993:2	4.579	4.542	7.677	5.217	2003:1	4.663	4.667	2.683	3.563
1993:3	4.578	4.545	6.827	5.133	2003:2	4.660	4.670	2.367	3.440
1993:4	4.575	4.549	6.353	4.900	2003:3	4.658	4.672	2.140	3.460
1994:1	4.579	4.557	5.880	4.830					

D.2 Figures from Chapter 6

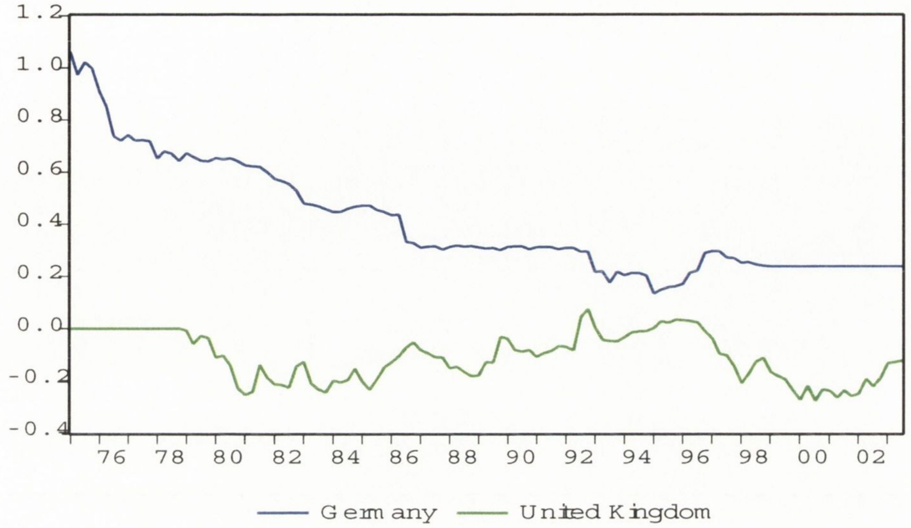

Figure D.1: The logs of the DM-IR£ and sterling-IR£ exchange rates.

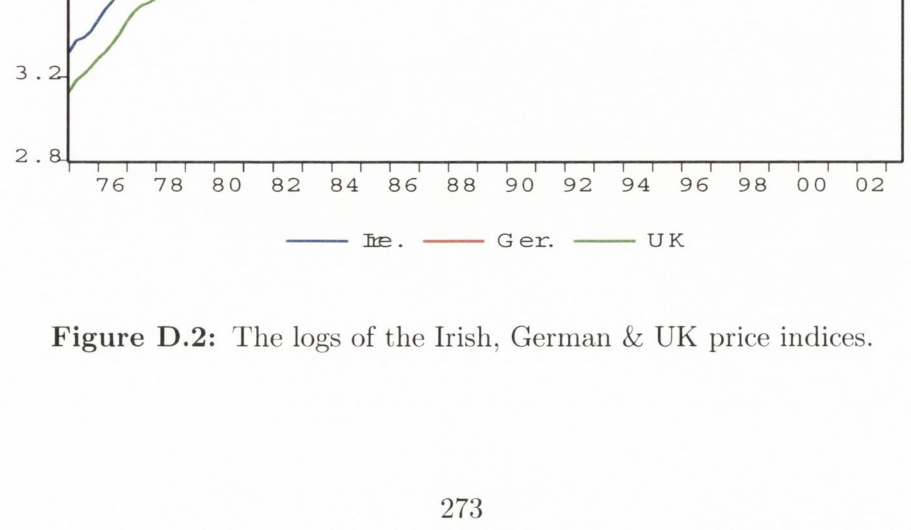

Figure D.2: The logs of the Irish, German & UK price indices.

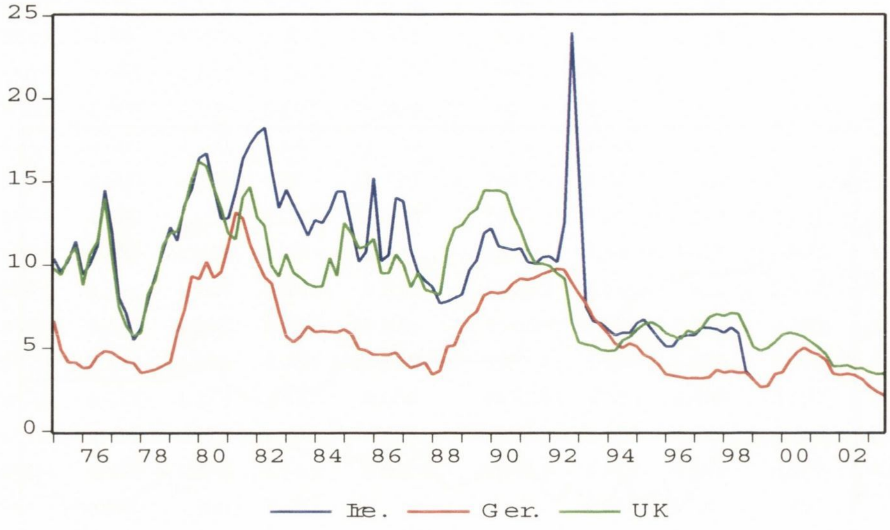

Figure D.3: The Irish, German & UK short-term interest rates.

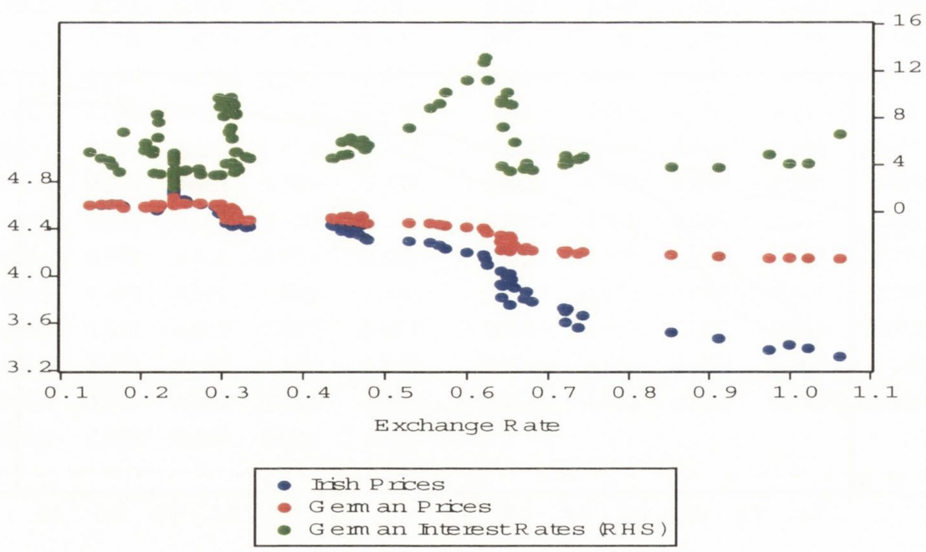

Figure D.4: Prices and German interest rate versus exchange rate.

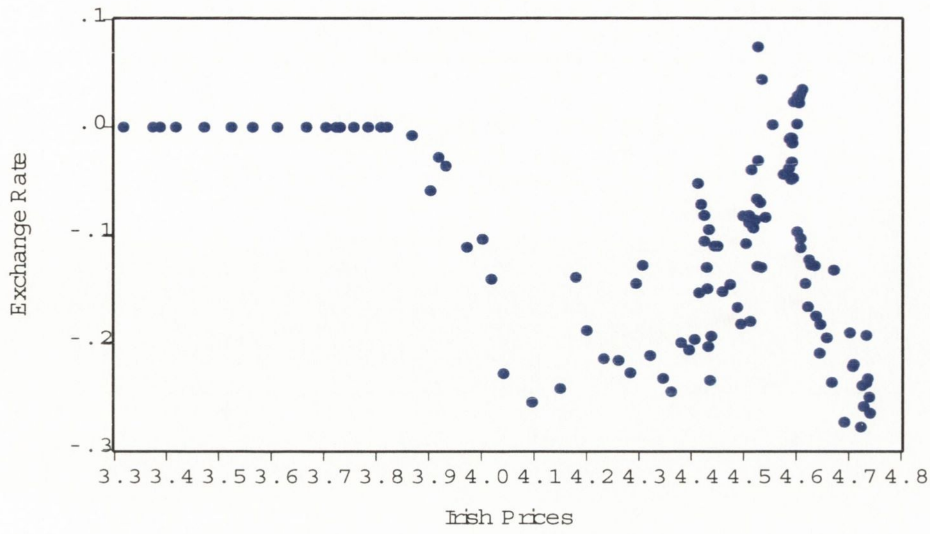

Figure D.5: Irish prices versus exchange rate.

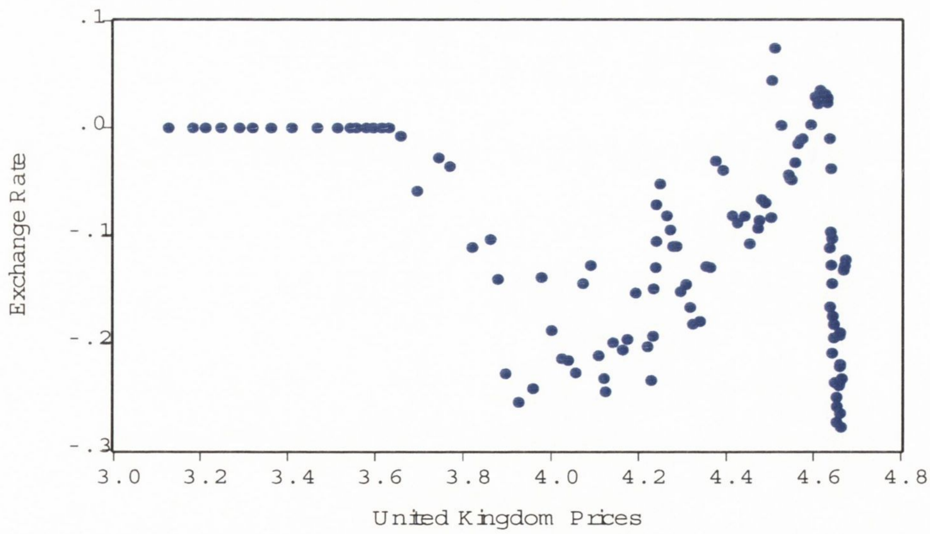

Figure D.6: UK prices versus exchange rate.

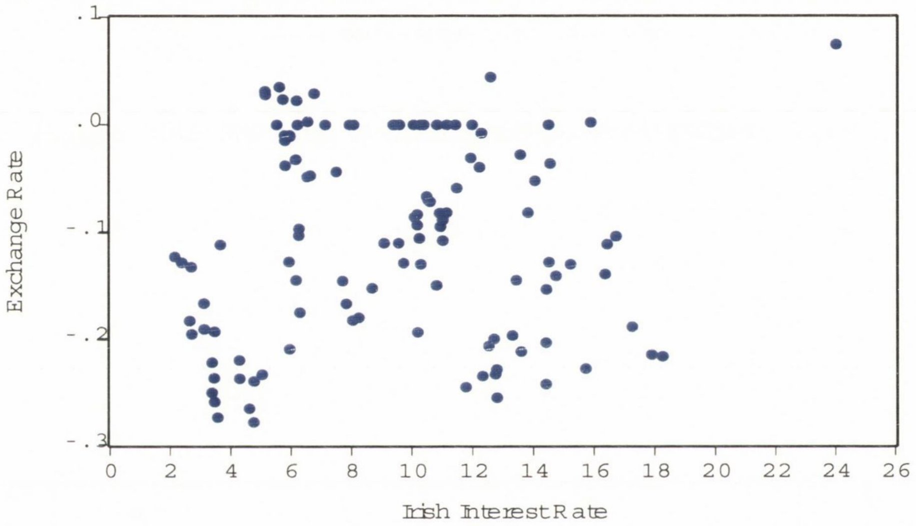

Figure D.7: Irish interest rate versus exchange rate.

D.3 Preliminary Analysis

Table D.3: Unit root tests, Ireland, Germany & UK.

VARIABLES	ADF	p-value	No. of Lags
		IRELAND & GERMANY	
Nominal exchange rate	-1.119	0.266	7
Irish price level	-2.155	0.034	4
German price level	-1.933	0.056	2
Irish interest rate	-1.085	0.250‡	2
German interest rate	-0.936	0.309‡	1
		IRELAND & UNITED KINGDOM	
Nominal exchange rate	-1.221	0.203‡	0
Irish price level	-2.155	0.034	4
UK price level	-1.722	0.088	8
Irish interest rate	-1.085	0.250‡	2
UK interest rate	-0.645	0.436‡	10

‡Trend and constant not included. MacKinnon (1996) p-values used.

Table D.4: Seasonal unit root tests, Ireland (T = 115).

Variable	Test Statistics [critical values]					Lag Length
	t_{π_1}	t_{π_2}	F_{34}	F_{234}	F_{1234}	
HEGY test without deterministic terms						
$e_t^{DM/IR£}$	−3.387 [−1.91]	−5.912 [−1.93]	76.246 [3.11]	146.904 [2.78]	187.824 [2.55]	0
$e_t^{Stg£/IR£}$	−1.135 [−1.91]	−7.138 [−1.93]	38.512 [3.11]	100.853 [2.78]	75.700 [2.55]	0
p_t^{Ire}	0.007 [−1.91]	−2.734 [−1.93]	6.017 [3.11]	7.025 [2.78]	5.273 [2.55]	5
i_t^{Ire}	−1.116 [−1.91]	−7.391 [−1.93]	26.652 [3.11]	83.998 [2.78]	63.790 [2.55]	0
HEGY test with intercept and time trend						
$e_t^{DM/IR£}$	−3.601 [−3.40]	−5.662 [−1.93]	78.263 [3.05]	139.566 [2.74]	122.428 [4.19]	0
$e_t^{Stg£/IR£}$	−2.182 [−3.40]	−7.145 [−1.93]	38.661 [3.05]	100.673 [2.74]	76.893 [4.19]	0
p_t^{Ire}	−2.681 [−3.40]	−8.505 [−1.93]	35.030 [3.05]	286.926 [2.74]	300.699 [4.19]	0
i_t^{Ire}	−3.128 [−3.40]	−7.580 [−1.93]	27.510 [3.05]	87.094 [2.74]	68.634 [4.19]	0
HEGY test with intercept and seasonal dummy variables						
$e_t^{DM/IR£}$	−3.664 [−2.84]	−6.191 [−2.83]	68.397 [6.57]	140.843 [5.95]	163.697 [5.56]	0
$e_t^{Stg£/IR£}$	−2.230 [−2.84]	−6.970 [−2.83]	40.925 [6.57]	104.682 [5.95]	80.297 [5.56]	0
p_t^{Ire}	−4.024 [−2.84]	−4.189 [−2.83]	7.296 [6.57]	11.596 [5.95]	12.768 [5.56]	9
i_t^{Ire}	−1.637 [−2.84]	−7.158 [−2.83]	28.355 [6.57]	85.939 [5.95]	64.681 [5.56]	0
HEGY test with intercept, time trend and seasonal dummy variables						
$e_t^{DM/IR£}$	−3.455 [−3.39]	−6.117 [−2.82]	71.072 [6.55]	144.389 [5.93]	126.485 [6.31]	0
$e_t^{Stg£/IR£}$	−2.187 [−3.39]	−6.939 [−2.82]	40.556 [6.55]	103.587 [5.93]	79.099 [6.31]	0
p_t^{Ire}	−3.014 [−3.39]	−4.159 [−2.82]	7.165 [6.55]	11.394 [5.93]	11.136 [6.31]	9
i_t^{Ire}	−3.140 [−3.39]	−7.280 [−2.82]	29.053 [6.55]	88.190 [5.93]	69.457 [6.31]	0

Note: critical values for the 5 per cent level of significance, taken from Franses and Hobijn (1997).

Table D.5: Seasonal unit root tests, Germany & UK (T = 115).

Variable	Test Statistics [critical values]					Lag Length
	t_{π_1}	t_{π_2}	F_{34}	F_{234}	F_{1234}	
	Hegy test without deterministic terms					
p_t^{Ger}	1.914 [−1.91]	−9.041 [−1.93]	28.933 [3.11]	658.690 [2.78]	854.313 [2.55]	0
p_t^{UK}	0.052 [−1.91]	−2.578 [−1.93]	2.063 [3.11]	3.856 [2.78]	2.892 [2.55]	11
i_t^{Ger}	−0.965 [−1.91]	−6.215 [−1.93]	58.522 [3.11]	336.624 [2.78]	253.665 [2.55]	0
i_t^{UK}	−0.645 [−1.91]	−2.916 [−1.93]	15.890 [3.11]	16.685 [2.78]	12.699 [2.55]	7
	Hegy test with intercept and time trend					
p_t^{Ger}	−2.245 [−3.40]	−8.781 [−1.93]	28.696 [3.05]	503.000 [2.74]	389.947 [4.19]	0
p_t^{UK}	−2.733 [−3.40]	−2.641 [−1.93]	1.956 [3.05]	3.902 [2.74]	4.998 [4.19]	11
i_t^{Ger}	−2.572 [−3.40]	−6.341 [−1.93]	57.309 [3.05]	341.586 [2.74]	259.467 [4.19]	0
i_t^{UK}	−3.096 [−3.40]	−3.454 [−1.93]	25.908 [3.05]	26.385 [2.74]	23.100 [4.19]	3
	Hegy test with intercept and seasonal dummy variables					
p_t^{Ger}	−2.698 [−2.84]	−7.826 [−2.83]	33.473 [6.57]	535.983 [5.95]	550.008 [5.56]	0
p_t^{UK}	−3.786 [−2.84]	−3.053 [−2.83]	26.872 [6.57]	25.902 [5.95]	30.411 [5.56]	3
i_t^{Ger}	−2.071 [−2.84]	−6.408 [−2.83]	53.398 [6.57]	352.151 [5.95]	264.977 [5.56]	0
i_t^{UK}	−1.769 [−2.84]	−3.453 [−2.83]	27.257 [6.57]	27.694 [5.95]	21.619 [5.56]	3
	Hegy test with intercept, time trend and seasonal dummy variables					
p_t^{Ger}	−2.117 [−3.39]	−7.884 [−2.82]	32.493 [6.55]	537.680 [5.93]	416.800 [6.31]	0
p_t^{UK}	−2.849 [−3.39]	−3.036 [−2.82]	26.622 [6.55]	25.654 [5.93]	24.598 [6.31]	3
i_t^{Ger}	−2.560 [−3.39]	−6.437 [−2.82]	52.446 [6.55]	348.839 [5.93]	264.975 [6.31]	0
i_t^{UK}	−3.108 [−3.39]	−3.393 [−2.82]	27.830 [6.55]	27.972 [5.93]	24.318 [6.31]	3

Note: critical values for the 5 per cent level of significance, taken from Franses and Hobijn (1997).

D.4 Testing for Cointegration

D.4.1 The Engle-Granger 2-step method

Table D.6: Engle-Granger levels models, Ireland, Germany & UK.

Variables	Ireland-Germany		Ireland-United Kingdom	
Constant	2.854 (0.549)	1.804 (0.575)	0.859 (0.108)	0.833 (0.108)
Price levels				
Irish	−0.568 (0.083)	−0.672 (0.081)	−0.875 (0.111)	−1.029 (0.123)
Foreign	0.007 (0.200)	0.329 (0.203)	0.670 (0.095)	0.825 (0.110)
Interest rates				
Irish		0.005 (0.002)		0.007 (0.003)
Foreign		0.002 (0.003)		−0.003 (0.003)
AEG test on residuals	−2.475 [−3.817]	−2.835 [−4.5398]	−2.653 [−3.8172]	−2.728 [−4.540]
CRDW test	0.186 [0.48]	0.245 [0.68]	0.239 [0.48]	0.250 [0.68]

Note: standard errors in parentheses. 5 per cent AEG and CRDW critical values in square brackets.

Table D.7: Error-correction models, Ireland, Germany & UK.

Variables	Ireland-Germany		Ireland-United Kingdom	
Constant	−0.004 (−0.003)	−0.004 (−0.003)	0.004 (0.005)	0.001 (0.004)
Δ Price levels				
Irish	−0.686 (0.157)	−0.667 (0.164)	−1.105 (0.282)	−1.020 (0.284)
Foreign	1.021 (0.428)	0.927 (0.502)	0.831 (0.361)	0.715 (0.357)
Δ Interest Rates				
Irish		0.0004 (0.001)		0.005 (0.001)
Foreign		0.001 (0.004)		0.00006 (0.003)
ECM	−0.108 (0.039)	−0.107 (0.040)	−0.133 (0.049)	−0.124 (0.052)
ECM test critical values	−3.244 [0.134]	−3.787 [0.326]	−3.244 [0.148]	−3.787 [0.444]

Note: standard errors in parentheses, p-values in square brackets.

D.4.2 Johansen's maximum likelihood approach

Table D.8: Testing the order of the VAR, Ireland-Germany.

Based on 112 observations from 1975 Q4 to 2003 Q3. Order of VAR = 3. List of variables included in the unrestricted VAR: $e_t^{DM/IR£}$, p_t^{Ire}, p_t^{Ger}. List of deterministic variables: c, sc_1, sc_2, sc_3.

ORDER	LL	AIC	SBC	χ^2	LR TEST	PROB.	ADJUSTED LR TEST	PROB.
3	1099.100	1060.100	1007.100	-	-	-	-	-
2	1090.600	1060.600	1019.800	χ^2 (9) =	17.093	0.047	15.109	0.088
1	1046.300	1025.300	996.717	χ^2 (18) =	105.708	0.000	93.430	0.000
0	384.828	372.828	356.517	χ^2 (27) =	1428.600	0.000	1262.800	0.000

Table D.9: Test of restrictions for deterministic variables, Ireland-Germany.

Based on 113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2.

VARIABLE / STATISTIC	MAX. LL WITHOUT VARIABLE	MAX. LL WITH VARIABLE	LR TEST χ^2		PROB.
c	1090.500	1086.000	χ^2 (3) =	9.037	0.029
t	1090.500	1087.900	χ^2 (3) =	5.340	0.149
sc_1	1090.500	1080.300	χ^2 (3) =	20.433	0.000
sc_2	1090.500	1089.100	χ^2 (3) =	2.967	0.397
sc_3	1090.500	1089.600	χ^2 (3) =	1.847	0.605
$sc_1 + sc_2 + sc_3$	1090.500	1074.200	χ^2 (9) =	32.586	0.000

Note: Max. LL - Maximised value of loglikelihood.

Table D.10: Testing the order of the VAR including interest rates, Ireland-Germany.

Based on 112 observations from 1975 Q4 to 2003 Q3. Order of VAR = 3. List of variables included in the unrestricted VAR: $e_t^{DM/IR£}$, p_t^{Ire}, p_t^{Ger}, i_t^{Ire}, i_t^{Ger}. List of deterministic variables: c, t, sc_1, sc_2, sc_3.

ORDER	LL	AIC	SBC	χ^2	LR TEST	PROB.	ADJUSTED LR TEST	PROB.
3	838.307	738.307	602.382	-	-	-	-	-
2	822.812	747.812	645.869	χ^2 (25) =	30.989	0.189	25.456	0.437
1	765.700	715.700	647.737	χ^2 (50) =	145.215	0.000	119.284	0.000
0	61.762	36.762	2.781	χ^2 (75) =	1553.100	0.000	1275.800	0.000

Table D.11: Test of restrictions for deterministic variables including interest rates, Ireland-Germany.

Based on 113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2.

VARIABLE / STATISTIC	MAX. LL WITHOUT VARIABLE	MAX. LL WITH VARIABLE	LR TEST χ^2		PROB.
c	819.268	812.776	χ^2 (5) =	12.985	0.024
t	819.268	808.274	χ^2 (5) =	21.988	0.001
sc_1	819.268	805.464	χ^2 (5) =	27.609	0.000
sc_2	819.268	816.156	χ^2 (5) =	6.223	0.285
sc_3	819.268	818.418	χ^2 (5) =	1.701	0.889
$sc_1 + sc_2 + sc_3$	819.268	798.479	χ^2 (15) =	41.579	0.000

Note: Max. LL - Maximised value of loglikelihood.

Table D.12: Testing the order of the VAR, Ireland-UK.

Based on 112 observations from 1975 Q4 to 2003 Q3. Order of VAR = 3. List of variables included in the unrestricted VAR: $e_t^{Stg£/IR£}$, p_t^{Ire}, p_t^{UK}. List of deterministic variables: c, t, sc_1, sc_2, sc_3.

ORDER	LL	AIC	SBC	χ^2	LR TEST	PROB.	ADJUSTED LR TEST	PROB.
3	1029.700	987.726	930.637	-	-	-	-	-
2	1021.300	988.296	943.440	χ^2 (9) =	16.860	0.051	14.753	0.098
1	984.437	960.437	927.815	χ^2 (18) =	90.578	0.000	79.256	0.000
0	396.519	381.519	361.130	χ^2 (27) =	1266.400	0.000	1108.100	0.000

Table D.13: Test of restrictions for deterministic variables, Ireland-UK.

Based on 113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2.

VARIABLE / STATISTIC	MAX. LL WITHOUT VARIABLE	MAX. LL WITH VARIABLE	LR TEST χ^2		PROB.
c	1024.900	1015.500	χ^2 (3) =	18.751	0.000
t	1024.900	1019.000	χ^2 (3) =	11.882	0.008
sc_1	1024.900	1012.600	χ^2 (3) =	24.600	0.000
sc_2	1024.900	1023.200	χ^2 (3)=	3.371	0.338
sc_3	1024.900	1023.400	χ^2 (3) =	3.083	0.379
$sc_1 + sc_2 + sc_3$	1024.900	1005.100	χ^2 (9) =	39.667	0.000

Note: Max. LL - Maximised value of loglikelihood.

Table D.14: Testing the order of the VAR including interest rates, Ireland-UK.

Based on 112 observations from 1975 Q4 to 2003 Q3. Order of VAR = 3. List of variables included in the unrestricted VAR: $e_t^{Stg£/IR£}$, p_t^{Ire}, p_t^{UK}, i_t^{Ire}, i_t^{UK}. List of deterministic variables: c, t, sc_1, sc_2, sc_3.

ORDER	LL	AIC	SBC	χ^2	LR TEST	PROB.	ADJUSTED LR TEST	PROB.
3	694.636	594.636	458.711	-	-	-	-	-
2	682.069	607.069	505.126	χ^2 (25) =	25.132	0.455	20.644	0.712
1	633.013	583.013	515.050	χ^2 (50) =	123.246	0.000	101.238	0.000
0	-98.878	-123.878	-157.860	χ^2 (75) =	1587.000	0.000	1303.600	0.000

Table D.15: Test of restrictions for deterministic variables including interest rates, Ireland-UK.

Based on 113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2.

VARIABLE / STATISTIC	MAX. LL WITHOUT VARIABLE	MAX. LL WITH VARIABLE	LR TEST χ^2		PROB.
c	682.180	666.472	χ^2 (5) =	31.416	0.000
t	682.180	667.005	χ^2 (5) =	30.349	0.000
sc_1	682.180	668.940	χ^2 (5) =	26.479	0.000
sc_2	682.180	679.107	χ^2 (5) =	6.145	0.292
sc_3	682.180	681.229	χ^2 (5) =	1.902	0.863
$sc_1 + sc_2 + sc_3$	682.180	662.219	χ^2 (15) =	39.921	0.000

Note: Max. LL - Maximised value of loglikelihood.

Table D.16: Number of cointegrating relations by model, Ireland, Germany & UK.

Test Type	No Inpts No Trends	Rest'd Inpts No Trends	Unrest'd Inpts No Trends	Unrest'd Inpts Rest'd Trends	Unrest'd Inpts Unrest'd Trends
			Ireland & Germany excluding interest rates		
Trace	1	1	1	0	0
Max. Eig.	1	1	1	0	0
			Ireland & Germany including interest rates		
Trace	2	2	2	1	1
Max. Eig.	2	2	1	1	1
			Ireland & United Kingdom excluding interest rates		
Trace	1	1	1	1	1
Max. Eig.	1	1	1	1	0
			Ireland & United Kingdom including interest rates		
Trace	2	2	2	2	3
Max. Eig.	0	1	1	2	1

Note: 0.05 per cent critical values based on Osterwald-Lenum (1992).

Table D.17: Johansen's cointegration, Ireland-Germany.

113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: $e_t^{DM/IR£}$, p_t^{Ire}, p_t^{Ger}, c. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)†

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
$r = 0$	$r \geq 1$	39.203	34.870	31.930	45.680
$r \leq 1$	$r \geq 2$	13.347	20.180	17.880	-
$r \leq 2$	$r = 3$	5.903	9.160	7.530	-

Cointegration Rank Test (Maximal Eigenvalue)†

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
$r = 0$	$r = 1$	25.856	22.040	19.860
$r \leq 1$	$r = 2$	7.444	15.870	13.810
$r \leq 2$	$r = 3$	5.903	9.160	7.530

†Cointegration with restricted intercepts and no trends in the VAR.
Note: the correction factor is 1.310.

Table D.18: Johansen's cointegration including interest rates, Ireland-Germany.

113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: $e_t^{DM/IR£}$, p_t^{Ire}, p_t^{Ger}, i_t^{Ire}, i_t^{Ger}, t. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

COINTEGRATION RANK TEST (TRACE)†

HYPOTHESES		TRACE STATISTIC	0.05 CRITICAL VALUE	0.10 CRITICAL VALUE	MODIFIED 0.05 CRITICAL VALUE
$r = 0$	$r \geq 1$	111.587	87.170	82.880	98.328
$r \leq 1$	$r \geq 2$	57.298	63.000	59.160	-
$r \leq 2$	$r \geq 3$	31.448	42.340	39.340	-
$r \leq 3$	$r \geq 4$	15.809	25.770	23.080	-
$r \leq 4$	$r = 5$	6.057	12.390	10.550	-

COINTEGRATION RANK TEST (MAXIMAL EIGENVALUE)†

HYPOTHESES		MAXIMAL EIGENVALUE STATISTIC	0.05 CRITICAL VALUE	0.10 CRITICAL VALUE
$r = 0$	$r = 1$	54.290	37.860	35.040
$r \leq 1$	$r = 2$	25.850	31.790	29.130
$r \leq 2$	$r = 3$	15.639	25.420	23.100
$r \leq 3$	$r = 4$	9.751	19.220	17.180
$r \leq 4$	$r = 5$	6.057	12.390	10.550

†Cointegration with unrestricted intercepts and restricted trends in the VAR.
Note: the correction factor is 1.128.

Table D.19: Johansen's cointegration, Ireland-UK.

113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: $e_t^{Stg£/IR£}$, p_t^{Ire}, p_t^{UK}, t. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

Cointegration Rank Test (Trace)†

Hypotheses		Trace Statistic	0.05 Critical Value	0.10 Critical Value	Modified 0.05 Critical Value
$r = 0$	$r \geq 1$	57.532	42.340	39.340	70.030
$r \leq 1$	$r \geq 2$	21.695	25.770	23.080	-
$r \leq 2$	$r = 3$	4.788	12.390	10.550	-

Cointegration Rank Test (Maximal Eigenvalue)†

Hypotheses		Maximal Eigenvalue Statistic	0.05 Critical Value	0.10 Critical Value
$r = 0$	$r = 1$	35.838	25.420	23.100
$r \leq 1$	$r = 2$	16.907	19.220	17.180
$r \leq 2$	$r = 3$	4.788	12.390	10.550

†Cointegration with unrestricted intercepts and restricted trends in the VAR.
Note: the correction factor is 1.654.

Table D.20: Johansen's cointegration including interest rates, Ireland-UK.

113 observations from 1975 Q3 to 2003 Q3. Order of VAR = 2. List of variables included in the cointegrating vector: $e_t^{Stg£/IR£}$, p_t^{Ire}, p_t^{UK}, i_t^{Ire}, i_t^{UK}, t. List of $I(0)$ variables included in the VAR: sc_1, sc_2, sc_3.

COINTEGRATION RANK TEST (TRACE)†

HYPOTHESES		TRACE STATISTIC	0.05 CRITICAL VALUE	0.10 CRITICAL VALUE	MODIFIED 0.05 CRITICAL VALUE
$r = 0$	$r \geq 1$	127.997	87.170	82.880	85.427
$r \leq 1$	$r \geq 2$	77.194	63.000	59.160	61.740
$r \leq 2$	$r \geq 3$	41.665	42.340	39.340	41.493
$r \leq 3$	$r \geq 4$	21.103	25.770	23.080	-
$r \leq 4$	$r = 5$	4.707	12.390	10.550	-

COINTEGRATION RANK TEST (MAXIMAL EIGENVALUE)†

HYPOTHESES		MAXIMAL EIGENVALUE STATISTIC	0.05 CRITICAL VALUE	0.10 CRITICAL VALUE
$r = 0$	$r = 1$	50.803	37.860	35.040
$r \leq 1$	$r = 2$	35.530	31.790	29.130
$r \leq 2$	$r = 3$	20.562	25.420	23.100
$r \leq 3$	$r = 4$	16.395	19.220	17.180
$r \leq 4$	$r = 5$	4.707	12.390	10.550

†Cointegration with unrestricted intercepts and restricted trends in the VAR.
Note: the correction factor is 0.980.

D.4.3 Common factor analysis

Table D.21: COMFAC model, Ireland-Germany.

Dependent variable is $e_t^{DM/IR£}$. 114 observations used for estimation from 1975 Q2 to 2003 Q3.

REGRESSOR	COEFFICIENT	STANDARD ERROR	T-RATIO	PROB.
c	0.519	0.304	1.705	0.091
$e_{t-1}^{DM/IR£}$	0.891	0.044	20.414	0.000
p_t^{Ire}	-0.444	0.220	-2.020	0.046
p_{t-1}^{Ire}	0.433	0.215	2.020	0.046
p_t^{Ger}	0.727	0.546	1.331	0.186
p_{t-1}^{Ger}	-0.824	0.542	-1.522	0.131
i_t^{Ire}	$0.59E-3$	0.001	0.452	0.652
i_{t-1}^{Ire}	$-0.86E-3$	0.001	-0.675	0.501
i_t^{Ger}	$0.18E-3$	0.005	0.039	0.969
i_{t-1}^{Ger}	$0.38E-3$	0.005	0.085	0.933

R-squared	0.987	R-bar-squared	0.986
S.E. of regression	0.024	F-stat. F (9, 104)	895.112 [0.000]
Mean of dependent variable	0.404	S.D. of dependent variable	0.207
Residual sum of squares	0.061	Equation loglikelihood	267.268
Akaike info. criterion	257.268	Schwarz Bayesian criterion	243.587
Durbin's h-statistic	-0.206 [0.837]		

TEST STATISTICS	LM VERSION		PROB.	F VERSION		PROB.
A:Serial correlation	χ^2 (4) =	2.887	0.577	F (4, 100) =	0.650	0.628
B:Functional form	χ^2 (1) =	1.730	0.188	F (1, 103) =	1.587	0.211
C:Normality	χ^2 (2) =	137.414	0.000	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	8.974	0.003	F (1, 112) =	9.570	0.002

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's RESET test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table D.22: Wald tests of restrictions, Ireland-Germany

114 observations used for estimation from 1975 Q2 to 2003 Q3.

Restriction(s) for Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	0.741	0.389
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	2.347	0.126
$\eta_1 + \eta_0\beta_1 = 0$	$\chi^2(1) =$	0.112	0.738
$\mu_1 + \mu_0\beta_1 = 0$	$\chi^2(1) =$	0.119	0.730
$\gamma_1 + \gamma_0\beta_1 = 0,$ $\delta_1 + \delta_0\beta_1 = 0,$ $\eta_1 + \eta_0\beta_1 = 0,$ $\mu_1 + \mu_0\beta_1 = 0$	$\chi^2(4) =$	2.866	0.580

Table D.23: COMFAC model including interest rates, Ireland-Germany.

Dependent variable is $e_t^{DM/IR£}$. 114 observations used for estimation from 1975 Q2 to 2003 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	0.491	0.261	1.882	0.063
$e_{t-1}^{DM/IR£}$	0.891	0.039	22.834	0.000
p_t^{Ire}	-0.479	0.208	-2.305	0.023
p_{t-1}^{Ire}	0.465	0.204	2.284	0.024
p_t^{Ger}	0.840	0.444	1.891	0.061
p_{t-1}^{Ger}	-0.927	0.430	-2.157	0.033

R-squared	0.987	R-bar-squared	0.987
S.E. of regression	0.024	F-stat. F (5, 108)	1664.500 [0.000]
Mean of dependent variable	0.404	S.D. of dependent variable	0.207
Residual sum of squares	0.062	Equation loglikelihood	266.975
Akaike info. criterion	260.975	Schwarz Bayesian criterion	252.766
Durbin's h-statistic	-0.234 [0.815]		

Test Statistics	Lm Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	2.720	0.606	F (4, 104) =	0.635	0.638
B:Functional form	χ^2 (1) =	1.554	0.213	F (1, 107) =	1.479	0.227
C:Normality	χ^2 (2) =	139.058	0.000	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	8.996	0.003	F (1, 112) =	9.595	0.002

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table D.24: Wald tests of restrictions, including interest rates, Ireland-Germany.

114 observations used for estimation from 1975 Q2 to 2003 Q3.

Restriction(s) for Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	0.995	0.318
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	3.575	0.059
$\gamma_1 + \gamma_0\beta_1 = 0$, $\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(2) =$	4.233	0.120

Table D.25: COMFAC model, Ireland-UK.

Dependent variable is $e_t^{Stg£/IR£}$. 114 observations used for estimation from 1975 Q2 to 2003 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	0.117	0.094	1.250 [0.214]	
$e_{t-1}^{Stg£/IR£}$	0.879	0.052	17.047	0.000
p_t^{Ire}	-1.058	0.296	-3.571	0.001
p_{t-1}^{Ire}	1.057	0.274	3.860	0.000
p_t^{UK}	0.544	0.476	1.142	0.256
p_{t-1}^{UK}	-0.570	0.475	-1.199	0.233
i_t^{Ire}	0.005	0.002	2.760	0.007
i_{t-1}^{Ire}	-0.004	0.002	-2.375	0.019
i_t^{UK}	-0.002	0.003	-0.487	0.627
i_{t-1}^{UK}	$-0.58E-3$	0.003	-0.170	0.865

R-squared	0.885	R-bar-squared	0.875
S.E. of regression	0.033	F-stat. F (9, 104)	88.806 [0.000]
Mean of dependent variable	-0.108	S.D. of dependent variable	0.093
Residual sum of squares	0.112	Equation loglikelihood	233.265
Akaike info. criterion	223.265	Schwarz Bayesian criterion	209.584
Durbin's h-statistic	-0.476 [0.634]		

Test Statistics	Lm Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	3.213	0.523	F (4, 100) =	0.725	0.577
B:Functional form	χ^2 (1) =	1.537	0.215	F (1, 103) =	1.408	0.238
C:Normality	χ^2 (2) =	18.595	0.000	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	1.749	0.186	F (1, 112) =	1.745	0.189

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table D.26: Wald tests of restrictions, Ireland-UK.

114 observations used for estimation from 1975 Q2 to 2003 Q3.

Restriction(s) for Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	3.533	0.060
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	1.261	0.261
$\eta_1 + \eta_0\beta_1 = 0$	$\chi^2(1) =$	0.002	0.965
$\mu_1 + \mu_0\beta_1 = 0$	$\chi^2(1) =$	1.343	0.246
$\gamma_1 + \gamma_0\beta_1 = 0$, $\delta_1 + \delta_0\beta_1 = 0$, $\eta_1 + \eta_0\beta_1 = 0$, $\mu_1 + \mu_0\beta_1 = 0$	$\chi^2(4) =$	5.512	0.239

Table D.27: COMFAC model including interest rates, Ireland-UK.

Dependent variable is $e_t^{Stg£/IR£}$. 114 observations used for estimation from 1975 Q2 to 2003 Q3.

Regressor	Coefficient	Standard Error	T-Ratio	Prob.
c	0.147	0.091	1.619	0.108
$e_{t-1}^{Stg£/IR£}$	0.870	0.048	17.992	0.000
p_t^{Ire}	-1.152	0.287	-4.014	0.000
p_{t-1}^{Ire}	1.113	0.273	4.083	0.000
p_t^{UK}	0.381	0.463	0.824	0.412
p_{t-1}^{UK}	-0.377	0.452	-0.834	0.406

R-squared	0.875	R-bar-squared	0.869
S.E. of regression	0.034	F-stat. F (5, 108)	150.633 [0.000]
Mean of dependent variable	-0.108	S.D. of dependent variable	0.093
Residual sum of squares	0.121	Equation loglikelihood	228.393
Akaike info. criterion	222.393	Schwarz Bayesian criterion	214.185
Durbin's h-statistic	0.381 [0.703]		

Test Statistics	Lm Version		Prob.	F Version		Prob.
A:Serial correlation	χ^2 (4) =	2.973	0.562	F (4, 104) =	0.696	0.596
B:Functional form	χ^2 (1) =	5.100	0.024	F (1, 107) =	5.011	0.027
C:Normality	χ^2 (2) =	18.910	0.000	Not applicable	-	-
D:Heteroscedasticity	χ^2 (1) =	1.420	0.233	F (1, 112) =	1.413	0.237

Note: p-values in square brackets.
A:Lagrange multiplier test of residual serial correlation.
B:Ramsey's Reset test using the square of the fitted values.
C:Based on a test of skewness and kurtosis of residuals.
D:Based on the regression of squared residuals on squared fitted values.

Table D.28: Wald tests of restrictions, including interest rates, Ireland-UK.

114 observations used for estimation from 1975 Q2 to 2003 Q3.

Restriction(s) for Wald Test		Wald Statistic	Prob.
$\gamma_1 + \gamma_0\beta_1 = 0$	$\chi^2(1) =$	2.880	0.090
$\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(1) =$	0.431	0.511
$\gamma_1 + \gamma_0\beta_1 = 0,$ $\delta_1 + \delta_0\beta_1 = 0$	$\chi^2(2) =$	3.503	0.173

D.5 Testing for Fractional Integration

Table D.29: Fractional integration analysis, Ireland, Germany & UK.

VARIABLES	EML	NLS	GPH	GSP
		FADF‡		
	COMMON SERIES			
IRISH price level	1.46 (0.043)	1.50 (0.070)	1.01 (0.11)	0.89 (0.07)
	-	-	-	4.5
IRISH interest rate	0.79 (0.100)	0.78 (0.100)	0.97 (0.10)	0.80 (0.06)
	-3.22	-3.21	-3.35	-3.23
	IRELAND & GERMANY			
Nominal exchange rate	1.49 (0.135)	1.89 (0.097)	0.94 (0.11)	0.82 (0.07)
	-	-	-5.48	-5.51
GERMAN price level	1.46 (0.045)	1.57 (0.094)	1.022 (0.11)	0.918 (0.07)
	-	-	-	2.89
GERMAN interest rates	0.69 (0.240)	0.65† (0.230)	1.12 (0.11)	1.03 (0.07)
	-1.49	-1.48	-	-
	IRELAND & UNITED KINGDOM			
Nominal exchange rate	0.95 (0.086)	0.95 (0.087)	0.88 (0.11)	0.91 (0.07)
	-1.60	-1.60	-1.608	-1.60
UK price level	1.48 (0.023)	1.55 (0.064)	0.99 (0.11)	0.87 (0.07)
	-	-	5.03	4.69
UK interest rate	1.07 (0.094)	1.08 (0.096)	1.00 (0.11)	0.94 (0.07)
	-	-	-	-2.53

†Trend and constant not included. McKinnon (1996) p-values used.
‡Based on the EML estimator of d.
- indicates FADF test not applicable.
Note: standard errors in parentheses.

D.6 Nonlinear Inference

Table D.30: Hamilton analysis, Ireland, Germany & UK.

	Ireland & Germany			Ireland & United Kingdom	
	Estimates (standard errors)				
Linear					
c	0.332 (1.488)	0.769 (1.121)	c	1.176 (0.751)	0.907 (0.213)
p_t^{Ire}	−0.896 (0.191)	−0.836 (0.152)	p_t^{Ire}	−1.439 (0.308)	−1.093 (0.239)
p_t^{Ger}	0.892 (0.502)	0.724 (0.390)	p_t^{UK}	1.164 (0.320)	0.882 (0.218)
i_t^{Ire}		−0.0004 (0.002)	i_t^{Ire}		0.009 (0.004)
i_t^{Ger}		0.007 (0.005)	i_t^{UK}		−0.009 (0.004)
Nonlinear					
σ	0.019 (0.002)	0.010 (0.004)	σ	0.021 (0.003)	0.009 (0.004)
ζ	3.987 (0.817)	5.859 (2.551)	ζ	9.572 (2.109)	8.148 (4.368)
p_t^{Ire}	4.265 (0.375)	4.609 (1.103)	p_t^{Ire}	0.480 (0.116)	2.777 (1.214)
p_t^{Ger}	11.068 (0.733)	16.971 (3.021)	p_t^{UK}	−1.864 (0.044)	10.454 (1.846)
i_t^{Ire}		−0.032 (0.023)	i_t^{Ire}		0.118 (0.039)
i_t^{Ger}		−0.146 (0.052)	i_t^{UK}		$-2.26E-7$ (0.040)

Appendix E

Appendix to Chapter 7

E.1 Data used in Chapter 7

Table E.1: Forward exchange rate anomaly data, Australia.

DATE	AUS	LASR	LAST	DATE	AUS	LASR	LAST
03/01/97	2.152	0.330	0.323	10/10/97	2.187	0.343	0.405
10/01/97	2.150	0.333	0.318	17/10/97	2.192	0.343	0.401
17/01/97	2.132	0.329	0.324	24/10/97	2.317	0.367	0.399
24/01/97	2.103	0.324	0.322	31/10/97	2.365	0.376	0.384
31/01/97	2.099	0.323	0.316	07/11/97	2.430	0.384	0.388
07/02/97	2.128	0.328	0.318	14/11/97	2.426	0.388	0.384
14/02/97	2.115	0.326	0.325	21/11/97	2.398	0.383	0.384
21/02/97	2.080	0.318	0.321	28/11/97	2.446	0.391	0.389
28/02/97	2.101	0.323	0.332	05/12/97	2.453	0.393	0.390
07/03/97	2.031	0.308	0.331	12/12/97	2.479	0.397	0.394
14/03/97	2.006	0.303	0.337	19/12/97	2.537	0.407	0.401
21/03/97	2.040	0.310	0.343	26/12/97	2.528	0.406	0.399
28/03/97	2.083	0.319	0.347	02/01/98	2.508	0.402	0.405
04/04/97	2.112	0.325	0.349	09/01/98	2.487	0.398	0.406
11/04/97	2.065	0.315	0.356	16/01/98	2.457	0.393	0.416
18/04/97	2.105	0.324	0.357	23/01/98	2.500	0.401	0.406
25/04/97	2.088	0.320	0.356	30/01/98	2.371	0.378	0.409
02/05/97	2.073	0.315	0.342	06/02/98	2.445	0.391	0.413
09/05/97	2.079	0.318	0.333	13/02/98	2.421	0.387	0.415
16/05/97	2.112	0.325	0.329	20/02/98	2.417	0.386	0.413
23/05/97	2.124	0.328	0.327	27/02/98	2.402	0.383	0.419
30/05/97	2.143	0.332	0.341	06/03/98	2.430	0.388	0.427
06/06/97	2.136	0.331	0.336	13/03/98	2.461	0.394	0.442
13/06/97	2.170	0.338	0.344	20/03/98	2.490	0.399	0.440
20/06/97	2.194	0.343	0.353	27/03/98	2.483	0.398	0.440
27/06/97	2.221	0.348	0.350	03/04/98	2.506	0.402	0.432
04/07/97	2.245	0.353	0.346	10/04/98	2.532	0.406	0.425
11/07/97	2.271	0.358	0.345	17/04/98	2.566	0.412	0.414
18/07/97	2.252	0.355	0.342	24/04/98	2.544	0.408	0.425
25/07/97	2.247	0.354	0.365	01/05/98	2.538	0.407	0.427
01/08/97	2.200	0.345	0.377	08/05/98	2.553	0.410	0.432
08/08/97	2.147	0.334	0.384	15/05/98	2.581	0.414	0.435
15/08/97	2.151	0.335	0.387	22/05/98	2.568	0.412	0.440
22/08/97	2.134	0.332	0.391	29/05/98	2.587	0.416	0.471
29/08/97	2.195	0.344	0.392	05/06/98	2.709	0.435	0.463
05/09/97	2.158	0.337	0.397	12/06/98	2.736	0.439	0.450
12/09/97	2.216	0.348	0.395	19/06/98	2.663	0.428	0.454
19/09/97	2.219	0.349	0.393	26/06/98	2.731	0.439	0.463
26/09/97	2.204	0.346	0.406	03/07/98	2.647	0.426	0.455
03/10/97	2.199	0.345	0.402	10/07/98	2.639	0.424	0.445

Continued on next page.

Date	Aus	Lasr	Last	Date	Aus	Lasr	Last
17/07/98	2.600	0.418	0.430	28/05/99	2.449	0.389	0.400
24/07/98	2.653	0.427	0.434	04/06/99	2.452	0.390	0.398
31/07/98	2.677	0.430	0.430	11/06/99	2.418	0.384	0.397
07/08/98	2.700	0.434	0.420	18/06/99	2.434	0.386	0.400
14/08/98	2.713	0.436	0.418	25/06/99	2.398	0.380	0.402
21/08/98	2.790	0.448	0.412	02/07/99	2.354	0.372	0.402
28/08/98	2.923	0.468	0.416	09/07/99	2.331	0.368	0.398
04/09/98	2.834	0.455	0.426	16/07/99	2.372	0.375	0.406
11/09/98	2.781	0.447	0.429	23/07/99	2.423	0.385	0.413
18/09/98	2.839	0.456	0.429	30/07/99	2.479	0.395	0.405
25/09/98	2.875	0.461	0.438	06/08/99	2.471	0.393	0.410
02/10/98	2.849	0.457	0.433	13/08/99	2.460	0.391	0.402
09/10/98	2.774	0.446	0.417	20/08/99	2.535	0.404	0.403
16/10/98	2.682	0.431	0.418	27/08/99	2.509	0.400	0.405
23/10/98	2.712	0.436	0.412	03/09/99	2.486	0.396	0.403
30/10/98	2.676	0.430	0.420	10/09/99	2.490	0.397	0.404
06/11/98	2.611	0.419	0.405	17/09/99	2.500	0.398	0.399
13/11/98	2.592	0.416	0.398	24/09/99	2.512	0.401	0.399
20/11/98	2.561	0.411	0.410	01/10/99	2.521	0.402	0.392
27/11/98	2.596	0.416	0.409	08/10/99	2.515	0.401	0.401
04/12/98	2.686	0.458	0.412	15/10/99	2.574	0.411	0.393
11/12/98	2.704	0.434	0.410	22/10/99	2.552	0.407	0.393
18/12/98	2.698	0.431	0.412	29/10/99	2.571	0.411	0.398
25/12/98	2.735	0.438	0.409	05/11/99	2.542	0.406	0.400
01/01/99	2.704	0.433	0.401	12/11/99	2.507	0.399	0.405
08/01/99	2.584	0.413	0.412	19/11/99	2.532	0.404	0.406
15/01/99	2.600	0.416	0.403	26/11/99	2.531	0.404	0.415
22/01/99	2.605	0.417	0.394	03/12/99	2.525	0.403	0.414
29/01/99	2.609	0.418	0.388	10/12/99	2.547	0.406	0.412
05/02/99	2.508	0.400	0.388	17/12/99	2.501	0.399	0.412
12/02/99	2.517	0.402	0.385	24/12/99	2.503	0.399	0.418
19/02/99	2.549	0.407	0.386	31/12/99	2.461	0.392	0.417
26/02/99	2.576	0.412	0.391	07/01/00	2.496	0.398	0.420
05/03/99	2.558	0.409	0.394	14/01/00	2.450	0.390	0.425
12/03/99	2.566	0.410	0.384	21/01/00	2.487	0.396	0.425
19/03/99	2.581	0.412	0.384	28/01/00	2.518	0.402	0.428
26/03/99	2.552	0.407	0.381	04/02/00	2.511	0.400	0.416
02/04/99	2.526	0.403	0.374	11/02/00	2.532	0.404	0.413
09/04/99	2.548	0.407	0.370	18/02/00	2.536	0.405	0.416
16/04/99	2.492	0.397	0.374	25/02/00	2.581	0.412	0.412
23/04/99	2.473	0.394	0.388	03/03/00	2.596	0.415	0.416
30/04/99	2.427	0.386	0.396	10/03/00	2.569	0.410	0.408
07/05/99	2.438	0.387	0.393	17/03/00	2.587	0.413	0.399
14/05/99	2.430	0.386	0.395	24/03/00	2.621	0.419	0.401
21/05/99	2.419	0.384	0.405	31/03/00	2.627	0.420	0.401

Continued on next page.

Date	Aus	Lasr	Last	Date	Aus	Lasr	Last
07/04/00	2.643	0.422	0.408	16/02/01	2.733	0.437	0.432
14/04/00	2.654	0.424	0.409	23/02/01	2.775	0.443	0.435
21/04/00	2.657	0.425	0.413	02/03/01	2.787	0.445	0.448
28/04/00	2.682	0.428	0.410	09/03/01	2.877	0.459	0.427
05/05/00	2.574	0.411	0.409	16/03/01	2.894	0.462	0.423
12/05/00	2.616	0.418	0.411	23/03/01	2.879	0.460	0.436
19/05/00	2.599	0.415	0.402	30/03/01	2.908	0.464	0.441
26/05/00	2.609	0.416	0.413	06/04/01	2.899	0.463	0.433
02/06/00	2.599	0.415	0.401	13/04/01	2.833	0.453	0.445
09/06/00	2.570	0.410	0.410	20/04/01	2.777	0.444	0.441
16/06/00	2.493	0.397	0.408	27/04/01	2.814	0.450	0.448
23/06/00	2.519	0.401	0.419	04/05/01	2.765	0.442	0.440
30/06/00	2.522	0.402	0.426	11/05/01	2.722	0.435	0.443
07/07/00	2.568	0.410	0.433	18/05/01	2.720	0.435	0.436
14/07/00	2.569	0.410	0.441	25/05/01	2.734	0.437	0.435
21/07/00	2.584	0.412	0.443	01/06/01	2.780	0.448	0.436
28/07/00	2.560	0.408	0.441	08/06/01	2.635	0.421	0.445
04/08/00	2.576	0.411	0.444	15/06/01	2.657	0.425	0.452
11/08/00	2.595	0.414	0.433	22/06/01	2.727	0.436	0.470
18/08/00	2.537	0.404	0.438	29/06/01	2.766	0.442	0.481
25/08/00	2.572	0.410	0.428	06/07/01	2.773	0.443	0.473
01/09/00	2.539	0.404	0.431	13/07/01	2.757	0.441	0.462
08/09/00	2.564	0.409	0.421	20/07/01	2.821	0.451	0.452
15/09/00	2.562	0.408	0.434	27/07/01	2.818	0.450	0.451
22/09/00	2.651	0.423	0.425	03/08/01	2.757	0.440	0.458
29/09/00	2.733	0.436	0.430	10/08/01	2.773	0.443	0.451
06/10/00	2.702	0.431	0.422	17/08/01	2.697	0.431	0.442
13/10/00	2.756	0.440	0.428	24/08/01	2.729	0.436	0.437
20/10/00	2.739	0.437	0.424	31/08/01	2.757	0.441	0.439
27/10/00	2.768	0.442	0.428	07/09/01	2.816	0.450	0.439
03/11/00	2.761	0.441	0.426	14/09/01	2.850	0.455	0.445
10/11/00	2.735	0.437	0.430	21/09/01	3.000	0.477	0.457
17/11/00	2.745	0.438	0.442	28/09/01	2.977	0.474	0.457
24/11/00	2.684	0.428	0.440	05/10/01	2.927	0.467	0.448
01/12/00	2.671	0.426	0.442	12/10/01	2.897	0.462	0.441
08/12/00	2.656	0.424	0.459	19/10/01	2.834	0.453	0.447
15/12/00	2.703	0.432	0.466	26/10/01	2.857	0.456	0.438
22/12/00	2.664	0.425	0.460	02/11/01	2.884	0.460	0.446
29/12/00	2.691	0.430	0.465	09/11/01	2.842	0.453	0.445
05/01/01	2.624	0.419	0.465	16/11/01	2.744	0.438	0.443
12/01/01	2.659	0.425	0.453	23/11/01	2.725	0.435	0.442
19/01/01	2.640	0.422	0.451	30/11/01	2.744	0.438	0.438
26/01/01	2.683	0.429	0.452	07/12/01	2.775	0.443	0.433
02/02/01	2.661	0.425	0.440	14/12/01	2.801	0.447	0.434
09/02/01	2.701	0.432	0.432	21/12/01	2.844	0.454	0.428

Continued on next page.

Date	Aus	Lasr	Last	Date	Aus	Lasr	Last
28/12/01	2.844	0.454	0.426	08/11/02	2.811	0.448	0.443
04/01/02	2.782	0.444	0.432	15/11/02	2.815	0.449	0.435
11/01/02	2.771	0.443	0.430	22/11/02	2.809	0.448	0.426
18/01/02	2.791	0.446	0.429	29/11/02	2.777	0.443	0.416
25/01/02	2.731	0.436	0.427	06/12/02	2.812	0.448	0.416
01/02/02	2.777	0.443	0.436	13/12/02	2.818	0.449	0.432
08/02/02	2.776	0.443	0.429	20/12/02	2.864	0.456	0.422
15/02/02	2.774	0.443	0.424	27/12/02	2.862	0.456	0.418
22/02/02	2.788	0.445	0.417	03/01/03	2.843	0.453	0.419
01/03/02	2.742	0.438	0.414	10/01/03	2.768	0.441	0.413
08/03/02	2.726	0.435	0.405	17/01/03	2.743	0.437	0.409
15/03/02	2.716	0.434	0.412	24/01/03	2.761	0.440	0.411
22/03/02	2.681	0.428	0.419	31/01/03	2.810	0.448	0.407
29/03/02	2.671	0.426	0.432	07/02/03	2.767	0.441	0.395
05/04/02	2.704	0.432	0.439	14/02/03	2.734	0.436	0.401
12/04/02	2.692	0.430	0.444	21/02/03	2.669	0.425	0.395
19/04/02	2.686	0.429	0.454	28/02/03	2.603	0.414	0.406
26/04/02	2.680	0.428	0.465	07/03/03	2.618	0.417	0.394
03/05/02	2.731	0.436	0.461	14/03/03	2.669	0.425	0.400
10/05/02	2.688	0.429	0.458	21/03/03	2.654	0.423	0.399
17/05/02	2.648	0.422	0.451	28/03/03	2.622	0.417	0.397
24/05/02	2.625	0.418	0.449	04/04/03	2.608	0.415	0.389
31/05/02	2.591	0.412	0.449	11/04/03	2.605	0.415	0.395
07/06/02	2.553	0.406	0.461	18/04/03	2.574	0.409	0.387
14/06/02	2.644	0.421	0.450	25/04/03	2.589	0.412	0.384
21/06/02	2.618	0.417	0.451	02/05/03	2.552	0.406	0.395
28/06/02	2.723	0.434	0.457	09/05/03	2.495	0.396	0.396
05/07/02	2.736	0.436	0.458	16/05/03	2.506	0.398	0.387
12/07/02	2.788	0.444	0.456	23/05/03	2.501	0.397	0.383
19/07/02	2.848	0.453	0.450	30/05/03	2.522	0.400	0.392
26/07/02	2.914	0.464	0.445	06/06/03	2.523	0.401	0.393
02/08/02	2.928	0.465	0.450	13/06/03	2.507	0.398	0.384
09/08/02	2.860	0.455	0.444	20/06/03	2.499	0.397	0.385
16/08/02	2.822	0.450	0.448	27/06/03	2.487	0.395	0.387
23/08/02	2.803	0.447	0.449	04/07/03	2.457	0.389	0.388
30/08/02	2.815	0.448	0.442	11/07/03	2.486	0.394	0.381
06/09/02	2.867	0.456	0.447	18/07/03	2.465	0.390	0.385
13/09/02	2.842	0.453	0.446	25/07/03	2.446	0.387	0.384
20/09/02	2.851	0.454	0.453	01/08/03	2.471	0.391	0.381
27/09/02	2.861	0.456	0.453	08/08/03	2.475	0.392	0.372
04/10/02	2.879	0.458	0.453	15/08/03	2.426	0.383	0.369
11/10/02	2.849	0.454	0.445	22/08/03	2.427	0.384	0.372
18/10/02	2.819	0.449	0.437	29/08/03	2.451	0.388	0.376
25/10/02	2.812	0.448	0.439	05/09/03	2.469	0.391	0.370
01/11/02	2.814	0.448	0.448	12/09/03	2.429	0.384	0.374

Continued on next page.

Date	Aus	Lasr	Last	Date	Aus	Lasr	Last
19/09/03	2.435	0.385	0.380	11/06/04	2.628	0.419	0.415
26/09/03	2.466	0.391	0.378	18/06/04	2.671	0.426	0.408
03/10/03	2.448	0.388	0.376	25/06/04	2.613	0.416	0.400
10/10/03	2.416	0.382	0.374	02/07/04	2.568	0.409	0.398
17/10/03	2.421	0.383	0.374	09/07/04	2.574	0.410	0.390
24/10/03	2.428	0.384	0.375	16/07/04	2.565	0.408	0.390
31/10/03	2.398	0.379	0.378	23/07/04	2.590	0.413	0.394
07/11/03	2.365	0.372	0.381	30/07/04	2.600	0.414	0.390
14/11/03	2.351	0.370	0.380	06/08/04	2.585	0.412	0.386
21/11/03	2.363	0.372	0.379	13/08/04	2.573	0.410	0.383
28/11/03	2.385	0.376	0.384	20/08/04	2.516	0.400	0.377
05/12/03	2.357	0.371	0.385	27/08/04	2.552	0.406	0.377
12/12/03	2.361	0.371	0.390	03/09/04	2.576	0.410	0.394
19/12/03	2.403	0.379	0.388	10/09/04	2.580	0.411	0.407
26/12/03	2.404	0.379	0.389	17/09/04	2.575	0.410	0.408
02/01/04	2.371	0.373	0.384	24/09/04	2.533	0.403	0.399
09/01/04	2.383	0.376	0.382	01/10/04	2.485	0.395	0.393
16/01/04	2.360	0.371	0.382	08/10/04	2.444	0.388	0.392
23/01/04	2.369	0.373	0.385	15/10/04	2.470	0.392	0.390
30/01/04	2.401	0.379	0.393	22/10/04	2.476	0.393	0.391
06/02/04	2.411	0.381	0.393	29/10/04	2.457	0.390	0.386
13/02/04	2.399	0.378	0.408	05/11/04	2.434	0.386	0.388
20/02/04	2.416	0.382	0.408	12/11/04	2.412	0.382	0.379
27/02/04	2.414	0.381	0.407	19/11/04	2.375	0.375	0.381
05/03/04	2.432	0.385	0.426	26/11/04	2.405	0.381	0.385
12/03/04	2.466	0.391	0.419	03/12/04	2.499	0.397	0.387
19/03/04	2.455	0.389	0.428	10/12/04	2.553	0.406	0.386
26/03/04	2.442	0.387	0.415	17/12/04	2.556	0.407	0.385
02/04/04	2.417	0.382	0.412	24/12/04	2.502	0.398	0.384
09/04/04	2.414	0.382	0.410	31/12/04	2.453	0.389	0.388
16/04/04	2.429	0.384	0.409	07/01/05	2.469	0.392	0.388
23/04/04	2.426	0.387	0.411	14/01/05	2.463	0.391	0.387
30/04/04	2.463	0.390	0.416	21/01/05	2.441	0.387	0.392
07/05/04	2.539	0.404	0.414	28/01/05	2.437	0.386	0.390
14/05/04	2.560	0.407	0.407	04/02/05	2.445	0.388	0.387
21/05/04	2.556	0.407	0.403	11/02/05	2.386	0.377	0.385
28/05/04	2.575	0.410	0.407	18/02/05	2.411	0.381	0.385
04/06/04	2.656	0.423	0.411	25/02/05	2.446	0.388	0.379

Continued on next page.

Date	Aus	Lasr	Last	Date	Aus	Lasr	Last
04/03/05	2.435	0.386	0.381	29/04/05	2.451	0.389	
11/03/05	2.433	0.385	0.375	06/05/05	2.441	0.387	
18/03/05	2.423	0.383	0.372	13/05/05	2.438	0.386	
25/03/05	2.428	0.384		20/05/05	2.423	0.383	
01/04/05	2.450	0.388		27/05/05	2.400	0.379	
08/04/05	2.443	0.387		03/06/05	2.402	0.380	
15/04/05	2.461	0.390		10/06/05	2.384	0.376	
22/04/05	2.452	0.389					

Source: Thomson Financial Datastream.

Aus - Australia dollar-sterling exchange rate.

Lasr - logarithm of Australian dollar-sterling spot rate.

Last - logarithm of Australian dollar-sterling spot 3-month rate.

Table E.2: Forward exchange rate anomaly data, Canada.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
05/01/95	2.194	0.340	0.349	02/11/95	2.130	0.327	0.318
12/01/95	2.224	0.348	0.341	09/11/95	2.143	0.331	0.323
19/01/95	2.234	0.351	0.346	16/11/95	2.111	0.323	0.326
26/01/95	2.256	0.352	0.341	23/11/95	2.119	0.325	0.329
02/02/95	2.231	0.347	0.343	30/11/95	2.080	0.317	0.325
09/02/95	2.184	0.338	0.331	07/12/95	2.100	0.322	0.322
16/02/95	2.224	0.347	0.328	14/12/95	2.113	0.325	0.319
23/02/95	2.230	0.346	0.333	21/12/95	2.092	0.321	0.320
02/03/95	2.241	0.349	0.337	28/12/95	2.115	0.326	0.316
09/03/95	2.298	0.358	0.342	04/01/96	2.094	0.322	0.316
16/03/95	2.261	0.351	0.347	11/01/96	2.106	0.324	0.314
23/03/95	2.250	0.350	0.344	18/01/96	2.070	0.317	0.311
30/03/95	2.244	0.350	0.336	25/01/96	2.081	0.319	0.314
06/04/95	2.244	0.349	0.340	01/02/96	2.077	0.318	0.309
13/04/95	2.215	0.344	0.334	08/02/96	2.098	0.323	0.317
20/04/95	2.218	0.344	0.337	15/02/96	2.115	0.326	0.316
27/04/95	2.215	0.343	0.336	22/02/96	2.118	0.327	0.317
04/05/95	2.214	0.343	0.336	29/02/96	2.096	0.322	0.319
11/05/95	2.131	0.322	0.338	07/03/96	2.094	0.322	0.326
18/05/95	2.125	0.327	0.324	14/03/96	2.082	0.319	0.321
25/05/95	2.196	0.341	0.321	21/03/96	2.089	0.321	0.325
01/06/95	2.192	0.341	0.315	28/03/96	2.072	0.317	0.322
08/06/95	2.189	0.340	0.317	04/04/96	2.071	0.317	0.327
15/06/95	2.203	0.342	0.321	11/04/96	2.048	0.312	0.328
22/06/95	2.213	0.343	0.325	18/04/96	2.047	0.312	0.327
29/06/95	2.184	0.341	0.327	25/04/96	2.058	0.314	0.329
06/07/95	2.189	0.340	0.324	02/05/96	2.038	0.310	0.330
13/07/95	2.172	0.337	0.323	09/05/96	2.078	0.319	0.325
20/07/95	2.180	0.338	0.322	16/05/96	2.063	0.316	0.329
27/07/95	2.169	0.336	0.332	23/05/96	2.077	0.318	0.328
03/08/95	2.180	0.339	0.327	30/05/96	2.096	0.322	0.329
10/08/95	2.155	0.334	0.332	06/06/96	2.098	0.323	0.331
17/08/95	2.092	0.319	0.324	13/06/96	2.089	0.321	0.329
24/08/95	2.094	0.321	0.325	20/06/96	2.104	0.324	0.331
31/08/95	2.081	0.319	0.320	27/06/96	2.103	0.324	0.330
07/09/95	2.075	0.317	0.323	04/07/96	2.117	0.327	0.329
14/09/95	2.118	0.327	0.324	11/07/96	2.128	0.329	0.326
21/09/95	2.125	0.328	0.324	18/07/96	2.110	0.325	0.332
28/09/95	2.136	0.330	0.327	25/07/96	2.134	0.330	0.331
05/10/95	2.112	0.325	0.323	01/08/96	2.133	0.330	0.340
12/10/95	2.104	0.324	0.324	08/08/96	2.121	0.327	0.339
19/10/95	2.099	0.322	0.321	15/08/96	2.131	0.329	0.343
26/10/95	2.152	0.334	0.316	22/08/96	2.119	0.327	0.352

Continued on next page.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
29/08/96	2.124	0.328	0.353	10/07/97	2.298	0.365	0.348
05/09/96	2.143	0.332	0.345	17/07/97	2.280	0.362	0.353
12/09/96	2.128	0.329	0.353	24/07/97	2.293	0.364	0.357
19/09/96	2.122	0.328	0.359	31/07/97	2.243	0.355	0.368
26/09/96	2.132	0.330	0.359	07/08/97	2.180	0.343	0.370
03/10/96	2.120	0.329	0.370	14/08/97	2.187	0.344	0.381
10/10/96	2.106	0.326	0.360	21/08/97	2.193	0.345	0.379
17/10/96	2.135	0.332	0.353	28/08/97	2.221	0.351	0.377
24/10/96	2.131	0.331	0.348	04/09/97	2.177	0.342	0.378
31/10/96	2.168	0.339	0.337	11/09/97	2.189	0.344	0.371
07/11/96	2.176	0.341	0.341	18/09/97	2.224	0.351	0.369
14/11/96	2.203	0.346	0.345	25/09/97	2.231	0.353	0.380
21/11/96	2.248	0.355	0.340	02/10/97	2.203	0.347	0.372
28/11/96	2.248	0.355	0.349	09/10/97	2.216	0.350	0.367
05/12/96	2.196	0.345	0.343	16/10/97	2.229	0.352	0.369
12/12/96	2.235	0.352	0.337	23/10/97	2.244	0.355	0.371
19/12/96	2.256	0.356	0.342	30/10/97	2.328	0.371	0.379
26/12/96	2.265	0.358	0.349	06/11/97	2.349	0.375	0.379
02/01/97	2.307	0.366	0.358	13/11/97	2.369	0.379	0.370
09/01/97	2.273	0.360	0.353	20/11/97	2.387	0.382	0.373
16/01/97	2.227	0.351	0.356	27/11/97	2.365	0.378	0.370
23/01/97	2.168	0.339	0.355	04/12/97	2.364	0.377	0.370
30/01/97	2.159	0.337	0.356	11/12/97	2.341	0.373	0.364
06/02/97	2.183	0.342	0.354	18/12/97	2.349	0.372	0.375
13/02/97	2.178	0.341	0.357	25/12/97	2.381	0.380	0.374
20/02/97	2.173	0.340	0.353	01/01/98	2.338	0.372	0.376
27/02/97	2.206	0.347	0.355	08/01/98	2.299	0.365	0.376
06/03/97	2.183	0.342	0.350	15/01/98	2.324	0.369	0.384
13/03/97	2.160	0.338	0.356	22/01/98	2.373	0.378	0.380
20/03/97	2.181	0.342	0.357	29/01/98	2.379	0.380	0.380
27/03/97	2.228	0.351	0.366	05/02/98	2.370	0.377	0.379
03/04/97	2.260	0.357	0.357	12/02/98	2.344	0.373	0.373
10/04/97	2.243	0.354	0.365	19/02/98	2.315	0.367	0.374
17/04/97	2.260	0.357	0.362	26/02/98	2.322	0.369	0.376
24/04/97	2.248	0.355	0.366	05/03/98	2.324	0.369	0.377
01/05/97	2.246	0.355	0.353	12/03/98	2.323	0.369	0.378
08/05/97	2.224	0.351	0.346	19/03/98	2.348	0.374	0.386
15/05/97	2.253	0.356	0.342	26/03/98	2.367	0.377	0.390
22/05/97	2.224	0.351	0.346	02/04/98	2.355	0.375	0.386
29/05/97	2.245	0.355	0.351	09/04/98	2.372	0.378	0.383
05/06/97	2.226	0.351	0.341	16/04/98	2.418	0.386	0.385
12/06/97	2.240	0.354	0.342	23/04/98	2.371	0.378	0.390
19/06/97	2.267	0.359	0.347	30/04/98	2.377	0.379	0.392
26/06/97	2.279	0.362	0.350	07/05/98	2.356	0.375	0.394
03/07/97	2.293	0.364	0.347	14/05/98	2.348	0.373	0.393

Continued on next page.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
21/05/98	2.348	0.373	0.395	01/04/99	2.412	0.383	0.368
28/05/98	2.358	0.375	0.407	08/04/99	2.410	0.383	0.360
04/06/98	2.374	0.378	0.412	15/04/99	2.402	0.381	0.365
11/06/98	2.376	0.379	0.404	22/04/99	2.384	0.378	0.373
18/06/98	2.441	0.390	0.402	29/04/99	2.362	0.374	0.380
25/06/98	2.433	0.389	0.409	06/05/99	2.376	0.377	0.384
02/07/98	2.416	0.386	0.414	13/05/99	2.360	0.374	0.381
09/07/98	2.386	0.381	0.413	20/05/99	2.358	0.373	0.376
16/07/98	2.420	0.387	0.421	27/05/99	2.354	0.372	0.376
23/07/98	2.442	0.391	0.420	03/06/99	2.369	0.375	0.378
30/07/98	2.459	0.393	0.412	10/06/99	2.341	0.370	0.382
06/08/98	2.473	0.396	0.402	17/06/99	2.322	0.366	0.376
13/08/98	2.449	0.392	0.410	24/06/99	2.327	0.367	0.382
20/08/98	2.473	0.396	0.415	01/07/99	2.316	0.365	0.381
27/08/98	2.593	0.415	0.411	08/07/99	2.289	0.360	0.387
03/09/98	2.576	0.413	0.407	15/07/99	2.322	0.366	0.389
10/09/98	2.531	0.405	0.408	22/07/99	2.377	0.376	0.395
17/09/98	2.560	0.410	0.412	29/07/99	2.433	0.386	0.385
24/09/98	2.544	0.408	0.416	05/08/99	2.418	0.384	0.383
01/10/98	2.596	0.416	0.414	12/08/99	2.386	0.378	0.379
08/10/98	2.637	0.423	0.398	19/08/99	2.406	0.381	0.377
15/10/98	2.624	0.421	0.402	26/08/99	2.383	0.377	0.374
22/10/98	2.612	0.419	0.400	02/09/99	2.407	0.382	0.372
29/10/98	2.588	0.415	0.400	09/09/99	2.411	0.383	0.381
05/11/98	2.515	0.402	0.394	16/09/99	2.396	0.380	0.378
12/11/98	2.550	0.408	0.386	23/09/99	2.411	0.383	0.376
19/11/98	2.572	0.412	0.389	30/09/99	2.415	0.384	0.370
26/11/98	2.523	0.404	0.380	07/10/99	2.425	0.386	0.377
03/12/98	2.545	0.408	0.393	14/10/99	2.450	0.390	0.380
10/12/98	2.557	0.409	0.394	21/10/99	2.487	0.396	0.377
17/12/98	2.572	0.412	0.394	28/10/99	2.410	0.383	0.373
24/12/98	2.587	0.414	0.393	04/11/99	2.397	0.381	0.366
31/12/98	2.548	0.408	0.388	11/11/99	2.374	0.376	0.367
07/01/99	2.489	0.397	0.379	18/11/99	2.363	0.374	0.367
14/01/99	2.526	0.403	0.383	25/11/99	2.368	0.375	0.371
21/01/99	2.501	0.399	0.379	02/12/99	2.360	0.374	0.361
28/01/99	2.510	0.401	0.378	09/12/99	2.392	0.380	0.364
04/02/99	2.447	0.389	0.375	16/12/99	2.377	0.377	0.363
11/02/99	2.416	0.384	0.373	23/12/99	2.379	0.377	0.363
18/02/99	2.429	0.386	0.376	30/12/99	2.334	0.369	0.363
25/02/99	2.410	0.382	0.372	06/01/00	2.396	0.380	0.365
04/03/99	2.451	0.390	0.377	13/01/00	2.387	0.379	0.366
11/03/99	2.474	0.394	0.372	20/01/00	2.379	0.377	0.368
18/03/99	2.473	0.393	0.366	27/01/00	2.343	0.371	0.368
25/03/99	2.463	0.392	0.367	03/02/00	2.304	0.363	0.367

Continued on next page.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
10/02/00	2.322	0.367	0.357	21/12/00	2.245	0.351	0.352
17/02/00	2.331	0.369	0.349	28/12/00	2.246	0.352	0.353
24/02/00	2.331	0.369	0.346	04/01/01	2.237	0.350	0.352
02/03/00	2.289	0.361	0.350	11/01/01	2.235	0.350	0.350
09/03/00	2.301	0.363	0.351	18/01/01	2.233	0.349	0.348
16/03/00	2.318	0.366	0.344	25/01/01	2.200	0.343	0.347
23/03/00	2.323	0.367	0.344	01/02/01	2.205	0.344	0.341
30/03/00	2.320	0.366	0.351	08/02/01	2.183	0.340	0.341
06/04/00	2.293	0.361	0.353	15/02/01	2.216	0.346	0.343
13/04/00	2.323	0.367	0.350	22/02/01	2.222	0.347	0.343
20/04/00	2.322	0.367	0.343	01/03/01	2.238	0.350	0.341
27/04/00	2.327	0.368	0.347	08/03/01	2.271	0.357	0.328
04/05/00	2.310	0.364	0.345	15/03/01	2.246	0.352	0.323
11/05/00	2.234	0.350	0.348	22/03/01	2.227	0.349	0.330
18/05/00	2.224	0.347	0.346	29/03/01	2.253	0.354	0.331
25/05/00	2.217	0.346	0.343	05/04/01	2.251	0.353	0.328
01/06/00	2.224	0.347	0.333	12/04/01	2.244	0.352	0.335
08/06/00	2.229	0.348	0.333	19/04/01	2.219	0.347	0.337
15/06/00	2.225	0.348	0.321	26/04/01	2.219	0.347	0.341
22/06/00	2.226	0.348	0.321	03/05/01	2.193	0.341	0.343
29/06/00	2.249	0.352	0.340	10/05/01	2.187	0.341	0.336
06/07/00	2.250	0.352	0.340	17/05/01	2.192	0.342	0.343
13/07/00	2.220	0.347	0.343	24/05/01	2.179	0.339	0.352
20/07/00	2.217	0.346	0.343	31/05/01	2.193	0.342	0.351
27/07/00	2.228	0.348	0.339	07/06/01	2.110	0.325	0.355
03/08/00	2.217	0.346	0.344	14/06/01	2.122	0.328	0.358
10/08/00	2.223	0.347	0.341	21/06/01	2.156	0.334	0.361
17/08/00	2.210	0.345	0.345	28/06/01	2.144	0.332	0.365
24/08/00	2.199	0.342	0.340	05/07/01	2.110	0.325	0.365
31/08/00	2.144	0.332	0.341	12/07/01	2.143	0.332	0.357
07/09/00	2.123	0.327	0.344	19/07/01	2.175	0.339	0.357
14/09/00	2.095	0.322	0.344	26/07/01	2.185	0.341	0.350
21/09/00	2.114	0.325	0.351	02/08/01	2.199	0.343	0.363
28/09/00	2.201	0.343	0.352	09/08/01	2.180	0.339	0.368
05/10/00	2.171	0.337	0.353	16/08/01	2.204	0.343	0.361
12/10/00	2.213	0.345	0.349	23/08/01	2.227	0.349	0.355
19/10/00	2.180	0.339	0.348	30/08/01	2.235	0.350	0.354
26/10/00	2.173	0.337	0.344	06/09/01	2.257	0.355	0.348
02/11/00	2.227	0.348	0.341	13/09/01	2.293	0.358	0.355
09/11/00	2.193	0.341	0.345	20/09/01	2.295	0.361	0.359
16/11/00	2.209	0.344	0.348	27/09/01	2.321	0.367	0.362
23/11/00	2.163	0.335	0.347	04/10/01	2.308	0.365	0.364
30/11/00	2.178	0.338	0.344	11/10/01	2.253	0.354	0.362
07/12/00	2.203	0.343	0.355	18/10/01	2.268	0.357	0.360
14/12/00	2.227	0.348	0.352	25/10/01	2.246	0.353	0.360

Continued on next page.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
01/11/01	2.317	0.367	0.352	12/09/02	2.454	0.391	0.390
08/11/01	2.314	0.366	0.355	19/09/02	2.439	0.388	0.396
15/11/01	2.265	0.357	0.357	26/09/02	2.452	0.391	0.393
22/11/01	2.252	0.354	0.356	03/10/02	2.492	0.398	0.405
29/11/01	2.238	0.352	0.358	10/10/02	2.480	0.396	0.398
06/12/01	2.236	0.348	0.353	17/10/02	2.435	0.388	0.392
13/12/01	2.257	0.356	0.352	24/10/02	2.409	0.383	0.393
20/12/01	2.271	0.358	0.352	31/10/02	2.443	0.389	0.400
27/12/01	2.311	0.366	0.355	07/11/02	2.446	0.385	0.398
03/01/02	2.291	0.362	0.359	14/11/02	2.485	0.396	0.393
10/01/02	2.301	0.364	0.360	21/11/02	2.485	0.397	0.383
17/01/02	2.305	0.365	0.357	28/11/02	2.435	0.388	0.374
24/01/02	2.267	0.358	0.357	05/12/02	2.445	0.389	0.372
31/01/02	2.237	0.352	0.358	12/12/02	2.446	0.390	0.376
07/02/02	2.253	0.355	0.360	19/12/02	2.475	0.395	0.365
14/02/02	2.259	0.356	0.355	26/12/02	2.466	0.394	0.365
21/02/02	2.257	0.356	0.350	02/01/03	2.507	0.400	0.366
28/02/02	2.259	0.356	0.350	09/01/03	2.494	0.398	0.358
07/03/02	2.238	0.352	0.350	16/01/03	2.461	0.392	0.359
14/03/02	2.254	0.355	0.354	23/01/03	2.468	0.394	0.362
21/03/02	2.244	0.353	0.361	30/01/03	2.513	0.401	0.360
28/03/02	2.262	0.356	0.365	06/02/03	2.487	0.397	0.349
04/04/02	2.279	0.360	0.369	13/02/03	2.460	0.392	0.349
11/04/02	2.276	0.359	0.371	20/02/03	2.395	0.380	0.346
18/04/02	2.270	0.358	0.383	27/02/03	2.356	0.373	0.355
25/04/02	2.266	0.357	0.396	06/03/03	2.358	0.373	0.346
02/05/02	2.277	0.359	0.394	13/03/03	2.382	0.377	0.354
09/05/02	2.277	0.359	0.385	20/03/03	2.314	0.365	0.351
16/05/02	2.254	0.355	0.382	27/03/03	2.303	0.363	0.352
23/05/02	2.224	0.349	0.378	03/04/03	2.313	0.365	0.349
30/05/02	2.238	0.351	0.379	10/04/03	2.288	0.360	0.351
06/06/02	2.227	0.349	0.388	17/04/03	2.289	0.360	0.345
13/06/02	2.259	0.355	0.387	24/04/03	2.322	0.366	0.355
20/06/02	2.292	0.362	0.390	01/05/03	2.284	0.359	0.354
27/06/02	2.296	0.362	0.393	08/05/03	2.230	0.349	0.355
04/07/02	2.330	0.369	0.396	15/05/03	2.232	0.349	0.348
11/07/02	2.357	0.374	0.394	22/05/03	2.231	0.349	0.347
18/07/02	2.420	0.385	0.391	29/05/03	2.268	0.356	0.344
25/07/02	2.477	0.395	0.385	05/06/03	2.220	0.347	0.337
01/08/02	2.459	0.392	0.387	12/06/03	2.253	0.353	0.337
08/08/02	2.418	0.385	0.385	19/06/03	2.264	0.355	0.341
15/08/02	2.387	0.379	0.399	26/06/03	2.248	0.352	0.351
22/08/02	2.361	0.374	0.397	03/07/03	2.235	0.350	0.351
29/08/02	2.413	0.384	0.386	10/07/03	2.254	0.353	0.344
05/09/02	2.452	0.391	0.389	17/07/03	2.217	0.346	0.345

Continued on next page.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
24/07/03	2.245	0.352	0.343	03/06/04	2.488	0.399	0.370
31/07/03	2.261	0.355	0.348	10/06/04	2.471	0.396	0.362
07/08/03	2.253	0.353	0.350	17/06/04	2.507	0.402	0.364
14/08/03	2.217	0.346	0.339	24/06/04	2.427	0.388	0.362
21/08/03	2.226	0.348	0.345	01/07/04	2.403	0.384	0.359
28/08/03	2.200	0.343	0.348	08/07/04	2.430	0.388	0.350
04/09/03	2.167	0.337	0.352	15/07/04	2.438	0.390	0.355
11/09/03	2.184	0.340	0.359	22/07/04	2.411	0.385	0.353
18/09/03	2.201	0.344	0.368	29/07/04	2.391	0.382	0.351
25/09/03	2.233	0.350	0.367	05/08/04	2.387	0.381	0.350
02/10/03	2.235	0.350	0.364	12/08/04	2.408	0.385	0.345
09/10/03	2.214	0.346	0.369	19/08/04	2.359	0.376	0.345
16/10/03	2.207	0.344	0.375	26/08/04	2.338	0.372	0.346
23/10/03	2.210	0.345	0.378	02/09/04	2.309	0.366	0.359
30/10/03	2.221	0.348	0.385	09/09/04	2.283	0.361	0.374
06/11/03	2.224	0.348	0.388	16/09/04	2.298	0.364	0.377
13/11/03	2.185	0.341	0.396	23/09/04	2.285	0.362	0.375
20/11/03	2.213	0.346	0.398	30/09/04	2.277	0.360	0.368
27/11/03	2.235	0.351	0.401	07/10/04	2.226	0.350	0.364
04/12/03	2.251	0.354	0.390	14/10/04	2.235	0.352	0.355
11/12/03	2.308	0.364	0.379	21/10/04	2.264	0.357	0.362
18/12/03	2.344	0.371	0.386	28/10/04	2.221	0.349	0.364
25/12/03	2.323	0.367	0.391	04/11/04	2.208	0.346	0.369
01/01/04	2.306	0.364	0.383	11/11/04	2.196	0.344	0.366
08/01/04	2.336	0.370	0.385	18/11/04	2.226	0.350	0.368
15/01/04	2.356	0.374	0.381	25/11/04	2.214	0.347	0.375
22/01/04	2.388	0.380	0.381	02/12/04	2.274	0.359	0.375
29/01/04	2.411	0.384	0.385	09/12/04	2.339	0.371	0.366
05/02/04	2.440	0.389	0.390	16/12/04	2.379	0.379	0.365
12/02/04	2.486	0.397	0.391	23/12/04	2.359	0.375	0.357
19/02/04	2.507	0.401	0.390	30/12/04	2.305	0.365	0.360
26/02/04	2.488	0.398	0.396	06/01/05	2.302	0.364	0.361
04/03/04	2.424	0.387	0.399	13/01/05	2.248	0.354	0.370
11/03/04	2.370	0.377	0.393	20/01/05	2.294	0.363	0.376
18/03/04	2.428	0.388	0.401	27/01/05	2.322	0.368	0.376
25/03/04	2.398	0.382	0.394	03/02/05	2.324	0.369	0.376
01/04/04	2.428	0.388	0.386	10/02/05	2.308	0.366	0.369
08/04/04	2.422	0.387	0.389	17/02/05	2.314	0.367	0.365
15/04/04	2.388	0.380	0.389	24/02/05	2.366	0.376	0.365
22/04/04	2.392	0.381	0.386	03/03/05	2.366	0.376	0.354
29/04/04	2.424	0.387	0.385	10/03/05	2.304	0.365	0.360
06/05/04	2.456	0.393	0.380	17/03/05	2.297	0.364	0.353
13/05/04	2.451	0.392	0.384	24/03/05	2.263	0.357	
20/05/04	2.415	0.386	0.378	31/03/05	2.273	0.359	
27/05/04	2.476	0.396	0.370	07/04/05	2.278	0.360	

Continued on next page.

Date	Cnd	Lcsr	Lcst	Date	Cnd	Lcsr	Lcst
14/04/05	2.324	0.369		19/05/05	2.307	0.366	
21/04/05	2.354	0.374		26/05/05	2.295	0.363	
28/04/05	2.375	0.378		02/06/05	2.254	0.355	
05/05/05	2.357	0.375		09/06/05	2.274	0.359	
12/05/05	2.316	0.367		16/06/05	2.244	0.353	

Source: Thomson Financial Datastream.

Cnd - Australia dollar-sterling exchange rate.

Lcsr - logarithm of Canadian dollar-sterling spot rate.

Lcst - logarithm of Canadian dollar-sterling spot 3-month rate.

Table E.3: Forward exchange rate anomaly data, Japan.

Date	Jap	LJSR	LJST	Date	Jap	LJSR	LJST
03/01/97	156.685	2.199	2.141	31/10/97	160.100	2.211	2.208
10/01/97	154.252	2.192	2.125	07/11/97	159.039	2.209	2.211
17/01/97	154.711	2.193	2.118	14/11/97	154.748	2.196	2.215
24/01/97	156.887	2.201	2.121	21/11/97	155.084	2.196	2.215
31/01/97	155.521	2.197	2.130	28/11/97	153.133	2.193	2.206
07/02/97	151.781	2.186	2.120	05/12/97	153.908	2.193	2.206
14/02/97	152.182	2.187	2.131	12/12/97	153.477	2.193	2.206
21/02/97	152.259	2.187	2.136	19/12/97	154.437	2.195	2.212
28/02/97	151.228	2.186	2.123	26/12/97	158.028	2.205	2.209
07/03/97	145.397	2.166	2.131	02/01/98	159.541	2.209	2.214
14/03/97	141.839	2.155	2.128	09/01/98	159.787	2.210	2.218
21/03/97	139.588	2.149	2.132	16/01/98	159.105	2.208	2.213
28/03/97	140.468	2.152	2.124	23/01/98	159.196	2.208	2.207
04/04/97	136.983	2.142	2.131	30/01/98	159.304	2.208	2.196
11/04/97	131.829	2.126	2.142	06/02/98	160.786	2.212	2.201
18/04/97	129.430	2.117	2.149	13/02/98	162.050	2.216	2.206
25/04/97	133.844	2.131	2.144	20/02/98	160.159	2.211	2.210
02/05/97	132.700	2.129	2.149	27/02/98	157.989	2.205	2.216
09/05/97	130.193	2.122	2.166	06/03/98	158.738	2.207	2.228
16/05/97	134.130	2.134	2.177	13/03/98	158.129	2.205	2.226
23/05/97	135.487	2.137	2.171	20/03/98	161.419	2.214	2.222
30/05/97	132.593	2.129	2.178	27/03/98	159.754	2.210	2.226
06/06/97	132.450	2.127	2.182	03/04/98	160.898	2.213	2.234
13/06/97	133.779	2.132	2.196	10/04/98	162.313	2.216	2.233
20/06/97	133.551	2.131	2.207	17/04/98	161.052	2.213	2.231
27/06/97	131.565	2.130	2.200	24/04/98	159.330	2.208	2.224
04/07/97	133.406	2.132	2.205	01/05/98	155.090	2.197	2.225
11/07/97	137.475	2.144	2.200	08/05/98	157.330	2.203	2.216
18/07/97	137.938	2.146	2.199	15/05/98	159.468	2.209	2.223
25/07/97	138.113	2.146	2.199	22/05/98	159.296	2.208	2.224
01/08/97	143.030	2.162	2.208	29/05/98	162.996	2.218	2.224
08/08/97	144.546	2.166	2.211	05/06/98	166.294	2.227	2.234
15/08/97	149.626	2.181	2.199	12/06/98	165.775	2.226	2.233
22/08/97	146.550	2.172	2.196	19/06/98	164.724	2.223	2.234
29/08/97	150.362	2.184	2.194	26/06/98	166.720	2.228	2.233
05/09/97	150.585	2.184	2.193	03/07/98	170.164	2.237	2.241
12/09/97	156.427	2.200	2.193	10/07/98	168.859	2.233	2.240
19/09/97	157.410	2.204	2.196	17/07/98	166.187	2.226	2.250
26/09/97	154.766	2.197	2.202	24/07/98	165.665	2.225	2.256
03/10/97	157.306	2.202	2.209	31/07/98	163.967	2.221	2.264
10/10/97	156.131	2.200	2.211	07/08/98	163.936	2.221	2.274
17/10/97	155.623	2.198	2.211	14/08/98	164.945	2.223	2.263
24/10/97	157.479	2.203	2.205	21/08/98	165.700	2.225	2.271

Continued on next page.

DATE	JAP	LJSR	LJST	DATE	JAP	LJSR	LJST
28/08/98	166.821	2.228	2.275	09/07/99	187.110	2.279	2.296
04/09/98	167.958	2.231	2.274	16/07/99	191.428	2.289	2.295
11/09/98	169.191	2.234	2.272	23/07/99	191.674	2.289	2.294
18/09/98	167.983	2.231	2.280	30/07/99	189.983	2.286	2.303
25/09/98	170.342	2.237	2.282	06/08/99	187.326	2.280	2.313
02/10/98	172.708	2.244	2.298	13/08/99	179.482	2.261	2.329
09/10/98	171.959	2.242	2.291	20/08/99	185.024	2.274	2.328
16/10/98	175.278	2.250	2.290	27/08/99	188.460	2.282	2.331
23/10/98	177.965	2.256	2.292	03/09/99	188.701	2.283	2.335
30/10/98	183.142	2.269	2.290	10/09/99	185.955	2.277	2.330
06/11/98	184.084	2.272	2.296	17/09/99	189.429	2.284	2.330
13/11/98	181.780	2.266	2.306	24/09/99	190.375	2.287	2.337
20/11/98	183.716	2.271	2.298	01/10/99	192.052	2.290	2.332
27/11/98	186.918	2.278	2.300	08/10/99	193.602	2.294	2.339
04/12/98	182.645	2.268	2.295	15/10/99	194.026	2.295	2.333
11/12/98	184.061	2.272	2.291	22/10/99	193.954	2.295	2.322
18/12/98	187.320	2.279	2.290	29/10/99	197.557	2.303	2.317
25/12/98	188.778	2.282	2.302	05/11/99	201.930	2.313	2.317
01/01/99	195.846	2.298	2.307	12/11/99	210.676	2.332	2.302
08/01/99	192.630	2.291	2.311	19/11/99	211.081	2.333	2.315
15/01/99	193.639	2.293	2.313	26/11/99	208.904	2.328	2.324
22/01/99	193.819	2.292	2.315	03/12/99	212.589	2.336	2.318
29/01/99	194.622	2.295	2.316	10/12/99	209.325	2.329	2.321
05/02/99	198.586	2.304	2.312	17/12/99	205.866	2.321	2.334
12/02/99	199.516	2.306	2.288	24/12/99	213.117	2.336	2.339
19/02/99	197.255	2.301	2.272	31/12/99	210.152	2.330	2.349
26/02/99	196.649	2.300	2.279	07/01/00	210.908	2.332	2.348
05/03/99	192.568	2.291	2.279	14/01/00	210.859	2.332	2.335
12/03/99	192.520	2.291	2.267	21/01/00	204.472	2.318	2.343
19/03/99	193.233	2.292	2.269	28/01/00	201.837	2.313	2.344
26/03/99	197.756	2.305	2.282	04/02/00	201.570	2.312	2.340
02/04/99	198.492	2.304	2.280	11/02/00	198.438	2.305	2.340
09/04/99	202.550	2.313	2.280	18/02/00	203.948	2.317	2.345
16/04/99	201.230	2.310	2.286	25/02/00	208.081	2.326	2.354
23/04/99	201.893	2.311	2.288	03/03/00	204.625	2.319	2.357
30/04/99	203.094	2.314	2.283	10/03/00	209.031	2.328	2.361
07/05/99	201.239	2.310	2.288	17/03/00	213.856	2.338	2.377
14/05/99	191.042	2.288	2.263	24/03/00	212.411	2.335	2.365
21/05/99	185.278	2.275	2.278	31/03/00	219.377	2.349	2.365
28/05/99	186.053	2.277	2.281	07/04/00	217.511	2.345	2.356
04/06/99	186.058	2.277	2.287	14/04/00	214.206	2.338	2.359
11/06/99	179.208	2.261	2.277	21/04/00	214.999	2.340	2.363
18/06/99	182.528	2.269	2.285	28/04/00	216.627	2.343	2.369
25/06/99	186.666	2.278	2.292	05/05/00	216.738	2.344	2.375
02/07/99	185.890	2.276	2.290	12/05/00	215.142	2.341	2.381

Continued on next page.

Date	Jap	LJSR	LJST	Date	Jap	LJSR	LJST
19/05/00	218.151	2.347	2.370	30/03/01	188.792	2.281	2.281
26/05/00	220.478	2.351	2.375	06/04/01	190.812	2.286	2.282
02/06/00	222.401	2.355	2.360	13/04/01	189.620	2.284	2.276
09/06/00	225.856	2.362	2.341	20/04/01	190.015	2.284	2.273
16/06/00	225.053	2.360	2.347	27/04/01	189.962	2.284	2.267
23/06/00	230.090	2.370	2.357	04/05/01	194.382	2.294	2.269
30/06/00	225.901	2.362	2.359	11/05/01	193.115	2.292	2.267
07/07/00	224.456	2.359	2.347	18/05/01	198.901	2.305	2.262
14/07/00	225.097	2.361	2.309	25/05/01	192.017	2.289	2.251
21/07/00	227.082	2.364	2.302	01/06/01	191.934	2.289	2.246
28/07/00	229.817	2.370	2.296	08/06/01	187.872	2.280	2.251
04/08/00	231.658	2.373	2.281	15/06/01	188.750	2.281	2.229
11/08/00	233.511	2.376	2.309	22/06/01	189.926	2.284	2.231
18/08/00	228.900	2.368	2.307	29/06/01	188.359	2.281	2.244
25/08/00	232.350	2.374	2.302	06/07/01	188.126	2.280	2.248
01/09/00	226.074	2.362	2.304	13/07/01	186.120	2.275	2.245
08/09/00	223.987	2.358	2.296	20/07/01	183.905	2.270	2.246
15/09/00	222.059	2.354	2.292	27/07/01	182.247	2.266	2.240
22/09/00	225.321	2.361	2.293	03/08/01	183.001	2.268	2.234
29/09/00	227.102	2.364	2.287	10/08/01	183.191	2.269	2.232
06/10/00	205.723	2.321	2.267	17/08/01	177.611	2.255	2.234
13/10/00	199.032	2.307	2.261	24/08/01	173.963	2.246	2.229
20/10/00	195.726	2.300	2.275	31/08/01	173.234	2.244	2.212
27/10/00	193.682	2.295	2.277	07/09/01	175.822	2.255	2.221
03/11/00	189.763	2.286	2.268	14/09/01	166.027	2.226	2.224
10/11/00	198.714	2.306	2.272	21/09/01	167.848	2.231	2.216
17/11/00	199.493	2.308	2.286	28/09/01	172.702	2.244	2.217
24/11/00	197.130	2.307	2.291	05/10/01	175.453	2.251	2.226
01/12/00	196.854	2.301	2.289	12/10/01	174.962	2.249	2.241
08/12/00	192.314	2.291	2.292	19/10/01	173.953	2.247	2.237
15/12/00	191.236	2.288	2.282	26/10/01	169.129	2.235	2.242
22/12/00	191.787	2.289	2.284	02/11/01	170.171	2.237	2.241
29/12/00	188.851	2.283	2.289	09/11/01	167.782	2.231	2.246
05/01/01	183.873	2.271	2.284	16/11/01	169.079	2.234	2.240
12/01/01	185.287	2.274	2.289	23/11/01	166.289	2.227	2.252
19/01/01	184.509	2.272	2.281	30/11/01	161.008	2.213	2.239
26/01/01	188.146	2.280	2.288	07/12/01	165.320	2.225	2.228
02/02/01	181.470	2.265	2.292	14/12/01	164.047	2.221	2.217
09/02/01	184.222	2.271	2.293	21/12/01	161.005	2.213	2.227
16/02/01	192.028	2.289	2.300	28/12/01	162.597	2.217	2.225
23/02/01	192.362	2.290	2.294	04/01/02	167.781	2.231	2.227
02/03/01	193.618	2.293	2.288	11/01/02	171.679	2.241	2.230
09/03/01	192.499	2.290	2.285	18/01/02	170.914	2.239	2.217
16/03/01	190.412	2.286	2.284	25/01/02	171.205	2.240	2.223
23/03/01	190.704	2.286	2.287	01/02/02	171.222	2.240	2.231

Continued on next page.

Date	Jap	Ljsr	Ljst	Date	Jap	Ljsr	Ljst
08/02/02	173.228	2.245	2.225	20/12/02	163.872	2.220	2.243
15/02/02	172.902	2.244	2.216	27/12/02	167.758	2.230	2.248
22/02/02	175.683	2.251	2.198	03/01/03	169.918	2.236	2.255
01/03/02	167.741	2.231	2.202	10/01/03	171.102	2.239	2.253
08/03/02	166.593	2.228	2.206	17/01/03	170.923	2.239	2.248
15/03/02	163.235	2.220	2.208	24/01/03	169.958	2.236	2.244
22/03/02	165.125	2.224	2.202	31/01/03	167.764	2.230	2.243
29/03/02	164.773	2.223	2.199	07/02/03	167.437	2.229	2.242
05/04/02	164.773	2.224	2.207	14/02/03	167.955	2.231	2.243
12/04/02	165.569	2.226	2.209	21/02/03	166.306	2.227	2.244
19/04/02	162.708	2.218	2.209	28/02/03	166.879	2.228	2.231
26/04/02	165.739	2.226	2.220	07/03/03	172.622	2.243	2.229
03/05/02	167.795	2.232	2.215	14/03/03	172.779	2.243	2.226
10/05/02	163.956	2.222	2.214	21/03/03	173.749	2.246	2.236
17/05/02	160.644	2.213	2.216	28/03/03	172.466	2.242	2.244
24/05/02	155.575	2.199	2.207	04/04/03	177.194	2.254	2.244
31/05/02	158.407	2.207	2.190	11/04/03	176.478	2.252	2.248
07/06/02	157.860	2.205	2.186	18/04/03	171.989	2.241	2.243
14/06/02	157.788	2.205	2.177	25/04/03	173.442	2.245	2.246
21/06/02	155.725	2.199	2.178	02/05/03	172.652	2.243	2.250
28/06/02	157.053	2.203	2.195	09/05/03	171.462	2.240	2.244
05/07/02	159.313	2.209	2.200	16/05/03	174.408	2.247	2.238
12/07/02	161.719	2.215	2.194	23/05/03	169.516	2.235	2.239
19/07/02	159.150	2.208	2.192	30/05/03	168.955	2.234	2.243
26/07/02	163.306	2.220	2.195	06/06/03	165.671	2.225	2.235
02/08/02	160.510	2.212	2.200	13/06/03	166.676	2.228	2.242
09/08/02	159.839	2.210	2.185	20/06/03	169.745	2.236	2.235
16/08/02	160.427	2.212	2.189	27/06/03	173.149	2.244	2.237
23/08/02	155.946	2.199	2.194	04/07/03	172.677	2.243	2.248
30/08/02	152.832	2.190	2.194	11/07/03	173.280	2.245	2.246
06/09/02	151.494	2.187	2.204	18/07/03	173.786	2.246	2.244
13/09/02	149.078	2.180	2.208	25/07/03	174.105	2.247	2.243
20/09/02	148.752	2.179	2.216	01/08/03	176.314	2.252	2.248
27/09/02	154.891	2.196	2.226	08/08/03	171.861	2.241	2.247
04/10/02	157.117	2.203	2.234	15/08/03	170.327	2.237	2.244
11/10/02	155.258	2.197	2.240	22/08/03	172.213	2.242	2.240
18/10/02	154.309	2.195	2.239	29/08/03	171.427	2.239	2.244
25/10/02	152.551	2.190	2.236	05/09/03	172.827	2.243	2.247
01/11/02	154.879	2.196	2.232	12/09/03	172.768	2.243	2.258
08/11/02	151.003	2.185	2.225	19/09/03	170.276	2.235	2.270
15/11/02	152.980	2.191	2.232	26/09/03	171.939	2.240	2.272
22/11/02	153.605	2.192	2.222	03/10/03	176.240	2.251	2.281
29/11/02	155.728	2.198	2.224	10/10/03	172.531	2.242	2.281
06/12/02	156.637	2.201	2.242	17/10/03	173.546	2.244	2.278
13/12/02	160.892	2.212	2.242	24/10/03	172.969	2.243	2.283

Continued on next page.

Date	Jap	Ljsr	Ljst	Date	Jap	Ljsr	Ljst
31/10/03	176.149	2.251	2.275	16/07/04	180.443	2.261	2.288
07/11/03	175.514	2.249	2.278	23/07/04	181.023	2.262	2.285
14/11/03	173.676	2.244	2.279	30/07/04	185.213	2.272	2.281
21/11/03	172.331	2.241	2.281	06/08/04	182.940	2.267	2.280
28/11/03	173.703	2.244	2.282	13/08/04	178.246	2.255	2.279
05/12/03	174.427	2.246	2.274	20/08/04	179.697	2.259	2.286
12/12/03	180.719	2.258	2.259	27/08/04	180.166	2.260	2.279
19/12/03	183.917	2.269	2.273	03/09/04	182.824	2.266	2.290
26/12/03	187.247	2.272	2.278	10/09/04	184.745	2.271	2.288
02/01/04	189.245	2.281	2.284	17/09/04	186.408	2.275	2.285
09/01/04	188.997	2.281	2.274	24/09/04	189.861	2.283	2.283
16/01/04	187.872	2.278	2.276	01/10/04	190.885	2.284	2.281
23/01/04	189.306	2.281	2.275	08/10/04	190.874	2.287	2.285
30/01/04	186.174	2.274	2.272	15/10/04	191.132	2.286	2.279
06/02/04	187.044	2.276	2.273	22/10/04	190.362	2.284	2.281
13/02/04	188.699	2.280	2.270	29/10/04	189.364	2.282	2.289
20/02/04	189.283	2.281	2.257	05/11/04	188.577	2.280	2.296
27/02/04	188.664	2.280	2.260	12/11/04	188.613	2.280	2.293
05/03/04	184.347	2.270	2.259	19/11/04	191.108	2.286	2.278
12/03/04	181.097	2.262	2.265	26/11/04	186.202	2.274	2.266
19/03/04	185.553	2.273	2.268	03/12/04	194.148	2.292	2.270
26/03/04	186.801	2.276	2.262	10/12/04	192.487	2.289	2.274
02/04/04	188.667	2.280	2.265	17/12/04	191.918	2.287	2.269
09/04/04	185.976	2.274	2.262	24/12/04	189.946	2.283	2.275
16/04/04	186.723	2.276	2.260	31/12/04	189.202	2.281	2.270
23/04/04	186.072	2.274	2.264	07/01/05	189.901	2.283	2.270
30/04/04	184.456	2.270	2.276	14/01/05	187.328	2.277	2.276
07/05/04	185.645	2.273	2.270	21/01/05	189.266	2.281	2.276
14/05/04	184.155	2.270	2.262	28/01/05	192.645	2.289	2.281
21/05/04	179.186	2.258	2.259	04/02/05	195.339	2.295	2.280
28/05/04	179.596	2.259	2.257	11/02/05	194.261	2.292	2.273
04/06/04	179.489	2.259	2.263	18/02/05	187.712	2.277	2.281
11/06/04	182.765	2.267	2.270	25/02/05	183.721	2.268	2.283
18/06/04	182.617	2.266	2.274	04/03/05	185.825	2.273	2.287
25/06/04	179.911	2.260	2.284	11/03/05	187.195	2.276	2.288
02/07/04	181.015	2.262	2.284	18/03/05	185.301	2.272	2.300
09/07/04	180.659	2.261	2.287	25/03/05	187.211	2.276	2.293

Continued on next page.

Date	Jap	Ljsr	Ljst	Date	Jap	Ljsr	Ljst
01/04/05	185.112	2.271	2.297	13/05/05	186.636	2.275	2.279
08/04/05	185.687	2.273	2.287	20/05/05	191.076	2.285	2.277
15/04/05	187.919	2.278	2.274	27/05/05	192.005	2.287	2.266
22/04/05	188.793	2.280	2.281	03/06/05	192.096	2.288	2.262
29/04/05	188.887	2.280	2.289	10/06/05	194.504	2.293	2.267
06/05/05	183.458	2.268	2.285				

Source: Thomson Financial Datastream.

Jap - Japanese yen-sterling exchange rate.

Ljsr - logarithm of Japanese yen-sterling spot rate.

Ljst - logarithm of Japanese yen-sterling spot 3-month rate.

E.2 Figures from Chapter 7

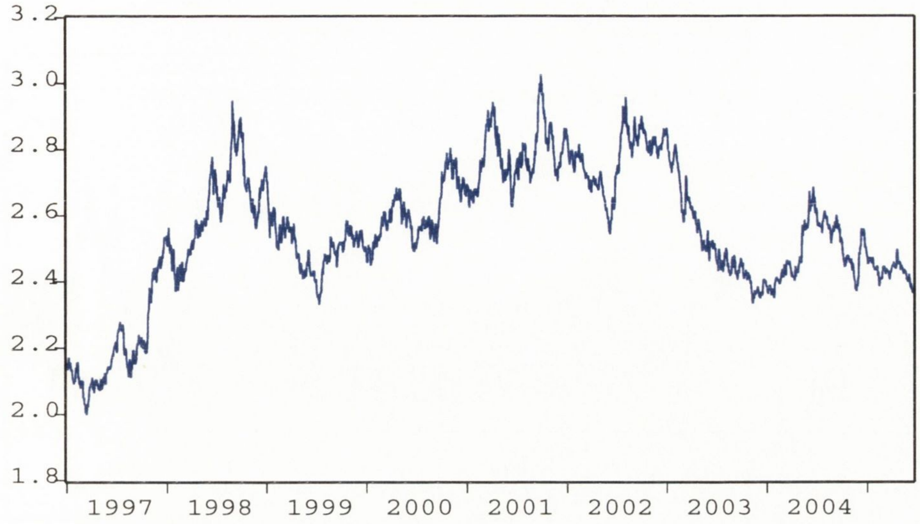

Figure E.1: The Australian dollar-sterling exchange rate.

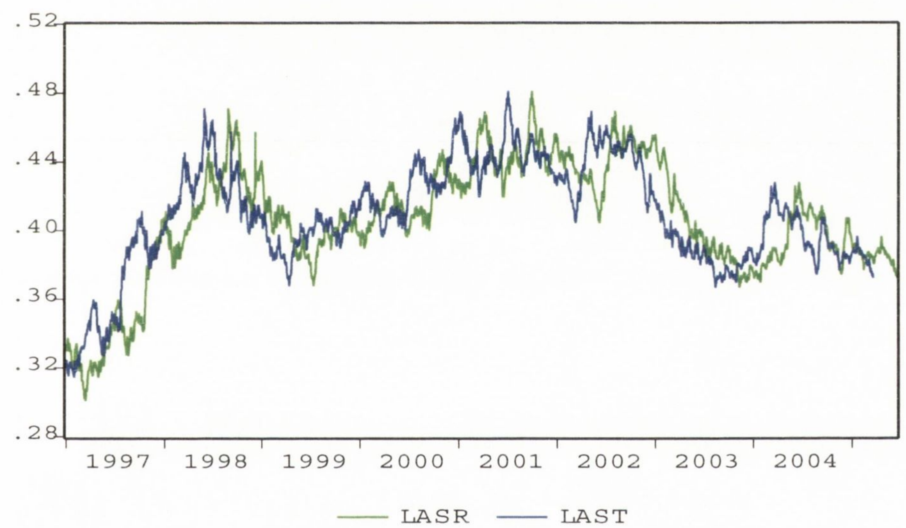

Figure E.2: The logs of the Australian dollar-sterling spot (LASR) and 3-month (LAST) rates.

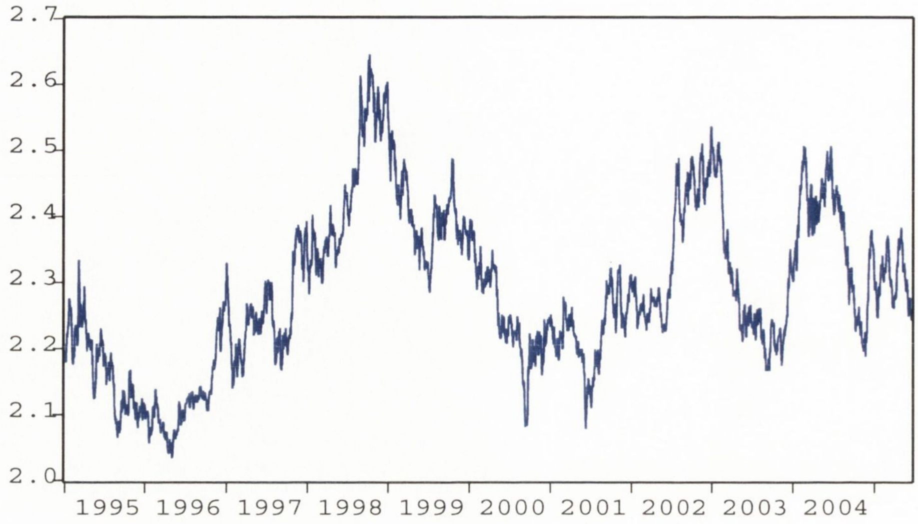

Figure E.3: The Canadian dollar-sterling exchange rate.

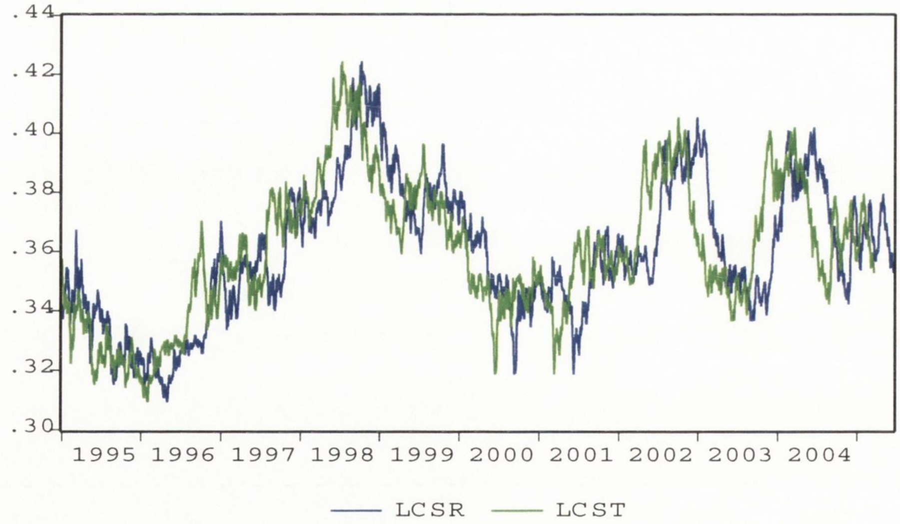

Figure E.4: The logs of the Canadian dollar-sterling spot (LCSR) and 3-month (LCST) rates.

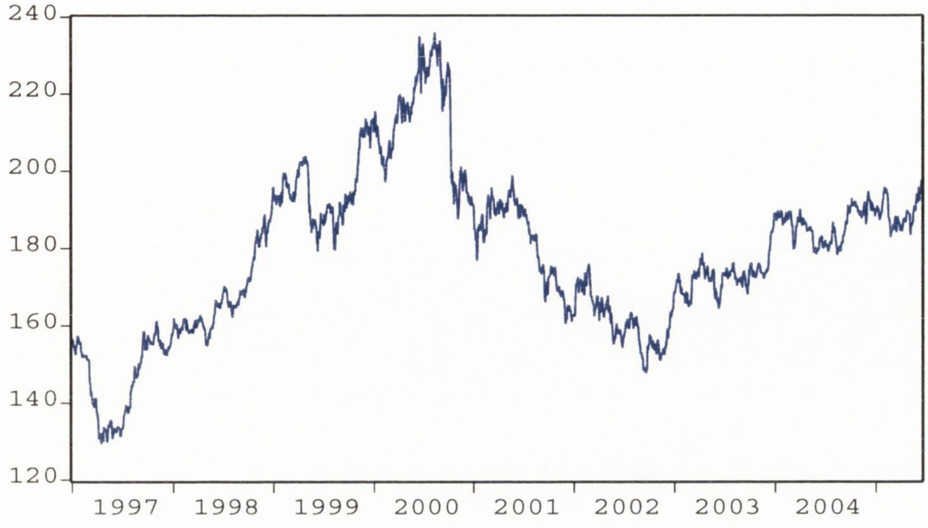

Figure E.5: The Japanese yen-sterling exchange rate.

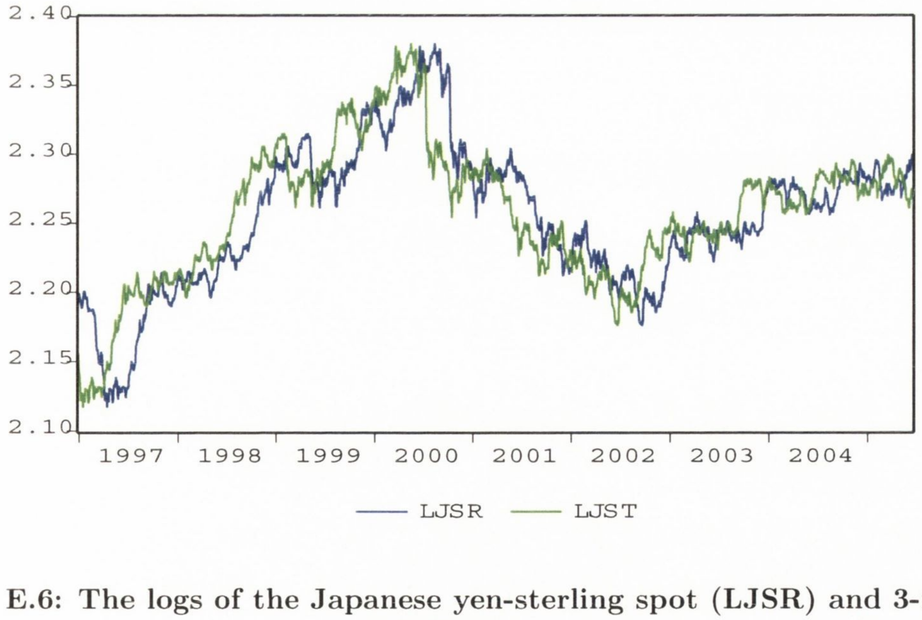

Figure E.6: The logs of the Japanese yen-sterling spot (LJSR) and 3-month (LJST) rates.

E.3 Results

Table E.4: Unit root tests, Australia, Canada & Japan.

Variables	Adf	p-value	No. of Lags
	Australian Dollar		
Forward rate	0·089	0·711	0
Spot rate	0·085	0·710	2
Forward premium	−1·315	0·175	20
Spot premium	−4·521	0·000	2
	Canadian Dollar		
Forward rate	−0·143	0·635	8
Spot rate	−0·102	0·649	0
Forward premium	−0·858	0·345	23
Spot premium	−5·446	0·000	0
	Japanese Yen		
Forward rate	0·623	0·851	7
Spot rate	0·581	0·842	5
Forward premium	−0·373	0·550	24
Spot premium	−4·424	0·000	5

Table E.5: Fractional integration analysis, Australia, Canada & Japan.

Variables	Eml	Nls	Gph	Gsp	Fdf†	Fadf†
			Australian Dollar			
Forward rate	0.99 (0.017)	0.99 (0.017)	0.99 (0.020)	0.97 (0.015)	0.31	-1.07
Spot rate	0.97 (0.017)	0.97 (0.017)	0.99 (0.020)	0.97 (0.015)	-0.42	-1.33
Forward premium	0.56 (0.028)	0.59 (0.024)	0.21 (0.020)	0.38 (0.015)	-35.3	-27.3
Spot premium	0.95 (0.017)	0.95 (0.017)	0.90 (0.020)	0.91 (0.015)	-1.03	-3.05
			Canadian Dollar			
Forward rate	0.87 (0.052)	0.87 (0.052)	0.85 (0.018)	0.87 (0.014)	-0.30	-0.94
Spot rate	0.95 (0.024)	0.94 (0.024)	0.91 (0.018)	0.96 (0.014)	0.19	-0.30
Forward premium	0.41 (0.009)	0.41 (0.009)	0.34 (0.019)	0.47 (0.014)	-42.9	-24.7
Spot premium	0.92 (0.030)	0.92 (0.030)	0.92 (0.019)	0.93 (0.014)	1.33	-3.69
			Japanese Yen			
Forward rate	0.99 (0.029)	0.99 (0.029)	1.01 (0.020)	0.97 (0.015)	0.16	-0.95
Spot rate	1.03 (0.017)	1.03 (0.017)	1.01 (0.020)	0.97 (0.015)	0.12	0.96
Forward premium	0.65 (0.028)	0.67 (0.026)	0.45 (0.020)	0.60 (0.015)	-35.0	-22.3
Spot premium	0.95 (0.032)	0.95 (0.032)	0.99 (0.020)	0.97 (0.015)	-3.42	-4.23

Note: standard errors in parentheses.
†Based on the Eml estimator of d.

Table E.6: Engle-Granger levels models, Australia, Canada & Japan.

Regressions	α	β	R^2	CRDW	AEG†
		Australian Dollar			
s_t on $f_{t,k}$	−0.002 (0.0003)	0.995 (0.001)	0.998	0.508 [0.20]	No
s_{t+k} on $f_{t,k}$	0.117 (0.005)	0.717 (0.005)	0.628	0.040 [0.20]	Yes
$\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$	−0.001 (0.001)	2.289 (0.301)	0.026	0.071 [0.20]	Yes
		Canadian Dollar			
s_t on $f_{t,k}$	−0.004 (0.0004)	1.014 (0.001)	0.997	0.321 [0.20]	No
s_{t+k} on $f_{t,k}$	0.096 (0.005)	0.738 (0.013)	0.532	0.037 [0.20]	Yes
$\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$	−0.002 (0.0004)	−1.967 (0.231)	0.026	0.059 [0.20]	Yes
		Japanese Yen			
s_t on $f_{t,k}$	−0.007 (0.001)	1.005 (0.001)	0.999	0.459 [0.20]	No
s_{t+k} on $f_{t,k}$	0.386 (0.025)	0.832 (0.025)	0.724	0.027 [0.20]	Yes
$\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$	−0.008 (0.003)	−0.917 (0.470)	0.002	0.028 [0.20]	Yes

Note: standard errors in parentheses. CRDW critical values in square brackets.
†Yes - significant at 5 per cent level. No - not significant at 5 per cent level.

Table E.7: Fractional integration analysis, residuals, Australia, Canada & Japan.

REGRESSIONS	EML	NLS	GPH	GSP	FDF†	FADF†
			AUSTRALIAN DOLLAR			
s_t on $f_{t,k}$	0.34 (0.01)	0.34 (0.01)	0.23 (0.020)	0.39 (0.015)	−38·56	−25·70
s_{t+k} on $f_{t,k}$	0.97 (0.02)	0.97 (0.02)	0.89 (0.020)	0.91 (0.015)	−0·32	−2·90
$\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$	0.84 (0.02)	0.84 (0.02)	0.80 (0.020)	0.85 (0.015)	−6·17	−5·23
			CANADIAN DOLLAR			
s_t on $f_{t,k}$	0.21 (0.02)	0.22 (0.02)	0.06 (0.018)	0.14 (0.014)	−48·57	−34·57
s_{t+k} on $f_{t,k}$	0.85 (0.03)	0.85 (0.03)	0.60 (0.017)	0.67 (0.014)	−23·34	−11·26
$\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$	0.91 (0.02)	0.91 (0.02)	0.84 (0.018)	0.88 (0.014)	−5·62	−3·64
			JAPANESE YEN			
s_t on $f_{t,k}$	0.47 (0.16)	0.50 (0.14)	0.33 (0.021)	0.43 (0.015)	−38·36	−25·47
s_{t+k} on $f_{t,k}$	0.77 (0.07)	0.77 (0.07)	0.99 (0.020)	0.98 (0.015)	2·45	−2·53
$\Delta_k s_{t+k}$ on $\{f_{t,k} - s_t\}$	0.95 (0.03)	0.95 (0.03)	0.99 (0.020)	0.98 (0.015)	2·90	−2·31

Note: standard errors in parentheses.
†Based on the EML estimator of d.

Table E.8: Hamilton analysis, Australia, Canada & Japan.

	Rates			Premiums	
			Estimates (standard errors)		
			Australia		
Linear					
c	-0.002 (0.002)				
$f_{t,k}$	1.009 (0.005)	1.004 (0.001)			
t	-0.0007 (0.0002)	-0.0007 (0.0002)			
Nonlinear					
σ	0.001 (0.00005)	0.001 (0.00005)			
ζ	0.422 (0.092)	0.436 (0.093)			
$f_{t,k}$	2.460 (7.523)	2.240 (5.609)			
t	1.810 (0.443)	1.873 (0.429)			
			Canada		
Linear			Linear		
c	-0.001 (0.002)		c	2.741 (6.090)	
$f_{t,k}$	1.004 (0.004)	1.003 (0.002)	$f_{t,k} - s_t$	1.305 (0.578)	1.267 (0.580)
t	0.0002 (0.0003)	0.0001 (0.0002)	t	-0.246 (2.144)	0.503 (0.727)
Nonlinear			Nonlinear		
σ	0.001 (0.00002)	0.001 (0.00002)	σ	0.128 (0.967)	0.060 (0.538)
ζ	1.974 (0.245)	-1.629 (0.192)	ζ	84.867 (639.699)	182.803 (1649.346)
$f_{t,k}$	-0.0001 (0.064)	-0.0002 (0.064)	$f_{t,k} - s_t$	-0.014 (0.036)	-0.015 (0.034)
t	1.169 (0.102)	1.866 (0.152)	t	-12.893 (0.599)	12.888 (0.598)
			Japan		
Linear					
c	0.003 (0.010)				
$f_{t,k}$	1.001 (0.005)	1.003 (0.0003)			
t	-0.0004 (0.0003)	-0.0004 (0.0003)			
Nonlinear					
σ	0.0005 (0.00002)	0.0005 (0.00002)			
ζ	-1.571 (0.216)	-1.568 (0.219)			
$f_{t,k}$	6.771 (1.420)	6.864 (1.959)			
t	1.637 (0.125)	1.640 (0.176)			