



uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES



FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES

Hui Liu

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Economics)
GRADE / DEGREE

Department of Economics

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Three Essays on Unit Root Tests and Cointegration: Theory and Applications

TITRE DE LA THÈSE / TITLE OF THESIS

Gabriel Rodriguez

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Gilles Grenier

Marcel Voia (absent)

Richard Luger

Simon Power

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Three Essays On Unit Root Tests And Cointegration: Theory And Applications

by

Hui Liu

Department of Economics
Faculty of Social Sciences

A thesis presented to the School of Graduate Studies and Research of the
University of Ottawa in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Economics

May, 2006

© Hui Liu, Ottawa, Canada, 2006



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-18592-6
Our file *Notre référence*
ISBN: 978-0-494-18592-6

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Dedication

To my mother, my father, my brother and my husband, for their great support and love.

Table of Contents

Abstract	
Acknowledgments	
Introduction	1
Chapter 1. Human Activities and Global Warming: A Cointegration Analysis	
1.1 Introduction	7
1.2 A Brief Review of the Literature	9
1.3 Methodological Issues	12
1.3.1 The I(1) Model	12
1.3.2 The I(2) Model	15
1.4 Empirical Analysis	18
1.4.1 The Data and Preliminary Issues	18
1.4.2 The First Case	20
1.4.2 The Second Case	23
1.4.3 The Third Case	25
1.5 Conclusions	26
References	28
Tables and Figures	34
Chapter 2. Unit Root Tests and Structural Change When the Initial Observation Is Drawn From Its Unconditional Distribution	
2.1 Introduction	42
2.2 The Model and Asymptotic Theory	43
2.2.1 The Model	43
2.2.2 Asymptotic Theory Assuming Known Break Point	44
2.2.3 The Selection of Break Point	47
2.2.4 Asymptotic Results	49
2.3 Small Sample Monte Carlo Evidence	49
2.3.1 The Selection of k	49
2.3.2 Size and Power in Finite Samples	50
2.4 Empirical Applications	51
2.5 Concluding Remarks	52

References	53
Appendix.....	56
Tables and Figures.....	65
Chapter 3. Unit Root Test Using GLS Detrended Data and Covariates in Structural Change Models	
3.1 Introduction.....	79
3.2 The Model and Asymptotic Theory.....	80
3.3 A Feasible Test.....	84
3.4 Evaluation of the Test.....	86
3.5 Conclusions	87
References	88
Appendix.....	90
Tables and Figures.....	96
Conclusions	101

Abstract

This dissertation includes three essays on unit root tests and cointegration.

The first essay uses econometric tools to explore the possible causes of global warming. By examining the time series properties of the global mean temperature, solar irradiance and greenhouse gases, it finds static long-run steady-state and dynamic long-run steady-state relations between temperature and radiative forcing of solar irradiance and a set of three greenhouse gas series. Estimates of the adjustment coefficients indicate that temperature series is error correcting around 5% to 65% of the disequilibria each year, depending on the type of the long-run relation. The estimates of the $I(1)$ and $I(2)$ trends indicate that they are driven by linear combinations of three greenhouse gases and their loadings indicate strong impact on the temperature series. The equilibrium temperature change for a doubling of carbon dioxide is between 2.15 and 3.40 Celsius degree, which is in agreement with past literature and the report of the Intergovernmental Panel on Climate Change using fifteen different general circulation models.

The second essay examines the effect of random initial condition on the power of unit root tests. It develops the test statistics in the context of unknown structural change models using generalized least square detrended data when the initial observation is drawn from its unconditional distribution. It derives the limiting distributions of M-tests, ADF test and a family of feasible point optimal tests. Using two methods to estimate the break point, this study calculates the power envelopes and asymptotic power functions, and compares them with the case where the initial condition is fixed. Finite sample Monte Carlo simulations under various forms of error processes are performed using different lag length selection methods. Empirical applications are also provided.

The third essay employs a multivariate framework to improve the power of unit root tests. In the context of unknown structural change models, it shows that when testing for a unit root against stationarity with an unknown structural break, substantial power gains can be achieved by incorporating extra information contained in an arbitrary number of covariates. The power gains are dependent on the long-run correlations between the shocks of covariates and quasi-differences of potentially integrated time series. The higher correlation between covariates and quasi-differences of time series, the higher power gains can be realized. A feasible statistic to estimate the unknown correlation is also derived. In finite sample experiments, Monte Carlo simulation results confirm large power improvements without size deteriorations.

Acknowledgments

My special thanks goes to my supervisor Professor Gabriel Rodríguez. Without his support, the completion of this thesis would be impossible. I am truly grateful for his insightful comments and advice, which guided me out of the darkness. What I have learned the most from him is his incredible devotion to his work, which will always encourage me to pursue the better, and the greater in the future!

Introduction

Since the mid-1980s there has been an extraordinary development of research on unit root tests and cointegration. Whether a variable has a unit root has important implications in macroeconomics. When a time series has a unit root (nonstationary or difference-stationary), a shock has permanent effect; while for a stationary (or trend-stationary) time series, a shock has transitory effect. Various methods of testing whether a time series is stationary or not have been developed. When variables are nonstationary, the traditional ordinary least square regression tends to give spurious relations. Therefore cointegration was developed in order to analyze the relationship between nonstationary variables. Surveys on the different issues concerning unit root testing and cointegration may be found in Campbell and Perron (1991), Stock (1994), Watson (1994) and Hubrich et al. (1998).

One of the most popular unit root test statistics is the Dickey-Fuller (DF test hereafter) proposed by Fuller (1976), assuming identically and independently distributed (*i.i.d.* hereafter) errors. However, macroeconomic time series often exhibit serial correlations, which will cause bias in the DF statistic. To account for this problem, Said and Dickey (1984) proposed the Augmented DF test (ADF test hereafter). Both statistics test the null hypothesis of a random walk against the alternative of a stationary time series. Nelson and Plosser (1982) applied the ADF statistic to 14 US macroeconomics time series, and found a rejection of the null hypothesis only for the unemployment rates. The macroeconomic implication of this result is that shocks have permanent effects. However, these results were challenged by Perron (1989). He argued that the lack of modeling a broken intercept and/or slope was responsible for the non-rejection of a unit root in the data set studied by Nelson and Plosser, since a stationary process with a structural break mimic the behavior of a random walk. Perron applied the ADF test to the same data set assuming an exogenous break, and he found more evidence of rejection. Christiano (1992) criticized the treatment of structural break as exogenous by arguing that a known break point causes the problem of data mining. Following his paper, Zivot and Andrews (1992) and Perron (1997) proposed methods to estimate break point endogenously.

The key issue related to unit root tests is their low power. To address this issue, Elliott, Rothenberg and Stock (ERS hereafter, 1996) proposed using generalized least square (GLS hereafter) detrending to eliminate the deterministic components before testing. This approach brings significant power gains in both large and small samples. Perron and Rodríguez (2003) extended this approach to the context of structural change models and they

developed the limiting distributions of a set of broadly used unit root statistics, including the feasible point optimal statistic (ERS, 1996).

Another contribution in the spirit of improving the power of unit root statistics has been achieved by Hansen (1995). He proposed a multivariate unit root test by including a stationary covariate in the traditional ADF test. The resulting covariate ADF (CADF hereafter) statistic showed substantial power gains, especially when the covariate is highly correlated with the potentially correlated time series. Following a similar approach, Elliott and Jansson (2003) extended their point optimal test by including an arbitrary number of covariates and they saw further power gains.

More researchers are interested in the testing of cointegration between nonstationary variables as the literature of unit root test develops. The notion of cointegration establishes that, even when a set of time series is nonstationary, their linear combination might be stationary; see Engle and Granger (1987). The first type of statistics proposed to test for cointegration is named as residual-based tests. Assuming a set of n variables represented by y_t , a residual-based approach proceeds in a two-step fashion. In the first step, we run a static or long-run regression between variables. In the second step, a unit root statistic is applied to the residual from the first step. If the null hypothesis of a unit root (in the residual) is not rejected, then non-existence of cointegration will be concluded. In fact, any unit root statistic may be used to verify the presence of cointegration. The critical values depend on the type of deterministic components included and the number of right-hand-side regressors in the first-step; see Phillips and Ouliaris (1990). The shortcoming of this type of statistics is its identification problem, i.e., which variable should be put on the left-hand-side of the regression?

The second type of statistics developed by Johansen (1988) used a multivariate approach. Therefore they are called system cointegration tests. The advantage of this type of statistics is that all variables are considered as endogenous, and more than one cointegrating vector is potentially allowed in a set of n variables. The most popular case analyzed in the literature is the cointegration in an $I(1)$ framework where the maximum integration order of each variable is one; see Johansen (1995b). However, in recent years, the possibility of modeling time series with double unit roots has been considered and therefore the system cointegration test in an $I(2)$ framework has been developed; see Johansen (1992, 1995a). The advantage of this framework is that it brings richer results in the sense that long, medium and short term relationships are all identified.

This thesis deals with issues related to unit root tests and cointegration. The first chapter uses $I(1)$ and $I(2)$ framework of cointegration to analyze the influence of greenhouse gases on global mean temperature deviation. The

second chapter is an extension of Perron and Rodríguez (2003) by using quasi-differences of the data in unknown structural change models, assuming the initial condition is drawn from its unconditional distribution. The third chapter follows the approach proposed by Hansen (1995) and develops a multivariate unit root test in the context of unknown structural change models.

In the first paper, a system cointegration test has been applied to analyze the cause of global warming. This topic is inspired by the observation that both global mean temperature deviation and greenhouse gases (GHG hereafter) have been increasing during the past 150 years. In the existing literature, the unit root test and cointegration in an $I(1)$ framework have been used to analyze the time series property of global temperature variable and its relationship with greenhouse gases (GHG hereafter). With the application of unit root tests, most studies have agreed that global temperature is an $I(1)$ variable while the picture for GHG is mixed (results indicate it is likely that GHG has more than one unit root). Furthermore, studies have come to the conclusion that GHG has been a driving force for temperature (global warming). But because the possibility that the GHG may have more than one unit roots, the test of cointegration in an $I(1)$ framework becomes insufficient. This chapter intends to solve this problem using the newly developed cointegration test in an $I(2)$ framework. This approach is more complex but richer in terms of the relationships that we may find. In fact, we can identify not only the standard (static) long-run steady-state relations but also the dynamic long-run steady-state relations which are given by the linear combinations between the levels of the variables and their growth rates. The empirical results show that global temperature, solar irradiance and GHG series have different integration orders. Accordingly, three different systems are proposed and are tested. The results indicate that the temperature series is responding around 5%-65% to the disequilibrium error each year and the equilibrium temperature change for a doubling of carbon dioxide is between 2.15 and 3.40 Celsius degrees, depending on which type of steady-state relation is considered.

The second chapter examines the effect of random initial condition on the power of unit root tests. The existing literature on unit root statistics has assumed a fixed or random but bounded initial value (fixed initial condition hereafter). The convenience of this assumption is that the initial value will disappear at a rate of $T^{-1/2}$ asymptotically. But this simple assumption cannot describe the behavior of initial observation in a realistic way. Elliott (1999) assumed the initial value is drawn from its unconditional distribution (random initial condition hereafter) and found loss of power in unit root statistics derived upon. In practice, many time series appear to have a structural break.

But none of the previous papers examined the impact of a random initial value when structural break is present. Therefore, the second chapter of this thesis intends to fill this gap by examining the impact of a random initial value when there is an unknown structural break in the time series. I adopt the unknown structural change models in Perron and Rodríguez (2003) and assume random initial condition. Two methods proposed by Zivot and Andrews (1992) and Perron (1997) are used to select the break point. The power envelope and asymptotic power functions are calculated. Simulation results show that the asymptotic power functions are tangent to power envelope at one point and never fall far below it. In small sample analysis, I use four information criteria and a sequential t-statistic method to select the lag length in order to account for the serial correlations in the time series. Both large and small sample results show that the power of the tests has declined due to the change in the assumption of the initial condition. An empirical application is provided at the end of this chapter.

The third chapter of this thesis adopts the multivariate approach proposed by Hansen (1995) to improve the power of unit root tests when there is an unknown structural change. Traditionally, unit root tests have been applied to univariate time series. Incorporating the useful information in the stationary covariate, a variable correlated to the potentially integrated time series, may increase the power of unit root tests. The reason is that the use of stationary covariates results in a new error variance that is smaller than the error variance of a univariate time series. This leads to smaller confidence intervals and more powerful test statistics than those of the conventional unit root tests. Inspired by this idea, the third chapter derives the limiting distribution of a feasible point optimal statistic when there is an arbitrary number of stationary covariates and an unknown structural break. It compares the maximum power of this test with the power envelope derived by Perron and Rodríguez (2003), where no covariates are present. Both large and small sample simulation results show that substantial power gains are obtained, especially when the covariates are highly correlated to the quasi-differences of the time series to be tested.

References

- [1] Campbell, J. Y. and P. Perron (1991), "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots" in Blanchard, O.J., Fischer, S. eds., *NBER Macroeconomics Annual* 6, 141-201.

- [2] Christiano, L. J. (1992), "Searching for a Break in GNP," *Journal of Business and Economics Statistics* **10**, 237-250.
- [3] Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* **74**, 427-431.
- [4] Elliott, G. (1999), "Efficient Tests for a Unit Root When the Initial Observations is Drawn from its Unconditional Distribution," *International Economic Review* **40**, 767-783.
- [5] Elliott, G. and M. Jansson (2003) "Testing for Unit Roots with Stationary Covariate," *Journal of Econometrics* **115**, 75-89.
- [6] Elliott, G., T. Rothenberg and J. H. Stock (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica* **64**, 813-839.
- [7] Engle, R. F. and C. W. J. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica* **55**, 251-276.
- [8] Fuller, W. A. (1976) "Introduction to Statistical Time Series," New York: John Wiley.
- [9] Hansen, B. E. (1995) "Rethinking the Univariate Approach to Unit Root Testing," *Econometric Theory* **11**, 1148-1171.
- [10] Hubrich, K., H. Lütkepohl and P. Saikkonen (1998), "A Review of the Systems Cointegration," working paper, Humboldt Universitaet Berlin, Sonderforschungsbereich **373**, 1998-101.
- [11] Johansen, S., (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control* **12**, 231-254.
- [12] Johansen, S., (1992), "A Representation of Vector Autoregressive Process of Order 2", *Econometric Theory* **8**, 188-202.
- [13] Johansen, S., (1995a), "A Statistical Analysis of Cointegration for I(2) Variables", *Econometric Theory* **11**, 25-59.

- [14] Johansen, S., (1995b), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
- [15] Nelson, C. R. and C. I. Plosser (1982), "Trends and Random Walks in Macroeconomics Time Series: Some Evidence and Implications," *Journal of Monetary Economics* **10**, 139-162.
- [16] Perron, P. (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica* **57**, 1361-1401
- [17] Perron, P. (1997), "Further Evidence of Breaking Trend Functions in Macroeconomics Variables," *Journal of Econometrics* **80**, 355-385.
- [18] Perron, P. and G. Rodríguez (2003), "GLS Detrending, Efficient Unit Root Tests and Structural Change," *Journal of Econometrics* **115**, 1-27.
- [19] Phillips, P. C. B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica* **58**, 165-193.
- [20] Said, S. E. and Dickey, D. A. (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika* **71**, 599-608.
- [21] Stock, J. H. (1994), "Unit roots, Structural Breaks, and Trends" in McFadden, D.L., Engle, R.F., White, H. (Eds.) *Handbook of Econometrics*, Vol IV, Elsevier Science, Amsterdam, pp2740-2841.
- [22] Watson, M. W. (1994), "Vector Autoregressions and Cointegration," in Engle, R. F. and D. L. McFadden (Eds.) *Handbook of Econometrics*, vol. IV, Elsevier Science, Amsterdam, chapter 47.
- [23] Zivot, E. and D. W. K. Andrews (1992), "Further Evidence on the Great Crash, The Oil-Price Shock and the Unit Root Hypothesis," *Journal of Business and Economics Statistics* **10**, 251-270.

Chapter 1

Human Activities and Global Warming: A Cointegration Analysis

May 7, 2006

1 Introduction

One of the main conclusions achieved by the Intergovernmental Panel on Climate Change (IPCC, 2001) is that global mean temperature has been increasing for the past 150 years. Furthermore, the same reference attributes the responsibility of the changes to human activities generated GHG. A basic argument in this diagnostic is that both temperature and GHG have increased since the pre-industrial periods; see also Santer et al (1996). There are two sources for this evidence. They are the physically-based simulation models of climate and the statistical analysis of historical data. The simulation models of climate are based on physical motion equations trying to describe the principal issues governing the behaviour of temperature. They also include the radiative forcing of GHG and tropospheric sulfates. This kind of models are often referred as general circulation models (GCMs).

On the other hand, the statistical analysis deals directly with the historical record of the temperature, solar irradiance and GHG. In some cases, this approach uses simple and standard statistical or econometric tools in identifying for the effects of human activities (GHG) on the temperature series. However, because all these time series exhibit strong trends, classical tools will indicate spurious positive relation among these variables. In consequence, it is important to identify clearly the time series properties of the data before starting any other kind of analysis. In this aspect, the principal goal is to identify the existence of unit root in the data, that is, the existence of stochastic trends. This is the starting point of the research agenda of Stern and Kaufmann (1997, 1999, 2000) and Kaufmann and Stern (2002)¹. Using different statistical tests for the identification of unit root, they found that temperature series is an $I(1)$ process², and that GHG contain two unit roots. After analyzing the causality between different sets of time series, Stern and Kaufmann (1997) investigated the existence of long-run steady-state relations (i.e. cointegration between sets of variables) using multivariate technique proposed by Johansen (1988, 1995b).

¹See next section for a brief but more complete survey of the literature.

²A time series y_t is integrated of order d , which is denoted by $I(d)$ if d differences are needed to transform it into a stationary time series. Therefore, $I(1)/I(2)$ means that the series has to be differenced one/two times to achieve stationarity.

They concluded that there exists a long-run relation between the set of variables and that temperature series is reacting to the disequilibria towards this steady-state relation at around 40-50% each year. However the authors believe that this value is too high and it may be a consequence of the existence of I(2) trends in the system. Kaufmann and Stern (2002) recognizes the necessity to take into account I(2) trends in a more adequate statistical framework.

The statistical framework for I(2) is proportionated by the approach suggested by Johansen (1992, 1995a); see also Paruolo (1996), Rahbek et al. (1999), Paruolo and Rahbek (1999). It is more complex but richer in terms of the long-run relations that we may find. In fact, there are the standard (static) long-run steady-state relations but there also exist the dynamic long-run steady-state relations which are given by the linear combinations between the levels of the variables and their first differences. It is also possible to find medium-run steady-state relations. In this paper, we follow this approach in order to address the I(2) feature of the system.

The empirical results show that global temperature and solar irradiance series are I(1) processes, carbon dioxide is an I(2) process, and methane and nitrous dioxide seem to contain explosive roots. According to this evidence, three different systems are proposed, estimated and analyzed. The results indicate that temperature series is error correcting around 5%-65% of the disequilibria each year and the equilibrium temperature change for a doubling of carbon dioxide is between $2.15^{\circ}C$ and $3.40^{\circ}C$, depending on the type of steady-state relation that is considered. It is worth to note that these values are in agreement with other values found in the literature (see Kaufmann and Stern, 2002) and the average value calculated by the Intergovernmental Panel on Climate Change (IPCC, 2001).

The rest of this paper is organized as follows. Section 2 presents a brief review of the literature. Section 3 deals with the methodological issues while Section 4 presents the empirical results. Section 5 concludes.

2 A Brief Review of the Literature

In order to place this work in the context of the literature on climate change, a brief summary of some of the relevant papers is provided below. When we observe pictures of global mean temperature deviation series, greenhouse gas concentrations, and solar irradiance, all give us the information that they have increased in the last one hundred and fifty years; see also Stern and Kaufmann (2000). The presence of strong trends implies that the use of standard statistical or econometric tools will indicate a significant and positive association between sets of variables analyzed. Unfortunately, the use of standard statistical or econometric tools is misleading in this context. Therefore, a careful analysis of the statistical properties of the time series appears as a necessary condition for an adequate empirical analysis. This is the starting point of the research applied by Stern and Kaufmann (1997, 1999, 2000), Kaufmann and Stern (1997, 2002). Apart from them, there is few research that explicitly argued in favor of the use of econometric time series methods, such as Tol (1994), Tol and de Vos (1998), Schönwiese (1994). Some other exceptions in the analysis of the time series properties of global temperature are Blomfield (1992), Blomfield and Nychka (1992), Woodward and Gray (1993, 1995). Before them, most of the research analyzing the relationship between temperature and forcing variables have used simple regression models as in Lean et al. (1995), or frequency domain methods as in Kuo et al (1990), and Thomson (1995, 1997).

The analysis of the time series properties of temperature, solar irradiance, GHG and other variables has been performed in Stern and Kaufmann (1997). One of the principal conclusions of their research is that global temperature series has a unit root or in another term, this time series has a stochastic trend, denoted by $I(1)$. At the same time, GHG variables have been found to contain one or two unit roots. The econometric tools used were the well-known unit root statistics proposed by Dickey and Fuller (1979), Said and Dickey (1984), Phillips and Perron (1988), and Schmidt and Phillips (1992). As a consequence of the size and power problems of these unit root statistics in detecting for more than one unit root³, there isn't a complete

³See Haldrup and Lildholdt (2002).

and clear picture regarding the integration order of the GHG. However, the evidence in favor of two unit roots is present in more than one test statistics. This is also recognized by Stern and Kaufmann (1999, 2000), and Kaufmann and Stern (2002).

When the time series are non-stationary there is room to test for cointegration, that is, the possible existence of long-run relations between sets of variables analyzed. In other words, it is possible to find linear combinations of the variables that annihilates the stochastic trends. The rest of linear combinations are still stochastic trends and they drive the behaviour of the time series. The cointegrating relations are also named as static long-run relations because they represent a steady-state relation between sets of variables. Because in the short term the variables are not in the steady-state, they will react to these disequilibria. In the case of temperature, for example, it is possible to find a steady-state relation with GHG and natural factors (solar irradiance, for example). In the short term, the temperature (and the other variables) reacts to the disequilibria exhibiting an error correction behaviour.

The only two trials to deal with the identification of cointegration relations have been Stern and Kaufmann (1997), and Kaufmann and Stern (2002). Using the methodology of Johansen (1988, 1995b)⁴, they arrive to the conclusion that there exists a long-run relationship between time series of temperature, solar irradiance and a set of GHG. The estimates of the short-run dynamic model indicate that temperature is correcting around 40%-60% of the disequilibria each year. The authors consider that this value is too high and they attribute its cause to the existence of I(2) trends.

The acceptance that time series of GHG contain I(2) trends is also present in Stern and Kaufmann (1999, 2000). In both papers, the authors apply a multivariate structural time series approach to model temperature, natural factors and GHG⁵. The advantage of this kind of models is that different alternatives for the deterministic components are allowed. Also, it is possible to allow for I(1) and I(2) trends. The model is also flexible in including cyclical components. The basic conclusions are that most time

⁴See the I(1) model below for further methodological details.

⁵This methodology is based on Harvey (1989).

series analyzed contain a stochastic trend while the GHG containing stochastic I(2) trends. Therefore, the two independent stochastic trends in the data are associated to the radiative forcing due to GHG, solar irradiance, and tropospheric sulfate aerosols that are found in the northern hemisphere, respectively. A more recent research but applied to the temperature of Australia is Lenten and Moosa (2003). According to their results, temperatures in six Australian locations are I(1) processes, the cyclical component is not significant, seasonality is deterministic and the irregular (or noise) component is very significant.

However, some different results are proportionated by Kelly (2000). He finds that temperature series contains a unit root but that GHG are stationary around a time trend. This impressive result implies that temperature rise is due to long run cycles, and that the relationship with GHG is spurious. These results are in complete opposition to most of the research detailed before. However, applying first difference to temperature series, Kelly (2000) estimated a regression between growth rates of temperature and GHG, finding support to previous results regarding the significant influence of these gases on temperature series.

One important but uncertain parameter of the general circulation models (GCMs) is the temperature sensitivity which is measured as the equilibrium temperature change per unit of the change in radiative forcing. As Kelly (2000) argues, an alternative measure, directly proportional to the climate sensitivity, is the total equilibrium (steady-state) temperature change from a doubling of GHG, frequently denoted as ΔT_{2x} . Kelly (2000) finds that this parameter is between $1.27^{\circ}C$ and $1.33^{\circ}C$; while for Kaufmann and Stern (2002) this value is around $2.0^{\circ}C$ to $2.5^{\circ}C$, in agreement with the general circulation models. It is worthwhile to mention that the IPCC (2001) argues that the average value of ΔT_{2x} is $3.5^{\circ}C$ after considering fifteen different models⁶.

In this paper we still use time series tools to analyze the existence of steady-state relations between temperature series, solar irradiance and a set of GHG. However, unlike the traditional approach of using the I(1) model,

⁶The standard deviation is $0.92 C^{\circ}$ and the range goes from $2.0 C^{\circ}$ and $5.1 C^{\circ}$.

as Stern and Kaufmann (1997) and Kaufmann and Stern (2002), we use econometric tools allowing for the presence of I(2) trends in the identification of the long-run relationships. The use of more sophisticated econometric tools taking into account the presence of I(2) trends has been recognized by Stern and Kaufmann (1999). The presence of two unit roots in time series allows us to use the I(2) model, see Johansen (1992, 1995a,b). In this model the number of cointegrating relations are detected, so are the number of I(1) and I(2) stochastic trends that still in the system and drive the behaviour of some variables. In terms of the cointegrating relations, unlike the analysis of the I(1) model, the I(2) model allows for the existence of two types of cointegration. The first type is a linear combination of the variables in their levels, which is the same definition as that in the I(1) model. The second type of cointegration is the possibility that there are linear combinations between the levels of the variables and their growth rates. It is named polynomial cointegration and also dynamic steady-state relations. Further details of I(1) and I(2) models are presented in the next section.

3 Methodological Issues

In this section we provide elements help to understand the methodology applied in the empirical analysis. Regression with nonstationary time series implies spurious regression except when there exists a (or more than one) linear combination between these variables that reduce the dimension of the space spanned by them. This is known as cointegration.

3.1 The I(1) Model

Let y_t , a vector of n variables, be represented by the following $VAR(k)$:

$$y_t = \sum_{i=1}^k \Pi y_{t-i} + \Phi D_t + \epsilon_t \quad (1)$$

where it is assumed that ϵ_t is a sequence of *i.i.d.* error with zero mean and covariance matrix Ω . In most cases it is also assumed that the errors are Gaussian which is denoted by $\epsilon_t \sim Niid(0, \Omega)$. The variable D_t contains the possible deterministic components of the process, such as a constant,

a time trend, seasonal dummies and intervention dummies. This is the model proposed by Johansen (1988, 1995b) and is widely used in empirical applications⁷.

The system (1) is reparameterized as a vector error correction model (VECM):

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \Phi D_t + \epsilon_t \quad (2)$$

with $\Pi = -I + \sum_{i=1}^k \Pi_i$, $\Gamma_i = \sum_{j=i+1}^k \Pi_j$. Notice that the matrix $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$.

I(1) cointegration occurs when the matrix Π is of reduced rank, $r < n$ where Π may be factorized into $\Pi = \alpha\beta'$, α and β are both full rank matrices of dimension $n \times r$; the matrix α contains the adjustment coefficients and β the cointegration vectors. These vectors have the property that $\beta'y_t$ is stationary, even though y_t itself is non-stationary. Notice that there also exist full rank matrices α_\perp and β_\perp of dimension $n \times (n-r)$ which are orthogonal to α and β , such that $\alpha'_\perp \alpha = 0$ and $\beta'_\perp \beta = 0$, and the $rank(\beta_\perp, \beta) = n$.

An alternative representation of the cointegrated VAR model is in terms of the common stochastic trend representation, see Stock and Watson (1988). According to that, the y_t vector is represented by

$$y_t = \Phi D_t + C \sum_{i=1}^t \epsilon_i + C(L)\epsilon_t \quad (3)$$

where $C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1} \alpha'_\perp$ and $C(L)\epsilon_t$ corresponds to a n -dimensional $I(0)$ component. Using this representation it is possible to observe that although y_t is n -dimensional, the vector series is driven by just $n-r$ common stochastic $I(1)$ trends which are $\alpha'_\perp \sum_{i=1}^t \epsilon_i$. In terms of observable variables the $I(1)$ directions are calculated as $\beta'_\perp y_t$ which are just a particular linear combinations of the stochastic trends.

To test the rank of matrix Π , Johansen (1995a, b) developed maximum likelihood cointegration testing method using the reduced rank regression

⁷There are a large number of empirical applications using this statistical framework. Two very detailed and influential applications are Johansen and Juselius (1992, 1994).

technique based on canonical correlations. The procedure consists of obtaining an $n \times 1$ vector of residuals r_{0t} and r_{1t} from auxiliary regressions (regressions of Δy_t and y_{t-1} on a constant and the lagged $\Delta y_{t-1} \dots \Delta y_{t-k+1}$). These residuals are used to obtain the $(n \times n)$ residual product matrices:

$$S_{ij} = (1/T) \sum_{t=1}^T r_{it} r'_{jt}, \quad (4)$$

for $i, j = 0, 1$. The next step is to solve the following eigenvalue problem

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0 \quad (5)$$

which gives the eigenvalues $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_n$ and the corresponding eigenvectors $\hat{\beta}_1$ through $\hat{\beta}_n$, which are also the cointegrating vectors. A test for the rank of matrix Π can now be performed by testing how many eigenvalues λ equal to unity. One test statistic for the resulting number of cointegration relations is the Trace statistic (see Johansen, 1988), which is a likelihood ratio test defined by

$$Trace = -T \sum_{i=r+1}^n \log(1 - \hat{\lambda}_i) \quad (6)$$

Another useful test is given by testing the significance of the estimated eigenvalues themselves

$$\lambda_{\max} = -T \log(1 - \hat{\lambda}_i) \quad (7)$$

In trace test, the null hypothesis is $r = 0$ (no cointegration) against the alternative hypothesis that $r > 0$ (cointegration). The λ_{\max} statistic tests the null hypothesis that $r = r_0$ versus the alternative hypothesis that $r = r_0 + 1$, where $r_0 = 0, 1, \dots, n-1$. For further details regarding the construction of these statistics, see Johansen (1995b).

In summary, the I(1) model allows to identify the rank of cointegration (r) and the number of stochastic trends driving the y_t vector. Because in this model it is assumed that y_t contains variables with order of integration no larger than one, it is clear that the number of unit root left in the system is $n - r$. Notice that the estimate of β (the cointegration vector) is not identified in the sense that any linear combination of β is also a cointegration

vector. In this sense, researchers have to identify these vectors by imposing restrictions that are in most cases suggested by the economic theory. For example, in a vector containing four variables, a possible restriction is that the coefficient associated to the second variable is zero. If the statistic (distributed as a χ^2) does not reject the null hypothesis, it means that the second variable is long-run excluded from the cointegration vector. Another example is the restriction of long-run homogeneity between a set of variables. If the null hypothesis is not rejected, it means that the unit vector (i.e. $[1, -1, \dots, -1]$) may be used in the subsequent analysis. The estimated adjustment coefficients (α) are also tested for restrictions, but in this case the restrictions are associated with the null hypothesis of weak exogeneity.

3.2 The I(2) Model

In the I(1) model, the matrix Π is of reduced rank and the matrix $\alpha'_\perp \Gamma \beta_\perp$ is of full rank. For the system to be I(2) it is also required that the matrix $\alpha'_\perp \Gamma \beta_\perp$ be of reduced rank $s_1 < n - r$. Following Johansen (1992, 1995a,b), the model (2) can be reparameterized as

$$\Delta^2 y_t = \Pi y_{t-1} - \Gamma \Delta y_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 y_{t-i} + \Phi D_t + \epsilon_t \quad (8)$$

where the matrix Γ is included as a parameter and $\Psi_i = -\sum_{j=i+1}^{k-1} \Gamma_j$.

With the matrix $\alpha'_\perp \Gamma \beta_\perp$ of reduced rank (s_1), it is possible, as in the I(1) model, to define parameter matrices ξ and η such that the second reduced rank condition is $\alpha'_\perp \Gamma \beta_\perp = \zeta \eta'$ with ζ and η both matrices of dimension $(n - r) \times s_1$.

As in the I(1) model, the number of cointegration relations (or I(0) relations) is denoted by r . However, unlike I(1) models, $n - r$ does not only represent the number of I(1) trends in I(2) models. It contains both I(1) and I(2) trends, denoted as s_1 and s_2 respectively. Consequently there are parameters describing the I(0), I(1) and I(2) directions of the variables and an important goal of the I(2) model is their identification. Following the notation of Juselius (1999), these matrices are $\beta, \beta_{\perp 1}$ and $\beta_{\perp 2}$ associated

with dimensions r, s_1 and $n - r - s_1 = s_2$ respectively⁸. Paruolo (1996) denotes r, s_1, s_2 as the integration indices of the VAR.

In a similar way as in the I(1) model, the common trend representation of I(2) model is given by

$$y_t = \Phi D_t + C_2 \sum_{j=1}^t \sum_{i=1}^j \epsilon_i + C_1 \sum_{i=1}^t \epsilon_i + C^*(L) \epsilon_t \quad (9)$$

where $C_2 = \beta_{\perp 2} (\alpha'_{\perp 2} \Theta \beta_{\perp 2})^{-1} \alpha'_{\perp 2}$ and $C^*(L)$ is a matrix polynomial with all roots strictly outside of the unit circle. The first clear observation is that y_t has s_2 common I(2) trends given by $\alpha'_{\perp 2} \sum_{j=1}^t \sum_{i=1}^j \epsilon_i$. In terms of observable variables, it is given by $\beta'_{\perp 2} y_t$ which are just linear combinations of the stochastic trends $\alpha'_{\perp 2} \sum_{j=1}^t \sum_{i=1}^j \epsilon_i$.

Furthermore, as mentioned by Haldrup (1999), the combinations $\beta' y_t$ can cointegrate to I(0) level and/or have the property that they potentially cointegrate with $\beta'_{\perp 2} \Delta y_t$, which is I(1) by construction. This is named as polynomial cointegration. These r relations are $\beta' y_t - \delta \beta'_{\perp 2} \Delta y_t$ which define the I(0) directions. However, not all the r I(0) relations need include the differenced I(2) components. In fact, after defining δ_{\perp} such that $\delta'_{\perp} \delta = 0$, there will be $r - s_2$ non-polynomially cointegrated relations given by $\delta'_{\perp} \beta' y_t$; and s_2 polynomial cointegrating relations given by $\delta' \beta' y_t - \delta' \delta \beta'_{\perp 2} \Delta y_t$.⁹

The approach suggested by Johansen (1992, 1995) to identify the number of I(2) trends is conducted as a combination of regression and reduced rank regression. It is performed in a similar way as the determination of the cointegration rank in the I(1) model. The difference is that now two reduced rank conditions need to be examined. This is more complicated in the sense that the second reduced rank condition depends on the first one. Instead of

⁸The notation and technical details of the I(2) model are complex. Further details regarding the calculation of matrices $\beta_{\perp 1}$ and $\beta_{\perp 2}$ are in Juselius (1999), also Haldrup (1999) using a slightly different notation. In summary, because ζ and η exist, their complements, ζ_{\perp} and η_{\perp} , also exist. Then it is possible to define $\alpha_{\perp} = \{\alpha_{\perp 1}, \alpha_{\perp 2}\}$ and $\beta_{\perp} = \{\beta_{\perp 1}, \beta_{\perp 2}\}$, where $\alpha_{\perp 1} = \alpha_{\perp} (\alpha'_{\perp} \alpha_{\perp})^{-1} \zeta$, $\alpha_{\perp 2} = \alpha_{\perp} \zeta_{\perp}$, $\beta_{\perp 1} = \beta_{\perp} (\beta'_{\perp} \beta_{\perp})^{-1} \eta$ and $\beta_{\perp 2} = \beta_{\perp} \eta_{\perp}$. Therefore, $\beta, \beta_{\perp 1}$ and $\beta_{\perp 2}$ are mutually orthogonal and thus jointly describe a basis for the n -dimensional space. The α has a similar property.

⁹The possibility to decompose r (in r_0 and r_1 , say) exists only when $r > s_2$. In this case, the r cointegrating relations can be divided into $r_0 = r - s_2$ directly stationary CI(2,2) relations and $r_1 = s_2$ polynomially cointegrating relations.

a joint estimation of the indices r and s_1 , Johansen (1992, 1995a) suggests a two step procedure. The first step is to solve the reduced rank problem associated with the matrix $\Pi = \alpha\beta'$. It calculates the estimates of α_r , β_r , $\alpha_{\perp r}$, and $\beta_{\perp r}$ for each value of $r = 0, 1, \dots, n - 1$. The second step deals with the second reduced rank condition problem by replacing the unknown matrices $\alpha_{\perp r}$, and $\beta_{\perp r}$ with the estimates from the first step. The problem is solved for $s_1 = 0, \dots, n - r - 1$. Then what remains to be determined is which combinations of r and s_1 should be chosen. Because steps 1 and 2 proportionate an array of different values for r and s_1 corresponding to different sub-models, the selection of r and s_1 can be performed in the following way¹⁰. The array is read starting from the left-corner. If the null hypothesis is rejected, the next element (continuing to the right) is read and so on. After the first row is done, and if no acceptance is observed, we read the second row of the array. We continue until the first non-rejection is found. The associated values of r , s_1 and s_2 are selected as the integration indices. Complete technical details can be found in Paruolo (1996), Johansen (1992, 1995a).

It must be emphasized that only the space spanned by the cointegrating vectors is identified; the single cointegration relations are unidentified. This issue is also present in the I(1) model but the problem is more complex in I(2) models¹¹.

What is perhaps more interesting is the fact that in an I(2) model there exist more than one steady-state relations. Recall that in an I(1) model, the cointegration relation $\beta'y_t$ represents the static long-run steady-state relation. In the present case, we have two additional steady-state relation. One is medium-run steady-state relations represented by $\beta'_{\perp 1}\Delta y_t$. The other one is dynamic long-run steady-state relations represented by the polynomial cointegrating relation (see Juselius, 1999, 2003). In consequence, there exist different adjustment coefficients associated with each of these steady-state relations.

¹⁰The test statistic is denoted as $H_{r,s}$ and it is the same notation used in the empirical analysis.

¹¹Regarding the adjustment coefficients, for example, the notion of weak exogeneity is now different. Paruolo and Rahbek (1999) have proposed a sequential approach to test for weak exogeneity in the I(2) model.

One should keep in mind that tools for analyzing I(2) models and statistics used to select integration indexes are not yet fully developed in econometric literature. This is the reason why the selection of the number of I(2) trends should be combined with other tools. One of them, as suggested by Juselius (1999), is the calculus of the roots of the companion matrix. If y_t vector contains variables that are integrated no more than order one, then the total number of unit (or close to unit) root of the companion matrix must be $n - r$. If there exists more than $n - r$ roots close to the unit circle, it constitutes an indicator for the presence of I(2) trends. When there are I(2) trends, the number of roots close to the unit circle is $s_1 + 2s_2$.

Another useful indicator for the presence of I(2) trends is to compare the graphs of $\beta' y_t$ and $\beta' R_{1t}$, where R_{1t} is a vector of residuals from regressing y_{t-1} on lagged short-run effects (Δy_{t-i} , $i = 1, 2, \dots, k - 2$) and D_t . If the first graph looks non-stationary whereas the second graph looks stationary, it can be considered as a strong support for the presence of I(2) trends.

Recently a few empirical applications in the field of economics have used the I(2) framework. Without the intention to be exhaustive, some of these references are Rahbek et al. (1999), Kongsted (2003), Holtemöller (2002), Vostroknutova (2003), Juselius (1999, 2003), Fiess and MacDonald (2001) and Haldrup (1999). It is worthwhile to mention that some of these papers, after identifying for the presence of I(2) trends, proceed with transforming the variables in such a form that a standard I(1) analysis can be performed (see Fiess and MacDonald, 2001). Almost all references mentioned study nominal variables such as money supply, nominal wages, and fundamentally prices. We do not know that a similar methodology has been applied to climate series.

4 Empirical Analysis

4.1 The Data and Preliminary Issues

The data used in this paper include the time series of global mean temperature deviation (denoted by *temp*), the concentrations of methane (*ch₄*),

nitrous dioxide (n_2o), carbon dioxide (co_2)¹². All data are from the web site of Goddard Institute for Space Studies available at www.giss.nasa.gov. The data is transformed into radiative forcing¹³, which affects the temperature directly and indirectly. The formulae are tabulated in the Intergovernmental Panel of Climate Change (IPCC, 2001). Therefore the radiative forcing of GHG at time t are denoted by $rfch_{4t}$, rfn_2o_t , $rfco_2t$. We also include radiative forcing of solar irradiance ($rfsun_t$) as another variable in the system.

Figure 1 presents the evolution of the time series for the period under study, 1856-2001. It is a similar period as analyzed by the previous literature but with more recent information. A clear observation appearing from Figure 1 is the fact that all GHG variables present strong upward trends that have been observed in the previous literature.

As that argued in the literature (Stern and Kaufmann, 2000), an essential preliminary step in the analysis of these variables is the identification of the time series properties. One way to approach this issue is the application of univariate unit root tests. Stern and Kaufmann (2000) proceeded with using three different univariate unit root tests. However, as documented by Haldrup and Lildholdt (2002), these statistics are incorrect. Because when tested for $I(2)$ and the underlying series is indeed integrated of order two, these statistics give rise to an excessive rejection of the null hypothesis

¹²A preliminary version of the paper included data of chlorofluorocarbons (CFC_{11} and CFC_{12}). Graphical inspection of both time series indicates zero values until 1930s, after that they increased very fast. These two variables were found to be $I(2)$ processes in Stern and Kaufmann (2000). We decided exclude both time series for the following reasons: i) the visual analysis indicates a particular behaviour that may distorts the analysis; ii) their importance in terms of total GHG has reduced; iii) increasing the dimension of the system of equation to be estimated, therefore reducing the number of freedom degrees given our sample size; and iv) identification of cointegration relations is complex and unlike the field of economics, we are not sure what are the physical relationships and interactions between all these GHG series, solar irradiance and temperature time series. Therefore, by excluding both variables, we reduce the possibility to find too many cointegration relations which are always difficult to identify and interpret.

¹³Radiative forcing is the change of the net irradiance caused by factors such as GHG, water vapor, solar radiation. GHG are of particular interest, as they are most likely to change radiative forcings over the next decade. The net irradiance is the difference of the irradiance that the earth absorbed minus the irradiance the earth emitted, expressed in Wm^{-2} , where W is a measure of energy (Watt) and m indicates meters.

of a unit root in favor of the stationary and explosive alternatives. This size distortion is caused by the fact that the test statistics have a different distribution originated by one additional unit root. In consequence, the recommendation of Haldrup and Lildholdt (2002) is to test I(2) against I(1) prior to testing I(1) against I(0). These authors concluded that all basic univariate unit root tests suffer from size distortion.

In a set of results not reported here¹⁴, we applied double unit root test suggested by Dickey and Pantula (1987) and also the statistic proposed by Hasza and Fuller (1979). The results indicated that the carbon dioxide series contains two unit roots, that is, it is an I(2) process. The temperature and solar irradiance series were found to be I(1) processes, the other GHG series under study (nitrous dioxide and methane) appeared containing explosive roots. The last result is not totally clear after the application of all statistics. The difficulty with this case is the fact that explosive roots mimic the behaviour of I(2) processes, see Haldrup and Lildholdt (2002).

In summary, most of the results confirm the previous results found in the literature, see Stern and Kaufmann (2000). Although it appears to be the case, we adopt another strategy in this paper. This approach consists of applying multivariate technique to detect the number of I(0), I(1) and I(2) trends in the system. As we shall see in the following analysis, there are more than one possible cases in the selection of the integration indices¹⁵.

4.2 The First Case

The first case is represented by the system with our 5 variables, ie, $y_t = \{temp_t, rfsun_t, rfcO_2_t, rfch_4_t, rfn_2o_t\}$. Table 1 presents the results obtained from the application of the approach of Johansen (1988, 1995b) in determining the rank of matrix Π ¹⁶. Both Trace and λ_{max} statistics indicate

¹⁴But they are available upon request. I applied standard ADF test (Dickey and Fuller, 1979; Said and Dickey, 1984); Phillips-Perron test (Phillips and Perron, 1988), and ADF test using GLS detrended data as suggested by ERS (1996).

¹⁵Analysis of I(1) and I(2) models was performed using CATS for RATS, see Hansen and Juselius (1995). A slightly modified version of the program of Jorgensen et al. (1999) was used. I also thank H. C. Kongsted who proportionated his program used in Kongsted (2003).

¹⁶In all cases, we consider linear trends in the data. An intercept and a time trend are also allowed in the cointegration space. In terms of the I(2) framework, this is the model

that $r = 3$, that is, there exist three cointegration relations. Because $n = 5$, we have $s_1 + s_2 = n - r = 2$.

Table 2 presents the results of the statistic $H_{r,s}$ for selecting s_2 . The first row represents the value of the statistic, the second and third rows (in italic) are the critical values at 95.0% and 97.5% respectively. As we explained in the previous section, reading of this table starts from the left-corner and we continue to the right. If no acceptance is found, we continue to the second row of values ($r = 1$). The procedure stops when a non-rejection is found. In the present case, the results indicate that $s_2 = 0$ and consequently $s_1 = 2$. Then, there are two I(1) trends and there is no I(2) trends in the system. However, note that using critical values at 97.5%, it is possible to find $s_2 = 1$ and consequently $s_1 = 1$.

As suggested by Juselius (1999), a good way to complement this information is the calculus of the eigenvalues of the companion matrix. The eight largest eigenvalues obtained from the unrestricted VAR are: 1.0292, 1.0292, 0.9641, 0.9641, 0.9384, 0.9384, 0.7163, 0.7163. It appears to exist two explosive roots and four eigenvalues close to the unit circle. In other words, there appears to exist four unit roots in the system. Now, because the first step (in the I(1) framework, see Table 1) indicates $r = 3$, we proceed with imposing this restriction and now the eight largest eigenvalues of the restricted VAR are: 1.0283, 1.0283, 1.000, 1.000, 0.9378, 0.9378, 0.7179, 0.7179. There are two unit roots as a consequence of the restriction of $r = 3$ but there are two more eigenvalues close to the unit circle. It indicates, again, the existence of four unit roots¹⁷. Remember that in the I(2) framework the total number of unit roots is given by $s_1 + 2s_2$, which in the present case indicates that $s_1 = 0$ and $s_2 = 2$.

In summary, we have some different information using the $H_{r,s}$ statistic and the eigenvalues of the companion matrix. We decide to work with both alternatives. Then, they are $r = 3, s_1 = 1, s_2 = 1$ and $r = 3, s_1 = 0, s_2 = 2$, respectively. In the following, they are named as Case 1 and Case 2 respectively.

Following the notation used in the previous section, Table 3a presents

suggested by Rahbek et al. (1999).

¹⁷Notice that the two explosive roots seem to confirm the univariate analysis.

cointegration relations and adjustment parameters corresponding to Case 1. Notice that there are two cointegrating relations including only the levels of the variables ($\widehat{\beta}_{0,i}y_t$, $i = 1, 2$). Furthermore, the relation $\widehat{\beta}_1y_t - \widehat{\kappa}_1\Delta y_t$ denotes the polynomial cointegrating relation or what we denoted in the previous section as the dynamic steady-state relation. The bottom panel gives the corresponding loadings ($\widehat{\alpha}_{0,1}$, $\widehat{\alpha}_{0,2}$, $\widehat{\alpha}_1$), and the coefficients that reflect the composition of the stochastic I(1) and I(2) trends are $\widehat{\alpha}_{\perp 1}$ and $\widehat{\alpha}_{\perp 2}$. The table also presents vectors $\widehat{\beta}_{\perp 1}$ and $\widehat{\beta}_{\perp 2}$ that denote the loadings (adjustment parameters) to the stochastic I(1) and I(2) trends.

Table 3a specifies that around 46.6% of the disequilibrium in the static long-run steady-state is corrected by the temperature series each year. This is similar as what was found by Stern and Kaufmann (2000). On the other hand, the result that temperature series does not react to the second static long-run relation is also interesting. The response of the temperature series to the dynamic steady-state relation is 19.1%. As we know, this dynamic steady-state relation (polynomial cointegration relations) includes the levels and the growth rates of the variables. Therefore, the result indicates the response of temperature series to disequilibria towards the steady-state. Regarding to the medium-run steady-state relation ($\widetilde{\beta}'_{\perp 1}\Delta y_t$), we observe that it is composed by temperature, solar irradiance, carbon dioxide and methane series.

Regarding the stochastic I(1) and I(2) trends, we have the following observations. Firstly, it appears that both I(1) and I(2) trends are driven by three GHG ($\widehat{\alpha}_{\perp 1}$, $\widehat{\alpha}_{\perp 2}$), whereas the influence of methane and nitrous dioxide is higher in stochastic I(2) trend ($\widehat{\alpha}_{\perp 2}$). In other words, permanent shocks to the three GHG seem to have generated the I(2) trend. Secondly, the respective loadings (or adjustment parameters) to these stochastic I(1) and I(2) trends indicate that temperature series, solar irradiance and carbon dioxide are influenced by them ($\widehat{\beta}_{\perp 1}$, $\widehat{\beta}_{\perp 2}$). The results also show that temperature, carbon dioxide, and solar irradiance are more influenced by the stochastic I(1) trend ($\widehat{\beta}_{\perp 1}$). However, the stochastic I(2) trend seems to affect strongly the temperature series ($\widehat{\beta}_{\perp 2}$).

Regarding the equilibrium temperature change for a doubling of carbon dioxide (ΔT_{2x}), the results indicate that it is 3.4 °C; see the first long-run

steady-state relation. This value is higher than those found in the literature but it is in agreement with the average of estimates from fifteen general circulation models coupled to mixed-layer ocean models reported in IPCC (2001).

Table 3b proportionates similar information but for the case $r = 3, s_1 = 0, s_2 = 2$. In this case the results show that there is only one static long-run steady-state relation and two dynamic steady-state relations. The loading corresponding to the long-run steady-state relation seem to indicate that temperature series is not error correcting which may indicate that the disequilibria are not corrected at all. The response to the dynamic steady-state is, as before, higher and with the correct sign. In fact, it appears that temperature series is correcting 4.8% and 64.6% (each year) of the disequilibria presented in these relations.

Because in this case $s_1 = 0$, there are not I(1) trends in the system. The two I(2) trends seem to be driven by carbon dioxide and nitrous dioxide in the first case ($\hat{\alpha}_{12,1}$) and by the three GHG in the second case ($\hat{\alpha}_{12,2}$). Observing the coefficients in $\hat{\beta}_{12,1}$ and $\hat{\beta}_{12,2}$, it is clear that the higher influence (in both cases) is on temperature series. All these results confirm the analysis of Kaufmann and Stern (2000).

4.3 The Second Case

Univariate and multivariate analysis seem to indicate that there are explosive roots in the system. Therefore, an alternative analysis is to separate the five-variable system into two sub-systems. The first system contains 3 variables: temperature, solar irradiance and carbon dioxide. The second system has two variables: methane and nitrous dioxide, the ones that seem to have explosive roots.

Tables 4a and 4b present the Trace and λ_{\max} statistics. Notice that the selection of the rank of matrix Π is valid even when there are explosive or I(2) trends, see Nielsen (2001, 2002). Table 4a (3-variable system) indicates that $r = 1$. The same result is found for Table 4b (2-variable system).

Tables 5a and 5b present the results from the $H_{r,s}$ statistic. The numbers in italic are critical values at 95% quantile. In the case of the 3-variable sys-

tem (Table 5a), we found $s_2 = 1$ and consequently $s_1 = 1$. The seven largest eigenvalues of the companion matrix corresponding to the unrestricted VAR are: 0.9942, 0.9297, 0.8402, 0.8402, 0.7796, 0.6543, 0.6543. It seems there are two roots close to the unit circle. When $r = 1$ is imposed, the seven largest eigenvalues are: 1.000, 1.000, 0.9546, 0.8494, 0.8494, 0.6844, 0.6844. We have two unit roots as the results of restriction $r = 1$. However, there is another root 0.9546 very close to the unit circle, and this indicates the presence of I(2) trends. Therefore, according to the results of the eigenvalues of the companion matrix, the total number of unit roots is $s_1 + 2s_2 = 3$. This is in agreement with the results of the statistic $H_{r,s}$. Because we have the evidences that solar irradiance and temperature are both I(1) variables from univariate tests. It seems that carbon dioxide is the one responsible for the existence of the I(2) trend.

For the system with 2 variables (Table 5b), the results indicate $s_2 = 1$, therefore, $s_1 = 0$. In this case, the existence of one I(2) trend is difficult to accept since our preliminary results (see the last sub-section) indicate that there are two explosive roots which should be attributed to the two variables in the 2-variable system. The five largest eigenvalues of the companion matrix from the unrestricted VAR are: 1.0393, 1.0393, 0.8774, 0.6901, 0.6901. When $r = 1$ is imposed, the eigenvalues are: 1.029, 1.029, 1.000, 0.714, 0.714. We have one unit root corresponding to the restriction of $r = 1$ and two explosive roots. Others are not close to the unit circle. This information indicates that there is no I(2) trend in this system. Whereas it seems to verify that explosive roots mimic the behaviour of I(2) trends, as reported by Haldrup and Lildholdt (2002). Therefore in this case we conclude with $r = 1$, $s_1 = 1$ and $s_2 = 0$.

Table 6 shows the estimates of $\hat{\beta}_1, \hat{\kappa}_1, \hat{\alpha}_1, \hat{\beta}_{\perp 1}, \hat{\beta}_{\perp 2}, \hat{\alpha}_{\perp 1}$ and $\hat{\alpha}_{\perp 2}$ ¹⁸. Each year the temperature series corrects 12.6% the disequilibria in the dynamic long-run steady-state relation (polynomial cointegration). It is interesting to observe that the I(1) trend is driven by temperature and radiative forcing of solar irradiance but not by radiative forcing of carbon dioxide. However, the I(2) trend is completely driven by this greenhouse gas. The magnitude of the

¹⁸There is no similar estimates in the case of the 2-variable system because there are not I(2) trends.

estimates of $\hat{\beta}_{\perp 1}$ tells us that the I(1) trend affects significantly all variables in the system, with the major effect on temperature series. Because this I(1) trend is driven by the temperature series itself and for the radiative forcing of solar irradiance, the estimates of $\hat{\beta}_{\perp 1}$ indicate that this trend corresponds to “inertial” (or persistent) factors (temperature itself) and natural factors (radiative forcing of solar irradiance). In the case of the magnitudes of $\hat{\beta}_{\perp 2}$, the effects are also appreciated on temperature series. This I(2) trend could correspond to the human factors, that is, the GHG effects, in this case represented by radiative forcing of carbon dioxide.

The cointegration relation detected in the 2-variable system indicates a relationship between radiative forcing of methane and nitrous dioxide. We introduce this relation (together with the polynomial cointegration relation found in the 3-variable system) in the error correction model to calculate the response of the temperature series to these steady-state relations. The results indicate that temperature series responds 50.18% to the static long-run relation between radiative forcing of methane and nitrous dioxide. These results are in agreement with those results found in the last sub-section. From Table 6, we found that $\Delta T_{2x} = 2.15 C^{\circ}$, which is in agreement with past literature, see Kaufmann and Stern (2002).

4.4 The Third Case

Another alternative is to transform the variables in such a way that an I(1) framework can be performed. Based on the results outlined before, temperature series and radiative forcing of solar irradiance are I(1) processes. Hence they enter the new system in levels. Radiative forcing of carbon dioxide is most likely an I(2) process and then enters the system in the first differences (Δrfc_2). For the case of radiative forcing of methane and nitrous dioxide, we found evidence of explosiveness. Therefore we apply filter Δ_{λ} ($\Delta_{\lambda}y_t = y_t - \lambda y_{t-1}$) to these variables with $\lambda = 1.029$. In summary, our 5-variable system is now composed of $y_t = \{temp_t, rfsun_t, \Delta rfc_2_t, \Delta_{\lambda}rfch_4_t, \Delta_{\lambda}rfn_2o_t\}$.

Table 7 presents the results from the application of the Trace and λ_{max} statistics and it suggests $r = 2$. In order to be sure that our system does not

contain I(2) trends, Table 8 presents the results of the $H_{r,s}$ statistics. The results confirm our claim. Therefore the integration indices are $r = 2$ and $s_1 = n - r = 3$. The eight largest eigenvalues of the companion matrix for the unrestricted VAR are: 0.9581, 0.9581, 0.923, 0.923, 0.861, 0.861, 0.848, 0.848. In the case of the restricted VAR ($r = 2$), these eigenvalues are: 1.000, 1.000, 1.000, 0.890, 0.890, 0.858, 0.858, 0.838. We have three unit roots corresponding to the restriction of $r = 2$ and no other roots close to the unit circle¹⁹. The results seem to confirm that the total number of unit root is three implying that there are not I(2) trends.

Table 9 presents the estimates of the two cointegrating vectors, the loading values and the I(1) trends. The null hypothesis for the long-run exclusion of solar irradiance in the first steady-state relation and the imposition of a coefficient of -0.5 associated to carbon dioxide was not rejected with a $\chi^2_{(1)} = 2.75$ corresponding to a p-value of 0.10. The long-run exclusion of the time trend was always strongly rejected.

The estimates of the adjustment parameters show that temperature series is error correcting 10.8% and 12.3% each year. The first I(1) trend is driven by three GHG with a larger weight on radiative forcing of methane. The second I(1) trend is almost completely driven by radiative forcing of nitrous dioxide but with a small participation of radiative forcing of solar irradiance. The last I(1) trend is driven by two GHG with a larger weight on radiative forcing of carbon dioxide and again, a small participation of radiative forcing of solar irradiance.

5 Conclusions

This paper applies multivariate I(1) and I(2) tools to identify the existence and the number of long-run steady-state relations between temperature series and radiative forcing of solar radiance and a set of three GHG. One of the results indicates that temperature and radiative forcing of solar irradiance series appear to be I(1) processes, radiative forcing of carbon dioxide is an I(2) process, and radiative forcing of methane and nitrous dioxide seem to

¹⁹In economics, some researchers consider that an eigenvalue of 0.89 is close to the unit circle. In our case, we prefer preclude this kind of possibilities.

contain explosive roots. In most variables (temperature, radiative forcing of solar irradiance and carbon dioxide), our results confirm previous evidences in the literature.

Given the complex structure and properties of the time series, we analyzed three alternative cases. In the first case, a 5-variable system is considered. The second case studies variables with I(1) and I(2) characteristics and the ones with explosive behaviour separately. The last system considers a transformation of the variables in such a way that the I(1) framework can be used. Overall, all systems show that temperature series is error correcting the disequilibria towards the static steady-state or dynamic steady-state relations. The degree of adjustment of the temperature depends on which cointegrating relation is considered with a range from 5% to 65%. The higher end of this range is comparable to the results found by Stern and Kaufmann (2000). In some cases, however, temperature series is not error correcting for the long-run disequilibria. It means that there is nothing in the system that allows for correcting these disequilibria. On the other hand, the high rate of adjustment with respect to other disequilibria could suggest abrupt changes in temperature in order to correct for these disequilibria.

Another interesting result is the composition of the I(1) and I(2) trends. According to our results, both trends are essentially composed by a linear combination of GHG that are affecting the temperature series strongly. Specifically, we find that the I(1) trend is driven by temperature and radiative forcing of solar irradiance, whereas the I(2) trend is driven by a linear combination of the three GHG or exclusively by the radiative forcing of carbon dioxide. The first component could be associated to inertial and/or natural factors. In the case of the second component, it could be related to human factors.

Finally, we find that the equilibrium temperature change for a doubling of carbon dioxide is between $2.15^{\circ}C$ and $3.40^{\circ}C$, which is agreement with previous literature and the report of the IPCC (2001) using fifteen different general circulation models.

References

- [1] Bloomfield, P., (1992), "Trends in Global Temperature", *Climatic Change* **21**, 1-16.
- [2] Bloomfield, P. and D. Nychka, (1992), "Climate Spectra and Detecting Climate Change", *Climatic Change* **21**, 275-287.
- [3] 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.
- [4] Dickey, D. A. and S. G. Pantula, (1987), "Determining the Order of Differencing in Autoregressive Process", *Journal of Business and Economic Statistics* **5**, 455-461.
- [5] Doornik, D. A. and D. F. Hendry, (2001), *Modelling Dynamic Systems using PcGive 10.0*, Volume II, United Kingdom, Timberlake Consultants Ltd.
- [6] Elliott, G., Rothenberg, T., Stock, J. H., (1996), "Efficient Tests for an Autoregressive Unit Root", *Econometrica* **64**, 813-839.
- [7] Fiess, N. and R. MacDonald, (2001), "The Instability of the Money Demand Function: An I(2) Interpretation", *Oxford Bulletin of Economics and Statistics*, **63** (4), 475-495.
- [8] Fomby, T. and T. Vogelsang, (2002), "Application of Size Robust Trend Analysis to Global Warming Temperature Series", *Journal of Climate* **15**, 117-123.
- [9] Galbraith, J. W. and C. Green, (1993), "Inference about Trend in Global Temperature Data", *Climatic Change* **22**, 209-221.
- [10] Haldrup, N. (1999), "A Review of the Econometric Analysis of I(2) Variables", in Oxley, L. and M. McAleer (eds.), *Practical Issues in Cointegration Analysis*, Blackwell, Oxford.

- [11] Haldrup, N. and P.M. Lildholdt, 2002, On the Robustness of Unit Root Tests in the Presence of Double Unit Roots. *Journal of Time Series Analysis*, **23(2)**, pp. 155-171
- [12] Hansen, H. and K. Juselius, (1995), "CATS in RATS, Manual to Cointegration Analysis of Time Series," Estima, Evanston, Illinois.
- [13] Harvey, A. C., (1989), *Forecasting, Structural Time Series Models, and the Kalman Filter*, Cambridge University Press.
- [14] Hasza, D. P. and W. A. Fuller, (1979), "Estimation of Autoregressive Processes with Unit Roots", *The Annals of Statistics* **7**, 1106-1120.
- [15] Holtemöller, O. (2002), "Money and Prices: An I(2) Analysis for the Euro Area", *Working Paper*, Humboldt-Universität zu Berlin.
- [16] Intergovernmental Panel on Climate Change (2001), "Climate Change 2001: the Intergovernmental Panel on Climate Change Scientific Assessment," in Houghton, J. T., Y. Ding, D.J. Griggs, et al (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. p358.
- [17] Johansen, S., (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control* **12**, 231-254.
- [18] Johansen, S., (1992), "A Representation of Vector Autoregressive Process of Order 2", *Econometric Theory* **8**, 188-202.
- [19] Johansen, S., (1995a), "A Statistical Analysis of Cointegration for I(2) Variables", *Econometric Theory* **11**, 25-59.
- [20] Johansen, S., (1995b), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
- [21] Johansen, S. and K. Juselius, (1992), "Testing Structural Hypothesis in a Multivariate Cointegration Analysis of the PPP and the UIP for UK", *Journal of Econometrics* **53**, 211-244.

- [22] Johansen, S. and K. Juselius, (1994), "Identification of the Long-Run and the Short-Run Structure. An Application to the ISLM Model", *Journal of Econometrics* **63**, 7-36.
- [23] Juselius, K., (1998), "A Structural VAR under Changing Monetary Regimes", *Journal of Business & Economic Statistics* **16**, 400-411.
- [24] Juselius, K., (1999), "Price Convergence in the Medium and Long Run: An I(2) Analysis of Six Price Indices", in Engle, R. F. and H. White (Editors), *Cointegration, Causality and Forecasting (A Festschrift in Honour of Clive W. J. Granger)*, New York: Oxford University Press.
- [25] Juselius, K., (2003), "Inflation, Money Growth, and I(2) Analysis", manuscript.
- [26] Kaufmann, A. and D. I. Stern, (1997), "Evidence for Human Influence on Climate from Hemispheric Temperature Relations", *Nature* **388**, 39-44.
- [27] Kaufmann, A. and D. I. Stern, (2002), "Cointegration Analysis of Hemispheric Temperature Relations", *Journal of Geophysical Research* **107** (D2), 10.1029/2000JD000174.
- [28] Kelly, D. L. (2000), "Unit Roots in the Climate: Is the Recent Warming Due to Persistent Shocks?", *Working Paper*, Department of Economics, University of Miami.
- [29] Kuo, C., C. Lindberg and D. J. Thompson, (1990) "Coherence Established Between Atmospheric Carbon Dioxide and Global Temperature", *Nature* **388**, 39-44.
- [30] Kongsted, H. C., (2003), "An I(2) Cointegration Analysis of Small-Country Import Price Determination", Institute of Economics, University of Copenhagen.
- [31] Lean, J., J. Beer, and R. Bradley, (1995), "Reconstruction of Solar Irradiance Since 1610: Implications for Climate Change. *Geophysical Research letters* **22**, 3195-3198.

- [32] Lenten, L. J. A. and I. A. Moosa, (2003), "An Empirical Investigation into Long-term Climate Change in Australis", *Environmental Modelling & Software* **18**, 59-70.
- [33] Nielsen, B., (2001), "The Asymptotic Distribution of Unit Root Tests of Unstable Autoregressive Processes", *Econometrica* **69**, 211-219.
- [34] Nielsen, B., (2002), "Cointegration Analysis of Explosive Processes", Discussion Paper, Nuffield College, Oxford.
- [35] Paruolo, P., (1996), "On the Determination of Integration Indices in I(2) Systems", *Journal of Econometrics* **72**, 313-356.
- [36] Paruolo, P. and A. Rahbek, (1999), "Weak Exogeneity in I(2) VAR Systems", *Journal of Econometrics* **93**, 281-308.
- [37] Phillips, P. C. B. and P. Perron, (1988), "Testing for a Unit Root in Time Series Regression", *Biometrika* **75**, 335-346.
- [38] Rahbek, A., H. C. Kongsted, and C. Jorgensen, (1999), "Trend Stationarity in the I(2) cointegration model", *Journal of Econometrics* **90**, 265-289.
- [39] Richards, G. R., (1993), "Change in Global Temperature: A Statistical Analysis", *Journal of Climate* **6**, 546-559.
- [40] Said, E. S. and D. A. Dickey, (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order", *Biometrika* **71**, 599-607.
- [41] Santer, B. D., K. E. Taylor, T. M. L. Wigley, T. C. Johns, P. D. Jones, D. J. Karoly, J. F. N. Mitchell, A. H. Oort, J. E. Penner, V. Ramaswamy, M. D. Schwartzkopf, R. J. Stouffer, and S. Tett (1996), "A Search for Human Influences on the Thermal Structure of the Atmosphere," *Nature* **382**, 39-46.
- [42] Schmidt, P. and P. C. B. Phillips, (1992), "LM Tests for a Unit Root in the Presence of Deterministic Trends", *Oxford Bulletin of Economics and Statistics* **54**, 257-287.

- [43] Schönwiese, C. -D., (1994), "Analysis and Prediction of Global Climate Temperature Change based on Multiforced Observational Statistics", *Environmental Pollution* **83**, 149-154.
- [44] Stern, D. I. and R. K. Kaufmann, (1997), "Time Series Properties of Global Climate Variables: Detection and Attribution of Climate Change", working paper in Ecological Economics 9702, Center for Resource and Environmental Studies, Australian National University.
- [45] Stern, D. I. and R. K. Kaufmann, (1999), "Econometric Analysis of Global Climate Change", *Environmental Modelling and Software* **14**, 597-605.
- [46] Stern, D. I. and R. K. Kaufmann, (2000), "Detecting A Global Warming Signal in Hemispheric Temperature Series: A Structural Time Series Analysis", *Climatic Change* **47**, 411-438.
- [47] Stock, J. H. and M. W. Watson, (1988), "Testing for Common Trends", *Journal of the American Statistical Association* **83**, 1097-1107.
- [48] Tol, R. DS. J., (1994), "Greenhouse Statistics-Time Series Analysis: Part II," *Theoretical and Applied Climatoloty* **49**, 63-74
- [49] Tol, R. S. and A. F. de Vos, (1998), "A Bayesian statistical analysis of the enhanced greenhouse effect", *Climatic Change* **38**, 87-112.
- [50] Thomson, D. J., (1995), "The Seasons, Global Temperature, and Precession", *Science* **268**, 59-68.
- [51] Thomson, D. J., (1997), "Dependence of Global Temperatures on Atmospheric CO₂ and Solar Irradiance", *Proceedings of the National Academy of Sciences* **94**, 8370-8377.
- [52] Vostroknutova, E., (2003), "Shock Therapy? An I(2) Cointegration Analysis of the Russian Stabilization", *EUI Working Paper ECO* No. 2003/16, Department of Economics, European University Institute.

- [53] Woodward, W. A. and H. L. Gray, (1993), "Global Warming and the Problem of Testing for Trend in Time Series Data", *Journal of Climate* 6, 953-962.
- [54] Woodward, W. A. and H. L. Gray, (1995), "Selecting a Model for Detecting the Presence of a Trend", *Journal of Climate* 8, 1929-1937.

Table 1. Testing for Cointegrating Ranks
 $Y_t = (temp_t, rfsun_t, rfcO_2t, rfch_4t, rfn_2O_t)'$

H_0	λ_{\max}	Trace	λ_{\max} 90% Critical Values	Trace 90% Critical Values
$r = 0$	83.6	196.2	34.8	82.7
$r = 1$	49.4	112.5	29.1	59.0
$r = 2$	41.9	63.1	23.1	39.1
$r = 3$	14.2	21.3	16.9	23.1
$r = 4$	7.1	7.1	10.5	10.6

Table 2. Testing for Integration Indices
 $Y_t = (temp_t, rfsun_t, rfcO_2t, rfch_4t, rfn_2O_t)'$

r	$H_{r,s}$					Q_r
0	418.0	312.6	252.0	220.8	198.2	196.2
	(198.2)	(167.9)	(142.2)	(119.8)	(101.5)	(87.2)
	(203.2)	(173.4)	(147.1)	(124.4)	(105.6)	(91.2)
1		268.6	171.4	136.8	114.8	112.5
		(137.0)	(113.0)	(92.2)	(75.3)	(62.8)
		(141.5)	(117.4)	(96.5)	(79.0)	(66.1)
2			163.4	91.2	69.0	63.1
			(86.7)	(68.2)	(53.2)	(42.7)
			(90.8)	(71.4)	(55.9)	(45.8)
3				57.9	36.2	21.3
				(47.6)	(34.4)	(25.4)
				(50.7)	(36.8)	(27.9)
4					15.3	7.1
					(19.9)	(12.5)
					(22.2)	(14.2)
$n - r - s = s_2$	5	4	3	2	1	0

Table 3a. Decomposing the system into I(0), I(1) and I(2) spaces; Case 1

$$Y_t = (temp_t, rfsun_t, rfcO_2t, rfch_4t, rfn_2ot)'$$

	$\hat{\beta}_{0,1}$	$\hat{\beta}_{0,2}$	$\hat{\beta}_1$	$\hat{\kappa}_1$	$\hat{\beta}_{1,1}$	$\hat{\beta}_{1,2}$
<i>temp</i>	1.000	1.000	1.000	-79.039	-3.921	-5.698
<i>rfsun</i>	-0.078	-4.639	-1.638	-26.868	-3.006	-1.937
<i>rfco₂</i>	-0.792	-2.161	7.303	-35.985	9.647	-2.594
<i>rfch₄</i>	0.928	5.565	-22.491	-21.236	2.074	-1.531
<i>rfn_{2o}</i>	-36.398	2.774	96.030	-1.872	-0.258	-0.135
<i>trend</i>	0.015	0.004	-0.050			
	$\hat{\alpha}_{0,1}$	$\hat{\alpha}_{0,2}$	$\hat{\alpha}_1$		$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{1,2}$
<i>temp</i>	-0.466	0.067	-0.191		-0.000	-0.001
<i>rfsun</i>	-0.034	0.096	-0.014		-0.001	-0.002
<i>rfco₂</i>	-0.015	0.010	-0.005		0.007	-0.007
<i>rfch₄</i>	-0.003	0.002	-0.002		0.011	0.095
<i>rfn_{2o}</i>	-0.000	0.002	0.000		-0.009	0.125

Table 3b. Decomposing the system into I(0), I(1) and I(2) spaces; Case 2

$$Y_t = (temp_t, rfsun_t, rfcO_2t, rfch_4t, rfn_2ot)'$$

	$\hat{\beta}_0$	$\hat{\beta}_{1,1}$	$\hat{\kappa}_1$	$\hat{\beta}_{1,2}$	$\hat{\kappa}_2$	$\hat{\beta}_{1,2,1}$	$\hat{\beta}_{1,2,2}$
<i>temp</i>	1	1	17.424	1	-20.510	8.404	-5.473
<i>rfsun</i>	-5.931	-0.238	-32.565	-0.959	3.306	7.540	-0.954
<i>rfco₂</i>	-3.127	-5.878	96.770	1.648	-28.565	-9.968	-3.149
<i>rfch₄</i>	6.308	12.513	22.507	-7.305	-8.875	-0.838	-1.444
<i>rfn_{2o}</i>	16.920	-88.710	-2.954	14.819	0.401	0.617	-0.054
<i>trend</i>	0.003	0.049		-0.007			
	$\hat{\alpha}_0$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{1,2}$			$\hat{\alpha}_{1,2,1}$	$\hat{\alpha}_{1,2,2}$
<i>temp</i>	0.100	-0.048	-0.646			0.0001	-0.002
<i>rfsun</i>	0.075	0.002	-0.031			0.001	-0.007
<i>rfco₂</i>	0.009	-0.002	-0.017			-0.018	0.050
<i>rfch₄</i>	0.002	0.000	-0.005			0.000	0.139
<i>rfn_{2o}</i>	0.000	-0.000	0.000			0.065	0.048

Table 4a. Testing for Cointegrating Ranks

$$Y_t = (temp_t, rfsun, rfcO_2t)'$$

H_0	λ_{\max}	Trace	λ_{\max} 90% Critical Values	Trace 90% Critical Values
$r = 0$	36.1	54.5	23.1	39.1
$r = 1$	14.7	18.4	16.9	23.0
$r = 2$	3.7	3.7	10.5	10.6

Table 4b. Testing for Cointegrating Ranks

$$Y_t = (rfch_{4t}, rfn_{2o_t})'$$

H_0	λ_{\max}	Trace	λ_{\max} 90% Critical Values	Trace 90% Critical Values
$r = 0$	43.3	52.2	16.9	23.0
$r = 1$	8.9	8.9	10.5	10.6

Table 5a. Testing for Integration Indices

$$Y_t = (\text{temp}_t, rfsun, rfcO_2t)'$$

r	$H_{r,s}$		Q_r	
0	213.5	114.9	58.0	54.5
	(86.7)	(68.2)	(53.2)	(42.7)
1		102.0	21.6	18.4
		(47.6)	(34.4)	(25.4)
2			6.4	3.7
			(19.9)	(12.5)
$n - r - s = s_2$	3	2	1	0

Table 5b. Testing for Integration Indices;

$$Y_t = (rfch_{4t}, rfn_{2O_t})'$$

r	$H_{r,s}$		Q_r	
0	77.3	53.2	52.2	
	(47.6)	(34.4)	(25.4)	
1		15.6	8.9	
		(19.9)	(12.5)	
$n - r - s = s_2$	2	1	0	

Table 6. Decomposing the systems into I(0), I(1) and I(2) spaces

$$Y_t = (temp_t, rfsun_t, rfcO_2t)'$$

	$\hat{\beta}_1$	$\hat{\kappa}_1$	$\hat{\beta}_{\perp 1}$	$\hat{\beta}_{\perp 2}$
<i>temp</i>	1.000	-0.819	10.505	2.606
<i>rfsun</i>	-3.341	-0.085	4.396	0.269
<i>rfco₂</i>	-0.500	-1.073	-8.366	3.414
<i>trend</i>	0.006			
	$\hat{\alpha}_1$	$\hat{\alpha}_{\perp 1}$	$\hat{\alpha}_{\perp 2}$	
<i>temp</i>	-0.126	0.046	0.000	
<i>rfsun</i>	0.069	0.085	-0.001	
<i>rfco₂</i>	-0.001	-0.001	-0.096	

Table 7. Testing for Cointegrating Ranks

$$Y_t = (temp_t, rfsun_t, \Delta rfcO_2t, \Delta_\lambda rfcH_4t, \Delta_\lambda rfn_2O_t)'$$

H_0	λ_{\max}	Trace	λ_{\max} 90% Critical Values	Trace 90% Critical Values
$r = 0$	57.2	131.4	34.8	82.7
$r = 1$	42.4	74.3	29.1	59.0
$r = 2$	17.0	31.9	23.1	39.1
$r = 3$	9.6	14.8	16.9	23.0
$r = 4$	5.2	5.2	10.5	10.6

Table 8. Testing for Integration Indices

$$Y_t = (temp_t, rfsun_t, \Delta rfcO_{2t}, \Delta_{\lambda} rfcH_{4t}, \Delta_{\lambda} rfn_{2O_t})'$$

r	$H_{r,s}$					Q_r
0	506.9 (198.2)	384.1 (167.9)	302.3 (142.2)	224.8 (119.8)	163.1 (101.5)	131.4 (87.2)
1		373.1 (137.0)	254.4 (113.0)	173.4 (92.2)	106.9 (75.3)	74.3 (62.8)
2			239.8 (86.7)	155.3 (68.2)	75.8 (53.2)	31.9 (42.7)
3				130.6 (47.6)	52.3 (34.4)	14.8 (25.4)
4					23.2 (19.9)	5.2 (12.5)
$n-r-s=s_2$	5	4	3	2	1	0

Table 9. Decomposing the system into I(0), I(1) and I(2) spaces

$$Y_t = (temp_t, rfsun_t, \Delta rfcO_{2t}, \Delta_{\lambda} rfcH_{4t}, \Delta_{\lambda} rfn_{2O_t})'$$

	$\hat{\beta}_{0,1}$	$\hat{\beta}_{0,2}$	$\hat{\beta}_{\perp 1,1}$	$\hat{\beta}_{\perp 1,2}$	$\hat{\beta}_{\perp 1,3}$
<i>temp</i>	1.000	1.000	-0.016	0.335	0.002
<i>rfsun</i>	0.000	-4.000	0.000	0.086	0.007
$\Delta rfcO_2$	-25.133	-0.500	-0.000	0.004	-0.004
$\Delta_{\lambda} rfcH_4$	-31.552	40.057	0.001	-0.000	0.000
$\Delta_{\lambda} rfn_{2O}$	1504.442	-196.466	0.000	-0.000	-0.000
<i>trend</i>	0.006	0.005			
	$\hat{\alpha}_{0,1}$	$\hat{\alpha}_{0,2}$	$\hat{\alpha}_{\perp 1,1}$	$\hat{\alpha}_{\perp 1,2}$	$\hat{\alpha}_{\perp 1,3}$
<i>temp</i>	-0.108	-0.123	0.001	-0.006	-0.008
<i>rfsun</i>	-0.000	0.060	-0.006	-0.015	-0.019
$\Delta rfcO_2$	-0.001	0.000	0.115	-0.003	0.993
$\Delta_{\lambda} rfcH_4$	0.000	0.001	0.993	-0.009	-0.115
$\Delta_{\lambda} rfn_{2O}$	-0.001	0.000	0.009	0.999	0.001

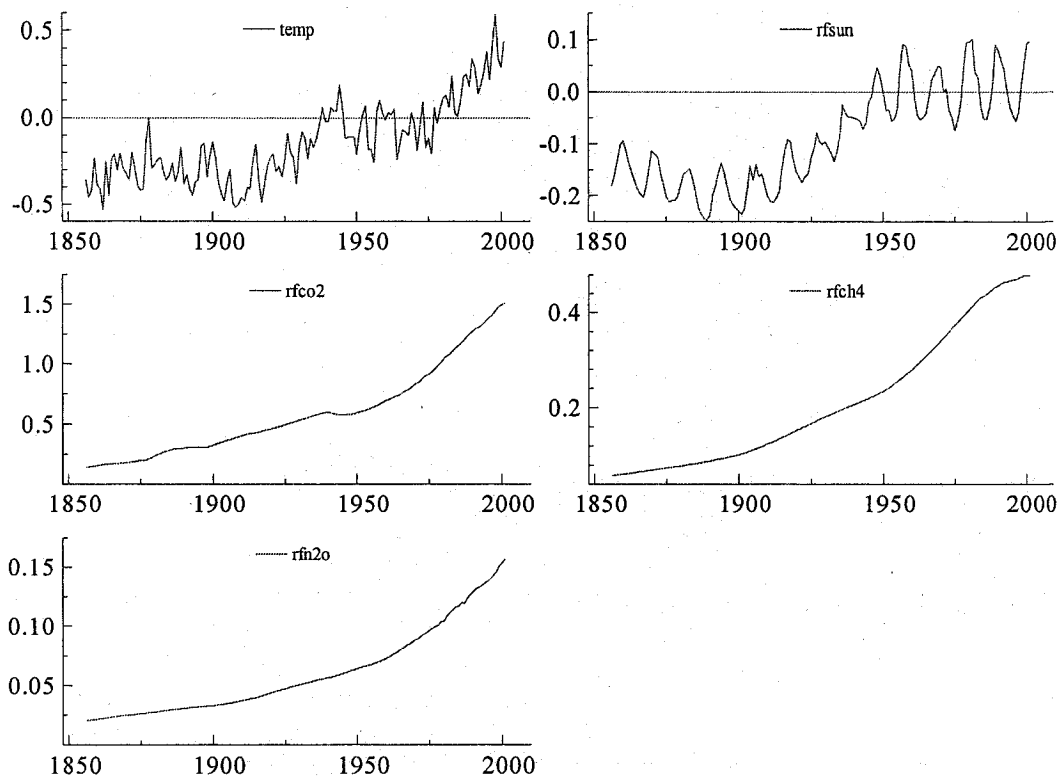


Figure 1. Global temperature deviations (in $^{\circ}C$, radiative forcing (in Wm^{-2}) of solar irradiance, GHG; 1856-2001.

Chapter 2

Unit Root Tests and Structural Change when the Initial Observation is Drawn from its Unconditional Distribution

May 7, 2006

1 Introduction

Since the seminal work of Nelson and Plosser (1982), large literature has contributed to the debate of whether macroeconomic time series are trend-stationary or difference-stationary. Perron (1989, 1997), Zivot and Andrews (1992), Christiano (1992), Banerjee, Lundsane and Stock (1992) contributed to the debate by introducing an unknown structural break into the model and proposed methods to estimate the break point. ERS (1996) derived a family of point optimal tests using GLS-detrending approach, which improved the power of unit root tests substantially. Perron and Rodríguez (2003) extended their approach to the context of unknown structural change models, and derived the statistics for point optimal tests, M-tests (Stock 1999) and ADF test (Dickey and Fuller 1979; Said and Dickey 1984).

However, few paper has studied the impact of initial observation on unit root tests. Traditionally, theoretical work assume the starting value of time series is zero or has finite expectations (fixed initial condition), therefore the effect of initial observation disappears asymptotically. Until now only two trials addressed the influence of initial condition explicitly. Elliott (1999) introduced a random initial condition by assuming the initial observation is drawn from its unconditional distribution under the alternative hypothesis, and he found loss of power in both large and small samples. Müller and Elliott (2003) extended this result by generalizing the initial condition and derived a family of point optimal tests over a weighting function of different initial conditions.

But no paper has explored the impact of random initial condition in the context of structural change models. This paper intends to fill this gap. Following Perron and Rodríguez (2003) and Elliott (1999), this experiment introduces both an unknown structural break and a random initial condition under the alternative hypothesis. Simulation results will show that the power of unit root tests discussed have decreased in both large and small samples.

The rest of this paper is organized as follows. Section 2 defines the model, derives the asymptotic theory and calculates the asymptotic results. Section 3 presents the finite sample results and evaluation. Section 4 gives an empirical application and Section 5 concludes. All proofs are presented in the Appendix.

2 The Model and Asymptotic Theory

2.1 The model

For the ease of comparison, this paper considers the same model and statistics as those in Perron and Rodríguez (2003), but with a different assumption on the initial observation. Hence the data generating process (DGP) is:

$$y_t = d_t + u_t \quad (1)$$

$$u_t = \alpha u_{t-1} + v_t \quad (2)$$

$$\alpha = 1 + cT^{-1} \quad (3)$$

Instead of assuming $u_0 = 0$, we adopt the following assumption from Elliott (1999):

Condition A (Initial condition assumption). *We assume that u_0 is zero when $\alpha = 1$, so $u_1 = v_1$, while u_1 has mean zero and variance $\sigma^2/(1 - \alpha^2)$ when $\alpha < 1$.*

Under this assumption, the initial observation does not disappear at the convergence rate of $T^{1/2}$ under the alternative hypothesis (see Lemma A.1 in Appendix for proof). Presumably, this will have an impact on the limiting distribution and hence the power performance of the test statistics. In the following section, we will use ξ in the limiting distributions to represent the effect of initial condition.

$d_t = \psi' z_t$ is the deterministic component. Two types of models are considered using different specifications on d_t . Model I includes an intercept, a time trend, a break in slope in the trend function and $z_t = \{1, t, 1(t \geq T_B)(t - T_B)\}$ where $\mathbf{1}(\cdot)$ is an indicator function, T_B is the date of the break point. A set of estimates of coefficients is denoted as $\hat{\varphi} = [\hat{\mu}_1, \hat{\beta}_1, \hat{\beta}_2]'$. Model II includes an intercept, a time trend, a break in both intercept and slope, therefore we have $z_t = \{1, 1(t \geq T_B), t, 1(t \geq T_B)(t - T_B)\}$. Accordingly, a set of estimate of coefficients is denoted as $\hat{\varphi} = [\hat{\mu}_1, \hat{\mu}_2, \hat{\beta}_1, \hat{\beta}_2]'$. Equation (3) is the local to unity framework established in Phillips (1987), Chan and Wei (1987), and it functions as a bridge connecting the stationary and non-stationary time series. When $c = 0$ we are under the null hypothesis of a random walk, while when $c < 0$ we are under the alternative hypothesis of stationary time series.

For the innovations $\{v_t\}$, we have the following assumptions:

Condition B (see Elliott, 1999; and Davidson, 1994; for a general treatment). The innovations $\{v_t\}$ satisfy (a) $E[v_t] = 0$ and $\hat{\gamma}_v(j) \xrightarrow{P} \gamma_v(j)$, for fixed j , where $\gamma_v(j) = E[v_t v_{t-j}]$ and $\hat{\gamma}_v(j) = T^{-1} \sum_{t=j+1}^T v_t v_{t-j}$; (b)

the functional central limit theorem can be applied to the partial sums $S_t = \sum_{j=1}^t v_j$ and $T^{-1/2}S_{[Tr]} \Rightarrow N(0, \sigma^2 r) \equiv \sigma W(r)$, where \Rightarrow signifies weak convergence, $W(r)$ is a standard Wiener process on the interval $[0, 1]$, $\sigma^2 = \lim_{T \rightarrow \infty} T^{-1}E(S_T^2)$ is the non-normalized spectral density at frequency zero¹, it is a finite positive number such that $\sigma^2 = \sum_{j=-\infty}^{\infty} \gamma_v(j)^2$.

2.2 Asymptotic Theory Assuming Known Break Point

To derive the asymptotic distributions using GLS detrending approach, we first transform the data using $y_t^{\bar{\alpha}} = \left[(1 - \bar{\alpha}^2)^{1/2} y_1, (1 - \bar{\alpha}L)y_t \right]$, $z_t^{\bar{\alpha}} = \left[(1 - \bar{\alpha}^2)^{1/2} z_1, (1 - \bar{\alpha}L)z_t \right]$, and $u_t^{\bar{\alpha}} = \left[(1 - \bar{\alpha}^2)^{1/2} u_1, (1 - \bar{\alpha}L)u_t \right]$, for $t = 1, \dots, T$, and $\bar{\alpha} = 1 + \bar{c}T^{-1}$. Notice that the treatment of the first observation changes compared to the case analyzed by ERS (1996), Perron and Rodríguez (2003). Secondly we estimate the coefficients of the deterministic component using OLS regression:

$$y_t^{\bar{\alpha}} = \hat{\psi}' z_t^{\bar{\alpha}} + u_t^{\bar{\alpha}} \quad (4)$$

The limiting distribution of $\Lambda^{-1}(\hat{\psi} - \psi)$ with a scaling matrix $\Lambda = \text{diag}(T^{1/2}, T^{-1/2}, T^{-1/2})$ is derived in Lemma A.2 and A.3 in the Appendix. Then we arrive at quasi-differenced $\{y_t\}$:

$$\tilde{y}_t = y_t - \hat{\psi}' z_t \quad (5)$$

Using the above detrended data, the asymptotic distributions of three test statistics are derived. The first one is the widely used ADF statistic (Dickey and Fuller, 1979; Said and Dickey, 1984), which tests whether $\hat{\alpha}_0 = 0$ in the following augmented regression:

$$\Delta \tilde{y}_t = \hat{\alpha}_0 \tilde{y}_{t-1} + \sum_{j=1}^{k-1} \hat{b}_j \Delta \tilde{y}_{t-j} + \hat{\epsilon}_{tk} \quad (6)$$

¹According to Elliott (1999), the process v_t is a potentially serial correlated stationary process. Under these assumptions with v_t Gaussian and further the initial value $u_0 = 0$ under both the null and alternative hypothesis, ERS (1996) showed that no uniformly most powerful test exists and derived the asymptotic power functions for the most powerful test against a sequence of local alternatives of $\alpha < 1$. Perron and Rodríguez (2003) proceeded with the same way. We will refer to this as the fixed initial condition case.

²A number of consistent estimators of σ^2 are available. See Stock (1994) for a discussion and review.

where the lagged first differences are used to account for the serial correlation in the time series.

The second group of statistics are based on M-tests proposed by Stock (1999), further analyzed by Perron and Ng (1996). This class of tests include a modified version of Phillips-Perron's (1988) Z_α test; a modified statistic developed by Bhargava (1986), Sargan and Bhargava (1983); and a modified Phillips-Perron's (1988) Z_{t_α} tests. Using the detrended \tilde{y}_t series, the M-test statistics can be written as:

$$MZ_{\alpha\xi}^{GLS} = (\tilde{y}_T^2 - T^{-1}s^2)(2T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2)^{-1}, \quad (7)$$

$$MSB_\xi^{GLS} = (T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2 / s^2)^{1/2}, \quad (8)$$

$$MZ_{t\xi}^{GLS} = (T^{-1}\tilde{y}_T^2 - s^2)(4s^2T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2)^{-1/2}. \quad (9)$$

Perron and Ng (1996) showed that the main advantage of M-tests is that they have less size distortions when the error term contains negative moving average dynamics and in other cases acceptable size distortions.

The third statistic is based on a family of point optimal tests developed by ERS (1996), in which they showed that no uniformly optimal tests exist for unit root testing. Based on Dufour and King (1991), they developed a feasible point optimal test, which has a power function tangent to the power envelope at one point of the alternative hypothesis. The statistic is defined as

$$P_T^{GLS} = \{S(\bar{\alpha}) - \bar{\alpha}S(1)\} / s^2 \quad (10)$$

where $S(\bar{\alpha}) = u^{\bar{\alpha}'}u^{\bar{\alpha}}$ and $S(1) = u^1u^1$ are the squared sum of residuals under the alternative and the null hypothesis respectively. s^2 is a consistent estimate of σ^2 , the spectral density at frequency zero. Perron and Rodríguez (2003) extended P_T^{GLS} test to the case of an unknown structural break and zero initial value. Here we adopt random initial condition and denote the resulting test statistics as $P_{T\xi}^{GLS}$, where ξ represents the effect of random initial condition.

To calculate the limiting distributions of M and P-tests, an estimator of the spectral density at frequency zero is required. We use the following

consistent autoregressive estimate s^2 proposed by Perron and Ng (1998):

$$s^2 = \hat{s}_{ek}^2 / [1 - \hat{b}(1)]^2,$$

where $\hat{s}_{ek}^2 = (T-k)^{-1} \sum_{t=k+1}^T \hat{e}_{tk}^2$, $\hat{b}(1) = \sum_{j=1}^k \hat{b}_j$, $\{\hat{e}_{tk}\}$ and \hat{b}_j are obtained from the augmented regression (6).

We first assume a known break point at T_B and $\delta = T_B/T$ (this assumption will be relaxed in Section 2.3, the following theorem 1 and 2 state the limiting distributions of the unit root statistics described above.

Theorem 1 *Let $\{y_t\}$ be generated by (1) to (3), the initial value is given by Condition A, $\{v_t\}$ satisfies Condition B, GLS-detrending is applied using $\bar{\alpha} = 1 + \bar{c}T^{-1}$, $\delta = T_B/T$ is the break point, s^2 is a consistent estimate of σ^2 , the spectral density at the frequency zero. Then the $M_{\alpha\xi}^{GLS}$ and ADF_{ξ}^{GLS} statistics for Model I and II have the following limiting distributions:*

$$\begin{aligned} MZ_{\alpha\xi}^{GLS}(\delta) &\Rightarrow \frac{0.5g_1(c, \bar{c}, \delta)}{g_2(c, \bar{c}, \delta)} \equiv J^{MZ_{\alpha\xi}^{GLS}}(c, \bar{c}, \delta) \\ MSB_{\xi}^{GLS}(\delta) &\Rightarrow (g_2(c, \bar{c}, \delta))^{1/2} \equiv J^{MSB_{\xi}^{GLS}}(c, \bar{c}, \delta) \\ MZ_{t\xi}^{GLS}(\delta) &\Rightarrow \frac{0.5g_1(c, \bar{c}, \delta)}{(g_2(c, \bar{c}, \delta))^{1/2}} \equiv J^{MZ_{t\xi}^{GLS}}(c, \bar{c}, \delta) \\ ADF_{\xi}^{GLS}(\delta) &\Rightarrow \frac{0.5g_1(c, \bar{c}, \delta)}{(g_2(c, \bar{c}, \delta))^{1/2}} \equiv J^{ADF_{\xi}^{GLS}}(c, \bar{c}, \delta) \end{aligned}$$

where $g_1(c, \bar{c}, \delta) = V_{\bar{c}\bar{c}}^{(1)}(1, \delta)^2 - 2V_{\bar{c}\bar{c}}^{(2)}(1, \delta) - 1$, $g_2(c, \bar{c}, \delta) = \int_0^1 V_{\bar{c}\bar{c}}^{(1)}(r, \delta) dr - 2 \int_0^1 V_{\bar{c}\bar{c}}^{(2)}(r, \delta) dr$, $V_{\bar{c}\bar{c}}^{(1)}(r, \delta) = W_c(r) - b_4 - r b_5$, $V_{\bar{c}\bar{c}}^{(2)}(r, \delta) = b_6(r - \delta)[W_c(r) - r b_5 - b_4 - (1/2)(r - \delta)b_6]$. The elements b_4, b_5, b_6 can be calculated using

$$\begin{aligned} &\begin{bmatrix} \bar{c}^2 - 2\bar{c} & \frac{1}{2}\bar{c}^2 - \bar{c} & -\bar{c}(1 - \delta) + \frac{1}{2}\bar{c}^2(1 - \delta)^2 \\ \frac{1}{2}\bar{c}^2 - \bar{c} & 1 + \frac{1}{3}\bar{c}^2 - \bar{c} & m \\ -\bar{c}(1 - \delta) + \frac{1}{2}\bar{c}^2(1 - \delta)^2 & m & d \end{bmatrix} \\ &\times \begin{pmatrix} \sigma b_1 \\ \sigma b_2 \\ \sigma b_3 \end{pmatrix} \\ &= \sigma \begin{pmatrix} b_4 \\ b_5 \\ b_6 \end{pmatrix} \end{aligned}$$

where $b_1 = -2\bar{c}\xi - \bar{c}(c - \bar{c}) \int_0^1 W_c(r) - \bar{c}W(1)$, $b_2 = (1 - \bar{c})W(1) + c \int_0^1 W_c(r) - \bar{c}(c - \bar{c}) \int_0^1 r W_c(r)$, $b_3 = (1 - \bar{c} + \delta\bar{c})W(1) - W(\delta) - \bar{c}(c - \bar{c}) \int_0^1 r W_c(r) +$

$\delta\bar{c}(c - \bar{c}) \int_{\delta}^1 W_c(r) + c \int_{\delta}^1 W_c(r)$, $m = 1 - \delta - \bar{c} + \bar{c}\delta - \frac{1}{2}\bar{c}^2\delta + \frac{1}{3}\bar{c}^2 + \frac{1}{6}\bar{c}^2\delta^3$, $d = 1 - \delta - \bar{c}(1 - \delta)^2 + \frac{1}{3}\bar{c}^2(1 - \delta)^3$, and $\xi \sim N(0, \frac{1}{2c})$ under the alternative hypothesis, with $W_c(r)$ the Ornstein-Uhlenbeck process and satisfies $dW_c(r) = cW_c(r)dr + dW(r)$ with $W_c(0) = 0$

Theorem 2 *Let $\{y_t\}$ be generated by (1) to (3), the initial value is given by Condition A, $\{v_t\}$ satisfies Condition B. GLS-detrending is applied using $\bar{\alpha} = 1 + \bar{c}T^{-1}$ to (5), $\delta = T_B/T$ is the break point, s^2 is a consistent estimate of σ^2 , the spectral density at frequency zero. Then the $P_{T\xi}^{GLS}$ test for Model I and II have the following limiting distribution*

$$\begin{aligned} P_{T\xi}^{GLS}(c, \bar{c}, \delta) &\Rightarrow -2\bar{c}\xi^2 - 2\bar{c} \int_0^1 W_c(r)dW(r) + (\bar{c}^2 - 2\bar{c}c) \int_0^1 W_c^2(r)dr \\ &\quad + \bar{M}(c, 0, \delta) - \bar{M}(c, \bar{c}, \delta) - \bar{c} \\ &\equiv J_{T\xi}^{PGLS}(c, \bar{c}, \delta) \end{aligned}$$

where $\bar{M}(c, \bar{c}, \delta) = \bar{A}(c, \bar{c}, \delta)' \bar{B}(\bar{c}, \delta) \bar{A}(c, \bar{c}, \delta)$, and $\bar{A}(c, \bar{c}, \delta)$, $\bar{B}(\bar{c}, \delta)$ are defined in the Appendix. Unlike Perron and Rodríguez (2003), the term ξ , representing the effect of random initial condition, enters the limiting distribution of the statistics.

2.3 The Selection of Break Point

In order to circumvent the data mining problem caused by pre-examination of the time series, two data-dependent methods estimating the break point endogenously are applied. One is the infimum method proposed by Zivot and Andrews (1992). The other one is the supremum method by Perron (1997).

The first method is to select the break point that gives the strongest rejection against the null hypothesis of $\alpha = 1$. Since among the statistics studied here a smaller one indicates a rejection of the null hypothesis, the break point δ^* can be selected using

$$J(c, \bar{c}) = \inf_{\delta \in [0,1]} J(c, \bar{c}, \delta)$$

where J represents the asymptotic distributions of the M_{ξ}^{GLS} , and the ADF_{ξ}^{GLS} statistics derived in Theorem 1. Using this method, the selection of δ^* for $P_{T\xi}^{GLS}$ is slightly different. According to Perron and Rodríguez

(2003), the break point can be selected using criterion:

$$P_{T\xi}^{GLS}(c, \bar{c}) = \left\{ \inf_{\delta \in [\epsilon, 1-\epsilon]} S(\bar{\alpha}, \delta) - \inf_{\delta \in [\epsilon, 1-\epsilon]} \bar{\alpha} S(1, \delta) \right\} / s^2 \quad (12)$$

where a truncation ϵ is needed for critical values to be bounded and $\epsilon = 0.15$ is used throughout the paper as standard literature suggested. Perron and Rodríguez recognized that when the break point is unknown “things are different. The principle is, however, still the same” (page 8, Perron and Rodríguez, 2003). Their argument, which was not clearly explained, was concerned with the fact that the infimum of the GLS squared sum of residuals are constructed in the same way as ERS (1996) suggested. Notice that this criterion may be taken as not optimal if we consider that there does not exist a break point under the null hypothesis. The reason is clear because we have an unidentified parameter under the null hypothesis (see Andrews, 1993). In this case we should take the Sup, Mean or Exp (see Andrews and Ploberger, 1994) of the likelihood ratio statistic for a specific break point. However, if we consider the existence of a break point under the null hypothesis, the procedure used by Perron and Rodríguez (2003) appears to be optimal because it is invariant with respect to all nuisance parameters of the model, including the break point. In other words, the minimization of the GLS square sum of residuals is performed taking into account of all corresponding parameters of the model³.

Applying (12) to Theorem 2, $P_{T\xi}^{GLS}$ test has the following limiting distribution:

$$\begin{aligned} P_{T\xi}^{GLS}(c, \bar{c}) &\Rightarrow -2\bar{c} \int_0^1 W_c(r) dW(r) + (\bar{c}^2 - 2\bar{c}c) \int_0^1 W_c^2(r) dr \quad (13) \\ &\quad + \sup \bar{M}(c, 0, \delta) - \sup \bar{M}(c, \bar{c}, \delta) - \bar{c} - 2\bar{c}\xi^2 \\ &\equiv J_{T\xi}^{PGLS}(c, \bar{c}). \end{aligned}$$

The second method of estimating break point is proposed by Perron (1997), where he recommended to choose the δ^* related to the largest absolute value of the t -statistic associated with the parameter of broken time trend. After selecting δ^* , we calculate the corresponding $M_\xi^{GLS}(\delta^*)$, and $ADF_\xi^{GLS}(\delta^*)$ statistics. However, there is no criterion available for feasible point optimal test when using supremum method to choose the break point.

³We thank important discussions with Pierre Perron regarding the adequacy of the constructed power envelope in Perron and Rodríguez (2003). Our conclusion is that, even in the case where their power envelope is not completely right, the correct or “optimal” power envelope should be very close, and the statistics proposed have power very close to this power envelope. In practical terms, there are no significant problems.

2.4 Asymptotic Results

Under the null hypothesis of $c = 0$, we use $T = 1000, 10,000$ replications to simulate asymptotic critical values when $\bar{c} = -1$ to -70 ($\bar{\alpha} = 0.999$ to 0.93). Then we let $c = \bar{c}$ to calculate maximal power at each c and graph the power envelope. To choose the \bar{c} for the *GLS* detrending of $\{y_t\}$, we select the one that achieves 50% of the maximum power as suggested by ERS (1996), and $\bar{c} = -24$. Intuitively, a lower \bar{c} here (compared to the $\bar{c} = -22.5$ in Perron and Rodríguez, 2003) indicates that it may take longer to reach the same level of power. In other words, the power envelope shifts down from the previous one calculated in Perron and Rodríguez (2003), and the loss of power is caused by the relaxation of assumption on the initial observation. Power envelopes for both cases are graphed in Figure 1.

Next, we use the critical values obtained when $\bar{c} = -24$, $T = 1000$ and $10,000$ replications to calculate the asymptotic power function for each statistic and the results are graphed in Figure 1 and 2. It is obvious that for each test, the curve of power function lies under the power envelope, but not far from it. Using infimum method to choose break point sometimes gives a slightly higher power function than supremum method. The results from Perron and Rodríguez (2003) are also graphed in the same figure as a comparison. We can see that the power of each test has dropped due to the change of initial condition.

3 Small Sample Monte Carlo Evidence

In practice, many time series have small sample size and asymptotic distribution is not a good approximation. Therefore it is necessary to simulate finite sample critical values and to evaluate the performance in terms of size and power. When doing this, there is always a question of how many lags should be chosen to account for the serial correlation and to maintain certain power. Now it is generally agreed that data-dependent methods bring a better test performance than choosing lag length k a priori (see Ng and Perron, 1995). In the following, we use four information criteria and a recursive method (Campbell and Perron, 1991) to choose k .

3.1 The Selection of k

Akaike and Bayesian Information Criteria (AIC and BIC hereafter) are two basic methods to choose the number of lags. They take the form of $AIC(k) = \arg \min_{k \in [0, k_{\max}]} \{\ln(\hat{s}_{ek}^2) + \frac{2k}{T^*}\}$ and $BIC(k) = \arg \min_{k \in [0, k_{\max}]} \{\ln(\hat{s}_{ek}^2) +$

$\frac{\ln(T^*)k}{T^*}$ }, where $T^* = T - k_{\max}$, $\hat{s}_{ek}^2 = (T - k_{\max})^{-1} \sum_{t=k_{\max}+1}^T \hat{e}_{tk}^2$. k_{\max} should be large enough to account for the potential serial correlations in the time series, and $k_{\max} = \text{int}[12 \times (T/100)^{1/4}]$ as recommended by Ng and Perron (2001). The shortcoming of these information criteria is that when there are strong negative moving average (MA hereafter) components in the error term, which is quite often in macroeconomic data, they tend to select a smaller k than necessary for unit root tests to have a decent size.

To account for the size distortions when there is strong negative MA errors, Ng and Perron (2001) proposed a modified version of AIC and BIC (MAIC and MBIC) by including a penalty factor $\hat{\tau}_T(k) = (\hat{s}_{ek}^2)^{-1} \hat{\alpha}_0^2 \sum_{t=k_{\max}+1}^T \hat{y}_{t-1}^2$, where $\hat{\alpha}_0$ can be estimated using the augmented autoregression (6). Ng and Perron (2001) have shown that when a strong negative MA error exists, $\hat{\tau}_T(k)$ increases as k decreases. Therefore, $\hat{\tau}_T(k)$ can be used as a penalty factor to correct underfitting, and MAIC, MBIC are of the form: $\text{MAIC}(k) = \arg \min_{k \in [0, k_{\max}]} \{\ln(\hat{s}_{ek}^2) + 2(\hat{\tau}_T(k) + k)/T^*\}$ $\text{MBIC}(k) = \arg \min_{k \in [0, k_{\max}]} \{\ln(\hat{s}_{ek}^2) + \ln T^*(\hat{\tau}_T(k) + k)/T^*\}$.

The last method is a sequential t -test or recursive method (t-sig hereafter) proposed by Campbell and Perron (1991). To apply this method, we start from $k_{\max} = \text{int}[(4 \times (T/100)^{1/4}]$ when estimating the augmented regression (6). If the t -statistic associated with the k_{\max} th lag is significant (p -value less than 0.1), then $k = k_{\max}$ is chosen. Otherwise we redo the regression with $k_{\max} - 1$ lags, and so on, until we find the lag that has a significant t -statistic. Note that if $k = 0$ and no rejection is found, we will select $k = 0$. The advantage of this method is that it has less size distortion than AIC and BIC when there is a strong negative MA error.

3.2 Size and Power in Finite Samples

The performance of AIC is poor and hence, not included. Furthermore, in order to save space, the results for MBIC are not included either⁴. The critical values for model I and II, using k selected by BIC, MAIC and t-sig, 1000 replications, are tabulated in Table 1 to 4. Table 1 and 2 calculate the critical values using infimum method to choose break point, whereas Table 3 and 4 present the critical values using supremum method. Based on these critical values, we simulate finite sample size and power when $T = 100$, and 200. We also include dynamics in the error term v_t , which follows an AR(1) and an MA(1) process, i.e., $v_t = \rho v_{t-1} + e_t$ and $v_t = e_t + \theta e_{t-1}$, where $e_t \sim i.i.d. N(0, 1)$. The simulation results are tabulated in Table 5 to 8. We

⁴All these results are available upon request.

summarize the following two main features from these results:

First, when comparing different methods of selecting k , MAIC has much acceptable size distortions than BIC when there is a strong negative MA(1) error. For example, when $\theta = -0.8$, $T = 200$ in Table 6, the size distortions using MAIC are 0.140, 0.137, 0.140, 0.167 and 0.103 respectively; While when using BIC, they are 0.874, 0.873, 0.871, 0.911, and 0.849 respectively. Second, when comparing different statistics, all tests except the ADF_{ξ}^{GLS} statistic have low power when there is a strong negative AR(1) error. For example, when $\rho = -0.8$, $T = 200$ in Table 6 and 8, the power for M_{ξ}^{GLS} and $P_{T\xi}^{GLS}$ tests are at most 5%, while the power for ADF_{ξ}^{GLS} test is about 50%.

4 Empirical Applications

Two time series from the Nelson-Plosser data set are examined: real wages (1900-1970) and common stock prices (1870-1970). A common characteristic that appears by pre-observing the data is that they both exhibit a change in level and slope (see Figure 3 and 4). Therefore, model II is selected to test whether the null hypothesis of a unit root will be rejected or not. The test results using information criteria to select lag length k are tabulated in Table 9a,b. We find that the date of break point selected are the same as those in Perron and Rodríguez (2003), and they are associated with major events. The fitted real wage series with a break in 1938 is graphed in Figure 3, and the fitted stock price series with a break in 1931 is graphed in Figure 4.

When using information criteria BIC and MAIC, most test statistics are able to reject the null hypothesis of a unit root at least at 5% significance level. But comparing the results with those of Perron and Rodríguez (2003), there exists less evidences in favor of the rejection of the null hypothesis of a unit root. For example, for real wages, it is rejected at 1% significance level in Perron and Rodríguez (2003) but at 2.5% in our case when using supremum method and MAIC for the MZ_t statistic. When using infimum method and ADF statistic for the real wages, it is rejected at 5% using BIC and 2.5% using MAIC in Perron and Rodríguez (2003). In the present study, it is rejected at 10.0% when using BIC and 5.0% when using MAIC. These evidence indicates that the power of unit root tests may have decreased due to the new assumption on the initial value.

The results using t-sig method are summarized in Table 10a,b. When using infimum method, the null hypothesis of a unit root is rejected in all

cases at least at 10.0%. However, when using supremum method, the null is not rejected for real wage time series.

To summarize, the empirical evidence suggests a rejection of the null hypothesis of a unit root but with less evidence in comparison with what has been found in Perron and Rodríguez (2003).

5 Concluding Remarks

Changing the initial condition in the DGP has different impact under the null and alternative hypothesis in unit root testing. Under the null, the change of initial value is equivalent to a mean shift. Therefore invariance principle can be applied for various initial conditions and the test statistics remain the same. But this proposition is not true under the alternative hypothesis, where invariant tests will have a different distribution as the initial condition changes, and hence, impacts on power performance are expected.

This paper examines M_{ξ}^{GLS} , ADF_{ξ}^{GLS} , $P_{T\xi}^{GLS}$ tests in the context of an unknown structural change when the initial observation is drawn from its unconditional distribution, in comparison with zero or fixed initial values as dealt by Perron and Rodríguez (2003). As a result, we find loss of power in both large and small samples. Therefore, it is recommended to assume the initial observation is drawn from an unconditional distribution in order to be more conservative. The finite sample size and power performance are also studied when a variety of dynamics are included in the error term. The performance of the tests are quite different when different criteria are used to select the order of autoregressions, but what we observe is consistent with what standard literature predicts.

References

- [1] Andrews, D.W.K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* **61**, 821-856.
- [2] Andrews, D. W. K., and W. Ploberger (1994), "Optimal Tests when a Nuisance Parameter is Present only under the Alternative," *Econometrica* **62**, 1383-1414.
- [3] Bhargava, A. (1986), "On the Theory of Testing for Unit Root in Observed Time Series," *Review of Economic Studies* **53**, 369-384.
- [4] Banerjee, A., R. Lumsdaine and J. H. Stock (1992), "Recursive and Sequential Tests of the Unit Root and Trend Break Hypothesis: Theory and International Evidence," *Journal of Business and Economics Statistics* **10**, 271-287.
- [5] Campbell, J. Y. and P. Perron (1991), "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots" in Blanchard, O.J., Fischer, S. eds., *NBER Macroeconomics Annual* **6**, 141-201.
- [6] Chan, N. H. and C. Z. Wei (1987), "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *Annals of Statistics* **15**, 1050-63.
- [7] Christiano, L. J. (1992), "Searching for a Break in GNP," *Journal of Business and Economics Statistics* **10**, 237-250.
- [8] Davidson J. (1994), "Stochastic Limit Theory," Oxford University Press.
- [9] Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* **74**, 427-431.
- [10] Dufour, J. -M. and King, M. (1991), "Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary Errors," *Journal of Econometrics* **47**, 115-143.
- [11] Elliott, G., T. Rothenberg and J. H. Stock (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica* **64**, 813-839.
- [12] Elliott, G. (1999), "Efficient Tests for a Unit Root When the Initial Observations is Drawn from its Unconditional Distribution," *International Economic Review* **40**, 767-783.

- [13] Müller, U. K. and G. Elliott (2003), "Tests for Unit Roots and the Initial Condition," *Econometrica* **71**, 1269-1286.
- [14] Nelson, C. R. and C. I. Plosser (1982), "Trends and Random Walks in Macroeconomics Time Series: Some Evidence and Implications," *Journal of Monetary Economics* **10**, 139-162.
- [15] Ng, S. and P. Perron (1995), "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association* **90**, 268-281.
- [16] Ng, S. and Perron, P. (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica* **69**, 1519-1554.
- [17] Perron, P. (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica* **57**, 1361-1401
- [18] Perron, P. (1997), "Further Evidence of Breaking Trend Functions in Macroeconomics Variables," *Journal of Econometrics* **80**, 355-385.
- [19] Perron, P. and S. Ng (1996), "Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties," *Review of Economics Studies* **63**, 435-463.
- [20] Perron, P. and S. Ng (1998): "An Autoregressive Spectral Density Estimator at Frequency Zero for Nonstationarity Tests," *Econometric Theory* **14**, 560-603.
- [21] Perron, P. and G. Rodríguez (2003), "GLS Detrending, Efficient Unit Root Tests and Structural Change," *Journal of Econometrics* **115**, 1-27.
- [22] Phillips, P. C. B. (1987), "Time Series Regression with Unit Roots," *Econometrica* **55**, 277-302.
- [23] Phillips, P. C. B. and P. Perron (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika* **75**, 335-346.
- [24] Said, S. E. and Dickey, D. A. (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika* **71**, 599-608.

- [25] Sargan, J. D. and Bhargava, A. (1983), "Testing Residuals from Least Squares Regression for being Generated by the Gaussian Random Walk" *Econometrica* **51**, 153-174.
- [26] Stock, J. H. (1999), "A Class of Tests for Integration and Cointegration," in Engle, R.F. and H. White (eds.), *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, Oxford University Press, 137-167.
- [27] Zivot, E. and D. W. K. Andrews (1992), "Further Evidence on the Great Crash, The Oil-Price Shock and the Unit Root Hypothesis," *Journal of Business and Economics Statistics* **10**, 251-270.

Appendix

Lemma A.1. Under the initial condition specified in Condition A, we have

$$\begin{aligned} T^{-1/2}u_1 &\Rightarrow \sigma N\left(0, \frac{1}{-2c}\right) \\ &\equiv \sigma\xi. \end{aligned}$$

and as $c \rightarrow 0$ the variance of ξ converges to zero.

Proof of Lemma A.1. Using the local to unity framework $\alpha = 1 + \frac{c}{T}$, we can show that $T(1 - \alpha^2) \Rightarrow -2c$. Applying to $u_1 \sim N\left(0, \frac{\sigma^2}{1 - \alpha^2}\right)$, we have $T^{-1/2}u_1 \Rightarrow \sigma\xi$, and $\xi \sim N\left(0, \frac{1}{-2c}\right)$. For continuity, we require that the variance converges to zero as $c \rightarrow 0$ following Elliott (1999).

Lemma A.2. Let $\{y_t\}$ be generated by (1) to (3), the deterministic components given by Model I, the initial condition is given by Condition A, and $\delta = T_B/T$ is the break point, then we have,

$$\begin{aligned} T^{-1/2}(\hat{\mu}_1 - \mu_1) &\Rightarrow \sigma b_4, \\ T^{1/2}(\hat{\beta}_1 - \beta_1) &\Rightarrow \sigma b_5, \\ T^{1/2}(\hat{\beta}_2 - \beta_2) &\Rightarrow \sigma b_6. \end{aligned}$$

where the definitions of b_4, b_5, b_6 are given in the following proof.

Proof of Lemma A.2. In matrix notation, we have:

$$\Lambda^{-1}(\hat{\psi}(\delta) - \psi) = [\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda]^{-1} [\Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'}]$$

where

$$\begin{aligned} z^{\bar{\alpha}'} &= \begin{bmatrix} (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & -\bar{c}T^{-1} & \dots & -\bar{c}T^{-1} \\ (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & \dots & -\bar{c}T^{-1}(t-1) + 1 & \dots \\ 0 & \dots & -\bar{c}T^{-1}(t-1) + \delta\bar{c} + 1 & \dots \end{bmatrix}, \\ u^{\bar{\alpha}'} &= \left[(-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} u_1, (c - \bar{c})T^{-1}u_1 + v_2, \dots, (c - \bar{c})T^{-1}u_{T-1} + v_T \right], \\ \Lambda &= \text{diag}(T^{1/2}, T^{-1/2}, T^{-1/2}). \end{aligned}$$

We first calculate matrix $D_{3 \times 3}$, the limiting distribution of $\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda$. We know that the limiting distributions of entries D_{11}, D_{12}, D_{22} are the same as those in Elliott (1999); the limiting distributions of D_{23}, D_{33} have been calculated in Perron and Rodríguez (2003). That is, $D_{11} = \bar{c}^2 - 2\bar{c}$, $D_{12} = \frac{1}{2}\bar{c}^2 - \bar{c}$, $D_{22} = 1 + \frac{1}{3}\bar{c}^2 - \bar{c}$, $D_{23} = 1 - \delta - \bar{c} + \bar{c}\delta - \frac{1}{2}\bar{c}^2\delta + \frac{1}{3}\bar{c}^2 + \frac{1}{6}\bar{c}^2\delta^3 \equiv m$,

$D_{33} = 1 - \delta - \bar{c}(1 - \delta)^2 + \frac{1}{3}\bar{c}^2(1 - \delta)^3 \equiv d$. Therefore we only need to calculate the limiting distribution of D_{13} :

$$\begin{aligned}
D_{13} &= \begin{bmatrix} (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & -\bar{c}/T & \cdots & -\bar{c}/T & -\bar{c}/T & \cdots & -\bar{c}/T \\ 0 & \cdots & 1 & 1 - \bar{c}/T & 1 - 2\bar{c}/T & \cdots & 1 - (T - T_B - 1)\bar{c}/T \end{bmatrix}' \\
&= -\frac{\bar{c}}{T} \left[\sum_{t=T_B+1}^T \left[1 - \frac{\bar{c}}{T}(t - T_B - 1) \right] \right] \\
&\Rightarrow -\bar{c}(1 - \delta) + \frac{1}{2}\bar{c}^2(1 - \delta)^2.
\end{aligned}$$

Next, we calculate the limiting distribution of $\Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'}$. We have

$$\begin{aligned}
\Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'} &= \begin{bmatrix} T^{1/2} & 0 & 0 \\ 0 & T^{-1/2} & 0 \\ 0 & 0 & T^{-1/2} \end{bmatrix} \\
&\times \begin{bmatrix} (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & -\bar{c}T^{-1} & \cdots & -\bar{c}T^{-1} \\ (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & \cdots & -\bar{c}T^{-1}(t-1) + 1 & \cdots \\ 0 & \cdots & -\bar{c}T^{-1}(t-1) + \delta\bar{c} + 1 & \cdots \end{bmatrix} \\
&\times \left[(-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} u_1, (c - \bar{c})T^{-1}u_1 + v_2, \cdots, (c - \bar{c})T^{-1}u_{T-1} + v_T \right]' \\
&= \begin{bmatrix} T^{\frac{1}{2}} \left(-\frac{2\bar{c}}{T} - \frac{\bar{c}^2}{T^2} \right) u_1 + T^{\frac{1}{2}} \sum_{t=2}^T (-\bar{c}T^{-1}) [(c - \bar{c})T^{-1}u_{t-1} + v_t] \\ T^{-\frac{1}{2}} \left(-\frac{2\bar{c}}{T} - \frac{\bar{c}^2}{T^2} \right) u_1 + T^{-\frac{1}{2}} \sum_{t=2}^T [-\bar{c}T^{-1}(t-1) + 1] [(c - \bar{c})T^{-1}u_{t-1} + v_t] \\ T^{-\frac{1}{2}} \sum_{t=T_B+1}^T [-\bar{c}T^{-1}(t-1) + \delta\bar{c} + 1] [(c - \bar{c})T^{-1}u_{t-1} + v_t] \end{bmatrix}.
\end{aligned}$$

The first element of the above 3×1 vector may be expressed as:

$$\begin{aligned}
&= -2\bar{c}T^{-1/2}u_1 - \bar{c}^2T^{-3/2}u_1 + T^{1/2} \sum_{t=2}^T (-\bar{c}T^{-1})(c - \bar{c})T^{-1}u_{t-1} \\
&\quad + T^{1/2} \sum_{t=2}^T (-\bar{c}T^{-1})v_t \\
&= -2\bar{c}T^{-1/2}u_1 - \bar{c}^2T^{-3/2}u_1 - T^{-3/2}\bar{c}(c - \bar{c}) \sum_{t=2}^T u_{t-1} \\
&\quad - \bar{c}T^{-1/2} \sum_{t=2}^T v_t \\
&\Rightarrow \sigma\{-2\bar{c}\xi - \bar{c}(c - \bar{c}) \int_0^1 W_c(r) - \bar{c}W(1)\} \\
&\equiv \sigma b_1. \tag{A.1}
\end{aligned}$$

By the same token, the second element may be written as

$$\begin{aligned}
&= -2T^{-3/2}\bar{c}u_1 - \bar{c}^2T^{-5/2}u_1 + \\
&\quad T^{-1/2} \sum_{t=2}^T \{-\bar{c}T^{-1}(t-1)(c - \bar{c})T^{-1}u_{t-1} - \bar{c}T^{-1}(t-1)v_t \\
&\quad + (c - \bar{c})T^{-1}u_{t-1} + v_t\} \\
&= -2T^{-3/2}\bar{c}u_1 - \bar{c}^2T^{-5/2}u_1 - \bar{c}(c - \bar{c})T^{-5/2} \sum_{t=2}^T (t-1)u_{t-1} \\
&\quad - \bar{c}T^{-3/2} \sum_{t=2}^T (t-1)v_t + (c - \bar{c})T^{-3/2} \sum_{t=2}^T u_{t-1} + T^{-1/2} \sum_{t=2}^T v_t \\
&\Rightarrow \sigma\{-\bar{c}(c - \bar{c}) \int_0^1 rW_c(r) - \bar{c}[W(1) - \int_0^1 W_c(r)] \\
&\quad + (c - \bar{c}) \int_0^1 W_c(r) + W(1)\} \\
&\equiv \sigma\{(1 - \bar{c})W(1) + c \int_0^1 W_c(r) - \bar{c}(c - \bar{c}) \int_0^1 rW_c(r)\} \\
&\equiv \sigma b_2. \tag{A.2}
\end{aligned}$$

Finally, the third element is given by:

$$\begin{aligned}
&= T^{-1/2} \sum_{t=T_B+1}^T \{-\bar{c}T^{-1}(t-1)(c-\bar{c})T^{-1}u_{t-1} - \bar{c}T^{-1}(t-1)v_t \\
&\quad + \delta\bar{c}(c-\bar{c})T^{-1}u_{t-1} + \delta\bar{c}v_t + (c-\bar{c})T^{-1}u_{t-1} + v_t\} \\
&= T^{-1/2} \sum_{t=T_B+1}^T \{-\bar{c}T^{-1}(c-\bar{c})T^{-1}tu_{t-1} + \bar{c}T^{-1}(c-\bar{c})T^{-1}u_{t-1} \\
&\quad - \bar{c}T^{-1}tv_t + \bar{c}T^{-1}v_t + \delta\bar{c}(c-\bar{c})T^{-1}u_{t-1} + \delta\bar{c}v_t \\
&\quad + (c-\bar{c})T^{-1}u_{t-1} + v_t\} \\
&\Rightarrow \sigma\{(1-\bar{c}+\delta\bar{c})W(1) - W(\delta) - \bar{c}(c-\bar{c})\int_{\delta}^1 rW_c(r) \\
&\quad + \delta\bar{c}(c-\bar{c})\int_{\delta}^1 W_c(r) + c\int_{\delta}^1 W_c(r)\} \\
&\equiv \sigma b_3. \tag{A.3}
\end{aligned}$$

Therefore, using D_{ij} , for $i, j = 1, 2, 3$ and (A.1), (A.2), and (A.3), we have:

$$\begin{aligned}
&\Lambda^{-1} \begin{pmatrix} \hat{\mu}_1 - \mu_1 \\ \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \end{pmatrix} \\
&= (\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda)^{-1} \Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'} \\
&\Rightarrow \left[\begin{array}{ccc} \bar{c}^2 - 2\bar{c} & \frac{1}{2}\bar{c}^2 - \bar{c} & -\bar{c}(1-\delta) + \frac{1}{2}\bar{c}^2(1-\delta)^2 \\ \frac{1}{2}\bar{c}^2 - \bar{c} & 1 + \frac{1}{3}\bar{c}^2 - \bar{c} & m \\ -\bar{c}(1-\delta) + \frac{1}{2}\bar{c}^2(1-\delta)^2 & m & d \end{array} \right]^{-1} \\
&\quad \times \begin{pmatrix} \sigma b_1 \\ \sigma b_2 \\ \sigma b_3 \end{pmatrix} \\
&\equiv \sigma \begin{pmatrix} b_4 \\ b_5 \\ b_6 \end{pmatrix},
\end{aligned}$$

the proof completes. ■

Lemma A.3. Let $\{y_t\}$ be generated by (1) to (3), the set of deterministic components is given by Model II, the initial condition is defined in Condition A, δ is the break point. Let $\hat{\psi}(\delta)$ be the estimates of the coefficients of (4),

then the results of Lemma A.3 still hold with the addition that $\hat{\mu}_2 - \mu_2 \Rightarrow \lim_{T \rightarrow \infty} \kappa(\bar{c}, \delta) v_{T_B} \equiv v^*$, where $\kappa(\bar{c}, \delta)$ is defined in the following proof.

Proof of Lemma A.3. For Model II, we have

$$\Lambda^{-1} \begin{pmatrix} \hat{\mu}_1 - \mu_1 \\ \hat{\mu}_2 - \mu_2 \\ \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \end{pmatrix} = (\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda)^{-1} \Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'},$$

where $\Lambda = \text{diag}(T^{1/2}, 1, T^{-1/2}, T^{-1/2})$, and

$$z^{\bar{\alpha}'} = \begin{bmatrix} (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & -\frac{\bar{c}}{T} & \cdots & \cdots & \cdots & -\frac{\bar{c}}{T} \\ 0 & \cdots & 1 & -\frac{\bar{c}}{T} & \cdots & -\frac{\bar{c}}{T} \\ (-2\bar{c}/T - \bar{c}^2/T^2)^{1/2} & 1 - \frac{\bar{c}}{T} & \cdots & 1 - \frac{(t-1)\bar{c}}{T} & \cdots & 1 - \frac{(T-1)\bar{c}}{T} \\ 0 & \cdots & \cdots & 1 - \frac{(t-1)\bar{c}}{T} + \delta\bar{c} & \cdots & 1 - \frac{(T-1)\bar{c}}{T} + \delta\bar{c} \end{bmatrix}.$$

As before, we first calculate the limiting distribution of $\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda$. From the proof of Lemma A.3, we know that $D_{11} \Rightarrow \bar{c}^2 - 2\bar{c}$, $D_{13} \Rightarrow \frac{\bar{c}^2}{2} - \bar{c}$, $D_{14} \Rightarrow -\bar{c}(1 - \delta) + \frac{\bar{c}^2}{2}(1 - \delta)^2$, $D_{33} \Rightarrow 1 + \frac{1}{3}\bar{c}^2 - \bar{c} \equiv a$, $D_{34} \Rightarrow m$, $D_{44} \Rightarrow d$. We only need to calculate the terms D_{12} , D_{22} , D_{23} , D_{24} :

$$\begin{aligned} D_{12} &= T^{1/2} \left\{ -\frac{\bar{c}}{T} + \frac{\bar{c}^2}{T^2} + \cdots + \frac{\bar{c}^2}{T^2} \right\} \\ &= T^{1/2} \left\{ -\frac{\bar{c}}{T} + \frac{\bar{c}^2}{T^2} (T - T_B) \right\} \\ &\Rightarrow 0, \end{aligned}$$

$$\begin{aligned} D_{22} &= 1 + \frac{\bar{c}^2}{T^2} + \cdots + \frac{\bar{c}^2}{T^2} \\ &= 1 + \frac{\bar{c}^2}{T^2} (T - T_B) \\ &\Rightarrow 1, \end{aligned}$$

$$\begin{aligned}
D_{23} &= T^{-1/2} \left\{ 1 - (T_B - 1) \frac{\bar{c}}{T} - \frac{\bar{c}}{T} \sum_{t=T_B+1}^T [1 - (t-1) \frac{\bar{c}}{T}] \right\} \\
&= T^{-1/2} \left\{ 1 - (T_B - 1) \frac{\bar{c}}{T} - \frac{\bar{c}}{T} (T - T_B) \right. \\
&\quad \left. + \frac{\bar{c}^2}{T^2} \sum_{t=T_B+1}^T (t-1) \right\} \\
&\Rightarrow 0,
\end{aligned}$$

$$\begin{aligned}
D_{24} &= T^{-1/2} \left\{ 1 - (T_B - 1) \frac{\bar{c}}{T} + \delta \bar{c} \right. \\
&\quad \left. - \frac{\bar{c}}{T} \sum_{t=T_B+1}^T [1 - (t-1) \frac{\bar{c}}{T} + \delta \bar{c}] \right\} \\
&= T^{-1/2} \left\{ 1 + \frac{\bar{c}}{T} - \frac{\bar{c}}{T} (T - T_B) + \frac{\bar{c}^2}{T^2} \sum_{t=T_B+1}^T (t-1) \right. \\
&\quad \left. - \frac{\delta \bar{c}^2}{T} (T - T_B) \right\} \\
&\Rightarrow 0.
\end{aligned}$$

Next we calculate the limiting distribution of $\Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'}$. Note that the first, the third and the fourth elements are already calculated in the proof for Lemma A.3. We only need to calculate the second entry which is:

$$\begin{aligned}
& \left[0 \quad \cdots \quad 1 \quad -\frac{\bar{c}}{T} \quad \cdots \quad -\frac{\bar{c}}{T} \right] \\
& \times \left[\left(-\frac{2\bar{c}}{T} - \frac{\bar{c}^2}{T^2} \right)^{1/2} u_1, \cdots, (c - \bar{c}) T^{-1} u_{t-1} + v_t, \cdots \right]' \\
&= (c - \bar{c}) T^{-1} u_{T_B-1} + v_{T_B} - \frac{\bar{c}}{T} \sum_{t=T_B+1}^T [(c - \bar{c}) T^{-1} u_{t-1} + v_t] \\
&\Rightarrow v_{T_B}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned} & (\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda)^{-1} \Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'} \\ \Rightarrow & \begin{bmatrix} \bar{c}^2 - 2\bar{c} & 0 & \frac{\bar{c}^2}{2} - \bar{c} & -\bar{c}(1-\delta) + \frac{\bar{c}^2}{2}(1-\delta)^2 \\ 0 & 1 & 0 & 0 \\ \frac{\bar{c}^2}{2} - \bar{c} & 0 & a & m \\ -\bar{c}(1-\delta) + \frac{\bar{c}^2}{2}(1-\delta)^2 & 0 & m & d \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} \sigma b_1 \\ v_{TB} \\ \sigma b_2 \\ \sigma b_3 \end{bmatrix}. \end{aligned}$$

Using matrix algebra, the second element of the resulting matrix is equal to:

$$\kappa(\bar{c}, \delta) v_{TB} = \frac{\kappa^*(\bar{c}, \delta) v_{TB}}{(\bar{c}^2 - 2\bar{c})ad - \left(\frac{\bar{c}^2}{2} - \bar{c}\right)^2 d},$$

with $\kappa^*(\bar{c}, \delta) = (\bar{c}^2 - 2\bar{c})ad + 2m(\bar{c}^2 - 2\bar{c})[-\bar{c}(1-\delta) + \frac{\bar{c}^2}{2}(1-\delta)^2] - [-\bar{c}(1-\delta) + \frac{\bar{c}^2}{2}(1-\delta)^2]^2 a - m^2(\bar{c}^2 - 2\bar{c}) - d(\frac{\bar{c}^2}{2} - \bar{c})^2$. That is, $\hat{\mu}_2 - \mu_2 \Rightarrow \lim_{t \rightarrow \infty} \kappa(\bar{c}, \delta) v_{TB} \equiv v^*$. ■

Proof of Theorem 1. We only show the proof of MZ_α^{GLS} for Model I in detail. The proof for Model II and other statistics follows analogously. First, we calculate the limiting distribution of $T^{-1}\tilde{y}_T^2$ as follows:

$$\begin{aligned} T^{-1}\tilde{y}_T^2 &= T^{-1} \left\{ u_T - [(\hat{\mu}_1 - \mu_1) + (\hat{\beta}_1 - \beta_1)T + (\hat{\beta}_2 - \beta_2)1(\cdot)(T - T\delta)] \right\}^2 \\ &= T^{-1} \{ u_T^2 + (\hat{\mu}_1 - \mu_1)^2 + (\hat{\beta}_1 - \beta_1)^2 T^2 \\ &\quad + (\hat{\beta}_2 - \beta_2)^2 1(\cdot)(T - T\delta)^2 - 2u_T(\hat{\mu}_1 - \mu_1) \\ &\quad + 2(\hat{\beta}_1 - \beta_1)T(\hat{\beta}_2 - \beta_2)1(\cdot)(T - T\delta) \\ &\quad - 2u_T(\hat{\beta}_1 - \beta_1)T - 2u_T(\hat{\beta}_2 - \beta_2)1(\cdot)(T - T\delta) \\ &\quad + 2(\hat{\mu}_1 - \mu_1)(\hat{\beta}_1 - \beta_1)T \\ &\quad + 2(\hat{\mu}_1 - \mu_1)(\hat{\beta}_2 - \beta_2)1(\cdot)(T - T\delta) \}. \end{aligned}$$

It is worth to note that the limiting distributions of some terms have been found already in Perron and Rodríguez (2003). The other terms are calculated as follow:

1. $T^{-1}(\hat{\mu}_1 - \mu_1)^2 = [T^{-1/2}(\hat{\mu}_1 - \mu_1)]^2 \Rightarrow \sigma^2 b_4^2.$
2. $-2T^{-1}u_T(\hat{\mu}_1 - \mu_1) = -2(T^{-1/2}u_T)[T^{-1/2}(\hat{\mu}_1 - \mu_1)] \Rightarrow -2\sigma^2 W_c(1)b_4.$
3. $2T^{-1}(\hat{\mu}_1 - \mu_1)(\hat{\beta}_1 - \beta_1)T = 2[T^{-1/2}(\hat{\mu}_1 - \mu_1)][T^{1/2}(\hat{\beta}_1 - \beta_1)] \Rightarrow 2\sigma^2 b_4 b_5.$
4. $2T^{-1}(\hat{\mu}_1 - \mu_1)(\hat{\beta}_2 - \beta_2)(T - T\delta) = 2(\hat{\mu}_1 - \mu_1)(\hat{\beta}_2 - \beta_2) - 2(\hat{\mu}_1 - \mu_1)(\hat{\beta}_2 - \beta_2)\delta = 2[T^{-1/2}(\hat{\mu}_1 - \mu_1)][T^{1/2}(\hat{\beta}_2 - \beta_2)] - 2T^{-1/2}(\hat{\mu}_1 - \mu_1)T^{1/2}(\hat{\beta}_2 - \beta_2)\delta \Rightarrow 2\sigma^2 b_4 b_6(1 - \delta).$

Therefore,

$$\begin{aligned}
T^{-1}\tilde{y}_T^2 &\Rightarrow \sigma^2 W_c^2(1) + \sigma^2 b_4^2 + \sigma^2 b_5^2 + \sigma^2 b_6^2(1 - \delta)^2 \\
&\quad - 2\sigma^2 W_c(1)b_4 + 2\sigma^2 b_5 b_6(1 - \delta) - 2\sigma^2 b_5 W_c(1) \\
&\quad - 2\sigma^2 b_6 W_c(1)(1 - \delta) + 2\sigma^2 b_4 b_5 + 2\sigma^2 b_4 b_6(1 - \delta) \\
&= \sigma^2 \{ [W_c^2(1) - 2b_5 W_c(1) + b_5^2 + b_4^2 - 2b_4 W_c(1) + 2b_4 b_5] \\
&\quad - [2b_6(1 - \delta)W_c(1) - 2b_5 b_6(1 - \delta) - b_6^2(1 - \delta)^2 - 2b_4 b_6(1 - \delta)] \} \\
&= \sigma^2 \{ [W_c(1) - b_4 - b_5]^2 - 2[b_6(1 - \delta)][W_c(1) - b_4 - b_5] \\
&\quad - \frac{1}{2}(1 - \delta)b_6 \} \\
&\equiv \sigma^2 \{ V_{\bar{c}\bar{c}}^{(1)}(1, \delta)^2 - 2V_{\bar{c}\bar{c}}^{(2)}(1, \delta) \}. \tag{A.4}
\end{aligned}$$

where $V_{\bar{c}\bar{c}}^{(1)}(1, \delta) = W_c(1) - b_4 - b_5$, and $V_{\bar{c}\bar{c}}^{(2)}(1, \delta) = [b_6(1 - \delta)][W_c(1) - b_4 - b_5 - \frac{1}{2}(1 - \delta)b_6]$.

Next we calculate the limiting distribution of $T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2$. Using the above results and by the Continuous Mapping Theorem (CMT), we have

$$2T^{-2} \sum_{t=1}^T \tilde{y}_t^2 \Rightarrow 2\sigma^2 \left\{ \int_0^1 V_{\bar{c}\bar{c}}^{(1)}(r, \delta)^2 dr - 2 \int_\delta^1 V_{\bar{c}\bar{c}}^{(2)}(r, \delta) dr \right\}. \tag{A.5}$$

Then by substituting (A.4), (A.5) into (7) and by using the fact that s^2 is a consistent estimate of σ^2 , the proof is complete. ■

Proof of Theorem 2. Here we only give the proof for Model I. Defining $\bar{M}_T(c, \bar{c}, \delta) = \left(u^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda \right) \left(\Lambda z^{\bar{\alpha}} z^{\bar{\alpha}'} \Lambda \right)^{-1} \left(\Lambda z^{\bar{\alpha}} u^{\bar{\alpha}'} \right)$, we have $S(\bar{\alpha}, \delta) = u^{\bar{\alpha}} u^{\bar{\alpha}'} - \bar{M}_T(c, 0, \delta)$ and $S(1) = u^1 u^{1'} - \bar{M}_T(0, 0, \delta)$. By definition, $u^{\bar{\alpha}} = [(1 -$

$\bar{\alpha}^2)^{1/2}u_1, (1 - \bar{\alpha}L)u_2, \dots, (1 - \bar{\alpha}L)u_T]$, and $u^1 = [0, (1 - L)u_2, \dots, (1 - L)u_T]$. Simple algebra shows that

$$\begin{aligned} & u^{\bar{\alpha}}u^{\bar{\alpha}'} - u^1u^{1'} \\ \Rightarrow & -2\bar{c}\xi^2 + (\bar{c}^2 - 2c\bar{c}) \int_0^1 W_c^2(r) dr - 2\bar{c} \int_0^1 W_c(r) dr. \end{aligned}$$

Notice that all elements of $\bar{M}_T(c, \bar{c}, \delta) = \left(u^{\bar{\alpha}}z^{\bar{\alpha}'}\Lambda\right) \left(\Lambda z^{\bar{\alpha}}z^{\bar{\alpha}'}\Lambda\right)^{-1} \left(\Lambda z^{\bar{\alpha}}u^{\bar{\alpha}'}\right)$ have already been calculated. Therefore,

$$\begin{aligned} & \bar{M}_T(c, \bar{c}, \delta) \\ = & \begin{bmatrix} \sigma b_1 & \sigma b_2 & \sigma b_3 \end{bmatrix} \\ & \times \begin{bmatrix} \bar{c}^2 - 2\bar{c} & \frac{1}{2}\bar{c}^2 - \bar{c} & -\bar{c}(1 - \delta) + \frac{1}{2}\bar{c}^2(1 - \delta)^2 \\ \frac{1}{2}\bar{c}^2 - \bar{c} & 1 + \frac{1}{3}\bar{c}^2 - \bar{c} & m \\ -\bar{c}(1 - \delta) + \frac{1}{2}\bar{c}^2(1 - \delta)^2 & m & d \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} \sigma b_1 \\ \sigma b_2 \\ \sigma b_3 \end{bmatrix}. \end{aligned}$$

The calculation of $\bar{M}_T(0, 0, \delta)$ follows the arguments of Perron and Rodríguez (2003) since the initial value is not changed under the null. Using infimum method to estimate the break point, the limiting distribution of the feasible point optimal statistic is given by

$$\begin{aligned} P_{T\xi}^{GLS}(c, \bar{c}) & \Rightarrow \sup_{\delta \in [\epsilon, 1-\epsilon]} \bar{M}(c, 0, \delta) - \sup_{\delta \in [\epsilon, 1-\epsilon]} \bar{M}(c, \bar{c}, \delta) \\ & - 2\bar{c}\xi^2 - 2\bar{c} \int_0^1 W_c(r) dW(r) + (\bar{c}^2 - 2c\bar{c}) \int_0^1 W_c(r)^2 dr - \bar{c} \\ & \equiv J_{P_{T\xi}}^{GLS}(c, \bar{c}). \end{aligned}$$

and this completes the proof. ■

Table 1. Critical Values for $P_{T\xi}^{GLS}$, M^{GLS} and ADF^{GLS} tests choosing T_B minimizing the statistics; Model I
($\bar{c} = -24$ when constructing the tests and s^2)

Test	Size	$T = \infty$	$T = 100$		$T = 100$			$T = 200$		
			$k = 0$	BIC	MAIC	t-sig	BIC	MAIC	t-sig	
MZ_α	.01	-43.210	-30.422	-43.919	-30.876	-261.061	-39.131	-35.537	-60.412	
	.025	-37.283	-27.589	-35.735	-27.956	-146.495	-33.697	-31.174	-51.096	
	.05	-33.366	-25.507	-30.209	-25.379	-75.534	-30.185	-28.461	-43.068	
	.10	-28.788	-22.994	-25.817	-22.492	-50.765	-27.080	-24.545	-32.537	
	.20	-24.105	-20.016	-21.408	-19.368	-34.711	-22.068	-20.815	-26.767	
MSB	.01	0.107	0.127	0.106	0.125	0.043	0.113	0.118	0.090	
	.025	0.115	0.133	0.118	0.132	0.058	0.120	0.126	0.098	
	.05	0.121	0.138	0.127	0.139	0.081	0.127	0.131	0.107	
	.10	0.130	0.145	0.138	0.147	0.098	0.135	0.142	0.123	
	.20	0.142	0.155	0.150	0.159	0.118	0.149	0.152	0.135	
MZ_t	.01	-4.621	-3.874	-4.674	-3.894	-11.408	-4.421	-4.213	-5.468	
	.025	-4.300	-3.691	-4.217	-3.701	-8.525	-4.067	-3.923	-5.014	
	.05	-4.064	-3.545	-3.863	-3.541	-6.143	-3.876	-3.769	-4.601	
	.10	-3.766	-3.367	-3.577	-3.340	-5.033	-3.657	-3.455	-4.020	
	.20	-3.449	-3.132	-3.240	-3.075	-4.129	-3.299	-3.205	-3.607	
ADF	.01	-4.621	-4.919	-5.122	-4.980	-5.297	-4.889	-4.790	-5.067	
	.025	-4.300	-4.562	-4.804	-4.538	-4.984	-4.530	-4.396	-4.675	
	.05	-4.064	-4.304	-4.488	-4.230	-4.655	-4.256	-4.172	-4.432	
	.10	-3.766	-3.997	-4.176	-3.865	-4.292	-4.046	-3.819	-4.149	
	.20	-3.449	-3.673	-3.755	-3.532	-3.929	-3.625	-3.479	-3.746	
$P_{T\xi}$.01	6.967	7.600	7.379	9.948	1.662	8.054	9.021	5.340	
	.025	8.065	9.149	8.950	11.340	2.935	9.260	10.008	6.399	
	.05	9.340	10.470	10.430	12.118	4.509	10.161	10.846	7.681	
	.10	10.866	12.135	12.016	13.479	7.129	11.755	12.962	10.003	
	.20	13.110	14.401	14.413	16.004	9.934	14.088	15.129	11.962	

Table 2. Critical Values for $P_{T\xi}^{GLS}$, M^{GLS} and ADF^{GLS} tests choosing T_B minimizing the statistics; Model II
($\bar{c} = -24$ when constructing the tests and s^2)

Test	Size	$T = \infty$	$T = 100$			$T = 200$			
			$k = 0$	BIC	MAIC	t-sig	BIC	MAIC	t-sig
MZ_α	.01	-43.210	-32.049	-92.017	-33.065	-529.130	-44.641	-36.079	-70.335
	.025	-37.283	-29.436	-43.719	-30.396	-220.145	-34.854	-32.217	-58.017
	.05	-33.366	-27.214	-36.687	-27.039	-119.088	-32.066	-29.773	-49.142
	.10	-28.788	-24.566	-30.082	-24.686	-73.058	-28.546	-26.219	-36.865
	.20	-24.105	-21.600	-24.610	-21.238	-47.797	-23.533	-21.852	-30.140
MSB	.01	0.107	0.123	0.073	0.122	0.030	0.105	0.117	0.084
	.025	0.115	0.129	0.106	0.127	0.047	0.117	0.124	0.092
	.05	0.121	0.134	0.116	0.133	0.064	0.124	0.129	0.100
	.10	0.130	0.141	0.127	0.141	0.082	0.131	0.136	0.115
	.20	0.142	0.150	0.141	0.151	0.101	0.144	0.148	0.127
MZ_t	.01	-4.621	-3.977	-6.763	-4.034	-16.265	-4.716	-4.233	-5.914
	.025	-4.300	-3.805	-4.653	-3.878	-10.489	-4.164	-3.971	-5.367
	.05	-4.064	-3.660	-4.254	-3.653	-7.710	-3.963	-3.854	-4.913
	.10	-3.766	-3.475	-3.836	-3.480	-6.014	-3.749	-3.596	-4.274
	.20	-3.449	-3.255	-3.465	-3.216	-4.832	-3.403	-3.277	-3.848
ADF	.01	-4.621	-5.109	-5.281	-5.072	-5.518	-5.065	-4.811	-5.187
	.025	-4.300	-4.758	-5.068	-4.685	-5.227	-4.661	-4.509	-4.841
	.05	-4.064	-4.492	-4.739	-4.381	-4.945	-4.385	-4.243	-4.570
	.10	-3.766	-4.183	-4.402	-4.103	-4.576	-4.121	-3.964	-4.260
	.20	-3.449	-3.845	-3.987	-3.716	-4.163	-3.754	-3.593	-3.894
$P_{T\xi}$.01	6.967	9.136	7.824	10.761	2.276	8.964	10.195	5.307
	.025	8.065	10.409	9.353	12.174	3.540	10.357	11.109	6.926
	.05	9.340	11.600	11.087	13.250	4.853	11.388	11.911	8.345
	.10	10.866	13.208	13.061	14.572	7.415	12.617	13.761	10.457
	.20	13.110	15.232	15.255	16.739	10.394	14.842	15.904	12.765

Table 3. Critical Values for M^{GLS} and ADF^{GLS} tests choosing T_B maximizing $|t_{\hat{\beta}_2}|$; Model I
($\bar{c} = -24$ when constructing the tests and s^2)

Test	Size	$T = \infty$	$T = 100$		$T = 100$			$T = 200$		
			$k = 0$	BIC	MAIC	t-sig	BIC	MAIC	t-sig	
MZ_α	.01	-42.432	-30.876	-40.049	-30.424	-187.405	-38.111	-34.804	-57.572	
	.025	-36.810	-28.678	-33.988	-27.585	-99.541	-33.506	-30.694	-50.102	
	.05	-32.689	-25.717	-29.498	-24.982	-66.232	-29.851	-27.964	-40.333	
	.10	-28.266	-23.204	-24.967	-21.966	-42.880	-26.705	-24.048	-31.524	
	.20	-23.748	-19.986	-20.751	-18.829	-30.353	-21.764	-20.405	-25.718	
MSB	.01	0.107	0.126	0.110	0.127	0.051	0.114	0.118	0.092	
	.025	0.115	0.131	0.121	0.132	0.070	0.121	0.126	0.099	
	.05	0.122	0.138	0.129	0.139	0.086	0.129	0.133	0.111	
	.10	0.131	0.145	0.139	0.149	0.107	0.136	0.143	0.125	
	.20	0.143	0.156	0.153	0.160	0.126	0.149	0.154	0.137	
MZ_t	.01	-4.579	-3.928	-4.440	-3.862	-9.653	-4.355	-4.171	-5.342	
	.025	-4.263	-3.750	-4.100	-3.675	-7.008	-4.059	-3.899	-4.970	
	.05	-4.019	-3.575	-3.784	-3.514	-5.751	-3.845	-3.734	-4.473	
	.10	-3.737	-3.381	-3.506	-3.291	-4.600	-3.634	-3.415	-3.926	
	.20	-3.423	-3.124	-3.836	-3.008	-3.875	-3.274	-3.162	-3.560	
ADF	.01	-4.579	-4.981	-5.103	-4.897	-5.205	-4.788	-4.606	-4.941	
	.025	-4.263	-4.615	-4.713	-4.435	-4.928	-4.506	-4.324	-4.597	
	.05	-4.019	-4.366	-4.427	-4.163	-4.543	-4.226	-4.081	-4.388	
	.10	-3.737	-3.960	-4.076	-3.814	-4.226	-3.989	-3.757	-4.079	
	.20	-3.423	-3.621	-3.681	-3.458	-3.840	-3.567	-3.419	-3.689	

Table 4. Critical Values for M^{GLS} and ADF^{GLS} tests choosing T_B maximizing $|t_{\beta_2}|$; Model II
($\bar{c} = -24$ when constructing the tests and s^2)

Test	Size	$T = \infty$	$T = 100$		$T = 100$			$T = 200$		
			$k = 0$	BIC	MAIC	t-sig	BIC	MAIC	t-sig	
MZ_α	.01	-42.432	-28.651	-38.583	-27.367	-108.695	-34.731	-30.784	-54.661	
	.025	-36.810	-26.634	-32.090	-24.442	-78.069	-31.155	-27.036	-40.413	
	.05	-32.689	-23.638	-27.184	-21.780	-49.497	-26.721	-24.692	-33.122	
	.10	-28.266	-20.865	-22.605	-19.560	-36.566	-22.881	-21.167	-27.747	
	.20	-23.748	-17.970	-18.664	-16.975	-25.375	-19.313	-17.837	-21.795	
MSB	.01	0.107	0.130	0.113	0.133	0.067	0.119	0.127	0.095	
	.025	0.115	0.136	0.123	0.142	0.079	0.126	0.135	0.110	
	.05	0.122	0.143	0.134	0.149	0.100	0.136	0.141	0.122	
	.10	0.131	0.153	0.147	0.158	0.115	0.146	0.152	0.134	
	.20	0.143	0.165	0.162	0.169	0.138	0.159	0.164	0.150	
MZ_t	.01	-4.578	-3.763	-4.385	-3.684	-7.367	-4.158	-3.909	-5.206	
	.025	-4.265	-3.608	-3.977	-3.483	-6.240	-3.909	-3.675	-4.494	
	.05	-4.012	-3.421	-3.660	-3.290	-4.972	-3.640	-3.491	-4.050	
	.10	-3.737	-3.208	-3.326	-3.102	-4.253	-3.351	-3.219	-3.687	
	.20	-3.423	-2.963	-3.028	-2.882	-3.555	-3.096	-2.957	-3.282	
ADF	.01	-4.579	-4.674	-4.823	-4.529	-4.987	-4.631	-4.352	-4.681	
	.025	-4.263	-4.403	-4.483	-4.121	-4.600	-4.286	-4.130	-4.434	
	.05	-4.019	-4.062	-4.220	-3.784	-4.289	-4.030	-3.803	-4.163	
	.10	-3.737	-3.750	-3.845	-3.548	-3.987	-3.724	-3.520	-3.801	
	.20	-3.423	-3.468	-3.512	-3.275	-3.627	-3.381	-3.223	-3.476	

Table 5. Size and Power when using Infimum Method for Model I; T=100
 ($\bar{c} = -24$ when constructing the tests and s^2)

	Criteria	Size					Power				
		MZ_α	MSB	MZ_t	ADF	$P_{T\xi}$	MZ_α	MSB	MZ_t	ADF	$P_{T\xi}$
<i>i.i.d.</i>	BIC	0.051	0.050	0.051	0.050	0.051	0.318	0.299	0.324	0.510	0.381
	MAIC	0.051	0.051	0.050	0.050	0.051	0.508	0.499	0.502	0.448	0.495
	t-sig	0.050	0.050	0.051	0.051	0.051	0.180	0.181	0.180	0.433	0.164
MA(1) Errors											
$\theta = -0.8$	BIC	0.937	0.935	0.938	0.976	0.932	0.996	0.996	0.996	1.000	0.995
	MAIC	0.399	0.397	0.400	0.378	0.339	0.827	0.829	0.825	0.829	0.707
	t-sig	0.107	0.108	0.107	0.798	0.088	0.297	0.299	0.297	0.961	0.263
$\theta = -0.4$	BIC	0.401	0.396	0.395	0.487	0.402	0.887	0.884	0.886	0.913	0.884
	MAIC	0.135	0.135	0.136	0.100	0.111	0.556	0.552	0.550	0.457	0.488
	t-sig	0.040	0.040	0.040	0.253	0.038	0.136	0.136	0.136	0.737	0.121
$\theta = 0.4$	BIC	0.214	0.212	0.213	0.095	0.201	0.729	0.720	0.732	0.524	0.735
	MAIC	0.100	0.105	0.094	0.006	0.048	0.134	0.129	0.137	0.065	0.101
	t-sig	0.066	0.066	0.066	0.061	0.064	0.189	0.190	0.187	0.404	0.191
$\theta = 0.8$	BIC	0.526	0.526	0.521	0.118	0.448	0.729	0.728	0.728	0.407	0.707
	MAIC	0.223	0.231	0.220	0.008	0.105	0.338	0.336	0.335	0.046	0.265
	t-sig	0.089	0.090	0.087	0.049	0.091	0.282	0.283	0.281	0.227	0.299
AR(1) Errors											
$\rho = -0.8$	BIC	0.018	0.017	0.018	0.047	0.013	0.040	0.039	0.040	0.456	0.033
	MAIC	0.000	0.000	0.000	0.043	0.000	0.008	0.007	0.008	0.330	0.010
	t-sig	0.020	0.020	0.019	0.041	0.015	0.072	0.073	0.072	0.386	0.057
$\rho = -0.4$	BIC	0.142	0.138	0.145	0.168	0.136	0.510	0.510	0.511	0.596	0.510
	MAIC	0.062	0.064	0.064	0.050	0.058	0.353	0.339	0.356	0.335	0.335
	t-sig	0.048	0.049	0.048	0.085	0.038	0.150	0.150	0.150	0.434	0.138
$\rho = 0.4$	BIC	0.155	0.152	0.149	0.053	0.131	0.538	0.527	0.543	0.296	0.562
	MAIC	0.103	0.110	0.098	0.013	0.055	0.074	0.073	0.075	0.021	0.048
	t-sig	0.076	0.079	0.075	0.046	0.059	0.170	0.170	0.167	0.262	0.165
$\rho = 0.8$	BIC	0.293	0.303	0.282	0.060	0.186	0.350	0.347	0.351	0.131	0.346
	MAIC	0.326	0.339	0.308	0.049	0.189	0.383	0.390	0.387	0.104	0.338
	t-sig	0.145	0.149	0.143	0.061	0.079	0.117	0.117	0.114	0.132	0.100

Table 6. Size and Power when using Infimum Method for Model I; T=200
($\bar{c} = -24$ when constructing the tests and s^2)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	ADF	$P_{T\xi}$	MZ_α	MSB	MZ_t	ADF	$P_{T\xi}$
<i>i.i.d.</i>	BIC	0.050	0.050	0.050	0.050	0.050	0.506	0.497	0.495	0.517	0.486
	MAIC	0.051	0.050	0.051	0.050	0.051	0.431	0.423	0.420	0.372	0.423
	t-sig	0.050	0.051	0.051	0.051	0.051	0.200	0.204	0.213	0.432	0.242
MA(1) Errors											
$\theta = -0.8$	BIC	0.874	0.873	0.871	0.911	0.849	0.996	0.996	0.996	0.999	0.988
	MAIC	0.140	0.137	0.140	0.167	0.103	0.394	0.390	0.391	0.421	0.281
	t-sig	0.425	0.427	0.432	0.706	0.418	0.813	0.813	0.816	0.973	0.755
$\theta = -0.4$	BIC	0.236	0.238	0.228	0.241	0.213	0.849	0.842	0.845	0.861	0.809
	MAIC	0.097	0.095	0.088	0.058	0.076	0.416	0.410	0.406	0.326	0.363
	t-sig	0.056	0.058	0.058	0.123	0.061	0.337	0.338	0.346	0.633	0.384
$\theta = 0.4$	BIC	0.177	0.175	0.173	0.104	0.156	0.719	0.708	0.711	0.567	0.699
	MAIC	0.103	0.099	0.099	0.037	0.080	0.450	0.443	0.441	0.219	0.378
	t-sig	0.064	0.066	0.068	0.060	0.064	0.338	0.342	0.347	0.378	0.373
$\theta = 0.8$	BIC	0.338	0.338	0.328	0.097	0.264	0.730	0.726	0.723	0.393	0.695
	MAIC	0.189	0.195	0.188	0.006	0.114	0.430	0.423	0.426	0.083	0.366
	t-sig	0.053	0.053	0.054	0.030	0.055	0.267	0.269	0.279	0.190	0.310
AR(1) Errors											
$\rho = -0.8$	BIC	0.005	0.005	0.005	0.053	0.004	0.027	0.025	0.029	0.475	0.021
	MAIC	0.001	0.001	0.001	0.040	0.000	0.013	0.013	0.013	0.276	0.010
	t-sig	0.011	0.012	0.012	0.047	0.010	0.055	0.055	0.055	0.373	0.051
$\rho = -0.4$	BIC	0.053	0.051	0.050	0.063	0.047	0.438	0.429	0.427	0.473	0.402
	MAIC	0.044	0.043	0.044	0.041	0.041	0.346	0.339	0.336	0.310	0.302
	t-sig	0.034	0.036	0.034	0.049	0.033	0.161	0.166	0.167	0.385	0.1830
$\rho = 0.4$	BIC	0.115	0.111	0.108	0.057	0.097	0.571	0.558	0.562	0.401	0.551
	MAIC	0.109	0.108	0.104	0.041	0.090	0.454	0.448	0.439	0.246	0.404
	t-sig	0.054	0.052	0.052	0.047	0.052	0.269	0.273	0.276	0.350	0.315
$\rho = 0.8$	BIC	0.145	0.149	0.138	0.066	0.102	0.403	0.398	0.393	0.234	0.380
	MAIC	0.153	0.154	0.136	0.042	0.093	0.365	0.371	0.357	0.162	0.345
	t-sig	0.091	0.096	0.092	0.051	0.061	0.200	0.200	0.204	0.192	0.215

Table 7. Size and Power when using Supremum Method for Model I; T=100
($\bar{c} = -24$ when constructing the tests and s^2)

		Size				Power			
	Criteria	MZ_α	MSB	MZ_t	ADF	MZ_α	MSB	MZ_t	ADF
<i>i.i.d.</i>	BIC	0.050	0.050	0.050	0.050	0.335	0.336	0.381	0.510
	MAIC	0.051	0.049	0.050	0.051	0.495	0.465	0.490	0.445
	t-sig	0.050	0.050	0.050	0.051	0.169	0.170	0.169	0.451
MA(1) Errors									
$\theta = -0.8$	BIC	0.931	0.931	0.935	0.969	0.993	0.993	0.994	1.000
	MAIC	0.333	0.328	0.334	0.329	0.707	0.706	0.706	0.711
	t-sig	0.080	0.080	0.080	0.770	0.266	0.266	0.265	0.934
$\theta = -0.4$	BIC	0.395	0.401	0.401	0.466	0.876	0.875	0.883	0.906
	MAIC	0.109	0.102	0.111	0.087	0.490	0.481	0.489	0.412
	t-sig	0.038	0.039	0.038	0.234	0.121	0.121	0.121	0.704
$\theta = 0.4$	BIC	0.190	0.192	0.202	0.088	0.718	0.714	0.734	0.519
	MAIC	0.047	0.048	0.047	0.005	0.098	0.082	0.097	0.071
	t-sig	0.066	0.068	0.064	0.060	0.194	0.194	0.193	0.423
$\theta = 0.8$	BIC	0.440	0.444	0.451	0.098	0.703	0.705	0.708	0.397
	MAIC	0.103	0.101	0.102	0.003	0.238	0.224	0.239	0.035
	t-sig	0.092	0.092	0.091	0.050	0.299	0.302	0.299	0.223
AR(1) Errors									
$\rho = -0.8$	BIC	0.012	0.011	0.012	0.042	0.030	0.029	0.030	0.417
	MAIC	0.000	0.000	0.000	0.038	0.006	0.005	0.008	0.279
	t-sig	0.014	0.014	0.014	0.042	0.057	0.057	0.057	0.362
$\rho = -0.4$	BIC	0.135	0.133	0.140	0.158	0.497	0.494	0.504	0.582
	MAIC	0.057	0.052	0.056	0.050	0.322	0.295	0.324	0.326
	t-sig	0.041	0.042	0.041	0.088	0.137	0.138	0.136	0.425
$\rho = 0.4$	BIC	0.134	0.137	0.134	0.047	0.537	0.535	0.558	0.296
	MAIC	0.053	0.053	0.048	0.006	0.044	0.036	0.046	0.019
	t-sig	0.065	0.065	0.064	0.042	0.165	0.165	0.164	0.285
$\rho = 0.8$	BIC	0.193	0.203	0.196	0.036	0.330	0.330	0.342	0.123
	MAIC	0.205	0.211	0.198	0.034	0.327	0.316	0.326	0.095
	t-sig	0.087	0.088	0.085	0.038	0.103	0.104	0.103	0.132

Table 8. Size and Power when using Supremum Method for Model I; T=200
($\bar{c} = -24$ when constructing the tests and s^2)

		Size				Power			
	Criteria	MZ_α	MSB	MZ_t	ADF	MZ_α	MSB	MZ_t	ADF
<i>i.i.d.</i>	BIC	0.049	0.052	0.051	0.051	0.506	0.523	0.505	0.510
	MAIC	0.048	0.050	0.051	0.050	0.423	0.421	0.410	0.366
	t-sig	0.050	0.051	0.050	0.051	0.240	0.241	0.237	0.422
MA(1) Errors									
$\theta = -0.8$	BIC	0.861	0.862	0.859	0.900	0.989	0.989	0.989	0.994
	MAIC	0.105	0.104	0.105	0.138	0.283	0.283	0.283	0.329
	t-sig	0.414	0.415	0.416	0.672	0.756	0.756	0.756	0.945
$\theta = -0.4$	BIC	0.228	0.238	0.229	0.232	0.826	0.827	0.823	0.832
	MAIC	0.084	0.086	0.079	0.062	0.376	0.377	0.370	0.308
	t-sig	0.060	0.061	0.059	0.113	0.386	0.387	0.380	0.599
$\theta = 0.4$	BIC	0.166	0.171	0.163	0.094	0.711	0.716	0.708	0.561
	MAIC	0.096	0.098	0.091	0.035	0.386	0.387	0.374	0.207
	t-sig	0.066	0.069	0.066	0.052	0.365	0.366	0.363	0.373
$\theta = 0.8$	BIC	0.289	0.293	0.284	0.088	0.699	0.705	0.697	0.374
	MAIC	0.118	0.126	0.115	0.007	0.364	0.359	0.360	0.091
	t-sig	0.054	0.057	0.053	0.030	0.289	0.291	0.289	0.179
AR(1) Errors									
$\rho = -0.8$	BIC	0.005	0.005	0.005	0.045	0.018	0.020	0.019	0.427
	MAIC	0.000	0.000	0.000	0.040	0.013	0.012	0.012	0.275
	t-sig	0.011	0.011	0.011	0.037	0.048	0.048	0.048	0.341
$\rho = -0.4$	BIC	0.050	0.051	0.049	0.053	0.428	0.441	0.427	0.452
	MAIC	0.043	0.044	0.041	0.040	0.324	0.327	0.311	0.303
	t-sig	0.033	0.034	0.033	0.037	0.182	0.187	0.182	0.363
$\rho = 0.4$	BIC	0.103	0.108	0.102	0.049	0.565	0.578	0.565	0.390
	MAIC	0.096	0.097	0.089	0.038	0.402	0.405	0.390	0.227
	t-sig	0.051	0.053	0.052	0.044	0.301	0.306	0.299	0.349
$\rho = 0.8$	BIC	0.113	0.126	0.109	0.052	0.377	0.385	0.376	0.220
	MAIC	0.105	0.113	0.099	0.039	0.342	0.347	0.338	0.170
	t-sig	0.067	0.073	0.063	0.036	0.201	0.205	0.200	0.186

Table 9a. Empirical Results using Information Criteria to select lag k and Infimum Method to choose Break Point T_B

Series	T	Criteria	MZ_α	k	T_B	MZ_t	k	T_B	ADF	k	T_B	P_T	k	T_B	$\hat{\alpha}$
Stock Prices	100	<i>BIC</i>	-49.89 ^b	1	1941	-4.95 ^b	1	1941	-5.25 ^b	1	1937	8.92 ^b	1	1931	0.65
		<i>MAIC</i>	-49.22 ^a	1	1937	-4.93 ^a	1	1937	-5.25 ^a	1	1937	13.26 ^d	2	1931	0.65
		<i>MBIC</i>	-49.22 ^a	1	1837	-4.93 ^a	1	1937	-5.25 ^a	1	1937	13.26 ^d	2	1931	0.65
Real Wages	71	<i>BIC</i>	-39.12 ^c	1	1938	-4.37 ^c	1	1938	-4.69	1	1938	9.43 ^c	1	1940	0.61
		<i>MAIC</i>	-39.12 ^a	1	1938	-4.37 ^a	1	1938	-4.67 ^d	1	1938	11.28 ^b	1	1940	0.61
		<i>MBIC</i>	-39.12 ^a	1	1938	-4.37 ^a	1	1938	-4.67 ^c	1	1938	11.28 ^b	1	1940	0.61

We use a, b, c, d to represent rejection at 1%, 2.5%, 5%, 10% of significance level.

Table 9b. Empirical Results using Information Criteria to select Lag k and Supremum Method to choose Break Point T_B

Series	T	Criteria	MZ_α	MZ_t	ADF	k	T_B	$\hat{\alpha}$
Stock Prices	100	<i>BIC</i>	-33.10 ^b	-4.05 ^b	-4.32 ^c	1	1931	0.73
		<i>MAIC</i>	-20.25 ^d	-3.21 ^d	-3.38	2	1931	0.77
		<i>MBIC</i>	-20.25 ^c	-3.21 ^c	-3.38	2	1931	0.77
Real Wages	71	<i>BIC</i>	-27.94 ^c	-3.67 ^c	-3.86 ^d	1	1933	0.69
		<i>MAIC</i>	-27.94 ^a	-3.67 ^b	-3.86 ^c	1	1933	0.69
		<i>MBIC</i>	-27.94 ^a	-3.67 ^b	-3.86 ^c	1	1933	0.69

We use a, b, c, d to represent rejection at 1%, 2.5%, 5%, 10% of significance level.

Table 10a. Empirical Results using Recursive Method to select Lag k and Infimum Method to choose T_B

Series	T	MZ_α	k	T_B	MZ_t	k	T_B	ADF	k	T_B	P_T	k	T_B	$\hat{\alpha}$
Stock Prices	100	-143.10 ^c	3	1948	-8.43 ^c	3	1948	-5.25 ^b	1	1936	5.94 ^d	3	1930	0.65
Real Wages	71	-11628.50 ^a	4	1941	-76.24 ^a	4	1941	-4.69 ^d	1	1938	4.05 ^c	3	1940	0.61

We use a, b, c, d to represent rejection at 1%, 2.5%, 5%, 10% of significance level.

Table 10b. Empirical Results using Recursive Method to select Lag k and supremum Method to choose T_B

Series	T	MZ_α	MZ_t	ADF	k	T_B	$\hat{\alpha}$
Stock Prices	100	-49.87 ^c	-4.98 ^c	-3.99 ^d	3	1930	0.73
Real Wages	71	-27.94	-3.67	-3.86	1	1933	0.69

We use a, b, c, d to represent rejection at 1%, 2.5%, 5%, 10% of significance level.

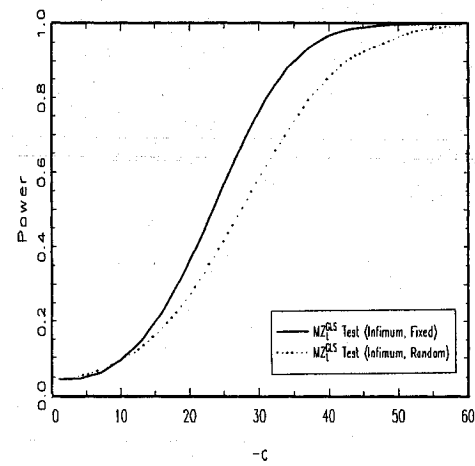
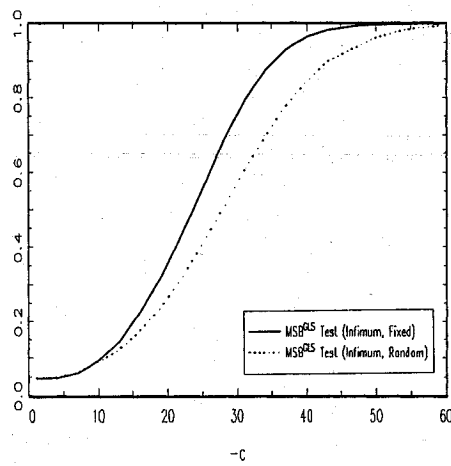
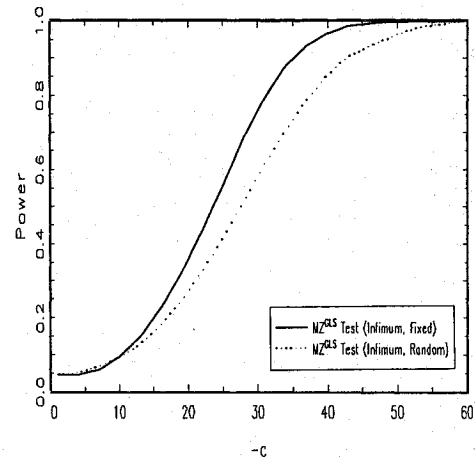
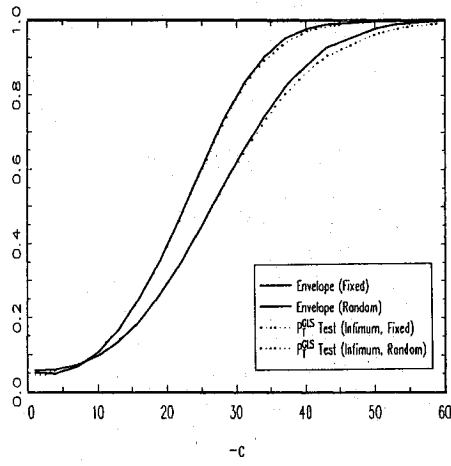


Figure 1. Gaussian Power Envelope and Asymptotic Power Functions; Infimum Method and Fixed and Random Initial Condition.

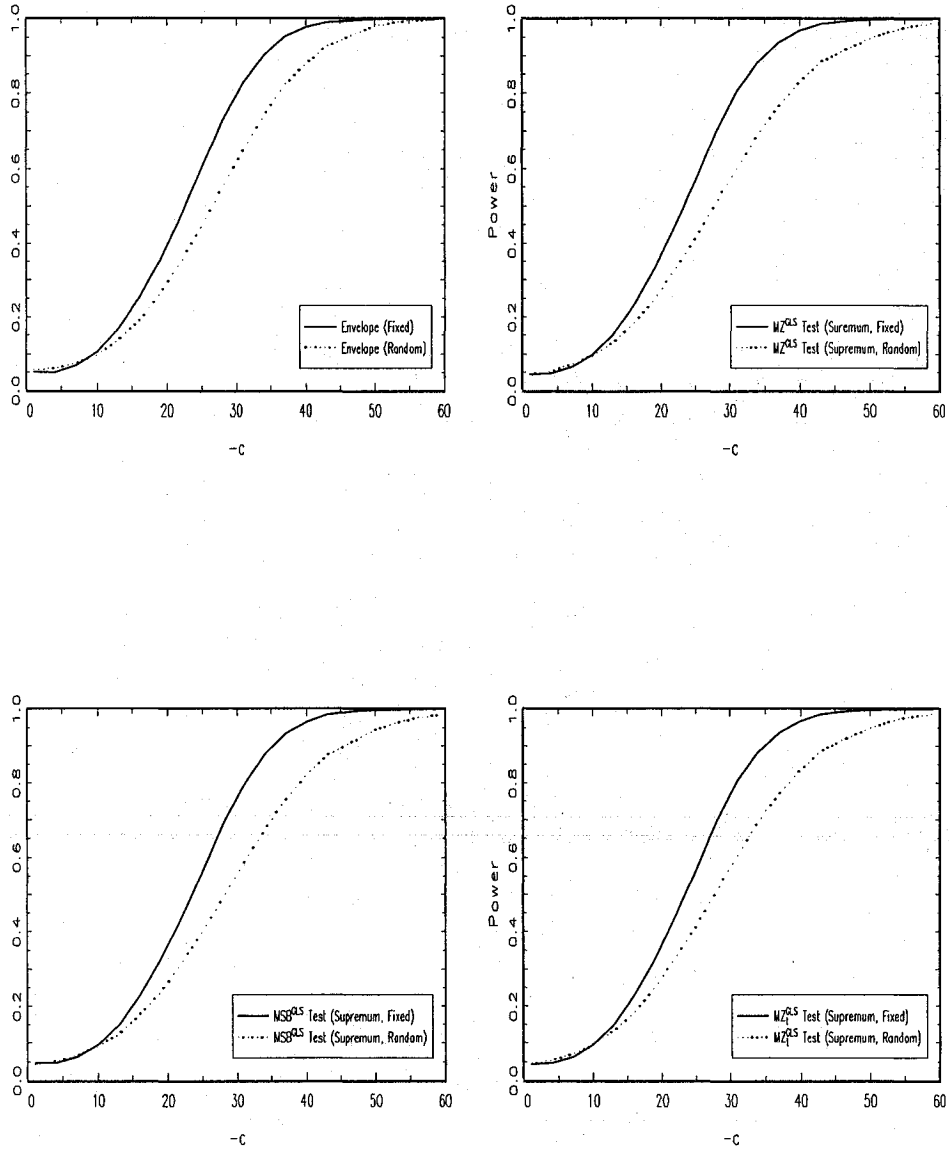


Figure 2. Gaussian Power Envelope and Asymptotic Power Functions; Supremum Method and Fixed and Random Initial Condition.

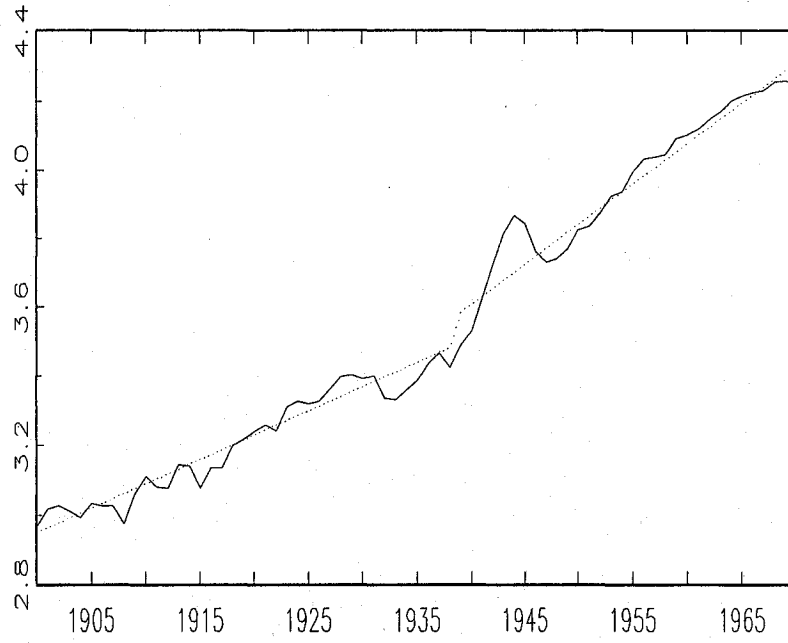


Figure 3. Logarithmes of Real Wages with a broken time trend; 1900-1970.

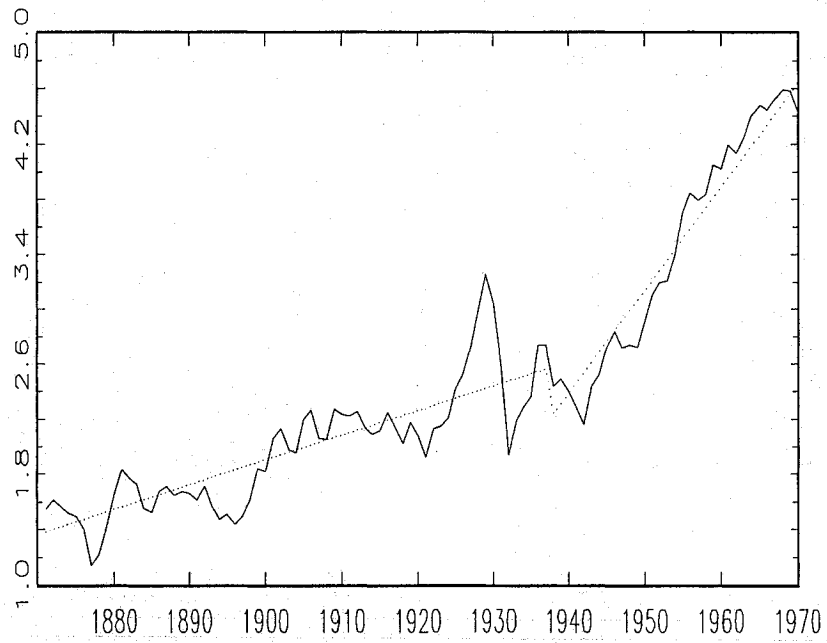


Figure 4. Logarithmes of Stock Prices with a broken time trend; 1871-1971.

Chapter 3

Unit Root Test Using GLS Detrended Data and Covariates in Structural Change Models

May 7, 2006

1 Introduction

There has been large literature on how to improve the power of unit root tests. Hansen (1995) showed that it can be costly by ignoring the information in related time series by using univariate unit root tests, as power gains can be achieved by including correlated stationary covariates in the regression equation. He reconstructed an ADF statistic (Dickey and Fuller, 1979; Said and Dickey, 1984) using multivariate framework. Comparing the power performance of the CADF (covariate ADF) and the ADF tests, he found substantial power gains in both large and small samples.

Recently, Elliott and Jansson (EJ 2003, hereafter) generalized the approach of Hansen (1995) by including different sets of deterministic component and an arbitrary number of covariates. They derived a family of multivariate point optimal tests and a feasible test when nuisance parameters are unknown. Their results have confirmed the importance of covariates in improving the power of unit root tests. In fact, the limiting distributions of the feasible point optimal tests are dependent on the long-run correlation between the covariates and the quasi-differences of the potentially integrated time series. The higher the correlation coefficient between the covariates and the quasi-differences, the larger improvements in the power of unit root tests.

This paper contributes to the existing literature in two ways. First it borrows the approach of EJ (2003) and applies it to the unknown structural change models. Two types of models are considered following Perron and Rodríguez (PR 2003, hereafter). Type I Model only includes a break in the slope of the trend function, while type II model considers a break in both intercept and time trend. The date of the break point is treated as unknown in both models¹. By introducing an arbitrary number of covariates to a family of P-tests derived in PR (2003), this paper will show that the tests of two types of models will converge to the same limiting distribution. In other words, the rule that a break in intercept is a slowly evolving deterministic component still applies even when covariates are present. The limiting distributions of the P test in PR (2003) are retrieved when the correlation

¹An ADF in the spirit of Hansen (1995) has been proposed by Shin (2001) but dealing with known break dates.

coefficient between covariates and the quasi-differences of the potentially integrated time series is zero. When the correlation is between zero and one, asymptotic and finite sample simulations show that large power gains are achieved when using covariates.

Second, this experiment allows us to derive a feasible test when nuisance parameters are unknown. It takes into account one additional nuisance parameter compared to that of EJ (2003) since the structural break is unknown. It will show that, although different from that of EJ (2003), the feasible test proposed here still achieves the power envelope asymptotically.

The rest of the paper is organized as follows. Section 2 defines the model and derives the limiting distribution of the test. Section 3 presents the feasible test when all the nuisance parameters are unknown. Section 4 presents results obtained from simulations and evaluates the test performance in both large and small samples. Section 5 concludes. Proofs are contained in the appendix.

2 The Model and Asymptotic Theory

The data generating process (DGP) considered is of the following form:

$$y_t = d_{yt} + u_{yt}, \quad (1)$$

$$x_t = d_{xt} + u_{xt}, \quad (2)$$

$$A(L) \begin{pmatrix} [1 - \rho L]u_{y,t} \\ u_{x,t} \end{pmatrix} \equiv A(L) u_t(\rho) = e_t \quad (3)$$

$$\rho = 1 + cT^{-1} \quad (4)$$

where (4) is the local to unity framework established in Phillips (1987), Chan and Wei (1987). $A(L)$ is a matrix polynomial of finite order k and L the lag operator, $t = 1, 2, \dots, T$; x_t , an $m \times 1$ vector, includes an arbitrary number of stationary covariates containing extra information of y_t (a scalar), the variable to be tested. The deterministic component of y_t is denoted as d_{yt} . Two cases are considered following PR (2003). The first case includes an

intercept, a time trend and a break in the slope of the trend function. Hence $d_{yt} = \beta_1 + \beta_3 t + \beta_5 (t - T_B) \mathbf{1}(t > T_B)$, where $\mathbf{1}(\cdot)$ is the indicator function and T_B the unknown break date. The second case considers an intercept, a time trend, and a break in both intercept and the slope. Therefore $d_{yt} = \beta_1 + \beta_3 t + \beta_5 (t - T_B) \mathbf{1}(t > T_B) + \beta_6 \mathbf{1}(t > T_B)$. The deterministic component of covariates include an intercept and a time trend in both cases, that is, $d_{xt} = \beta_2 + \beta_4 t$. Notice that the first case can be obtained by setting $\beta_6 = 0$ of the second case. Therefore Case 1 and 2 can be simplified using the restriction $[I_{2(m+2)} - S_i] \beta^i = 0$, where $i = 1, 2$, $S_1 = \begin{bmatrix} I_{2m+3} & 0 \\ 0 & 0 \end{bmatrix}$, and $S_2 = I_{2(m+2)}$.

Following EJ (2003), we have the following assumptions for the model: i) $|A(z)| = 0$ has roots outside of the unit circle; ii) $E_{t-1}(e_t) = 0$, $E_{t-1}(e_t e_t') = \Sigma$ and $\sup_t E \|e_t\|^{2+\eta} < \infty$ (a.s.) for some $\eta > 0$, where Σ is positive definite and $E_{t-1}(\cdot)$ denotes conditional expectation with respect to $\{e_{t-1}, e_{t-2}, \dots\}$; iii) u_0, u_1, \dots, u_{t-k} , are $O_p(1)$.

Furthermore, using the same notation as in EJ (2003), the expression

$$\Omega = A(1)^{-1} \Sigma A(1)^{-1'} = \begin{bmatrix} \omega_{yy} & \omega_{yx} \\ \omega'_{yx} & \Omega_{xx} \end{bmatrix} \quad (5)$$

is defined as the spectral density at the frequency zero (scaled by 2π) of $u_t(\rho)$. Therefore $R^2 = \omega_{yy}^{-1} \omega_{yx} \Omega_{xx}^{-1} \omega'_{yx}$ is a measure of the long-run correlation between shocks to x_t and quasi-differences of y_t at the frequency zero. The value of R^2 represents the contribution of covariates to the explanation of y_t , and the value of R^2 ranges from zero to unity. When $R^2 = 0$, covariates are irrelevant to the potentially integrated series and we have the traditional case where no covariates are present in testing for unit roots. When $R^2 = 1$, there is perfect correlation between covariates and quasi-differences of y_t , and the partial sums of x_t cointegrates with y_t^2 . In most cases, R^2 ranges from zero to one, and we are concerned with the performance of unit root tests as R^2 increases.

²If there exists cointegration, the model should be set up using another framework, unless the coefficients of the cointegrating vector are assumed to be known, which is the case analyzed by Elliott, Jansson and Pesavento (2005). Therefore, the range of correlation is $0 \leq R^2 < 1$ in this paper.

To summarize, we study a VAR model including covariates x_t and quasi-differences of y_t . Our interest is to test whether the parameter ρ is equal to unity, or in other words, whether y_t has a unit root under the null against the alternative that y_t has a root less than unity. Using an univariate framework, King (1980, 1988) and Dufour and King (1991) showed that the unit root hypothesis can be examined using Neyman-Pearson tests in a univariate framework. ERS (1996) further developed a family of point optimal tests using quasi-differenced data without considering structural break. PR (2003) extended this type of statistics to models with an unknown structural change. In this paper, we extend them using a multivariate framework originally proposed by Hansen (1995).

We first derive an optimal statistic (denoted by P test) assuming all the nuisance parameters are known except the date of structural break. Accordingly we have the following assumptions: i) $A(L) = I$, therefore $\Omega = \Sigma$; ii) e_t is normally distributed with $u_{y_0} = 0$; iii) $T_B = T\lambda$ for some $\lambda \in (0, 1)$. Based on these assumptions, we test the null hypothesis of $c = 0$ against the local alternative that $c = \bar{c} < 0$ where $c = T(\rho - 1)$ and $\bar{c} = T(\bar{\rho} - 1)$. PR (2003) showed that when the break point is unknown, the limiting distribution of the P test can be defined by:

$$P^i(1, \bar{\rho}) = \inf_{\lambda \in [0, 1]} \sum_{t=1}^T \hat{u}_t^i(\bar{\rho}, \lambda)' \Sigma^{-1} \hat{u}_t^i(\bar{\rho}, \lambda) - \inf_{\lambda \in [0, 1]} \sum_{t=1}^T \hat{u}_t^i(1, \lambda)' \Sigma^{-1} \hat{u}_t^i(1, \lambda) - \bar{c}. \quad (6)$$

where in the present case, $\hat{u}_t^i(r) = z_t(r) - d_t(r)' \hat{\beta}^i$ and $r = \bar{\rho}, 1$. When $t > 1$, we have:

$$\begin{aligned} z_t(r) &= [(1-rL)y_t, x_t']' \\ d_t(r)' &= \begin{bmatrix} 1-r & 0 & (1-rL)t & 0 & (1-rL)(t-T_B)\mathbf{1}(\cdot) & (1-r)\mathbf{1}(\cdot) \\ 0 & I_m & 0 & I_m t & 0 & 0 \end{bmatrix} \end{aligned}$$

when $t = 1$, we have:

$$\begin{aligned} z_1(r) &= [y_1, x_1']' \\ d_1(r)' &= \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & I_m & 0 & I_m & 0 & 0 \end{bmatrix}. \end{aligned}$$

Furthermore,

$$\hat{\beta}^i = [S_i(\sum_{t=1}^T d_t(r) \Sigma^{-1} d_t(r)') S_i]^{-1} [S_i \sum_{t=1}^T d_t(r) \Sigma^{-1} z_t(r)],$$

where D^- is the Moore Penrose inverse of D . The next Theorem establishes the limiting distribution of the statistic.

Theorem 1 *Let $\{y_t\}$ and $\{x_t\}$ be generated by (1) to (4), with $A(L) = I$, $e_t \sim i.i.d. N(0, \Sigma)$, assumptions (i) to (iii) hold. To test the null hypothesis of $\rho = 1$ ($c = 0$) against the alternative hypothesis of $\rho = \bar{\rho} = 1 + \bar{c}/T$ with $c = \bar{c}$ fixed as $T \rightarrow \infty$, the P-test has the following asymptotic distribution for Case $i = 1$ and 2 :*

$$P^i(1, \bar{\rho}) \Rightarrow \Lambda_1(c, \bar{c}, R^2) + \Lambda_2^i(c, \bar{c}, \lambda, R^2) - \bar{c} \quad (7)$$

where $\Lambda_1(c, \bar{c}, R^2) = (\bar{c}^2 - 2c\bar{c})(1 + \frac{R^2}{1-R^2}) \int W_{1c}^2 - 2\bar{c} \int W_{1c} dW_1 + 2\bar{c} \frac{R}{\sqrt{1-R^2}} \int W_{1c} dW_2$, W_1 and W_2 are independent univariate standard Brownian motions. W_{1c} is an Ornstein-Uhlenbeck process defined by $W_{1c}(s) = c \int_0^s e^{c(s-h)} W_1(h) dh + W_1(s)$. The expression $\Lambda_2(c, \bar{c}, \lambda, R^2)$ is defined in the appendix.

The above limiting distribution has several indications. First, the asymptotic power depends on \bar{c} , which corresponds to one particular point under the alternative hypothesis. This means that the asymptotic power functions are tangent to the highest power envelope at point $c = \bar{c}$. Second, the limiting distribution is nonstandard. Third, the distribution of the P-test also depends on the parameter R^2 . When $R^2 = 0$, there is no covariate correlated with the quasi-differences of y_t and consequently we retrieve the same asymptotic distribution as that derived in PR (2003). When R^2 is greater than zero, the limiting distribution is a function of R^2 , indicating that extra information contained in the covariates may make a difference on the performance of the test.

Using Theorem 1, $T = 1,000$, and 10,000 replications, the asymptotic power functions and power envelopes at different R^2 can be calculated. The simulation results for $R^2 = 0.0, 0.3, 0.5, 0.7, 0.9$ are graphed in Figure 1. We observe that the power envelopes reach their lower bound when

$R^2 = 0$. When R^2 is greater than zero, the power attainable increases considerably. For example, when $c = -10$, the maximum power of $R^2 = 0.0, 0.3, 0.5, 0.7, 0.9$ are 10.61%, 21.27%, 35.14%, 55.33%, and 85.95% respectively. That is, the power increases by 100% from $R^2 = 0.0$ to $R^2 = 0.3$; by 65.2% from $R^2 = 0.3$ to 0.5; by 57.5% from $R^2 = 0.5$ to 0.7; and by 55.3% from $R^2 = 0.7$ to 0.9. These evidence suggests that incorporating the information in covariates can substantially increase the power of unit root test in large sample. This result is consistent with what has been found in Hansen (1995) and EJ (2003).

3 A Feasible Test

In the previous section, we assumed all the nuisance parameters except the break point are known and the errors are normally distributed. In practice, R^2 , the correlation between x_t and the variable to be tested is not known, and it is likely that there is serial correlation in the time series. Therefore it is necessary to develop a feasible test to estimate all the unknown nuisance parameters. In this section, we derive this test by adopting the approach of EJ (2003) in the first three steps when constructing the test, and we modify the last step in order to estimate the break point. The construction of the test is as follows:

1. Estimate R^2 . It is performed by running a VAR model: $A(L)z_t(1, \lambda) = d_t + e_t$, where λ is the break point, d_t is the deterministic component specified according to the model under analysis. Then we use the residuals generated from the VAR model to construct $\hat{\Sigma} = T^{-1} \sum_{t=k+1}^T \hat{e}_t(\lambda) \hat{e}_t(\lambda)'$, $\hat{\Omega} = \hat{A}(1, \lambda)^{-1} \hat{\Sigma} \hat{A}(1, \lambda)^{-1'}$, and obtain $\hat{R}^2 = \hat{\omega}_{yx} \hat{\Omega}_{xx}^{-1} \hat{\omega}'_{yx} / \hat{\omega}_{yy}$ where $\hat{A}(1, \lambda) = I + \sum_{t=1}^k A_t$ is the $(i+1)th$ matrix element of $A(L)$.
2. Estimate the nuisance parameters for quasi-differencing and detrend the data under the null and alternative hypothesis, respectively. That is, construct $\tilde{u}_t^i(r, \lambda) = z_t(r, \lambda) - d_t(r, \lambda)' \tilde{\beta}^i(r, \lambda)$ when $r = 1, \rho$, and

$$\tilde{\beta}^i(r, \lambda) = \left[\sum_{t=1}^T d_t(r, \lambda) \hat{\Omega}^{-1} d_t(r, \lambda)' \right]^{-1} \left[\sum_{t=1}^T d_t(r, \lambda) \hat{\Omega}^{-1} z_t(r, \lambda) \right].$$

3. Run the VAR model $\tilde{A}(L)\tilde{u}_t^i(r, \lambda) = \tilde{e}_t^i(r, \lambda)$, using $\tilde{u}_t^i(r, \lambda)$ obtained from the previous step, and construct the variance-covariance matrices

$$\tilde{\Sigma}^i(r, \lambda) = T^{-1} \sum_{t=k+1}^T \tilde{e}_t^i(r, \lambda) \tilde{e}_t^i(r, \lambda)'$$

4. Select the break point and construct the test statistic. The break point is selected by choosing the variance-covariance matrices $\tilde{\tilde{\Sigma}}(1)$ and $\tilde{\tilde{\Sigma}}(\bar{\rho})$ minimizing $\sum_{t=k+1}^T \tilde{e}_{y,t}^i(r, \lambda) \tilde{e}_{y,t}^i(r, \lambda)'$ for $r = 1, \bar{\rho}$, respectively (see appendix for proof). Then we can construct the test statistic

$$\tilde{P}^i(1, \bar{\rho}) = T \{ \text{tr}[\tilde{\tilde{\Sigma}}(1)^{-1} \tilde{\tilde{\Sigma}}(\bar{\rho})] - [m + \bar{\rho}] \}. \quad (8a)$$

This test has the asymptotic power that achieves the power envelope bound at \bar{c} under assumptions (i)-(iii). The following Theorem establishes its limiting distribution.

Theorem 2 *Let $\{y_t\}$ and $\{x_t\}$ be generated by (1) to (4), assumptions (i) to (iii) hold, also assuming deterministic components are correctly specified for $i = 1, 2$. To test the null hypothesis of $\rho = 1$ ($c = 0$) against the alternative hypothesis of $\rho = \bar{\rho} = 1 + \bar{c}/T$ with $c = \bar{c}$ fixed as $T \rightarrow \infty$, the limiting distribution of the test defined in (8a) has asymptotic distribution:*

$$\tilde{P}^i(1, \bar{\rho}) \Rightarrow \Lambda_1(c, \bar{c}, R^2) + \Lambda_2^i(c, \bar{c}, \lambda, R^2) - \bar{c}. \quad (9)$$

As what we have observed before, the limiting distributions of the feasible test and hence the asymptotic results depend on the selection of \bar{c} for detrending and R^2 . We have shown that the feasible test asymptotically achieves the highest possible power at \bar{c} . Ideally we should calculate a power envelope corresponding to each R^2 and our choice of \bar{c} for quasi-differencing should be dependent on R^2 . However, in practice we may use $\bar{c} = -22.5$, which is the same value selected by PR (2003) when $R^2 = 0$, for GLS detrending. Because as R^2 increases, the power of the test increases dramatically and the choice of \bar{c} becomes less important; see EJ, 2003.

We present the asymptotic critical values in Table 1. The simulations of the limiting distribution are done by using 1,000 steps to approximate the Wiener process on $[0, 1]$ as the partial sums of *i.i.d.* $N(0, 1)$ random variables. As a comparison, we also present finite sample critical values for both cases analyzed in the paper. We use $T = 100$ and 200 with data generated by a random walk with zero initial value and *i.i.d.* $N(0, 1)$ errors. The lag length is set to zero in the estimation of the VAR model in step one.

4 Evaluation of the Test

Figure 2 graphs the asymptotic power of the feasible test for both cases and $R^2 = 0.0, 0.3, 0.5, 0.7, 0.9$. We first observe that the feasible test has power tangent to the power envelope at one point and never falls far below it. This suggests that very little asymptotic power is lost at points away from the tangent point, especially for lower values of R^2 . Overall, substantial power gains are realized due to covariates. Consider the asymptotic power gains from using covariates when $R^2 = 0.5$ and the local alternative $c = -10$. The asymptotic power rises by 217% (power is 34.52% when $R^2 = 0.5$, and 10.86% when $R^2 = 0.0$). Using the local alternative $c = -4$, power gains are smaller, but still an increase of approximately 87% (power is 8.98% when $R^2 = 0.5$, while it is 4.8% when $R^2 = 0.0$).

The test is also evaluated using small sample Monte Carlo analysis. To simulate the critical values, power and size, $\{y_t\}$ and $\{x_t\}$ are generated using (1) to (4) with normal errors and known variance-covariance matrix, whose variances equal to one and covariances equal to 0.3, 0.5, 0.7, 0.9 respectively. Using 10,000 replications, $T = 100, 200$, finite sample critical values are calculated and tabulated in Table 1 along with the asymptotic critical values. Power and size are presented in Table 2 (Case 1) and Table 3 (Case 2). There are two factors that influence the finite sample power performance. One is the number of observations, power increases as T increases. For example, when $\rho = 0.90$, $R^2 = 0.81$, the power increases 15% when T increases from 100 to 200 for Case 1 and 26% for Case 2. Another factor is the increase of correlation between the covariate and the variable to be tested. As R (or R^2) is closer to 1, power increases dramatically. For example, power

increases 19% as R increases from 0.3 to 0.5 in Table 2 when $\rho = 0.90$.

In terms of the overall power gains in finite sample, we have the following observations. For Case 1, when $T = 100$ and $\rho = 0.96$, the test has a power of 7.36% without covariates, while power increases to 11.57% when $R^2 = 0.49$, approximately a 57.2% power gain; while for Case 2, approximately a 52.4% power gain. For Case 1, when $T = 200$, $\rho = 0.96$, the power is 11.87% without covariates, while it increases to 29.19% when $R^2 = 0.49$, approximately a 146% power gain; while for Case 2, approximately a 120% power gain.

5 Conclusions

Following previous research (Hansen 1995; EJ 2003; Shin 2001; PR 2003), we introduce an arbitrary number of covariates to unknown structural change models. We have derived the test statistic that has asymptotic power function tangent to the power envelope at one point under the alternative hypothesis. We have found that the proposed test is a function of both \bar{c} and R^2 (the correlation of covariates and quasi-differences of y_t). Large sample simulations have shown substantial power gains when R^2 is greater than zero, especially when R^2 is close to 1. Monte Carlo simulations have also shown dramatic power improvements in finite samples without size deterioration. To be of more empirical relevance, we have developed a feasible test when there are serial correlations, and when R^2 and break point are unknown. Our study has shown that this test achieves the power envelope when $T \rightarrow \infty$.

References

- [1] Chan, N. H. and C. Z. Wei (1987), "Asymptotic Inference for Nearly Nonstationary AR(1) Processes," *Annals of Statistics* **15**, 1050-63.
- [2] Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* **74**, 427-431.
- [3] Dufour, J.-M. and M. King (1991), "Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary Errors," *Journal of Econometrics* **47**, 115-143.
- [4] Elliott, G. and M. Jansson (2003), "Testing for Unit Roots with Stationary Covariate," *Journal of Econometrics* **115**, 75-89.
- [5] Elliott, G., M. Jansson, and E. Pesavento (2005), "Optimal Power for Testing Potential Cointegration Vectors with Known Parameters for Nonstationarity," *Journal of Business & Economic Statistics* **23**, 34-48.
- [6] Hansen, B. E. (1995), "Rethinking the Univariate Approach to Unit Root Testing," *Econometric Theory* **11**, 1148-1171.
- [7] King, M. L. (1980), "Robust Test for Spherical Symmetry and Their Application to Least Squares Regression," *Annals of Statistics* **8**, 1265-1271.
- [8] King, M. L. (1988), "Towards a Theory of Point Optimal Testing," *Econometrics Review* **6**, 169-218.
- [9] Perron, P. and G. Rodríguez (2003), "GLS Detrending, Efficient Unit Root Tests and Structural Change," *Journal of Econometrics* **115**, 1-27.
- [10] Phillips, P. C. B. (1987), "Time Series Regression with Unit Roots," *Econometrica* **55**, 277-302.

- [11] Said, S. E. and D. A. Dickey (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika* **71**, 599-608.
- [12] Shin, S. (2001), "Covariate Unit Root Test Allowing for a Structural Change," Department of Economics Working Paper #01-14, The Ohio State University.

Appendix

Proof of Theorem 1. We derive the limiting distribution of the P-test for Case $i = 1, 2$. The P-test statistic is given by

$$P^i(1, \bar{\rho}) = \inf \sum_{t=1}^T \hat{u}_t^i(\bar{\rho})' \Sigma^{-1} \hat{u}_t^i(\bar{\rho}) - \inf \sum_{t=1}^T \hat{u}_t^i(1)' \Sigma^{-1} \hat{u}_t^i(1) - \bar{c},$$

where $\hat{u}_t^i(r) = z_t(r) - d_t(r)' \hat{\beta}^i$ for $r = \bar{\rho}, 1$.

When $t > 1$,

$$z_t(r) = \begin{bmatrix} (1-rL)y_t \\ x_t \end{bmatrix},$$

$$d_t(r)' = \begin{pmatrix} 1-r & 0 & (1-rL)t & 0 & (1-rL)(t-T_B)\mathbf{1}(\cdot) & (1-r)\mathbf{1}(\cdot) \\ 0 & I_m & 0 & I_m t & 0 & 0 \end{pmatrix};$$

while when $t = 1$,

$$z_1(r) = \begin{bmatrix} y_1 \\ x_1 \end{bmatrix},$$

$$d_1(r)' = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & I_m & 0 & I_m & 0 & 0 \end{pmatrix},$$

where $\mathbf{1}(\cdot)$ is the indicator function. Using OLS,

$$\hat{\beta}^i = [S_i (\sum_{t=1}^T d_t(r) \Sigma^{-1} d_t(r)') S_i']^{-1} [S_i \sum_{t=1}^T d_t(r) \Sigma^{-1} z_t(r)].$$

Therefore the P-test has the following form:

$$P^i(1, \bar{\rho}) = \sum_{t=1}^T e_t(\bar{\rho})' \Sigma^{-1} e_t(\bar{\rho}) - \sum_{t=1}^T e_t(1)' \Sigma^{-1} e_t(1) - \bar{c}$$

$$\inf [S_i N_T(1)]' [S_i D_T(1) S_i]^{-1} [S_i N_T(1)] -$$

$$\inf [S_i N_T(\bar{\rho})]' [S_i D_T(\bar{\rho}) S_i]^{-1} [S_i N_T(\bar{\rho})]$$

where $D_T(r) = \sum_{t=1}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} d_t(r)' \Psi_T^{-1}]$, $N_T(r) = \sum_{t=1}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} e_t(r)]$, Ψ_T is a scaling matrix. Because $s_1 \Sigma^{-1} s_1' = (1 + \bar{\delta}' \bar{\delta}) \omega_{yy}^{-1}$ and $s_1 \Sigma^{-1/2'} =$

$$\lim_{T \rightarrow \infty} \left[\sup_{\frac{2}{T} \leq s \leq 1} T^{1/2} \Psi_T^{-1} d_{[Ts]}(r) \Sigma^{-1/2'} \right] \Rightarrow \begin{pmatrix} 0 & 0 \\ 0 & I_m \\ 1 - c_r s & -(1 - c_r s) \bar{\delta}' \\ 0 & s I_m \\ 1 - c_r (s - \lambda) & -\bar{\delta}' [1 - c_r (s - \lambda)] \\ 0 & 0 \end{pmatrix}, \quad (\text{A.2})$$

where $\lambda = T_B/T$ is the break point. Combining the above results (A.1) and (A.2) we have

$$\begin{aligned} D_T(r) &= \sum_{t=1}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} d_t(r)' \Psi_T^{-1}] \\ &= \Psi_T^{-1} d_1(r) \Sigma^{-1} d_1(r)' \Psi_T^{-1} + \sum_{t=2}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} d_t(r)' \Psi_T^{-1}] \\ &\Rightarrow \begin{bmatrix} 1 + \bar{\delta}' \bar{\delta} & 0 & 0 \\ 0 & I_m & -(1 - c_r/2) \bar{\delta} \\ 0 & -(1 - c_r/2) \bar{\delta}' & (1 + c_r^2/3 - c_r)(1 + \bar{\delta}' \bar{\delta}) \\ 0 & I_m/2 & -(1/2 - c_r/3) \bar{\delta} \\ 0 & \bar{\delta}' a & (1 + \bar{\delta}' \bar{\delta}) v \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ I_m/2 & \bar{\delta} a & 0 \\ -(1/2 - c_r/3) \bar{\delta} & (1 + \bar{\delta}' \bar{\delta}) v & 0 \\ I_m/3 & \bar{\delta} b & 0 \\ \bar{\delta}' b & (1 + \bar{\delta}' \bar{\delta}) d & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (\text{A.3}) \end{aligned}$$

where

$$\begin{aligned} a &= 1 - \lambda - c_r \lambda + c_r \lambda^2/2, \\ b &= -\frac{1 - \lambda^2}{2} + \frac{c_r}{3} + \frac{c_r \lambda^3}{6} - \frac{c_r \lambda}{2}, \\ v &= 1 - \lambda - c_r + c_r \lambda - c_r^2 \lambda/2 + c_r^2 \lambda^3/2 + c_r^2 (1 - \lambda^3)/3, \\ d &= 1 - \lambda - c_r - c_r \lambda^2 + 2c_r \lambda + c_r^2/3 - c_r^2 \lambda^3/3 - c_r^2 \lambda + c_r^2 \lambda^2. \end{aligned}$$

Similarly, we can calculate the limiting distribution for $N_T(r) = \sum_{t=1}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} e_t(r)]$,

$$\begin{aligned}
N_T(r) &= \sum_{t=1}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} e_t(r)] \\
&= \sum_{t=1}^T \Psi_T^{-1} [d_t(r) \Sigma^{-1/2'}] [\Sigma^{-1/2} e_t(r)] \\
&= \Psi_T^{-1} [d_1(r) \Sigma^{-1/2'}] [\Sigma^{-1/2} e_t(r)] + \sum_{t=2}^T \Psi_T^{-1} [d_t(r) \Sigma^{-1/2'}] [\Sigma^{-1/2} e_t(r)] \\
&= \Psi_T^{-1} [d_1(r) \Sigma^{-1/2'}] [\Sigma^{-1/2} e_t(r)] + \\
&\quad T^{-1/2} \sum_{t=2}^T [T^{1/2} \Psi_T^{-1} d_t(r) \Sigma^{-1/2'}] [\Sigma^{-1/2} e_t(r)] \tag{A.4}
\end{aligned}$$

$$\Rightarrow \begin{pmatrix} \epsilon_{y,1} - \bar{\delta}' \epsilon_{x,1} \\ \epsilon_{x,t} + (\rho - r) \Sigma^{-1/2} s_1' u_{y,t-1} \\ \int_0^1 (1 - c_r s) [dW_1(s) - \bar{\delta}' V(s)] + \\ (c - c_r)(1 + \bar{\delta}' \bar{\delta}) \int_0^1 (1 - c_r s) W_{1c}(s) ds \\ \int_0^1 s dV(s) - (c - c_r) \bar{\delta} \int s W_{1c}(s) ds \\ G \\ 0 \end{pmatrix}, \tag{A.5}$$

where

$$\begin{aligned}
G &= \int_{\lambda}^1 (1 - c_r s) [dW_1(s) - \bar{\delta}' V(s)] + (c - c_r)(1 + \bar{\delta}' \bar{\delta}) \int_{\lambda}^1 (1 - c_r s) W_{1c} ds \\
&\quad + c_r \lambda \int_{\lambda}^1 d[W_1(s) - \bar{\delta}' V(s)] + c_r \lambda (c - c_r)(1 + \bar{\delta}' \bar{\delta}) \int_{\lambda}^1 W_{1c}(s) ds.
\end{aligned}$$

Thus the limiting distribution for Case 2 is:

$$P^i[1, \bar{\rho}] \Rightarrow \Lambda_1(c, \bar{c}, R^2) + \Lambda_2^i(c, \bar{c}, R^2) - \bar{c}$$

where expression $\Lambda_2^i(c, \bar{c}, R^2)$ is the limiting distribution of $\inf [S_i N_T(1)]' [S_i D_T(1) S_i]^{-1} [S_i N_T(1)] - \inf [S_i N_T(\bar{\rho})]' [S_i D_T(\bar{\rho}) S_i]^{-1} [S_i N_T(\bar{\rho})]$, which can be calculated using (A.3) and (A.5). Proof for Case 2 completes.

Next we want to show that Case 1 has the same limiting distribution as Case 2. For Case 1,

$$\begin{aligned} D_T(r) &= \sum_{t=1}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} d_t(r)' \Psi_T^{-1}] \\ &= \Psi_T^{-1} d_1(r) \Sigma^{-1} d_1(r)' \Psi_T^{-1} + \sum_{t=2}^T [\Psi_T^{-1} d_t(r) \Sigma^{-1} d_t(r)' \Psi_T^{-1}] \end{aligned}$$

When $t > 1$,

$$d_t(r)' = \begin{pmatrix} 1-r & 0 & (1-rL)t & 0 & (1-rL)(t-T_B) \mathbf{1}(\cdot) & 0 \\ 0 & I_m & 0 & I_m t & 0 & 0 \end{pmatrix},$$

and

$$\lim_{T \rightarrow \infty} \left[\sup_{\frac{2}{T} \leq s \leq 1} T^{1/2} \Psi_T^{-1} d_{[Ts]}(r) \Sigma^{-1/2'} \right] \Rightarrow \begin{pmatrix} 0 & 0 \\ 0 & I_m \\ 1-c_r s & -(1-c_r s) \bar{\delta}' \\ 0 & s I_m \\ 1-c_r(s-\lambda) & -\bar{\delta}' [1-c_r(s-\lambda)] \\ 0 & 0 \end{pmatrix}.$$

Therefore, $D_T(r)$ and $N_T(r)$ have the same asymptotic distributions for Case 1 and 2, and so is the limiting distribution for P test. Proof completes ■

Proof of Theorem 2. Following EJ (2003),

1. $\sum_{t=k+1}^T \tilde{e}_t^i(\bar{\rho}) \tilde{e}_t^i(\bar{\rho})' - \sum_{t=k+1}^T \tilde{e}_t^i(1) \tilde{e}_t^i(1)' = \sum_{t=k+1}^T \hat{e}_t^i(\bar{\rho}) \hat{e}_t^i(\bar{\rho})' - \sum_{t=k+1}^T \hat{e}_t^i(1) \hat{e}_t^i(1)' + o_p(1)$, where $\hat{e}_t^i(r) = A(L) \tilde{u}_t^i(r)$.
2. $\sum_{t=k+1}^T \hat{e}_t^i(r)' \Sigma^{-1} \hat{e}_t^i(r) - \sum_{t=k+1}^T e_t(r)' \Sigma^{-1} e_t(r) \Rightarrow -[S_i N(c, \bar{c}, R^2)]' [S_i D(\bar{c}, R^2) S_i]^{-1} [S_i N(c, \bar{c}, R^2)]$,
3. $\sum_{t=k+1}^T e_t(\bar{\rho})' \Sigma^{-1} e_t(\bar{\rho}) - \sum_{t=k+1}^T e_t(1)' \Sigma^{-1} e_t(1) \Rightarrow \Lambda_1^i(c, \bar{c}, R^2)$.

Using the identity $\text{tr}[S \times e \times e'] = e' \times S \times e$, $S = \Sigma^{-1} = \Sigma(1)^{-1}$, we can transform the equation 1:

$$\begin{aligned}
& \inf_{\lambda \in [0,1]} [\text{tr}[\Sigma^{-1} \sum_{t=k+1}^T \tilde{e}_t^i(\bar{\rho}, \lambda) \tilde{e}_t^i(\bar{\rho}, \lambda)']] - \inf_{\lambda \in [0,1]} [\text{tr}[\Sigma^{-1} \sum_{t=k+1}^T \tilde{e}_t^i(1, \lambda) \tilde{e}_t^i(1, \lambda)']] \\
&= \inf_{\lambda \in [0,1]} [\text{tr}[\Sigma^{-1} \sum_{t=k+1}^T \hat{e}_t^i(\bar{\rho}, \lambda) \hat{e}_t^i(\bar{\rho}, \lambda)']] - \inf_{\lambda \in [0,1]} [\text{tr}[\Sigma^{-1} \sum_{t=k+1}^T \hat{e}_t^i(1, \lambda) \hat{e}_t^i(1, \lambda)']] + o_p(1) \\
&= \inf_{\lambda \in [0,1]} \sum_{t=k+1}^T \hat{e}_t^i(\bar{\rho}, \lambda)' \Sigma^{-1} \hat{e}_t^i(\bar{\rho}, \lambda) - \inf_{\lambda \in [0,1]} \sum_{t=k+1}^T \hat{e}_t^i(1, \lambda)' \Sigma^{-1} \hat{e}_t^i(1, \lambda) + o_p(1). \quad (\text{A.6})
\end{aligned}$$

From equation 2 and 3, we have

$$\begin{aligned}
& \sum_{t=k+1}^T \hat{e}_t^i(\bar{\rho})' \Sigma^{-1} \hat{e}_t^i(\bar{\rho}) - \sum_{t=k+1}^T \hat{e}_t^i(1)' \Sigma^{-1} \hat{e}_t^i(1) \\
&= \sum_{t=1}^T e_t(\bar{\rho})' \Sigma^{-1} e_t(\bar{\rho}) - \sum_{t=1}^T e_t(1)' \Sigma^{-1} e_t(1) \\
&\quad - [S_i N(c, \bar{c}, R^2)]' [S_i D(\bar{c}, R^2) S_i]^{-1} [S_i N(c, \bar{c}, R^2)] \\
&\quad + [S_i N(c, 0, R^2)]' [S_i D(\bar{c}, R^2) S_i]^{-1} [S_i N(c, 0, R^2)] \quad (\text{A.7})
\end{aligned}$$

Now, substitute (A.7) into (A.6), we get

$$\begin{aligned}
& \inf_{\lambda \in [0,1]} [\text{tr}[\Sigma^{-1} \sum_{t=k+1}^T \tilde{e}_t^i(\bar{\rho}, \lambda) \tilde{e}_t^i(\bar{\rho}, \lambda)']] - \inf_{\lambda \in [0,1]} [\text{tr}[\Sigma^{-1} \sum_{t=k+1}^T \tilde{e}_t^i(1, \lambda) \tilde{e}_t^i(1, \lambda)']] \\
&\Rightarrow \sum_{t=1}^T e_t(\bar{\rho}, \lambda)' \Sigma^{-1} e_t(\bar{\rho}, \lambda) - \sum_{t=1}^T e_t(1, \lambda)' \Sigma^{-1} e_t(1, \lambda) \\
&\quad - \sup_{\lambda \in [0,1]} [S_i N(c, \bar{c}, R^2, \lambda)]' [S_i D(\bar{c}, R^2) S_i]^{-1} [S_i N(c, \bar{c}, R^2, \lambda)] \\
&\quad + \sup_{\lambda \in [0,1]} [S_i N(c, 0, R^2, \lambda)]' [S_i D(\bar{c}, R^2, \lambda) S_i]^{-1} [S_i N(c, 0, R^2, \lambda)] \\
&\equiv P^i(1, \bar{\rho}).
\end{aligned}$$

In our model, only y_t contains a structural break. Therefore we use $\tilde{\Sigma}(r) = \inf_{\lambda \in [0,1]} \sum_{t=k+1}^T \hat{e}_{y,t}^i(r, \lambda)' \Sigma^{-1} \hat{e}_{y,t}^i(r, \lambda)$ to select the break point, and the feasible test statistic is,

$$\tilde{P}^i(1, \bar{\rho}) = T \{ \text{tr}[\tilde{\Sigma}(1)^{-1} \tilde{\Sigma}(\bar{\rho})] - [m + \bar{\rho}] \} \blacksquare$$

Table 1. Asymptotic and Finite Critical Values for P test choosing T_B minimizing the statistic;
($\bar{c} = -22.5$ when constructing the test)

Size	$T = \infty$	$T = 100$		$T = 200$		
		Case 1	Case 2	Case 1	Case 2	
$R^2 = 0.0$.01	6.6206	7.0831	8.2017	6.8656	7.9770
	.025	7.7184	8.3496	9.6151	8.2074	9.3505
	.05	8.7784	9.6171	10.7726	9.4273	10.4317
	.10	10.2240	11.2630	12.2715	11.0206	12.0643
	.20	12.4409	13.2986	14.1866	13.0273	14.0473
$R^2 = 0.1$.01	6.3306	7.1161	7.7447	6.8647	7.4366
	.025	7.5704	8.3208	8.9815	8.1843	8.8018
	.05	8.8655	9.5971	10.3549	9.3553	10.0709
	.10	10.7418	11.1881	12.2124	11.0294	11.8583
	.20	13.3284	13.6536	14.7468	13.5258	14.5056
$R^2 = 0.2$.01	6.3366	7.0755	7.6763	6.7214	7.5043
	.025	7.7627	8.5707	9.2372	8.3291	9.0627
	.05	9.3670	10.0312	10.9802	9.8132	10.6892
	.10	11.5830	11.9907	13.1029	11.8267	12.7795
	.20	14.6210	14.8797	16.2305	14.7595	15.9071
$R^2 = 0.3$.01	6.3418	7.2735	8.1307	6.7738	7.4329
	.025	8.1433	9.1303	9.8774	8.7816	9.6521
	.05	10.1212	10.9046	11.8302	10.6501	11.5858
	.10	12.7136	13.2524	14.4177	12.9126	14.0307
	.20	16.2374	16.6372	18.0740	16.4752	17.6493
$R^2 = 0.4$.01	6.4881	7.8499	8.7390	7.2775	8.2742
	.025	8.8891	10.0639	10.9070	9.7519	10.5740
	.05	11.3466	12.2461	13.3519	11.8917	12.9659
	.10	14.3757	15.0278	16.2870	14.6182	15.8075
	.20	18.4936	19.0608	20.5460	18.7793	20.4266

Table 1 (continues). Asymptotic and Finite Critical Values for P test choosing T_B
 minimizing the statistic;
 ($\bar{c} = -22.5$ when constructing the test)

Size	$T = \infty$	$T = 100$		$T = 200$		
		Case 1	Case 2	Case 1	Case 2	
$R^2 = 0.5$.01	7.1952	9.1102	9.6599	8.4968	9.2217
	.025	10.0659	11.7718	12.8048	11.3839	12.3373
	.05	13.0346	14.3242	15.6562	13.9767	14.9083
	.10	16.7505	17.6416	19.1295	17.1750	18.7120
	.20	21.5898	22.5719	24.3645	22.2750	23.9609
$R^2 = 0.6$.01	8.3182	11.2399	12.0004	10.8230	12.0457
	.025	12.1112	14.5488	15.8906	14.0533	15.1350
	.05	15.7378	17.7247	19.2281	17.3959	18.5054
	.10	20.2065	21.9735	23.5982	21.3516	23.1323
	.20	26.1940	28.0226	30.1219	27.5030	29.5977
$R^2 = 0.7$.01	10.0794	15.3051	16.5035	15.1582	16.4011
	.025	15.6598	19.7340	21.1399	19.2412	20.3855
	.05	20.0063	24.0217	25.8066	23.3866	24.6636
	.10	26.0101	29.2556	31.4516	28.7144	30.6317
	.20	33.8148	37.2448	40.0105	36.5297	39.0657
$R^2 = 0.8$.01	13.9343	24.6019	26.0516	24.7058	26.6857
	.025	22.0148	30.7935	32.8000	30.3721	32.2811
	.05	28.8717	36.7978	39.1595	36.0741	38.1094
	.10	37.6966	44.5746	47.0921	43.7420	46.6231
	.20	48.6718	55.6362	59.6670	54.7744	58.7771
$R^2 = 0.9$.01	25.6335	54.3674	56.9218	55.6351	58.4940
	.025	43.7594	64.9650	69.1188	65.2765	68.7000
	.05	55.7402	76.4949	80.8711	75.554	80.1414
	.10	72.5938	90.7439	95.3365	89.8564	95.8853
	.20	92.7436	111.4939	119.3781	110.4932	118.0968

Table 2. Finite Sample Size and Power; Case 1
($\bar{c} = -22.5$ when constructing the test)

	$T = 100$					$T = 200$				
$R =$	0.0	0.30	0.50	0.70	0.90	0.0	0.30	0.50	0.70	0.90
$R^2 =$	0.0	0.09	0.25	0.49	0.81	0.0	0.09	0.25	0.49	0.81
ρ										
1.00	0.0501	0.0501	0.0501	0.0501	0.0501	0.0501	0.0501	0.0500	0.0500	0.0500
0.98	0.0565	0.0567	0.0610	0.0755	0.1033	0.0733	0.0752	0.0873	0.1206	0.2121
0.96	0.0736	0.0750	0.0795	0.1157	0.2082	0.1187	0.1378	0.1679	0.2919	0.6508
0.94	0.0896	0.0967	0.1127	0.1838	0.3956	0.2029	0.252	0.3214	0.5796	0.9661
0.92	0.1344	0.1462	0.1634	0.2894	0.6482	0.3455	0.3555	0.5306	0.8325	0.9997
0.90	0.1839	0.1948	0.2312	0.4284	0.8633	0.5218	0.5409	0.7441	0.9525	1.0000
0.88	0.2390	0.2592	0.3252	0.5825	0.9685	0.6796	0.7223	0.8821	0.9902	1.0000
0.86	0.3156	0.3777	0.4226	0.7150	0.9942	0.8030	0.8511	0.9498	0.9986	1.0000

Number of lags in VAR is zero; 10,000 replications.

Table 3. Finite Sample Size and Power; Case 2
($\bar{c} = -22.5$ when constructing the test)

	$T = 100$					$T = 200$				
$R =$	0.0	0.30	0.50	0.70	0.90	0.0	0.30	0.50	0.70	0.90
$R^2 =$	0.0	0.09	0.25	0.49	0.81	0.0	0.09	0.25	0.49	0.81
ρ										
1.00	0.0500	0.0501	0.0501	0.0500	0.0501	0.0501	0.0501	0.0500	0.0501	0.0501
0.98	0.0585	0.0595	0.0615	0.0775	0.0952	0.0628	0.0674	0.0832	0.1078	0.1882
0.96	0.0765	0.0775	0.0798	0.1166	0.1886	0.1239	0.1323	0.1706	0.2726	0.5875
0.94	0.0965	0.0978	0.1136	0.1788	0.3513	0.2269	0.2588	0.3209	0.5329	0.8996
0.92	0.1427	0.1440	0.1613	0.2738	0.5768	0.3646	0.4100	0.5417	0.7947	0.9726
0.90	0.1921	0.1954	0.2262	0.4048	0.7867	0.5269	0.6557	0.7408	0.9251	0.9914
0.88	0.2529	0.2616	0.3216	0.5549	0.9114	0.6597	0.7198	0.8669	0.9735	0.9961
0.86	0.3277	0.3519	0.4168	0.6798	0.9550	0.7750	0.8580	0.9473	0.9892	0.9992

Number of lags in VAR is zero; 10,000 replications.

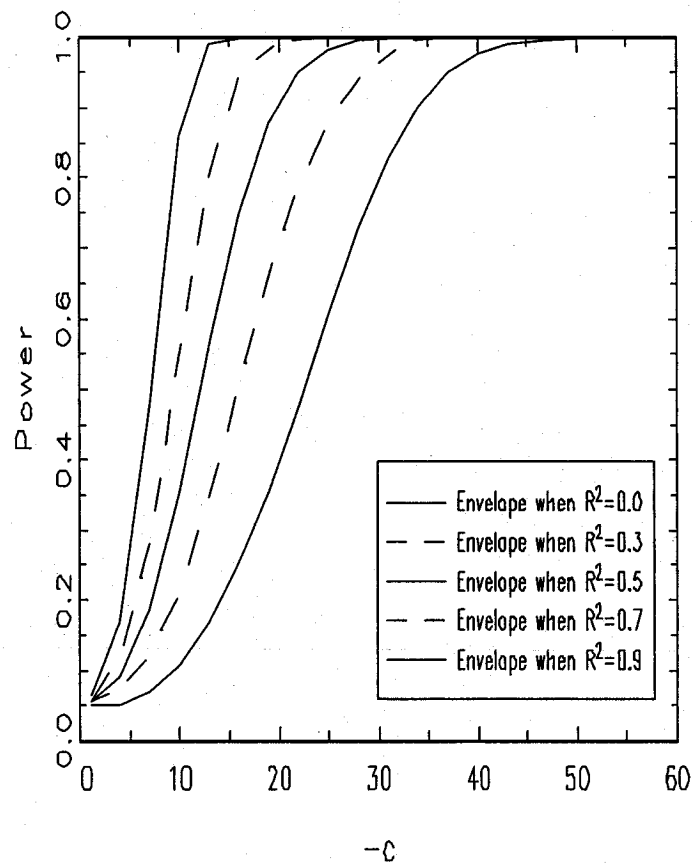


Figure 1. Power Envelopes for $R^2 = 0.0, 0.3, 0.5, 0.7, 0.9$.

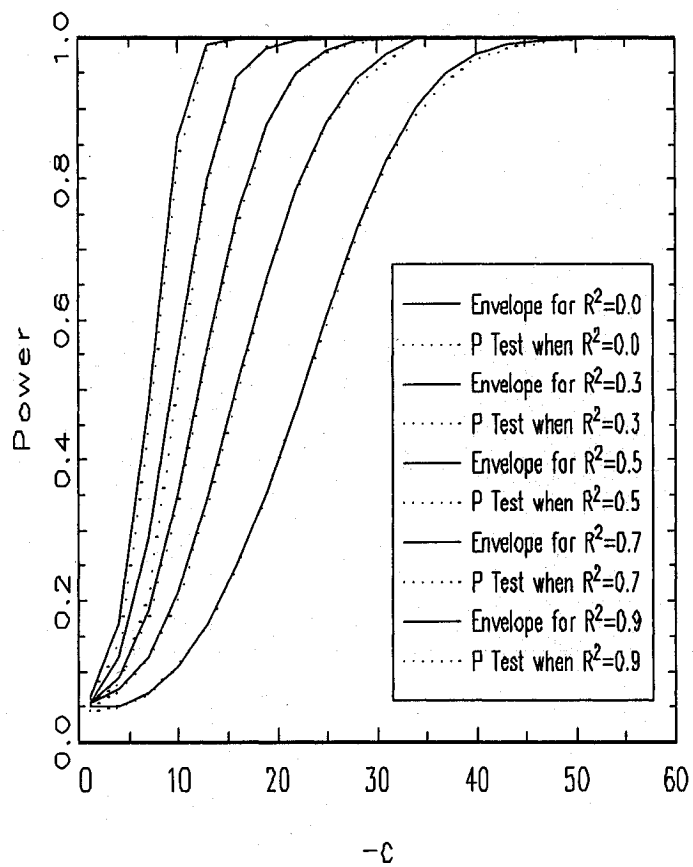


Figure 2. Power Envelopes and Asymptotic Power Functions of the Feasible Point Optimal Test

Conclusions

This thesis includes three essays dealing with issues related to unit root tests and cointegration.

The first chapter applies I(1) and I(2) framework of cointegration to analyze the cause of global warming. Due to the complexity of the time series property of global mean temperature deviation and GHG, I use three cases to study the system. The first case includes the levels of global temperature, the irradiance of the Sun and three major GHG. The results indicate the possibility that explosive roots may cause ambiguity on the detection of the number of unit roots in the system. Therefore the second case subdivides the five variables into two sub-systems. One includes global mean temperature, the radiative forcing of solar irradiance and carbon dioxide; the other one includes the radiative forcing of two GHG suspected of containing explosive roots. The results show a very clear long-run relationship between global temperature, solar irradiance and carbon dioxide. The third case performs a traditional I(1) analysis on all the variables, which are either integrated of order one or transformed to I(1) variable. The purpose of this trial is, on the one hand, to verify the results of the previous two cases, and on the other hand, to utilize cointegration in an I(1) framework (a well-developed approach) to study the system as a complimentary. The results of all cases show that the GHG have been driving up the global temperature, and the temperature sensitivity of a doubling of carbon dioxide is between 2.15 and 3.40 Celsius degrees.

The second chapter investigates the effect of random initial values on unit root tests in the context of unknown structural change models. Traditionally, the initial value of unit root tests has been assumed to be zero or with fixed expectations. Therefore it disappears and won't have any impact on the statistical theory of unit root tests asymptotically. But when assuming the initial value is drawn from its unconditional distribution under the alternative hypothesis, it does not disappear at a convergence rate of $T^{1/2}$, which will generate impacts on unit root tests even in large samples. In fact, Monte Carlo studies show that the power of unit root tests has dropped in both large and small samples. Therefore it is recommended to use random initial condition in unit root testing in order to be more conservative.

The third chapter proposes a multivariate unit root test for unknown structural change models by including an arbitrary number of stationary covariates correlated with the variable to be tested. Conventionally, unit root

tests have been performed on univariate time series. But this may cause loss of power due to the ignorance of information in related time series. Based on this idea, this chapter includes stationary covariates in models with an unknown structural change. It finds that the power of unit root tests is a function of the correlation R^2 between covariates and the variable to be tested. When R^2 is close to one, substantial power gains are realized. The worst case is when $R^2 = 0$, where we retrieve the power of univariate unit root tests. Therefore this trial offers a practical solution to increase the power of unit root tests considerably in the context of unknown structural changes.