Université de Montréal

Unit Root, Outliers and Cointegration Analysis with Macroeconomic Applications

par

Gabriel Rodríguez

Département des sciences économiques
Faculté des arts et des sciences

Thèse présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en **sciences économiques**

October, 1999

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

Unit Root, Outliers and Cointegration Analysis with Macroeconomic Applications

présentée par

Gabriel Rodríguez

a été evaluée par un Jury composé des personnes suivantes:

Jean-Marie Dufour
(Président du Jury)

Pierre Perron
(Directeur de Recherche)

Francisco Ruge-M.
(Membre du Jury)

John Galbraith
(Examinateur Externe)

Roch Roy
Représentant de la FES

Thèse acceptée le: 25 novembre 1999

Summary

In this thesis, we deal with three particular issues in the literature on nonstationary time series. The first essay deals with various unit root tests in the context of structural change. The second paper studies some residual based tests in order to identify cointegration. Finally, in the third essay, we analyze several tests in order to identify additive outliers in nonstationary time series.

The first paper analyzes the hypothesis that some time series can be characterized as stationary with a broken trend. We extend the class of M-tests and ADF test for a unit root to the case where a change in the trend function is allowed to occur at an unknown time. These tests (M^{GLS}, ADF^{GLS}) adopt the Generalized Least Squares (GLS) detrending approach to eliminate the set of deterministic components present in the model. We consider two models in the context of the structural change literature. The first model allows for a change in slope and the other for a change in slope as well as intercept. We derive the asymptotic distribution of the tests as well as that of the feasible point optimal test (P_T^{GLS}) which allows us to find the power envelope. The asymptotic critical values of the tests are tabulated and we compute the non-centrality parameter used for the local GLS detrending that permits the tests to have 50% asymptotic power at that value. Two methods to select the break point are analyzed. A first method estimates the break point that yields the minimal value of the statistic. In the second method, the break point is selected such that the absolute value of the t-statistic on the change in slope is maximized. We show that the M^{GLS} and P_T^{GLS} tests have an asymptotic power function close to the power envelope. An extensive simulation study analyzes the size and power of the tests in finite samples under various methods to select the truncation lag for the autoregressive spectral density estimator. In an empirical application, we consider two U.S. macroeconomic annual series widely used in the unit

root literature: real wages and common stock prices. Our results suggest a rejection of the unit root hypothesis. In other words, we find that these series can be considered as trend stationary with a broken trend.

Given the fact that using the GLS detrending approach allows us to attain gains in the power of the unit root tests, a natural extension is to propose this approach to the context of tests based on residuals to identify cointegration. This is the objective of the second paper in the thesis. In fact, we propose residual based tests for cointegration using local *GLS* detrending to eliminate separately the deterministic components in the series. We consider two cases, one where only a constant is included and one where a constant and a time trend are included. The limiting distributions of various residuals based tests are derived for a general quasi-differencing parameter $\bar{c}$ and critical values are tabulated for values of $\bar{c} = 0$ irrespective of the nature of the deterministic components and also for other values as proposed in the unit root literature. Simulations show that *GLS* detrending yields tests with higher power. Furthermore, using $\bar{c} = -7.0$ or $\bar{c} = -13.5$ as the quasi-differencing parameter, based on the two cases analyzed, is preferable.

The third paper is an extension of a recently proposed method to detect outliers which explicitly imposes the null hypothesis of a unit root. It works in an iterative fashion to select multiple outliers in a given series. We show, via simulation, that under the null hypothesis of no outliers, it has the right size in finite samples to detect a single outlier but when applied in an iterative fashion to select multiple outliers, it exhibits severe size distortions towards finding an excessive number of outliers. We show that this iterative method is incorrect and derive the appropriate limiting distribution of the test at each step of the search. Whether corrected or not, we also show that the outliers need to be very large for the method to have any decent power. We propose an alternative method based on first-differenced data that has considerably more power. The issues are illustrated using two US/Finland real exchange rate series.

Résumé

Nous analysons trois sujets particuliers dans la vaste littérature portant sur les séries chronologiques non stationnaires. Le premier essaie analyse plusieurs tests de racine unitaire dans le contexte de changement structurel. Le deuxième papier analyse des tests basés sur les résidus d'une régression statique pour identifier l'existence d'une rélation de cointégration. Le troisième essai analyse quelques tests pour identifier la présence d'observations à l'écart.

Le premier papier analyse l'hypothèse selon laquelle quelques séries chronologiques peuvent être caracterisées comme stationnaires autour d'une tendance brisée. Depuis le papier original de Nelson et Plosser (1982), l'hypothèse de racine unitaire a reçu beaucoup d'intérê, autant théorique qu'empirique (voir, par exemple, Campbell et Perron (1991), Stock (1994) pour des survols). Utilisant des tests développés par Dickey et Fuller (1979), Nelson et Plosser (1982) ont postulé que les chocs courants ont des effets permanents sur le niveau de la plupart des séries macroéconomiques. Ce résultat a été confirmé par des autres approches lesquelles ont trouvé qu'un choc typique a des composantes transitoires et permanentes importantes.

D'un autre point de vue, Perron (1989) argumentait, comme une alternative à l'hypothèse de racine unitaire, que les fluctuations macroéconomiques sont plutôt stationnaires si la possibilité d'une fonction de tendance avec des changements occasionels est prise en ligne de compte. Avec un seul changement dans l'intercepte et/ou la pente, il rejetait l'hypothèse de racine unitaire pour 11 des 14 séries analysées par Nelson et Plosser. Tel que discuté par Banerjee, Lumsdaine et Stock (1992), ce résultat peut être important pour les raisons suivantes. La première, c'est que elle donne une image alternative de la persistence des fluctuations macroéconomiques. La seconde raison est que l'approche peut donner un modèle alternatif avec une tendance qui évolue lentement

et qui peut être utile comme description des données. La troisième raison est que les implications pour l'inférence dans les modèles plus compliqués sont très différents.

Christiano (1992) avait mis en question les résultats de Perron (1989) sur la base que le point de bris ne doit pas être considéré comme exogène. En accord avec ce point de vue, Zivot et Andrews (1992), Banerjee, Lumsdaine et Stock (1992) et Perron (1997) ont considéré le point de bris comme étant inconnu.

Dans le premier essai, nous continuons à analyser le point de bris potentiel comme étant inconnu et nous contribuons à la literature avec deux points. Premièrement, nous utilisons les M-tests analyses par Perron et Ng (1996) et nous les étendons pour permettre l'existence d'un changement structural dans la fonction de tendance. Dans le deuxième, comme dans Dufour et King (1991) et Elliot, Rotemberg et Stock (1996, ERS), nous utilisons les Moindres Carrés Generalisés (MCG) local à l'unité pour éliminer les composantes déterministes dans les séries. Nos considérons deux modèles spécifiques. Le premier modèle inclut un bris dans la pente de la fonction de tendance, alors que le second modèle considére considère un bris dans l'intercept ainsi que dans la pente.

Les M-tests ont été proposés par Stock (1990), et analysés plus en détail en Perron et Ng (1996). Ces tests ont les plus petites distorsions de niveau parmi les autres tests de racine unitaire, même dans le cas que les erreurs ont de la forte corrélation négative. En même temps, l'utilisation de des MCG, pour éliminer les composantes déterministes dans la fonction de tendance dans la construction des M -tests permet d'avoir plus de puissance comme a été montré par Ng et Perron (1999), ce qui est équivalent au test DF^{GLS} proposé par ERS. Le résultat est une conséquence de l'utilisation de MCG qui permet une meilleur estimation des composantes déterministes.

Nous analysons deux méthodes pour sélectionner le point de bris. La première methode sélectione le point de bris comme le point associé à la valeur minimale de la statistique. Dans la deuxième méthode, le point de bris est sélectionné tel que la valeur absolue de la t-statistique sur le changement de la pente est maximisée. D'un autre côté, la construction des M-tests exige le choix d'un paramètre de troncature (k) qui est utilisé pour l'estimation de la densité spectrale à la fréquence zéro. Nous suivons la suggestion de Ng et Perron (1999) d'utiliser une version modifiée des critères AIC et BIC (appellées MIC). Cette modification permet de prendre en compte le fait que le biais dans la somme de coefficients autoregréssifs est très dépendant de k en échantillons finis, et la procédure ajuste la fonction de penalité de façon appropriée.

Nous dérivons aussi une version réalisable du test optimal en un point (P_T^{GLS}) proposé par ERS. La fonction de puissance asymptotique est dérivée et nous utilisons cette enveloppe de puissance pour choisir le paramètre de non-centralité ($\bar{c}$) qui est utilisé pour appliquer les MCG. Ce paramètre est choisi tel que la puissance asymptotique des tests est 50% contre l'hypothèse alternative . Nous montrons que les tests M^{GLS} et P_T^{GLS} ont une fonction de puissance asymptotique proche de l'enveloppe de puissance, en particulier quand nous utilisons la première méthode pour choisir le point de bris. Les propriétés de niveau et de puissance des tests sont analysées avec des simulations utilisant des procédures différentes pour choisir le paramètre de troncature. Pour les applications empiriques, nous utilisons les séries des salaires réels et les prix d'actions ordinaires aux Etats Unis. Nos résultats sont favorables à l'hypothèse de stationnarité autour d'une tendance brisée.

Etant donné le fait que les MCG permettent d'avoir plus de puissance pur les tests de racine unitaire, une extension naturelle est d'utiliser cete approche pour analyser des tests basés sur les résidus d'une régression statique et d'identifier l'existence de cointégration. Cela est l'objectif du deuxième essai. Même si ces tests ne sont applicables sous des conditions spécifiques, les tests basés sur les résidus proposés par Phillips et Ouliaris (1990) ont été très utilisés en économétrie appliquée, dans la plupart des cas parce qu'ils sont très faciles à calculer. Les tests ont été proposés pour vérifier l'hypothèse nulle de non cointégration dans le contexte d'une équation individuelle. Une autre caractéristique de ces tests est leur caractère intuitif à partir de la définition basique de cointégration telle que définie par Engle et Granger (1987). Si le système de variables est cointegré, alors il existe une combinaison linéaire (donnée par le vecteur de cointégration) qui est stationnaire. Dans ce cas, les résidus d'une simple régression statique sont stationnaires et, tel que demontré par Stock (1987), l'estimation de cette régression estimé par MCO donnera un estimateur consistent du vecteur de cointégration. En l'absence de cointégration, les résidus de la régression statique ne sont pas stationnaires et nous avons affaire au phénomène que Granger et Newbold (1974), et plus tard Phillips (1986), ont appelé régression illégitime (*spurious régression*). Ainsi, une stratégie évidente est de vérifier l'hypothèse nulle de non cointégration utilisant quelques tests de racine unitaire sur les résidus estimés de la régression statique.

Ainsi, le deuxième essai analyse des tests basés sur les résidus afin de vérifier l'existence de cointégration lorsque ces tests sont construits utilisant les *MCG* pour éliminer les composantes déterministes dans les séries. Nous considérons le test *ADF* et la classe de tests modifiés qui ont été analysés par Stock (1990) et Ng et Perron (1999). Nous dérivons la distribution asymptotique utilisant un paramètre de non centralité général et nous calculons les valeurs critiques pour deux cas: a) $\overline{c} = 0$ pour les deux types de composantes déterministes considérées; b) $\overline{c} = -7.0$ pour le cas où le modèle inclut seulement une constante et $\overline{c} = -13.5$ pour le cas où le modèle inclut aussi une tendance linéaire. Le deuxième cas est suggéré par l'analyse de ERS dans le contexte des tests pour racine unitaire. Nos simulations pour la puissance des tests indiquent que le fait d'utiliser des *MCG* pour éliminer les composantes déterministes de la fonction de tendance permet d'avoir des gains de puissance importants, surtout dans le cas où nous utilisons le paramètre de non-centralité tel que proposé par ERS.

La présence et les effets des observations à l'écart son le sujet du troisième essaie. Depuis Fox (1972), qui avait introduit le concept d'observation à l'écart (*additive and innovational outliers*), la recherche reliée à ce type d'observations non typiques en séries chronologiques a proliféré, autant dans le contexte statistique qu'économétrique. Le sujet de détection d'observations à l'écart lui même a reçu intérêt particulier (voir, par exemple, Hawkins (1980) pour un survol des méthodes proposées avant 1980). Un autre sujet d'intérêt dans la recherche a été l'estimation des modèles $ARMA$ en présence d'observations à l'écart. Dans ce cas, comme il est mentionné par Chen et Liu (1993), une approche commune est d'identifier les localisations et les types d'observations à l'écart et donc, de prendre en compte les effets de ces observations en utilisant des modèles d'intervention tel que proposé par Box et Tiao (1975). Cette approche exige l'usage d'itérations entre les étapes pour détecter les observations à l'àcart et l'estimation même du modèle (voir aussi Chang, Tiao et Chen (1988) et Tsay (1986)).

Dans le contexte des séries integrées (processus avec une racine unitaire autoregréssive), les effets des observations à l'écart ont été récemment l'objet de plusieurs recherches. Maintenant, il est connu que les observations à l'écart affectent les propriétés des tests de racine unitaire (Franses et Haldrup (1994)). Cela est effectué par l'introduction d'une composante de moyenne mobile négative dans la fonction de bruit, faisant en sorte que la plupart des tests de racine unitaire présentent des distorsions de niveau très importantes. En d'autres mots, ces tests rejettent l'hypothèse de racine unitaire trop souvant. Franses et Haldrup (1994) ont suggeré l'application du test de Dickey et Fuller (1979) par l'introduction de variables dichotomiques dans l'autoregréssion. Ces variables ont été choisies sur la base de la procédure de détection des observations à l'écart proposée par Chen et Liu (1993).

Dans un papier récent, Vogelsang (1999) a fait deux contributions au sujet des effets des observations à l'écart sur les tests de racine unitaire. Dans la première contribution, reconnaissant que les observations à l'écart introducent une composante de moyenne mobile négative, il propose l'utilisation des tests de racine unitaire proposés par Stock (1990) et Perron et Ng (1996), lesquels sont résistants à ce problème, dans le sens qu'ils ont un niveau exact proche du niveau nominal. Vogelsang (1999) montre, avec des simulations, que ces tests sont peu affectés par la présence d'observations à l'écart systématiques. Dans le contexte de séries integrées, la deuxième contribution de Vogelsang (1999) est la dérivation d'une distribution asymptotique pour la t-statistique sur la variable dichotomique pertinente.

Dans le troisième papier, nous faisons quelques contributions additionelles dans la direction de la deuxième contribution de Vogelsang (1999). Nous montrons avec des simulations que la procédure de Vogelsang (1999), sous l'hypothèse nulle de non existence d'observations à l'écart, a le niveau correct en échantillons finis pour détecter une observation à l'écart. Toutefois lorsque la procédure est appliquée d'une manière itérative pour sélectionner plusieurs observations à l'écart, elle présente de fortes distorsions en sélectionnant un nombre excessif d'observations à l'écart. Nous montrons qu'un estimateur alternatif de la variance des erreurs résout ce problème. Cependant, il existe un défaut dans la procédure itérative de Vogelsang (1999). En effet, contrairement à ce qu'il suppose, la distribution limite du test utilisé est différente dans chaque étape de la procédure de détection des observations à l'écart. Nous dérivons la distribution asymptotique appropriée et nous calculons quelques valeurs critiques. Quand nous analysons la correction, nous trouvons que la méthode a une puissance faible pour détecter des observations à l'écart sauf si la magnitude de l'observation à l'écart est très elevé. Comme une alternative, nous proposons une méthode basée sur les donnés en premières différences, qui possède une puissance supérieure. Dans les applications empiriques, nous utilisons deux séries de taux de change réel pour les Etats Units avec la Finland. Nos résultats sont favorables à la méthode utilisant les donnés en premières différences.

Introduction

In this thesis, we deal with three particular issues in the nonstationary time series literature. The first essay deals with various unit root tests in the context of structural change. The second paper studies residual based tests to identify cointegration. Finally, in the third essay, we analyze some tests to identify additive outliers in nonstationary time series.

The first paper analyzes the hypothesis that some time series can be characterized as stationary with a broken trend. Since the seminal paper of Nelson and Plosser (1982), the unit root hypothesis has received much attention from both the theoretical and empirical perspectives (see, e.g., Campbell and Perron (1991), Stock (1994) for surveys). Using tests developed by Dickey and Fuller (1979), Nelson and Plosser (1982) argued that current shocks have permanent effects on the level of most macroeconomic series. This finding was supported by other approaches which found that a typical shock has both important transitory and permanent components.

In contrast to this literature, Perron (1989) argued, counter to the unit root hypothesis, that macroeconomic fluctuations are probably stationary if an allowance is made for the trend function to exhibit occasional changes. Allowing for a single change in intercept and/or slope, he rejected the unit root hypothesis for 11 of the 14 series analyzed by Nelson and Plosser. As discussed in Banerjee, Lumsdaine and Stock (1992) this finding may be important for the following reasons. First, it offers an alternative picture of the persistence in macroeconomics series. Second, this approach can provide a parsimonious model for a slowly changing trend component that may be useful as a description of the data. Third, the implications for inference in more complex models are very different.

Christiano (1992) criticized the results of Perron (1989) on the basis that the break point should not be treated as exogenous since the imposition of a given break date involves an issue of data mining. Accordingly, Zivot and Andrews (1992), Banerjee, Lumsdaine and Stock (1992) and Perron (1997) considered unit root tests with unknown break points.

In the first paper, we continue to treat the potential break point as occurring at an unknown time and contribute to this literature in two ways. First, we use the M^{GLS} tests recently analyzed by Perron and Ng (1996) and extend them to permit a one time change in the trend function. Second, as in Dufour and King (1991) and Elliott, Rothemberg and Stock (1996, ERS), we use local to unity GLS detrending of the data. We consider two specific models. The first involves a break in the slope of the trend function and the second a break in both the intercept and slope.

The reasons for considering the M-tests, originally proposed by Stock (1990) and further analyzed by Perron and Ng (1996) is that these tests have much smaller size distortions than other classes of unit root tests even when the residuals have strong negative serial correlation. Also, using GLS detrending when constructing the M-tests allows substantial gains in power as showed by Ng and Perron (1999), similar to the DF^{GLS} test proposed by ERS. This is a consequence of the fact that a GLS framework permits a more accurate estimation of the deterministic component than is possible using an OLS approach.

Two methods to select break points are analyzed. A first method estimates the break point that yields the minimal value of the statistic. In the second method, the break point is selected such that the absolute value of the t-statistic on the change in slope is maximized. On the another hand, the implementation of the M^{GLS} tests requires a truncation lag (k) for the autoregressive spectral density estimate at frequency zero. We follow the suggestion of Ng and Perron (1999) and use a modified version of the AIC and BIC (labeled MIC). Their modification permits that the bias in the sum of the autoregressive coefficients is highly dependent on k in finite samples and adjusts the penalty function accordingly.

Given that a uniformly most powerful test is not attainable, we follow ERS and derive a feasible point optimal test (P_T^{GLS}). The asymptotic power function of this test is derived and we use this power envelope to choose the non-centrality parameter ($\bar{c}$) to perform the GLS detrending such that the asymptotic power of the tests is 50% against the local alternative $\bar{\alpha} = 1 + \bar{c}/T$. We show that the M^{GLS} and P^{GLS} tests have an asymptotic power function close to the power envelope particularly, when the first method to select the break point is used. Extensive simulations analyze the size and power properties of the tests. In the empirical applications, we consider two U.S. macroeconomic annual series widely used in the unit root literature: real wages and common stock prices. Our results tend to the rejection of the unit root hypothesis in favor of the hypothesis of stationarity with a broken trend.

Given the fact that using GLS detrending approach permits attaining higher power for unit root tests, a natural extension is to apply this approach to residual based tests to identify cointegration. This is the aim of the second paper. Even though they are applicable only under some specific conditions, residual based tests for cointegration, developed by Phillips and Ouliaris (1990), have been quite popular in applied work mostly because of their computational simplicity. The statistics introduced are designed to test the null hypothesis of no cointegration in a single equation setting assuming that the variables introduced as regressors are not cointegrated. These tests also are appealing because they follow intuitively from the basic definition of cointegration as laid out in Engle and Granger (1987). If the system of variables is cointegrated, then there exists a linear combination (given by the cointegrating vector) that is stationary. In this case, the residuals from a simple static regression are stationary and, as shown by Stock (1987), the regression estimated by *OLS* will provide a consistent estimate of the cointegrating vector. In the absence of cointegration, the residuals from the static regression are nonstationary for any choice of the parameter vector and we have what has been labelled by Granger and Newbold (1974), and later Phillips (1986), a spurious regression. Hence, an obvious testing strategy is to test the null hypothesis of no cointegration using unit root tests on the estimated residuals from the simple static regression.

Hence, our aim in the second paper is to analyze residual based tests for cointegration when they are constructed using *GLS* detrended or quasi-differenced data. We consider the standard *ADF* test and the class of modified unit root tests analyzed in Stock (1990) and Ng and Perron (1999). We derive their asymptotic distribution assuming a general quasi-differencing parameter $\bar{c}$ and tabulate critical values for two choices: a) $\bar{c} = 0$ irrespective of the nature of the deterministic components; b) $\bar{c} = -7.0$ for the constant only case and $\bar{c} = -13.5$ for the linear trend case, as suggested by ERS in the context of unit root tests. Our simulation results on the power of the tests reveal that important gains can indeed be achieved by using *GLS* detrended data, especially if the quasi-difference parameter is set as suggested by ERS.

The presence of atypical observations in some time series is the subject of the third paper. From Fox (1972), who introduced the notion of additive and innovational outliers, issues related to this type of atypical observations in time series have received considerable attention in the statistics and econometric literature. The outlier detection issue, itself, has received particular attention (see, e.g., Hawkins (1980) for a survey of methods before 1980). Another topic of interest has been the estimation of *ARMA* models in the presence of outliers. In this case, as mentioned by Chen and Liu (1993), a common approach is to identify the locations and the types of outliers and then to accommodate the effects of outliers using intervention models as proposed by Box and Tiao (1975). This approach requires iterations between stages of outlier detection and estimation of the model (see also, Chang, Tiao and Chen (1988) and Tsay (1986)).

In the context of integrated data (processes with an autoregressive unit root), the effects of additive outliers have recently been the object of sustained research. It is by now well recognized that outliers affect the properties of unit root tests (e.g., Franses and Haldrup (1994)). They do so by inducing a negative moving average component in the noise function which causes most unit root tests to exhibit substantial size distortions towards rejecting the null hypothesis too often. Franses and Haldrup (1994) suggested applying Dickey-Fuller (1979) unit root tests by incorporating dummy variables in the autoregression chosen on the basis of the outlier detection procedure proposed by Chen and Liu (1993).

In an interesting recent paper, Vogelsang (1999) makes two contributions to the issue about the effects of additive outliers on unit root tests. First, recognizing that outliers induce a negative moving average component, he suggests using unit the root tests developed by Stock (1990) and Perron and Ng (1996) that are robust, in terms of achieving exact size close to nominal size in small samples, even in the presence of a substantial negative moving average component. He shows via simulations that these unit root tests are only slightly affected by systematic outliers. Secondly, he recognized that one can take advantage of the null hypothesis of a unit root in devising an outlier detection procedure. This allows the derivation of a non-degenerate limiting distribution for the t-statistic on the relevant one-time dummy.

In the third paper, we make further contributions along the second suggestion of Vogelsang (1999). We show, via simulations, that Vogelsang's (1999) procedure, under the null hypothesis of no outliers, has the right size in finite samples to detect a single outlier but, when applied in an iterative fashion to select multiple outliers, it exhibits severe size distortions towards finding an excessive number of outliers. We show that an alternative estimate of the variance of the errors alleviates this problem but that there is a basic flaw in the iterative method suggested by Vogelsang (1999). In effect, contrary to what he implicitly assumes, the limiting distribution of the test used is different at each iteration of the outlier detection procedure. We derive the appropriate limiting distribution and tabulate some critical values. When so corrected, his method

is shown to have very low power in detecting outliers (even a single one without making the correction) unless the magnitude of the outlier is very large. As an alternative, we propose a method based on first-differenced data which has considerably more power. In an empirical application, we use two US/Finland real exchange rates widely used in the literature. The results are favorable to the method based on the first-differenced data.

Table of Contents

List of Tables

List of Figures

Acknowledgments

During the last four years, many people helped me arrive at this stage of my career, both from professionally and personally. I am very grateful to Pierre Perron, my research supervisor for both financial support and for his invaluable criticism and comments in all aspects of this thesis. Our extensive communication is a testimony of his availability to discuss and to answer my questions. Furthermore, I would like to acknowledge his friendship. I have not only learned from Pierre many things about econometrics and particularly about the structural change literature, but I have also learned about the philosophy of the research world, where sensations of frustration and success are always present. I would also like to thank J.-M. Dufour, who read and commented on my first paper, particularly when I presented it at a seminar at the University of Montreal. Additionally, I also received comments from Réné Garcia and Benoit Perron. Alain Guay contributed with valuable comments to my first paper when I presented it at the 38th meeting of the société des sciences économiques at Québec in May of 1998. Lynda Khalaf commented on my third paper when I presented it at the 39th meeting of the société des sciences économiques at Hull (Québec) in May of this year.

Participation at the conferences mentioned above (and others) was possible thanks to the financial support of the CRDE. This institution also supported me in my first year of my Ph.D. studies with a graduate fellowship. Similar acknowledgment goes to the Department of Economics at the University of Montreal for financial support in my two first years. Special thanks are due to Jocelyne Demers, Suzanne Larouche and Lyne Racine for their efficiency in administrative tasks.

The student environment was also stimulating during the last four years. Particular mention is deserved to the student seminars where I presented my three papers and received interesting questions and suggestions. The comments of Paul Johnson, David Servettaz and Norma Kozhaya were particularly helpful. I would also like to thank Gamal Atallah who corrected my french version of the Résumé.

Lastly, I would like to thank my family: my wife Milagros, my daughter Gabriela and my son Gabriel H. F. Their moral support and encouragement were very important in continuing this process. Their patience and comprehension were invaluable. My parents also provided very important encouragement and stimulation.

For Milagros

GLS Detrending, Efficient Unit Root Tests and Structural Change

Abstract

We extend the class of M-tests for a unit root analyzed by Perron and Ng (1996) and Ng and Perron (1997) to the case where a change in the trend function is allowed to occur at an unknown time. These tests (M^{GLS}) adopt the GLS detrending approach of Elliott, Rothemberg and Stock (1996, ERS). Following Perron (1989), we consider two models: one allowing for a change in slope and the other for both a change in intercept and slope. We derive the asymptotic distribution of the tests as well as that of the feasible point optimal tests (P_T^{GLS}) suggested by ERS. The asymptotic critical values of the tests are tabulated. Also, we compute the non-centrality parameter used for the local GLS detrending that permits the tests to have 50% asymptotic power at that value. We show that the M^{GLS} and P_T^{GLS} tests have an asymptotic power function close to the power envelope. An extensive simulation study analyzes the size and power in finite samples under various methods to select the truncation lag for the autoregressive spectral density estimator. An empirical application is also provided.

1 Introduction

Since the seminal paper of Nelson and Plosser (1982), the unit root hypothesis has received a lot of attention both from theoretical and empirical perspectives (see, e.g., Campbell and Perron (1991), Stock (1994) for surveys). Using tests developed by Dickey and Fuller (1979), Nelson and Plosser (1982) argued that current shocks have permanent effects on the level of most macroeconomic series. This finding was supported by other approaches which found that a typical shock has both important transitory and permanent components (see, e.g., Campbell and Mankiw (1987a, 1987b), Shapiro and Watson (1988), Clark (1987), Cochrane (1988) and Christiano and Eichembaum (1989)).

In contrast to this literature, Perron (1989) argued, as an alternative to the unit root hypothesis, that macroeconomic fluctuations are most likely stationary if allowance is made for the trend function to exhibit occasional changes. Allowing for a single change in intercept and/or slope, he rejected the unit root hypothesis for 11 of the 14 series analyzed by Nelson and Plosser. As discussed in Banerjee, Lumsdaine and Stock (1992) this finding may be important for the following reasons. First, it offers an alternative picture of the persistence in macroeconomics series. Second, this approach can provide a parsimonious model for a slowly changing trend component that may be useful as a data description. Third, the implications for inference in more complex models are very different.

Christiano (1992) criticized the results of Perron (1989) on the basis that the break point should not be treated as exogenous since the imposition of a given break date involves an issue of data mining. Accordingly, Zivot and Andrews (1992), Banerjee, Lumsdaine and Stock (1992) and Perron (1997) considered unit root tests with unknown break point.

In this paper, we continue to treat the potential break point as occurring at unknown time and contribute to this literature in two ways. First, we use the M^{GLS} tests recently analyzed by Perron and Ng (1996) and extend them to permit a one time change in the trend function. Second, as in Elliott, Rothemberg and Stock (1996) (hereafter ERS), we use local to unity GLS detrending of the data. We consider two specific models. The first involves a break in the slope of the trend function and the second a break in both the intercept and slope. In this setup, there is no need to analyze the case where only a change in the intercept is allowed since the tests have then the same asymptotic distribution as the case where the deterministic components include a constant and a time trend which was analyzed in ERS. This is a consequence of condition B of ERS since a change in intercept is a special case of what they refer to as a "slowly evolving deterministic component".

The reasons for considering the M-tests, originally proposed by Stock (1990) and further analyzed by Perron and Ng (1996) is that these tests have much smaller size distortions than other classes of unit root tests even when the residuals have strong negative serial correlation. Also, using GLS detrending when constructing the M-tests allows substantial gains in power as showed by Ng and Perron (1999), similar to the DF^{GLS} test proposed by ERS. This is a consequence of the fact that a GLS framework permits a more accurate estimation of the deterministic component than what is possible using an OLS approach.

The implementation of the M^{GLS} tests requires a truncation lag (k) for the autoregressive spectral density estimate at frequency zero. We follow the suggestion of Ng and Perron (1999) to use a modified version of the AIC and BIC (labeled MIC). Their modification allows for the fact that the bias in the sum of the autoregressive coefficients is highly dependent on k in finite samples and adjust the penalty function accordingly.

Given that a uniformly most powerful test is not attainable, we follow ERS and derive a feasible point optimal test (P_T^{GLS}). The asymptotic power function of this test is derived and we use this power envelope to choose the non-centrality parameter ($\bar{c}$) to perform the GLS detrending such that the asymptotic power of the tests is 50% against the local alternative $\bar{\alpha} = 1 + \bar{c}/T$. For our two models, we obtain $\bar{c} = -23$.

The rest of the paper is organized as follows. The model and some preliminary theoretical results are presented in Section 2. In section 3, we derive the asymptotic distribution of the M^{GLS}, ADF^{GLS} and P_T^{GLS} in both cases where the break point is known or unknown. Section 4 considers the asymptotic Gaussian power envelope and the limit distribution of the feasible point optimal test. The asymptotic critical values and the asymptotic power function of the various tests are presented in Section 5. Section 6 considers the size and power of the tests in finite samples using simulations. Section 7 presents an empirical application and Section 8 briefly concludes. An appendix contains technical derivations.

2 GLS detrending with structural change

The data generating process considered is of the form:

$$y_t \quad = \quad d_t + u_t, \qquad t = 1, ..., T, \tag{1}$$

$$u_t = \alpha u_{t-1} + v_t, \tag{2}$$

where $\{v_t\}$ is an unobserved stationary mean-zero process. We use the assumption that $u_0 = 0$ throughout, thought the results generally hold for the weaker requirement that $Eu_0^2 < \infty$. The noise function is $v_t = \sum_{i=0}^{\infty} \gamma_i \eta_{t-i}$ with $\sum_{i=0}^{\infty} i|\gamma_i| < \infty$ and where $\{\eta_t\}$ is a martingale difference sequence. The process v_t has a non-normalized spectral density at frequency zero given by $\sigma^2 = \sigma_\eta^2 \gamma(1)^2$, where $\sigma_\eta^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{\infty} E(\eta_t^2)$. Furthermore, $T^{-1/2} \sum_{t=1}^{[rT]} v_t \Rightarrow \sigma W(r)$, where $\Rightarrow$ denotes weak convergence in distribution and $W(r)$ is the Wiener process defined on $C\,[0,1]$ the space of continuous functions on the interval $[0,1]$. In (1), $d_t = \psi' z_t$, where z_t is a set of deterministic components to be discussed below.

For any series y_t, with deterministic components z_t, we define the transformed data $y_t^{\bar{\alpha}}$ and $z_t^{\bar{\alpha}}$ by:

$$y_t^{\bar{\alpha}} = (y_1, (1-\bar{\alpha}L)y_t), \quad t = 2, ..., T,$$

$$z_t^{\bar{\alpha}} = (z_1, (1-\bar{\alpha}L)z_t), \quad t = 2, ..., T,$$

We let $\hat{\psi}$ be the estimator that minimizes:

$$S(\psi) = \left(y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}}\right)' \left(y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}}\right). \tag{3}$$

2.1 Model I: Structural change in the slope

For this model, the set of deterministic components, z_t in (1), is given by:

$$z_t = \{1,\ t, 1(t \geq T_B)(t - T_B)\}, \tag{4}$$

where 1(.) is the indicator function and T_B is the time of the change. Without loss of generality, we assume that $T_B = T\delta$ for some $\delta \in (0, 1)$. In this case, $\hat{\psi}(\delta) = (\hat{\mu}_1, \hat{\beta}_1, \hat{\beta}_2)'$ is the vector of estimates that minimizes (3). The next theorem provides the asymptotic distribution of these coefficient estimates.

Theorem 1 *Suppose that y_t is generated by (1) with $\alpha = 1 + c/T$ and $\{z_t\}$ is given by (4). Let $\hat{\psi}(\delta)$ be the GLS estimates, from (3), of the coefficients of the trend function obtained using $\bar{\alpha} = 1 + \bar{c}/T$. Then:*

$$\begin{aligned}
\hat{\mu}_1 - \mu_1 &\Rightarrow v_1; \\
T^{1/2}\left(\hat{\beta}_1 - \beta_1\right) &\Rightarrow \sigma\left(\lambda_1 b_1 + \lambda_2 b_2\right) \equiv \sigma b_3; \\
T^{1/2}\left(\hat{\beta}_2 - \beta_2\right) &\Rightarrow \sigma(\lambda_2 b_1 + \lambda_3 b_2) \equiv \sigma b_4;
\end{aligned}$$

where $b_1 = (1-\bar{c})W_c(1) + \bar{c}^2 \int_0^1 rW_c(r)dr$, $b_2 = (1-\bar{c}+\delta\bar{c})W_c(1) + \bar{c}^2 \int_\delta^1 W_c(r)(r-\delta)dr - W_c(\delta)$, $\lambda_1 = d/\Theta$, $\lambda_2 = -m/\Theta$, $d = 1 - \delta - \bar{c} + 2\bar{c}\delta - \bar{c}\delta^2 - \bar{c}^2\delta + \bar{c}^2\delta^2 + (\bar{c}^2/3)(1 - \delta^3)$, $m = 1 - \delta - \bar{c} + \bar{c}\delta - (\bar{c}^2/2)\delta + (\bar{c}^2/2)\delta^3 + (\bar{c}^2/3)(1 - \delta^3)$, $a = 1 - \bar{c} + \bar{c}^2/3$, $\Theta = ad - m^2$ and $\lambda_3 = a/\Theta$. Also, $W_c(r)$ is the Ornstein-Uhlenbeck process that is the solution to the stochastic differential equation $dW_c(r) = cW_c(r)dr + dW(r)$ with $W_c(0) = 0$.

2.2 Model II: Structural change in intercept and slope

For Model II, the deterministic components in (1) are:

$$z_t = \{1,\ 1(t \geq T_B), t, 1(t \geq T_B)(t - T_B)\}\,. \tag{5}$$

In this case, the vector of coefficient estimates is $\hat{\psi}(\delta) = (\hat{\mu}_1, \hat{\mu}_2, \hat{\beta}_1, \hat{\beta}_2)'$. In this model, we have the same result as the last theorem because the effect of $\hat{\mu}_2 - \mu_2$ is negligible in large samples. This is because the change in intercept is a special case of a slowly evolving deterministic component in condition B of ERS. Hence, we have:

Theorem 2 *Suppose that y_t is generated by (1) with $\alpha = 1 + c/T$ and $\{z_t\}$ is given by (5). Let $\hat{\psi}(\delta)$ be the GLS estimates, from (3), of the coefficients of the trend function obtained using $\bar{\alpha} = 1 + \bar{c}/T$. Then, the result of Theorem 1 still apply with the addition that $\hat{\mu}_2 - \mu_2$ $\Rightarrow \lim_{T\to\infty} v_{[T\delta]+1} \equiv v^*$.*

3 The tests and their asymptotic distributions

3.1 The tests

The M-tests, originally proposed by Stock (1990), and further analyzed by Perron and Ng (1996), exploit the feature that a series converges with different rates of normalization under the null and the alternative hypothesis. They are defined by:

$$MZ_\alpha(\delta) = (T^{-1}\tilde{y}_T^2 - s^2)\left(2T^{-2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2\right)^{-1} \tag{6}$$

$$MSB(\delta) = \left(T^{-2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2/s^2\right)^{1/2} \tag{7}$$

$$MZ_t(\delta) = (T^{-1}\tilde{y}_T^2 - s^2)\left(4s^2T^{-2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2\right)^{-1/2} \tag{8}$$

where $\widetilde{y}_t$ is y_t after detrending, i.e.,

$$\widetilde{y}_t = y_t - \hat{\psi}' z_t, \tag{9}$$

where $\hat{\psi}$ minimizes the expression (3). The term s^2 is the autoregressive estimate of the spectral density at frequency zero of v_t, defined as:

$$s^2 = s_{ek}^2 / \left(1 - \hat{b}(1)\right)^2, \tag{10}$$

where $s_{ek}^2 = T^{-1} \sum_{t=k+1}^{T} \hat{e}_{tk}^2$, $\hat{b}(1) = \sum_{j=1}^{k} \hat{b}_j$, with $\hat{b}_j$ and $\{\hat{e}_{tk}\}$ obtained from the autoregression:

$$\Delta \widetilde{y}_t = b_0 \widetilde{y}_{t-1} + \sum_{j=1}^{k} b_j \Delta \widetilde{y}_{t-j} + e_{tk}. \tag{11}$$

The first statistic is a modified version of the Phillips and Perron (1988) Z_α test originally developed by Phillips (1987). The second statistic is a modified version of Bhargava's (1986) R_1 statistic which builds upon the work of Sargan and Bhargava (1983). The third statistic is a modified version of the Phillips and Perron (1988) Z_t test. As Perron and Ng (1996) showed, the MSB and Z_α tests are approximately related by:

$$Z_t \approx MSB \cdot Z_\alpha.$$

This relation suggests the MZ_t tests defined by (8) since it satisfies the relation

$$MZ_t = MSB \cdot MZ_\alpha.$$

Another test of interest is the so-called ADF test (Said and Dickey, 1984) which is the t-statistic for testing $b_0 = 0$ in the regression (11). We denote this test by $ADF^{GLS}(\delta)$. Our approach is an extension of Ng and Perron (1999) and Elliott, Rothemberg and Stock (1996) to the case where the trend function contains a structural change. In this case, the M^{GLS} and ADF^{GLS} tests will depend of the unknown break point (δ).

3.2 Asymptotic distributions of the tests

We start with a statement of the limiting distribution of the various tests in the case where the break point is considered known.

Theorem 3 *Let y_t be generated by (1) with $\alpha = 1 + c/T$. Let MZ_α, MSB and MZ_t be defined by (6),(7) and (8) with data obtained from local GLS detrending $(\tilde{y}_t)$ at $\bar{\alpha} = 1 + \bar{c}/T$. Also, let ADF^{GLS} be the t-statistic for testing $b_0 = 0$ in the regression (11). In all cases, s^2 is a consistent estimate of σ^2. For Models I and II, we have:*

$$
\begin{aligned}
MZ_\alpha^{GLS}(\delta) &\Rightarrow \frac{0.5K_1(c,\bar{c},\delta)}{K_2(c,\bar{c},\delta)} \equiv H^{MZ_\alpha^{GLS}}(c,\bar{c},\delta), \\
MSB^{GLS}(\delta) &\Rightarrow (K_2(c,\bar{c},\delta))^{1/2} \equiv H^{MSB^{GLS}}(c,\bar{c},\delta), \\
MZ_t^{GLS}(\delta) &\Rightarrow \frac{0.5K_1(c,\bar{c},\delta)}{(K_2(c,\bar{c},\delta))^{1/2}} \equiv H^{MZ_t^{GLS}}(c,\bar{c},\delta), \\
ADF^{GLS}(\delta) &\Rightarrow \frac{0.5K_1(c,\bar{c},\delta)}{(K_2(c,\bar{c},\delta))^{1/2}} \equiv H^{ADF^{GLS}}(c,\bar{c},\delta),
\end{aligned}
$$

where

$$
\begin{aligned}
K_1(c,\bar{c},\delta) &= V_{c\bar{c}}^{(1)}(1,\delta)^2 - 2V_{c\bar{c}}^{(2)}(1,\delta) - 1, \\
K_2(c,\bar{c},\delta) &= \int_0^1 V_{c\bar{c}}^{(1)}(r)^2 dr - 2\int_\delta^1 V_{c\bar{c}}^{(2)}(r,\delta)dr,
\end{aligned}
$$

and $V_{c\bar{c}}^{(1)}(r,\delta) = W_c(r) - rb_3$, $V_{c\bar{c}}^{(2)}(r,\delta) = b_4(r-\delta)[W_c(r) - rb_3 - (1/2)(r-\delta)b_4]$; with b_3, b_4 and $W_c(r)$ as defined in Theorem 1.

In practice, it is usually the case that an investigator wants to treat the break point as unknown. In this case, an estimate is needed. A method suggested by Zivot and Andrews (1992) is to consider estimating δ as the break point that yields the minimal value of the statistics, i.e. using $\inf_\delta J^{GLS}(\delta)$ where $J = MZ_\alpha$, MSB, MZ_t, and ADF. Using the continuous mapping theorem and arguments as in Perron (1997), we have, assuming no shift in the trend function under the null hypothesis:

$$\inf_{\delta\in[0,1]} J^{GLS}(\delta) \Rightarrow \inf_{\delta\in[0,1]} H^{J^{GLS}}(c,\bar{c},\delta), \tag{12}$$

for $J = MZ_\alpha$, MSB, MZ_t, and ADF with the functions $H(\cdot)$ defined in Theorem 3. Note that no truncation for the range of possible break points needs to be imposed. As discussed in Vogelsang and Perron (1997), the implied estimate of δ is not consistent for the true value of the break point when the data generating process contains a break. These authors also note that the tests statistic are not invariant (even asymptotically) to values of the coefficients of the change in the trend. Nevertheless, they argue that, in typical sample sizes, this is not a problem unless the changes are extremely large. Thus, these tests can still be used with the critical values derived assuming no shift under the null hypothesis.

An alternative method to select the break date, as used in Perron (1997), is to choose it such that the absolute value of the t-statistic on the coefficient of the change in slope is maximized. This procedure has been used by many authors, e.g. Christiano (1992), Banerjee, Lumsdaine and Stock (1992), Perron (1997) and Vogelsang and Perron (1997). Consider, for example Model I where the deterministic component is given by

$$d_t = \mu_1 + \beta_1 t + \beta_2 (t - T_B) 1(t > T_B).$$

Let $\widehat{\beta}_2(\delta)$ be the GLS estimate of β_2 and $t_{\widehat{\beta}_2}(\delta)$ be its associated t-statistic. The break point can be selected using the estimate

$$\hat{\delta} = \arg \max_{\delta \in (\varepsilon, 1-\varepsilon)} |t_{\widehat{\beta}_2}(\delta)|,$$

where ε is some small number imposing a trimming on the possible values of the break dates. As discussed in Vogelsang and Perron (1997), if under the null hypothesis we have $\beta_2 \neq 0$ and the true break point given by $T_B^0/T = \delta^0$, then $\hat{\delta}$ is a consistent estimate of δ^0 and the limiting distributions of the test statistics correspond to those in the case where the break date is known, i.e. the limit distributions given in Theorem 3 evaluated at δ^0. In practice, one can simply evaluate these limit distributions at the estimated value $\hat{\delta}$.

When, under the null hypothesis, $\beta_2 = 0$ in which case there is no change in the slope of the trend function, it is easy to show (using the results of Theorem 1) that

$$t_{\widehat{\beta}_2}(\delta) \Rightarrow b_4/(\lambda_3^{1/2}),$$

where b_4 and λ_3 are defined in Theorem 1. We then have

$$\hat{\delta} = \arg \max_{\delta \in (\varepsilon, 1-\varepsilon)} |t_{\widehat{\beta}_2}(\delta)| \Rightarrow \arg \max_{\delta \in (\varepsilon, 1-\varepsilon)} |b_4/(\lambda_3^{1/2})| \equiv \delta^*. \tag{13}$$

Hence, the limiting distributions of the statistics are given by

$$J^{GLS}(\hat{\delta}) \Rightarrow H^{J^{GLS}}(c, \bar{c}, \delta^*), \tag{14}$$

for $J = MZ_\alpha$, MSB, MZ_t, and ADF with the functions $H(\cdot)$ defined in Theorem 3.

In practice, it is difficult to know if there is a change in slope since any test of such hypothesis would depend on whether a unit root is present or not. Hence, a conservative procedure is to use the critical values corresponding to the case where it is assumed that no break is present, i.e. (14). This is the procedure we use in the following.

4 Feasible point optimal test and the power envelope

Elliott, Rothemberg and Stock (1996), following Dufour and King (1991), have considered the issue of developing tests with optimality properties under Gaussian errors. The case where the break point is assumed known follows closely their analysis. While a uniformly most powerful test is not attainable, it is possible to define a point optimal test against the alternative $\alpha = \bar{\alpha}$. If v_t is *i.i.d.*, this is provided by the likelihood ratio statistic, which simplifies to $L(\delta) = S(\bar{\alpha}, \delta) - S(1, \delta)$, where $S(\bar{\alpha}, \delta)$ and $S(1, \delta)$ are the sums of squared errors from a GLS regression with $\alpha = \bar{\alpha}$ and $\alpha = 1$, respectively. Varying the value of $\bar{\alpha}$, gives a family of point optimal tests and the power envelope. Under the assumption that the errors follow a normal distribution, the power function forms a Gaussian power envelope for testing $\alpha = 1$. ERS have proposed an asymptotic version of a feasible point optimal test (P_T^{GLS}) which takes into account that v_t is a serially correlated series. The P_T^{GLS} test is defined by:

$$P_T^{GLS}(c, \bar{c}, \delta) = \{S(\bar{\alpha}, \delta) - \bar{\alpha} S(1, \delta)\}/s^2. \qquad (15)$$

The next theorem provides the limiting distribution of the P_T^{GLS} test:

Theorem 4 *Let y_t be generated by (1) with $\alpha = 1 + c/T$. Let P_T^{GLS} be defined by (15) with data obtained from local GLS detrending ($\tilde{y}_t$) at $\bar{\alpha} = 1 + \bar{c}/T$. Also, let s^2 be a consistent estimate of σ^2. Then, the limit distribution for the P_T^{GLS} test is the same under Models I and II and is given by:*

$$P_T^{GLS}(\delta) \Rightarrow J_{1c}(r) + J_{2c}(r, \delta) \equiv H^{P_T^{GLS}}(c, \bar{c}, \delta), \qquad (16)$$

where

$$\begin{aligned} J_{1c}(r) &= \bar{c}^2 \int_0^1 W_c(r)^2 dr - \bar{c} W_c(1)^2, \\ J_{2c}(r, \delta) &= (1 - \delta)^{-1}[W_c(1)^2 - 2W_c(1)W_c(\delta) + \delta^{-1} W_c(\delta)^2] \\ &\quad -\lambda_1 b_1^2 - 2\lambda_2 b_1 b_2 - \lambda_3 b_2^2. \end{aligned}$$

The asymptotic expression (16) for the P_T^{GLS} test allows us to define the asymptotic power envelope for the two models. It is given by

$$\pi(c, \delta) = \Pr[H^{P_T^{GLS}}(c, c, \delta) < b^{P_T^{GLS}}(c, \delta)],$$

where $b^{P_T^{GLS}}(c,\delta)$ is such that

$$\Pr[H^{P_T^{GLS}}(0,c,\delta) < b^{P_T^{GLS}}(c,\delta)] = \upsilon,$$

with υ the size of the test. Note that in general, a different power envelope exists for each values of δ.

When δ is unknown, things are rather different. The principle is, however, the same. To maximize the likelihood function, the estimate of δ must minimize the sum of squares residuals $S(1,\delta)$ and $S(\bar{\alpha},\delta)$ under the null and alternative hypotheses, respectively. This leads to the statistic $\inf_{\delta\in[0,1]} P_T^{GLS}(\delta)$ which, using Theorem 4 and the continuous mapping theorem, has the following distribution

$$\inf_{\delta\in[0,1]} P_T^{GLS}(\delta) \Rightarrow \inf_{\delta\in[0,1]} H^{P_T^{GLS}}(c,\bar{c},\delta).$$

The asymptotic Gaussian power envelope is then defined as

$$\pi^*(c) = \Pr[\inf_{\delta\in[0,1]} H^{P_T^{GLS}}(c,c,\delta) < b_*^{P_T^{GLS}}(c)],$$

where $b_*^{P_T^{GLS}}(c)$ is such that

$$\Pr[\inf_{\delta\in[0,1]} H^{P_T^{GLS}}(0,c,\delta) < b_*^{P_T^{GLS}}(c)] = \upsilon, \tag{17}$$

with υ the size of the test.

Furthermore, the power envelope allows us to find the "optimal" non-centrality parameter ($\overline{c}$) for our models. ERS recommended to choose the value $\overline{c}$ such that the asymptotic power of the test is 50%, i.e. $\bar{c}$ is such that $\Pr[\inf_{\delta\in[0,1]} H^{P_T^{GLS}}(\bar{c},\bar{c}) < b_*^{P_T^{GLS}}(\bar{c})] = 0.5$. Using simulations, we found that $\bar{c} = -23$ and we, henceforth use this value in the rest of the paper.

For comparison with the other tests, we shall also consider the version of the feasible point optimal test when the break date is chosen maximizing the t-statistic on the coefficient of the change in slope; that is $P_T^{GLS}(\hat{\delta})$ with $\hat{\delta} = \arg\max|t_{\widehat{\beta}_2}(\delta)|$. It is straightforward to deduce that the asymptotic distribution of $P_T^{GLS}(\hat{\delta})$ is given by

$$P_T^{GLS}(\hat{\delta}) \Rightarrow H^{P_T^{GLS}}(c,\bar{c},\delta^*),$$

where δ^* is defined by (13). Note that this method of choosing the break point is ad hoc since it is not optimal in the sense of leading to a maximal value of the likelihood ratio as does the other method. We should, therefore, expect this version of the tests to have lower power compared to the version where the break point is chosen by minimizing the tests.

5 Critical values and asymptotic power functions

In this section, we obtain the asymptotic critical values for tests assuming $\bar{c} = -23$ is used to detrend the data. We simulate directly the asymptotic distributions 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. The limiting distributions are tabulated for the null hypothesis $c = 0$. For the finite sample distributions, we use $T = 100$ with data generated by a random walk with zero initial condition and *i.i.d.* $N(0,1)$ errors. Here k is set to 0 which is equivalent to using the true value of σ^2; the effects of selecting k on the finite sample critical values are investigated in the next section. In all cases 10,000 replications are used. The results are presented in Table 1.a for the case where the break point is selected by minimizing the tests, and in Table 1.b when the break point is selected maximizing the absolute value of the t-statistic on the coefficient of the change in slope. We omit the ADF^{GLS} test since it has the same asymptotic distribution as the MZ_t^{GLS} test. In general, the approximation to the finite sample distribution is adequate but somewhat less good for model 2 which contains a change in intercept that is asymptotically negligeable.

The asymptotic power functions of the tests are defined by:

$$\pi^*_{J^{GLS}}(c, \overline{c}) = \Pr[\inf_{\delta \in [0,1]} H^{J^{GLS}}(c, \bar{c}, \delta) < b^{J^{GLS}}(\bar{c})],$$

or

$$\pi^*_{J^{GLS}}(c, \overline{c}) = \Pr[H^{J^{GLS}}(c, \bar{c}, \delta^*) < b_*^{J^{GLS}}(\bar{c})],$$

for $J = MZ_\alpha$, MSB, MZ_t, ADF and P_T with $H^i(c, \bar{c})$ defined in Theorems 3 and 4, and δ^* is defined by (13). The constants $b^{J^{GLS}}(\bar{c})$ and $b_*^{J^{GLS}}(\bar{c})$ are such that $\Pr[\inf_{\delta \in (0,1)} H^{J^{GLS}}$ $(0, \bar{c}, \delta) < b^{J^{GLS}}(\bar{c})] = v$, and $\Pr[H^{J^{GLS}}(0, \bar{c}, \delta^*) < b_*^{J^{GLS}}(\bar{c})] = v$, the size of the tests. The asymptotic power functions are showed in figures 1 to 8. The solid line is the power envelope. As can see, the M^{GLS} tests, and especially the P_T^{GLS} test have power functions very close to the power envelope when the break point is selected by minimizing the tests. On the other hand, the power functions of the same tests with the break point selected maximizing the absolute value of the t-statistic of the coefficient on the change in slope are clearly lower than the power envelope. We should there expect this class of tests to have less desirable finite sample properties.

6 Size and power of the tests in finite samples

6.1 The size issue, the selection of k and information criteria

It is clear that all tests considered require the estimation of the augmented autoregression (11). Ng and Perron (1997) recommended using GLS detrended data using the same non-centrality parameter $\bar{c}$ for constructing the autoregression and the tests. They also tried using GLS detrended data under the null hypothesis ($\bar{c} = 0$) when the autoregression (11) is used to construct s^2 but found, in the linear trend case, that the properties of the tests were very similar. Our simulations showed, however, that in the case with a break in the slope of the trend function, using $\bar{c} = 0$ to GLS-detrend the data when estimating the autoregression (11) to construct s^2 led to tests with better finite sample properties[1]. Hence, in what follows $\bar{c} = -23$ is used to detrend the data when constructing the tests but $\bar{c} = 0$ is used to detrend the data when estimating the autoregression (11). Of course, when constructing the ADF^{GLS} test, the autoregresion (11) uses data detrended by GLS using the value $\bar{c} = -23$.

To see how the lag order, k, influences the behavior of the tests, we first consider the finite sample size of the MZ_α^{GLS} and P_T^{GLS} tests for given fixed values of the truncation lag k under a variety a data-generating processes. We consider simulations for Model

[1] To support this assertion, Annex 3 include a set of tables using $\overline{c} = -23$ to construct s^2.

I. Reported in Tables 2.a and 2.b are the sizes of the tests at selected values of θ when v_t is an MA(1) process , i.e. when $v_t = e_t + \theta e_{t-1}$ with $e_t \sim i.i.d.\ N(0,1)$. Tables 2.c and 2.d present results at selected values of ρ when v_t is an AR(1) process, i.e., when $v_t = \rho v_{t-1} + e_t$. We report results for $T = 100$ and $T = 200$. The nominal size of 5% is used as the benchmark.

Several features of the results for MA errors are of note here. First, for a θ of the same absolute value, a negative θ always requires a larger lag to obtain a more accurate size. Second, as the sample increases, so does the "optimal k", defined as the smallest k for which the exact size is "closest" to the nominal size of 5%. Third, for positive θ, the sizes of the tests are significantly better when k is even than when it is odd. Fourth, if we compare with the results of Ng and Perron (1999), the larger the number of deterministic terms, the more distant are the exact sizes from the nominal ones. For the MZ_α^{GLS} tests, it is to be noted that when $\theta = -0.8$, the smallest size achievable with a given fixed k is 0.191 (at $k = 5$) when $T = 100$ and 0.084 (at $k = 7$) when $T = 200$. These are quite large, but as we shall see some data dependent methods to select k are able to yield tests with sizes much closer to 5%, i.e. better than what is achievable with the "best" fixed k.

For the results with AR errors, size discrepancies also exist, albeit not as dramatic as in the MA case. For large negative AR errors, the tests are undersized, and for large positive errors, they are oversized. As analyzed in Perron and Ng (1998), the autoregressive spectral density estimator has more stable properties when there are over 100 observations. The size distortions are apparently much reduced as we increase the size of the sample.

Clearly, the selection of k is very important for the properties of the tests, especially in the presence of negative moving average errors. Various practical solutions have been suggested for this problem. In ERS, the authors use the BIC to select k but they set the lower bound to be 3. This is because the BIC would have chosen k to be 0 or 1 frequently if zero was the lower bound. An alternative method of forcing in larger k's is the sequential t test for the significance of the last lag considered in Ng and Perron (1995). For any chosen upper bound, say $k\max$, the procedure will select $k\max$ with a probability equalling the size of the test. The procedure thus has the ability to yield higher k's for the augmented autoregression than the BIC when there are negative moving average errors and, hence, smaller size distortions are obtained. But, the sequential procedure over-parameterizes the augmented autoregression in other cases. This, as does ERS's implementation of the BIC, leads to less satisfactory estimates and subsequently to power losses. Neither approach is fully satisfactory.

Recently, Ng and Perron (1999) have proposed a modification of the AIC and BIC, labeled MIC, which can be expressed as follows:

$$k_{mic} = \arg\min_k \log(s_{ek}^2) + \frac{C_T(\hat{\tau}_T(k) + k)}{T},$$

where

$$\hat{\tau}_T(k) = (s_{ek}^2)^{-1}\hat{b}_0^2 \sum_{t=1}^{T} \tilde{y}_{t-1}^2.$$

The $MAIC$ uses $C_T = 2$ and the $MBIC$ uses $C_T = \log(T)$. Ng and Perron (1999), based on theoretical considerations and simulations, recommended $MAIC$. The advantage of the MIC is that it takes into account the possible dependence of $\hat{b}_0$ on k. For example, with a large negative MA component, $\hat{b}_0$ decreases substantially as k increases. Hence, the term $\hat{\tau}_T(k)$ implies a higher penalty for a low value of k. In such cases, a larger k is selected. When there is little correlation in the residuals, $\hat{b}_0$ is quite insensitive to k and the $MAIC$ is basically equivalent to AIC.

To see some of the empirical properties of tests based on the MIC, we performed the following simulation experiment. For a given DGP, we constructed the MZ_α^{GLS} test for our first model at each $k \in [0, 10]$, and recorded the exact size of the tests. We then found the k, denoted k^*, as the first k with a size closest to within three standard errors of the nominal size of 0.05. For 1000 replications, the standard error in the simulated size of the tests is $(0.05(0.95)/1000)^{1/2} = 0.007$. Thus, k^* is the first k that fall in the range 0.029 and 0.071. If no such k exists, k^* is the k with the smallest absolute deviation from the nominal size of 5%. This procedure is used to obtain a set of k^* for our first model. We then obtain k_{bic} as the median value selected by BIC over the range 0 and 10. A similar procedure is used to obtain k_{aic}, k_{maic} and k_{mbic}. The simulations are based on MA(1) and AR(1) processes assuming the null hypothesis that $\alpha = 1$. The results of the simulation experiments are reported in Table 3. They reveal that k_{maic} and k_{mbic} select lag orders that are, indeed, very close to k^* for all parameter configurations, except when there is a large negative AR component in which case all methods select $k = 1$ instead of $k = 4$ or 5 which would yield M^{GLS} tests with better sizes[2].

[2] Of course, in the pure AR case the "optimal" lag for the ADF test is 1 and for this test the order selected by k_{maic} and k_{mbic} are indeed adequate.

6.2 Finite sample critical values with data dependent methods to select k

We now consider the finite sample critical values of the tests using the various data-dependent methods to select the truncation lag described above. While the asymptotic distributions provide good approximations to the finite sample distributions, it is sometimes the case that using a data dependent method to select k can induce distortions. Our simulations are based on 1000 replications of the DGP defined by (1) and (2) with $d_t = 0$, $\alpha = 1$ and $v_t \sim i.i.d.\ N(0,1)$. We present results for the cases where the lag length of the autoregression (11) is selected using the AIC, BIC, $MAIC$, $MBIC$ and the sequential t-sig methods for the ADF^{GLS} test or for constructing the autoregressive spectral density estimator for the other tests. In the simulations, the lower bound on the lag length is always zero to reduce the chance of over-parameterizing when a large k is not necessary. The upper bound is $k\max = int[10(T/100)^{1/4}]$ for the AIC, BIC, $MAIC$ and $MBIC$ methods and $k\max = int[4(T/100)^{1/4}]$ for the sequential t-sig method. We consider $T = 100$, $T = 150$ and $T = 200$. The results are presented in Table 4 (Model 1, choosing T_B minimizing the tests), Table 5 (Model 1, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$), Table 6 (Model 2, choosing T_B minimizing the tests), and Table 7 (Model 2, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$).

The results show the finite sample distributions of the tests involving the autoregressive spectral density estimator to be very sensitive to the method used to select the truncation lag k. In particular, the AIC and t-sig methods imply that the finite sample distributions of the M^{GLS} tests and the P_T^{GLS} test are much different in the left tail. Indeed, the finite sample critical values are much smaller than those of the corresponding asymptotic distributions. Hence, the use of the latter would imply very liberal tests. On the other hand, when k is chosen using the $MAIC$, $MBIC$ or BIC criteria, the finite sample distributions are somewhat less spread than the asymptotic ones. Here, the use of asymptotic critical values would imply slightly conservative tests. These problems are common for the two models and the two procedures to select the break point. The problems for the AIC and sequential t-sig methods are somewhat less severe when choosing the break point by maximizing the absolute value of the sequential t-statistic on the coefficient of the change in slope and comparatively more severe for Model 2 than for Model 1. Hence, the results in this base case show that the AIC or sequential t-sig methods should not be used to select k when constructing the autoregressive spectral density estimate for the M^{GLS} tests and the P_T^{GLS} test.

Overall, the finite sample distribution of the ADF^{GLS} test is little affected by the method to choose k. The results shows that inference with the asymptotic critical values would yield slightly liberal tests with AIC and sequential t-sig and slightly conservative tests with the criteria $MAIC$, $MBIC$ or BIC.

6.3 Size and power of tests

We now consider the size of the tests in finite samples using the various data-dependent methods to select the truncation lag described above. Our simulations are based on 1000 replications of the DGP defined by (1) with $d_t = 0$. We consider pure $MA(1)$ processes, i.e. with $v_t = (1 + \theta L)e_t$ and pure $AR(1)$ processes, i.e. $(1 + \rho L)v_t = e_t$, where $e_t \sim i.i.d.$ $N(0,1)$. For both the $MA(1)$ and $AR(1)$ cases, we consider θ and ρ in the range $[-0.8, 0.8]$. The autoregressive spectral density estimator defined by (10) is used. We present results for the cases where the lag length of the autoregression is selected using the AIC, BIC, $MAIC$, $MBIC$ and the sequential t-sig methods. In the simulations, the lower bound on this lag length is always zero to reduce the chance of over-parameterizing when a large k is not necessary. The upper bound is $k\max = int[10(T/100)^{1/4}]$ for the AIC, BIC, $MAIC$ and $MBIC$ methods and $k\max = int[4(T/100)^{1/4}]$ for the sequential t-sig method. We consider the sample sizes $T = 100$ and $T = 200$. Given the finite sample results documented in the previous section, we calculate the size and power of the tests using the 5% finite sample critical values for the case where the errors are *i.i.d.*. The power is evaluated at $\bar{\alpha} = 1 + \bar{c}/T$ for $\bar{c} = -23$ which implies that the asymptotic power is 50%.

The results for the case where the break point is chosen by minimizing the tests are presented in Table 8 for $T = 100$ and Table 9 for $T = 200$. In the case where the errors are *i.i.d.*, as expected, the power of the tests when the methods AIC or sequential t-sig are used to select k is very low. On the other hand, when the $MBIC$, $MAIC$ or BIC methods are used, the power is indeed close to the asymptotic value of 50%. It is somewhat higher when $MBIC$ is used and somewhat lower with $MAIC$. For the ADF^{GLS} the power is high for all methods to choose k but again higher if $MBIC$ is used. Given these results, we shall not discuss further the behavior of the tests with the AIC or the sequential t-sig methods.

Consider now the case where the errors have a negative MA component. For all tests, the use of the BIC to select k implies tests with severe distortions with exact size above 75% when the MA component is -0.8 and about 30% when it is -0.4. On the other hand, the $MBIC$ or $MAIC$ implies that the M^{GLS} tests and the P_T^{GLS} test have much less size distortions. The exact size is about 12% when the coefficient is -0.8 and about 6% when it is −0.4. But even when using the $MAIC$ or the $MBIC$, the ADF^{GLS} test still suffers from high size distortions (with $\theta = -.8$ it is 35.3% with $MAIC$ and $MBIC$). An interesting fact is that, the M^{GLS} tests and the P_T^{GLS} test, using the $MAIC$ lead to tests with much smaller size distortions than if the "best fixed k" had been used. For example, at $T = 100$, the exact size is .12 for the MZ_α^{GLS} with $MAIC$ as opposed to .19 with a fixed $k = 5$. When $T = 200$, the size distortions of the M^{GLS} tests and the P_T^{GLS} test disappears for all values of θ considered. The ADF^{GLS} test continues, however, to exhibit important size distortions at $\theta = -.8$.

When the errors have a positive moving average coefficient, the M^{GLS} tests and the P_T^{GLS} tests have the correct size only when the $MBIC$ is used, otherwise they are liberal. The ADF^{GLS} test has the correct size with either $MAIC$ or $MBIC$ but power is very low.

Consider now the case where the errors have a negative autoregressive component. Here, the M^{GLS} tests and the P_T^{GLS} test are very conservative and, hence, show basically no power. The ADF has the correct size and power is good. When the autoregressive coefficient is positive, the M^{GLS} tests and the P_T^{GLS} test are liberal. The ADF^{GLS} has better size but no power.

Overall, the results show that the ADF^{GLS} test with the truncation lag chosen using either the $MAIC$ or the $MBIC$ has better overall properties unless there is a negative MA component in the residuals, in which case the M^{GLS} tests and the P_T^{GLS} test are superior. The size distortions are, however, smaller and power also higher when $T = 200$.

The results pertaining to the case where the break point is selected by maximizing the absolute value of the t-statistic on the coefficient of the change in slope are presented in Table 10 for $T = 100$ and Table 11 for $T = 200$. They show size and power properties which are much less satisfactory than for the case where the break point is selected by minimizing the tests. First, in the case of *i.i.d.* errors, the power is substantially lower. Even at $T = 200$, the power is much lower than the asymptotic target of 50% as suggested by the asymptotic power analysis of Section 5. Secondly with negative MA errors, the M^{GLS} tests and the P_T^{GLS} test using the $MAIC$ or $MBIC$ are so conservative under the null hypothesis that they have no power. The ADF^{GLS} test have slight size distortions but also little power. Third, power is low for all tests when the errors have a positive MA component though here size distortions are smaller. Fourth, with negative AR errors, even the ADF^{GLS} test is very conservative and, accordingly, has little power. Finally, with positive AR errors, there are still liberal size distortions for the M^{GLS} tests and the P_T^{GLS} test; the ADF^{GLS} test has the correct size but little power.

Overall, the theoretical and simulation results show that the tests based on choosing the break point by minimizing the test are better in terms of size and power and should be recommended for practical applications.

7 Empirical applications

Among the macroeconomic time series considered by Nelson and Plosser (1982), Perron (1989) argued that two of them were likely affected by a significant change in slope for the samples analyzed, namely the Real Wages and Stock Prices series. The series are presented in Figures 9 and 10. We re-evaluate the claim made by Perron (1989) to the effect that the noise function of these series is stationary if allowance is made for such a change in slope using the tests described here. We applied the MZ_t, P_T and ADF tests using the BIC, $MAIC$ and $MBIC$ criteria to select the autoregressive order (imposing a minimal value of 1). The estimate s^2 is constructed from an autoregression with data detrended using the GLS method with $\bar{c} = 0$. For the construction of the tests (and to estimate the autoregression used for the ADF test), we use GLS detrended data with $\bar{c} = -23$.

The results are presented in Table 12.a for the case where the break date is selected minimizing the appropriate test and in Table 12.b for the case where the break date is selected maximizing the absolute value of the t-statistic on the coefficient of the change in slope (in which case the estimated break date is common for all tests). Using the former method, all tests points to a strong rejection at the 1% significance level for the Stock Price series with the break date estimated at either 1937 or 1945 depending on the specification used. For the Real Wage series, there is a rejection at least at the 5% significance level using the criterion $MAIC$ or $MBIC$ to select k and at the 10% using the BIC. The estimated break date is at 1938 or 1940 depending on the specification used. The estimate trend function is plotted in Figures 9 and 10 along with the series using $T_B = 1937$ for the Stock Price series and $T_B = 1938$ for the Real Wages series.

When choosing the break point maximizing the absolute value of the t-statistic on the coefficient of the change in slope, the results are different as predicted by the asymptotic and finite sample results discussed before. Here, using $MAIC$ or $MBIC$ to select k a rejection of the unit root is no longer possible for the Stock Prices series (although a rejection at the 2.5% and 5% levels is possible using BIC, for P_T and ADF tests, respectively). Similarly, the evidence against the unit root is considerably weaker for the Real Wages series. This accords with our asymptotic analysis which indicated that this method to choose the break point leads to tests with lower power.

8 Conclusions

This paper has considered a class of tests for testing the null hypothesis of a unit root in the presence of a one time change in the slope or the slope and intercept of a trend function. Our approach followed that of Elliott, Rothemberg and Stock (1996) in that we considered detrending the data using a local to unity GLS approach. We considered the required extensions of the ADF and the P_T as well as of the various M tests suggested by Perron and Ng (1996). We also followed the recent literature in investigating the properties of the tests when the break point is selected either by minimizing the tests or by maximizing the absolute value of the t-statistic on the coefficient of the change in slope. A novel aspect of our results is that the latter method to select the break point leads to tests that are asymptotically inferior in the sense that their local asymptotic power function lies well below the Gaussian power envelope. On the other hand, the former method leads to tests whose local asymptotic power function is very close to the envelope. This power issues was corroborated, for finite samples, using simulation experiments which also showed even comparatively worse size distortions. Hence, for applications, we recommend using either the ADF^{GLS} or the M^{GLS} tests and P_T^{GLS} test. The difference is that the ADF^{GLS} has worse size distortions in the negative MA case but better power in the negative AR case; and the M^{GLS} tests and P_T^{GLS} test have good size overall but very little power in the negative AR case. The choice between the two depends on the investigator's assessment of the likely importance of one or the other class of processes in the data considered. Our experiments also suggest that the use of the $MAIC$ to select the autoregressive truncation lag leads to tests with better properties overall.

The analysis pertaining to the tests where the break is selected using the maximal value of the t-statistic on the coefficient of the change in slope rests, however, on the assumption that the investigator does not know if a break has occurred. If the investigator has evidence that a break has occurred, even if the timing is unknown, the relative merits of this strategy could drastically change. In such a case, the method consistently estimates the break point and, hence, the limiting distribution is the same as that of the case where the break point is known. Methods to tests if a change in the trend function has occurred being agnostic about the presence or absence of a unit root have been considered by Perron (1991) and Vogelsang (1997). Hence, it is possible to make this procedure operational in practice. This, however, implies a different asymptotic power envelope for each case (since the power envelope then depends on the break point) and, hence, the derivation of an "optimal" non-centrality parameter $\bar{c}$ on a case by case basis. Hence, the method becomes heavily computer intensive. It is nevertheless, an interesting avenue for future research.

9 Appendix

Throughout, we use the following lemma which is by now standard.

Lemma A.1: Let $\{u_t\}$ be a near-integrated series generated by (2). Then, we have: a) $T^{-1/2}u_{[Tr]} \Rightarrow \sigma W_c(r)$; b) $T^{-3/2}\sum_{t=1}^{T} u_t \Rightarrow \sigma \int_0^1 W_c(r)dr$; c) $T^{-2}\sum_{t=1}^{T} u_t^2 \Rightarrow \sigma^2 \int_0^1 W_c^2(r)dr$; d) $T^{-1}\sum_{t=1}^{T} u_{t-1}v_t \Rightarrow \sigma^2\{\int_0^1 W_c(r)dW(r)+\gamma\}$ with $\gamma = (\sigma^2 - \sigma_v^2)/2\sigma^2$.

Proof of Theorem 1: In matrix notation, we have:

$$\hat{\psi}(\delta) - \psi = \left[\left(\Delta z - \bar{c}T^{-1}z_{-1}\right)'\left(\Delta z - \bar{c}T^{-1}z_{-1}\right)\right]^{-1} \left[\left(\Delta z - \bar{c}T^{-1}z_{-1}\right)'\left(\Delta u - \bar{c}T^{-1}u_{-1}\right)\right], \tag{A.1}$$

where

$$\begin{aligned} \Delta z &= (z_1, z_2 - z_1, \ldots.., z_T - z_{T-1}), \\ z_{-1} &= (0, z_1, z_2, \ldots., z_{T-1}), \\ \Delta u &= (u_1, u_2 - u_1, \ldots.., u_T - u_{T-1}), \\ u_{-1} &= (0, u_1, u_2, \ldots., u_{T-1}). \end{aligned}$$

Now define the scaling matrix $\Upsilon_T = diag(1, T^{1/2}, T^{1/2})$, we can write expression (A.1) as:

$$\Upsilon_T\left(\hat{\psi}(\delta) - \psi\right) = \Gamma_T(\delta)^{-1}\Psi_T(\delta), \tag{A.2}$$

where

$$\Gamma_T(\delta) = \Upsilon_T^{-1}\left[\left(\Delta z - \bar{c}T^{-1}z_{-1}\right)'\left(\Delta z - \bar{c}T^{-1}z_{-1}\right)\right]\Upsilon_T^{-1},$$

and

$$\Psi_T(\delta) = \Upsilon_T^{-1}\left[\left(\Delta z - \bar{c}T^{-1}z_{-1}\right)'\left(\Delta u - \bar{c}T^{-1}u_{-1}\right)\right].$$

We first consider the limit of each element of the matrix $\Gamma_T(\delta)$ denoted Γ_{ij} $(i,j = 1,2,3)$. We let $\Delta z_{(i)}$ and $z_{-1(i)}$ be the ith element of the vectors Δz and z_{-1}, respectively. We have:

1. $\Gamma_{11} = \left(\Delta z_{(1)} - \bar{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(1)} - \bar{c}T^{-1}z_{-1(1)}\right) \Rightarrow 1$;
2. $\Gamma_{12} = T^{-1/2}\left(\Delta z_{(1)} - \bar{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(2)} - \bar{c}T^{-1}z_{-1(2)}\right) \Rightarrow 0$;
3. $\Gamma_{13} = T^{-1/2}\left(\Delta z_{(1)} - \bar{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(3)} - \bar{c}T^{-1}z_{-1(3)}\right) \Rightarrow 0$;
4. $\Gamma_{22} = T^{-1}\left(\Delta z_{(2)} - \bar{c}T^{-1}z_{-1(2)}\right)'\left(\Delta z_{(2)} - \bar{c}T^{-1}z_{-1(2)}\right) \Rightarrow 1 - \bar{c} + \bar{c}^2/3 \equiv a$;
5. $\Gamma_{23} = T^{-1}\left(\Delta z_{(2)} - \bar{c}T^{-1}z_{-1(2)}\right)'\left(\Delta z_{(3)} - \bar{c}T^{-1}z_{-1(3)}\right)$
   $\Rightarrow 1 - \delta - \bar{c} + \bar{c}\delta - (\bar{c}^2/2)\delta + (\bar{c}^2/2)\delta^3 + (\bar{c}^2/3)(1-\delta^3) \equiv m$;

6. $\Gamma_{33} = T^{-1}\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right)'\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right)$
$\Rightarrow 1 - \delta - \overline{c} + 2\overline{c}\delta - \overline{c}\delta^2 - \overline{c}^2\delta + \overline{c}^2\delta^2 + (\overline{c}^2/3)(1-\delta^3) \equiv d.$

We next consider the limit of each element of the vector $\Psi_T(\delta)$, denoted Ψ_i $(i = 1,2,3)$. We have:

1. $\Psi_1 = \left(\Delta z_{(1)} - \overline{c}T^{-1}z_{-1(1)}\right)'\left(\Delta u - \overline{c}T^{-1}u_{-1}\right) \Rightarrow v_1;$
2. $\Psi_2 = T^{-1/2}\left(\Delta z_{(2)} - \overline{c}T^{-1}z_{-1(2)}\right)'\left(\Delta u - \overline{c}T^{-1}u_{-1}\right)$
$\Rightarrow \sigma[W_c(1)(1-\overline{c}) + \overline{c}^2\int_0^1 rW_c(r)dr] \equiv \sigma b_1;$
3. $\Psi_3 = T^{-1/2}\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right)'\left(\Delta u - \overline{c}T^{-1}u_{-1}\right)$
$\Rightarrow \sigma[W_c(1)(1-\overline{c}+\delta\overline{c}) + \overline{c}^2\int_\delta^1 W_c(r)(r-\delta)dr - W_c(\delta)] \equiv \sigma b_2.$

Hence, using the symmetry of $\Gamma_T(\delta)$,

$$\Upsilon_T\left(\hat{\psi}(\delta) - \psi\right) \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & m \\ 0 & m & d \end{bmatrix}^{-1} \begin{bmatrix} v_1 \\ \sigma b_1 \\ \sigma b_2 \end{bmatrix}.$$

The proof of the Theorem follows upon solving for the inverse.

Proof of Theorem 2: Here, the scaling matrix $\Upsilon_T = diag(1, T^{1/2}, T^{1/2})$. We first consider the limit of each element of the matrix $\Gamma_T(\delta)$ denoted Γ_{ij} $(i,j = 1,2,3,4)$. We let $\Delta z_{(i)}$ and $z_{-1(i)}$ be the ith element of the vectors Δz and z_{-1}, respectively. We have:

1. $\Gamma_{11} = \left(\Delta z_{(1)} - \overline{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(1)} - \overline{c}T^{-1}z_{-1(1)}\right) \Rightarrow 1;$
2. $\Gamma_{12} = T^{-1/2}\left(\Delta z_{(1)} - \overline{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(2)} - \overline{c}T^{-1}z_{-1(2)}\right) \Rightarrow 0;$
3. $\Gamma_{13} = T^{-1/2}\left(\Delta z_{(1)} - \overline{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right) \Rightarrow 0;$
4. $\Gamma_{14} = T^{-1/2}\left(\Delta z_{(1)} - \overline{c}T^{-1}z_{-1(1)}\right)'\left(\Delta z_{(4)} - \overline{c}T^{-1}z_{-1(4)}\right) \Rightarrow 0;$
5. $\Gamma_{22} = T^{-1}\left(\Delta z_{(2)} - \overline{c}T^{-1}z_{-1(2)}\right)'\left(\Delta z_{(2)} - \overline{c}T^{-1}z_{-1(2)}\right) \Rightarrow 1;$
6. $\Gamma_{23} = T^{-1}\left(\Delta z_{(2)} - \overline{c}T^{-1}z_{-1(2)}\right)'\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right) \Rightarrow 0;$
7. $\Gamma_{24} = T^{-1}\left(\Delta z_{(2)} - \overline{c}T^{-1}z_{-1(2)}\right)'\left(\Delta z_{(4)} - \overline{c}T^{-1}z_{-1(4)}\right) \Rightarrow 0;$
8. $\Gamma_{33} = T^{-1}\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right)'\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right) \Rightarrow 1 - \overline{c} + \overline{c}^2/3 \equiv a;$
9. $\Gamma_{34} = T^{-1}\left(\Delta z_{(3)} - \overline{c}T^{-1}z_{-1(3)}\right)'\left(\Delta z_{(4)} - \overline{c}T^{-1}z_{-1(4)}\right)$
$\Rightarrow 1 - \delta - \overline{c} + \overline{c}\delta - (\overline{c}^2/2)\delta + (\overline{c}^2/2)\delta^3 + (\overline{c}^2/3)(1-\delta^3) \equiv m;$
10. $\Gamma_{44} = T^{-1}\left(\Delta z_{(4)} - \overline{c}T^{-1}z_{-1(4)}\right)'\left(\Delta z_{(4)} - \overline{c}T^{-1}z_{-1(4)}\right)$
$\Rightarrow 1 - \delta - \overline{c} + 2\overline{c}\delta - \overline{c}\delta^2 - \overline{c}^2\delta + \overline{c}^2\delta^2 + (\overline{c}^2/3)(1-\delta^3) \equiv d.$

We next consider the limit of each element of the vector $\Psi_T(\delta)$, denoted Ψ_i $(i = 1, 2, 3, 4)$. We have:

1. $\Psi_1 = \left(\Delta z_{(1)} - \bar{c}T^{-1}z_{-1(1)}\right)'\left(\Delta u - \bar{c}T^{-1}u_{-1}\right) \Rightarrow v_1$;
2. $\Psi_2 = T^{-1/2}\left(\Delta z_{(2)} - \bar{c}T^{-1}z_{-1(2)}\right)'\left(\Delta u - \bar{c}T^{-1}u_{-1}\right) = \lim_{T\to\infty} v_{[T\delta]+1} \equiv v^*$;
3. $\Psi_3 = T^{-1/2}\left(\Delta z_{(3)} - \bar{c}T^{-1}z_{-1(3)}\right)'\left(\Delta u - \bar{c}T^{-1}u_{-1}\right)$
   $\Rightarrow \sigma[W_c(1)(1-\bar{c}) + \bar{c}^2\int_0^1 rW_c(r)dr] \equiv \sigma b_1$;
4. $\Psi_4 = T^{-1/2}\left(\Delta z_{(4)} - \bar{c}T^{-1}z_{-1(4)}\right)'\left(\Delta u - \bar{c}T^{-1}u_{-1}\right)$
   $\Rightarrow \sigma[W_c(1)(1-\bar{c}+\delta\bar{c}) + \bar{c}^2\int_\delta^1 W_c(r)(r-\delta)dr - W_c(\delta)] \equiv \sigma b_2.$

Hence, using the symmetry of $\Gamma_T(\delta)$,

$$\Upsilon_T\left(\hat{\psi}(\delta) - \psi\right) \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & a & m \\ 0 & 0 & m & d \end{bmatrix}^{-1} \begin{bmatrix} v_1 \\ v^* \\ \sigma b_1 \\ \sigma b_2 \end{bmatrix}.$$

The proof of the Theorem follows upon solving for the inverse.

Proof of Theorem 3: The proof uses the results of Theorem 1. We first show the proof for Model 1 and for the $MZ_\alpha^{GLS}(\delta)$ test. Proofs for other tests are simple extensions.

1.1. Consider Model 1 and expression for $MZ_\alpha^{GLS}(\delta)$ given in (6). We first have

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

After some algebra, we obtain:

(a) $T^{-1}u_T^2 \Rightarrow \sigma^2 W_c(1)^2$;
(b) $2T^{-1}u_T(\hat{\mu}_1 - \mu_1) \Rightarrow 0$;
(c) $2T^{-1}u_T\left(\hat{\beta}_1 - \beta_1\right)T \Rightarrow 2\sigma^2 b_3 W_c(1)$;
(d) $2T^{-1}u_T\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 2\sigma^2 b_4 W_c(1)(1-\delta)$;
(e) $T^{-1}(\hat{\mu}_1 - \mu_1)^2 \Rightarrow 0$;
(f) $2T^{-1}(\hat{\mu}_1 - \mu_1)\left(\hat{\beta}_1 - \beta_1\right)T \Rightarrow 0$;
(g) $T^{-1}\left(\hat{\beta}_1 - \beta_1\right)^2 T^2 \Rightarrow \sigma^2 b_3^2$;
(h) $2T^{-1}(\hat{\mu}_1 - \mu_1)\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 0$;
(i) $2T^{-1}\left(\hat{\beta}_1 - \beta_1\right)T\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 2\sigma^2 b_3 b_4(1-\delta)$;

(j) $T^{-1}\left(\hat{\beta}_2-\beta_2\right)^2 1(t>T\delta)(T-T\delta)^2 \Rightarrow \sigma^2 b_4^2(1-\delta)^2$.

Using these results, we have:

$$T^{-1}\widetilde{y}_T^2 \Rightarrow \sigma^2\left\{V_{c\bar{c}}^{(1)}(1,\delta)^2-2V_{c\bar{c}}^{(2)}(1,\delta)\right\}, \tag{A.3}$$

where

$$V_{c\bar{c}}^{(1)}(1,\delta)=W_c(1)-b_3,$$

and

$$V_{c\bar{c}}^{(2)}(1,\delta)=b_4(1-\delta)[W_c(1)-b_3-(1/2)(1-\delta)b_4].$$

Consider now the term $2T^{-2}\sum_{t=1}^{T}\widetilde{y}_t^2$, defined by

$$\begin{aligned}2T^{-2}\sum_{t=1}^{T}\widetilde{y}_t^2 &= 2T^{-2}\sum_{t=1}^{T}\{y_t-[\hat{\mu}_1+\hat{\beta}_1 t+\hat{\beta}_2 1(t>T\delta)(t-T\delta)]\}^2\\ &= 2T^{-2}\sum_{t=1}^{T}\{u_t-[(\hat{\mu}_1-\mu_1)+(\hat{\beta}_1-\beta_1)t+(\hat{\beta}_2-\beta_2)1(t>T\delta)\\ &\quad (t-T\delta)]\}^2\end{aligned}$$

After some algebra, we obtain:

(a) $2T^{-2}\sum_{t=1}^{T}u_t^2 \Rightarrow 2\sigma^2\int_0^1 W_c(r)^2 dr$;
(b) $4T^{-2}\left(\hat{\mu}_1-\mu_1\right)\sum_{t=1}^{T}u_t \Rightarrow 0$;
(c) $4T^{-2}\left(\hat{\beta}_1-\beta_1\right)\sum_{t=1}^{T}tu_t \Rightarrow 4\sigma^2\int_0^1 rb_3W_c(r)dr$;
(d) $4T^{-2}\left(\hat{\beta}_2-\beta_2\right)\sum_{t=1}^{T}1(t>T\delta)(t-T\delta)u_t \Rightarrow 4\sigma^2\int_\delta^1 b_4W_c(r)(r-\delta)dr$;
(e) $2T^{-1}(\hat{\mu}_1-\mu_1^2) \Rightarrow 0$;
(f) $4T^{-2}\left(\hat{\mu}_1-\mu_1\right)\left(\hat{\beta}_1-\beta_1\right)\sum_{t=1}^{T}t \Rightarrow 0$;
(g) $2T^{-2}\left(\hat{\beta}_1-\beta_1\right)^2\sum_{t=1}^{T}t^2 \Rightarrow 2\sigma^2\int_0^1 b_3^2r^2dr$;
(h) $4T^{-2}\left(\hat{\mu}_1-\mu_1\right)\left(\hat{\beta}_2-\beta_2\right)\sum_{t=1}^{T}1(t>T\delta)(t-T\delta) \Rightarrow 0$;
(i) $4T^{-2}\left(\hat{\beta}_1-\beta_1\right)\left(\hat{\beta}_2-\beta_2\right)\sum_{t=1}^{T}t1(t>T\delta)(t-T\delta) \Rightarrow 4\sigma^2\int_\delta^1 b_3b_4r(r-\delta)dr$;
(j) $2T^{-2}\left(\hat{\beta}_2-\beta_2\right)^2\sum_{t=1}^{T}1(t>T\delta)(t-T\delta)^2 \Rightarrow 2\sigma^2\int_\delta^1 b_4^2(r-\delta)^2dr$.

Using these results we have:

$$2T^{-2}\sum_{t=1}^{T}\widetilde{y}_t^2 \Rightarrow 2\sigma^2\{\int_0^1 V_{c\bar{c}}^{(1)}(r,\delta)^2dr-2\int_\delta^1 V_{c\bar{c}}^{(2)}(r,\delta)dr\}. \tag{A.4}$$

Using (A.3), (A.4) and the fact that s^2 is a consistent estimate of σ^2, the proof is complete.

1.2. Consider now the MSB^{GLS} test defined in (7). Using results from (A.4) and the fact that s^2 is a consistent estimator of σ^2, the proof is complete.

1.3. Consider now the $MZ_t{}^{GLS}$ test. From Perron and Ng (1996), expression given in (8) is alternatively written by $MZ_t{}^{GLS} = MZ_t{}^{GLS} * MSB^{GLS}$. Hence, using results obtained for these two tests, the proof is complete.

2.1. Now, consider model 2 and $MZ_\alpha{}^{GLS}$ test. We first have

$$\begin{aligned} T^{-1}\widetilde{y}_T^2 &= T^{-1}\{y_T - (\hat{\mu}_1 + \widehat{\mu}_2)1(t > T\delta) + \hat{\beta}_1 T + \hat{\beta}_2 1(t > T\delta) \\ & (T - T\delta)\}^2 \\ &= T^{-1}\{u_T - [(\hat{\mu}_1 - \mu_1) + (\widehat{\mu}_2 - \mu_2)1(t > T\delta) + (\hat{\beta}_1 - \beta_1)T \\ +(\hat{\beta}_2 - \beta_2)1(t &> T\delta)(T - T\delta)\}^2. \end{aligned}$$

After some algebra, we obtain:

(a) $T^{-1}u_T^2 \Rightarrow \sigma^2 W_c(1)^2$;
(b) $2T^{-1}u_T\,(\hat{\mu}_1 - \mu_1) \Rightarrow 0$;
(c) $2T^{-1}u_T\,(\hat{\mu}_2 - \mu_2)\,1(t > T\delta) \Rightarrow 0$;
(d) $2T^{-1}u_T\left(\hat{\beta}_1 - \beta_1\right)T \Rightarrow 2\sigma^2 b_3 W_c(1)$;
(e) $2T^{-1}u_T\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 2\sigma^2 b_4 W_c(1)(1-\delta)$;
(f) $T^{-1}\left(\hat{\mu}_1 - \mu_1\right)^2 \Rightarrow 0$;
(g) $T^{-1}\left(\hat{\mu}_1 - \mu_1\right)^2\left(\hat{\mu}_2 - \mu_2\right)1(t > T\delta) \Rightarrow 0$;
(h) $T^{-1}\left(\hat{\mu}_2 - \mu_2\right)^2 1(t > T\delta) \Rightarrow 0$;
(i) $2T^{-1}\left(\hat{\mu}_1 - \mu_1\right)\left(\hat{\beta}_1 - \beta_1\right)T \Rightarrow 0$;
(j) $2T^{-1}\left(\hat{\mu}_2 - \mu_2\right)1(t > T\delta)(\hat{\beta}_1 - \beta_1)T \Rightarrow 0$;
(k) $2T^{-1}\left(\hat{\mu}_1 - \mu_1\right)\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 0$;
(l) $2T^{-1}\left(\hat{\mu}_2 - \mu_2\right)1(t > T\delta)\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 0$;
(m) $T^{-1}\left(\hat{\beta}_1 - \beta_1\right)^2 T^2 \Rightarrow \sigma^2 b_3^2$;
(n) $2T^{-1}\left(\hat{\beta}_1 - \beta_1\right)T\left(\hat{\beta}_2 - \beta_2\right)1(t > T\delta)(T - T\delta) \Rightarrow 2\sigma^2 b_3 b_4(1-\delta)$;
(o) $T^{-1}\left(\hat{\beta}_2 - \beta_2\right)^2 1(t > T\delta)(T - T\delta)^2 \Rightarrow \sigma^2 b_4^2(1-\delta)^2$.

Using these results, we have same result as in (A.3). Consider now the term $2T^{-2}\sum_{t=1}^{T}\widetilde{y}_t^2$, defined by

$$\begin{aligned} 2T^{-2}\sum_{t=1}^{T}\widetilde{y}_t^2 &= 2T^{-2}\sum_{t=1}^{T}\{y_t - [\hat{\mu}_1 + \widehat{\mu}_2 1(t > T\delta) + \hat{\beta}_1 t + \hat{\beta}_2)1(t > T\delta)(t - T\delta)]\}^2 \\ &= 2T^{-2}\sum_{t=1}^{T}\{u_t - [(\hat{\mu}_1 - \mu_1) + (\hat{\mu}_2 - \mu_2)1(t > T\delta) + (\hat{\beta}_1 - \beta_1)t + \\ (\hat{\beta}_2 - \beta_2)1(t &> T\delta)(t - T\delta)]\}^2 \end{aligned}$$

After some algebra, we obtain:

(a) $2T^{-2}\sum_{t=1}^{T} u_t^2 \Rightarrow 2\sigma^2 \int_0^1 W_c(r)^2 dr;$

(b) $4T^{-2}(\hat{\mu}_1 - \mu_1)\sum_{t=1}^{T} u_t \Rightarrow 0;$

(c) $4T^{-2}(\hat{\mu}_2 - \mu_2)1(t > T\delta)\sum_{t=1}^{T} u_t \Rightarrow 0;$

(d) $4T^{-2}(\hat{\beta}_1 - \beta_1)\sum_{t=1}^{T} tu_t \Rightarrow 4\sigma^2 \int_0^1 rb_3 W_c(r)dr;$

(e) $4T^{-2}(\hat{\beta}_2 - \beta_2)\sum_{t=1}^{T} 1(t > T\delta)(t - T\delta)u_t \Rightarrow 4\sigma^2 \int_\delta^1 b_4 W_c(r)(r-\delta)dr;$

(f) $2T^{-1}(\hat{\mu}_1 - \mu_1)^2 \Rightarrow 0;$

(g) $4T^{-1}(\hat{\mu}_1 - \mu_1)(\hat{\mu}_2 - \mu_2)1(t > T\delta) \Rightarrow 0;$

(h) $2T^{-1}(\hat{\mu}_2 - \mu_2)^2 1(t > T\delta) \Rightarrow 0;$

(i) $4T^{-2}(\hat{\mu}_1 - \mu_1)(\hat{\beta}_1 - \beta_1)\sum_{t=1}^{T} t \Rightarrow 0;$

(j) $4T^{-2}(\hat{\mu}_1 - \mu_1)(\hat{\beta}_2 - \beta_2)\sum_{t=1}^{T} 1(t > T\delta)(t - T\delta) \Rightarrow 0;$

(k) $4T^{-2}(\hat{\mu}_2 - \mu_2)1(t > T\delta)(\hat{\beta}_1 - \beta_1)\sum_{t=1}^{T} t \Rightarrow 0;$

(l) $4T^{-2}(\hat{\mu}_2 - \mu_2)1(t > T\delta)(\hat{\beta}_2 - \beta_2)1(t > T\delta)\sum_{t=1}^{T} t \Rightarrow 0;$

(m) $2T^{-2}(\hat{\beta}_1 - \beta_1)^2\sum_{t=1}^{T} t^2 \Rightarrow 2\sigma^2 \int_0^1 b_3^2 r^2 dr;$

(n) $4T^{-2}(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2)\sum_{t=1}^{T} t1(t > T\delta)(t - T\delta) \Rightarrow 4\sigma^2 \int_\delta^1 b_3 b_4 r(r-\delta)dr;$

(o) $2T^{-2}(\hat{\beta}_2 - \beta_2)^2\sum_{t=1}^{T} 1(t > T\delta)(t - T\delta)^2 \Rightarrow 2\sigma^2 \int_\delta^1 b_4^2 (r-\delta)^2 dr.$

Using these results we obtain sam results as in the expression (A.4). Furthermore, using the fact that s^2 is a consistent estimate of σ^2, the proof is complete.

2.2. Consider now the MSB^{GLS} test defined in (7). Using results from (A.4) and the fact that s^2 is a consistent estimator of σ^2, the proof is complete.

2.3. Consider now the $MZ_t{}^{GLS}$ test. From Perron and Ng (1996), expression given in (8) is alternatively written by $MZ_t{}^{GLS} = MZ_t{}^{GLS} * MSB^{GLS}$. Hence, using results obtained for these two tests, the proof is complete.

Proof of Theorem 4. We first give the proof for Model I. Defining

$$Q_T(\alpha) = (u^{\alpha\prime} z^{\alpha})(z^{\alpha\prime} z^{\alpha})^{-1}(z^{\alpha\prime} u^{\alpha}), \tag{A.5}$$

we have $S(\bar{\alpha}) = u^{\bar{\alpha}}\,{}' u^{\bar{\alpha}} - Q_T(\bar{\alpha})$ and $S(1) = u^{1\prime}u^1 - Q_T(1)$. Hence, $s^2 P_T$ is given by:

$$\begin{aligned} s^2 P_T \;=\; & \bar{c}^2 T^{-2} u_{-1}' u_{-1} - 2\bar{c}T^{-1}\Delta u' u_{-1} - \bar{c}T^{-1}\Delta u'\Delta u \\ & - Q_T(\bar{\alpha}) + Q_T(1) + \bar{c}T^{-1} Q_T(1). \end{aligned} \tag{A.6}$$

Using the fact that $2\bar{c}T^{-1}\Delta u' u_{-1} = \bar{c}T^{-1}u_T^2 - \bar{c}T^{-1}\Delta u'\Delta u$, we have:

$$s^2 P_T = \bar{c}^2 T^{-2} u_{-1}' u_{-1} - \bar{c}T^{-1} u_T^2 - Q_T(\bar{\alpha}) + Q_T(1). \tag{A.7}$$

Using the results in the proof of Theorem 2:

$$\begin{aligned}
\bar{c}^2 T^{-2} u_{-1}' u_{-1} &\Rightarrow \sigma^2 \bar{c}^2 \int_0^1 W_c(r)^2 dr \\
\bar{c} T^{-1} u_T^2 &\Rightarrow \sigma^2 \bar{c} W_c(1)^2
\end{aligned}$$

$$\begin{aligned}
Q_T(\bar{\alpha}) &= (u^{\bar{\alpha}\prime} z^{\bar{\alpha}})(z^{\bar{\alpha}\prime} z^{\bar{\alpha}})^{-1}(z^{\bar{\alpha}\prime} u^{\bar{\alpha}}) \\
&\Rightarrow \begin{bmatrix} v_1 \\ \sigma b_1 \\ \sigma b_2 \end{bmatrix}' \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda_1 & \lambda_2 \\ 0 & \lambda_2 & \lambda_3 \end{bmatrix} \begin{bmatrix} v_1 \\ \sigma b_1 \\ \sigma b_2 \end{bmatrix} \\
&= v_1^2 + \sigma^2[\lambda_1 b_1^2 + 2\lambda_2 b_1 b_2 + \lambda_3 b_2^2]
\end{aligned} \tag{A.8}$$

$$\begin{aligned}
Q_T(1) &= (u^{1\prime} z^1)(z^{1\prime} z^1)^{-1}(z^{1\prime} u^1) \\
&\Rightarrow \begin{bmatrix} v_1 \\ \sigma W_c(1) \\ \sigma(W_c(1) - W_c(\delta)) \end{bmatrix}' \begin{bmatrix} 1 & 0 & 0 \\ 0 & \delta^{-1} & -\delta^{-1} \\ 0 & -\delta^{-1} & \delta^{-1}(1-\delta)^{-1} \end{bmatrix} \\
&\quad \cdot \begin{bmatrix} v_1 \\ \sigma W_c(1) \\ \sigma(W_c(1) - W_c(\delta)) \end{bmatrix} \\
&= v_1^2 + \sigma^2\{(1-\delta)^{-1}\delta^{-1}[W_c(1) - W_c(\delta)]^2 \\
&\quad +2\delta^{-1} W_c(1) W_c(\delta) - \delta^{-1} W_c(1)^2\}
\end{aligned} \tag{A.9}$$

The proof follows directly using these limits. For Model II, we only need to change (A.8) as follows:

$$\begin{aligned}
Q_T(\bar{\alpha}) &\Rightarrow \begin{bmatrix} v_1 \\ v^* \\ \sigma b_1 \\ \sigma b_2 \end{bmatrix}' \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & \lambda_2 \\ 0 & 0 & \lambda_2 & \lambda_3 \end{bmatrix} \begin{bmatrix} v_1 \\ v^* \\ \sigma b_1 \\ \sigma b_2 \end{bmatrix} \\
&\Rightarrow v_1^2 + (v^*)^2 + \sigma^2[\lambda_1 b_1^2 + 2\lambda_2 b_1 b_2 + \lambda_3 b_2^2]
\end{aligned}$$

and (A.9) as follows:

$$Q_T(1) \Rightarrow \begin{bmatrix} v_1 \\ v^* \\ \sigma W_c(1) \\ \sigma(W_c(1) - W_c(\delta)) \end{bmatrix}' \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \delta^{-1} & -\delta^{-1} \\ 0 & 0 & -\delta^{-1} & \delta^{-1}(1-\delta)^{-1} \end{bmatrix}$$

$$
\cdot \begin{bmatrix} v_1 \\ v^* \\ \sigma W_c(1) \\ \sigma(W_c(1) - W_c(\delta)) \end{bmatrix}
$$

$$
\Rightarrow \quad v_1^2 + (v^*)^2 + \sigma^2\{(1-\delta)^{-1}\delta^{-1}[W_c(1) - W_c(\delta)]^2 + 2\delta^{-1}W_c(1)W_c(\delta) - \delta^{-1}W_c(1)^2\}
$$

The proof is then complete upon substitution.

References

[1] 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.

[2] Bhargava, A. (1986), "On the Theory of Testing for Unit Root in Observed Time Series," *Review of Economic Studies* **53**, 369-384.

[3] Campbell, J. Y. and N. G. Mankiw (1987a), "Permanent and Transitory Components in Macroeconomics Fluctuations," *American Economic Review: Papers and Proceedings* **77**, 111-117.

[4] Campbell, J.Y. and N.G. Mankiw (1987b), "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics* **102**, 857-880.

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

[6] Christiano, L. J. (1992), "Searching for a Break in GNP," *Journal of Business and Economics Statistics* **10**, 237-250.

[7] Christiano, L. J. and M. Eichenbaum (1989), "Unit Roots in Real GNP: Do We know and Do We Care?" *Carnegie-Rochester Conference Series on Public Policy* **32**, 7-62.

[8] Clark, P. K. (1987), "The Cyclical Component of U.S. Economic Activity", *Quarterly Journal of Economics* **102**, 798-814.

[9] Cochrane, J. H. (1988), "How Big is the Random Walk in GNP?" *Journal of Political Economy* **96**, 893-920.

[10] 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.

[11] 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.

[12] Elliott, G., T. Rothemberg and J. H. Stock (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica* **64**, 813-839.

[13] 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.

[14] Ng, S. and P. Perron (1999), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," manuscript, Boston College.

[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. (1991), "A Test for a Change in a Polynomial Trend Function for a Dynamic Time Series," manuscript, Princeton University.

[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] Phillips, P. C. B. (1987), "Time Series Regression with Unit Roots," *Econometrica* **55**, 277-302.

[22] Phillips, P. C. B. and P. Perron (1988), "Testing for Unit Roots in Time Series Regression," *Biometrika* **75**, 335-346.

[23] Said, S. E. and D. A. Dickey (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika* **71**, 599-608.

[24] Sargan, J. D. and A. Bhargava (1983), "Testing Residuals from Least Squares Regression for being generated by the Gaussian Random Walk," *Econometrica* **51**, 153-174.

[25] Shapiro, M. and M. Watson (1988), "Sources and Business Cycles Fluctuations," in S. Fischer (ed.), *NBER Macroeconomics Annual*, Vol. 3, 111-148.

[26] Stock, J. H. (1990), "A Class of Tests for Integration and Cointegration," Unpublished manuscript, Kennedy School of Government, Harvard University.

[27] Stock, J. H. (1994), "Unit Roots, Structural Breaks and Trends," in D.L. McFaden and R.F. Engle (eds.), *Handbook of Econometrics*, Vol. IV, 2740-2841.

[28] Vogelsang, T. J. (1997), "Wald-type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series," *Econometric Theory* **13**, 818-849.

[29] Vogelsang, T. J and P. Perron (1997), "Additional Tests for a Unit Root Allowing the Possibility of Breaks in the Trend Function," forthcoming in *International Economic Review.*

[30] 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.

Annex 1

Table 1.a. Percentage Points of the M^{GLS} and P_T^{GLS} Tests under the Null Hypothesis ($c = 0$) Choosing the Break Point Minimizing the Tests

	MZ_α^{GLS}			MSB^{GLS}		
	$T=\infty$	$T=100$		$T=\infty$	$T=100$	
		Model 1	Model 2		Model 1	Model 2
1.0%	-40.89	-41.44	-45.56	0.110	0.109	0.104
2.5%	-35.48	-36.61	-40.51	0.118	0.116	0.110
5.0%	-31.64	-32.73	-35.81	0.125	0.122	0.117
10.0%	-27.46	-28.38	-31.29	0.134	0.131	0.125
20.0%	-22.51	-23.82	-26.36	0.147	0.143	0.136
30.0%	-19.57	-20.79	-23.22	0.158	0.153	0.145
40.0%	-17.08	-18.52	-20.69	0.169	0.162	0.153
50.0%	-15.13	-16.38	-18.66	0.179	0.171	0.161
60.0%	-13.21	-14.64	-16.62	0.191	0.181	0.170
70.0%	-11.44	-12.93	-14.72	0.205	0.192	0.180
80.0%	-9.53	-11.02	-12.70	0.223	0.208	0.193
90.0%	-7.46	-8.81	-10.20	0.250	0.231	0.214
95.0%	-6.01	-7.12	-8.56	0.275	0.253	0.233
97.5%	-4.97	-6.06	-7.19	0.299	0.275	0.249
99.0%	-4.10	-4.85	-6.03	0.324	0.301	0.273

	MZ_t^{GLS}			P_T^{GLS}		
	$T=\infty$	$T=100$		$T=\infty$	$T=100$	
		Model 1	Model 2		Model 1	Model 2
1.0%	-4.49	-4.53	-4.75	6.59	6.64	6.24
2.5%	-4.18	-4.24	-4.44	7.70	7.65	7.07
5.0%	-3.96	-4.01	-4.20	8.53	8.50	7.92
10.0%	-3.68	-3.73	-3.92	9.83	9.76	9.05
20.0%	-3.33	-3.41	-3.59	11.96	11.74	10.79
30.0%	-3.09	-3.18	-3.37	13.80	13.39	12.29
40.0%	-2.89	-3.00	-3.18	15.72	15.07	13.86
50.0%	-2.71	-2.82	-3.01	17.74	16.87	15.47
60.0%	-2.53	-2.66	-2.84	20.19	19.02	17.33
70.0%	-2.35	-2.50	-2.66	23.20	21.72	19.59
80.0%	-2.13	-2.30	-2.47	27.60	25.15	22.79
90.0%	-1.88	-2.04	-2.20	34.66	31.73	27.83
95.0%	-1.67	-1.83	-2.00	42.57	38.34	33.46
97.5%	-1.52	-1.69	-1.82	49.76	44.43	39.59
99.0%	-1.35	-1.49	-1.63	58.76	53.46	46.92

Table 1.b. Percentage Points of the M^{GLS} and P_T^{GLS} Tests under the Null Hypothesis ($c=0$)
Choosing the Break Point maximizing $|t_{\widehat{\beta_2}}|$

	MZ_α^{GLS}			MSB^{GLS}		
	$T=\infty$	$T=100$		$T=\infty$	$T=100$	
		Model 1	Model 2		Model 1	Model 2
1.0%	-41.01	-41.30	-42.33	0.110	0.110	0.109
2.5%	-34.96	-36.08	-36.56	0.119	0.117	0.116
5.0%	-30.75	-32.20	-32.65	0.127	0.124	0.123
10.0%	-26.41	-27.82	-28.23	0.137	0.133	0.132
20.0%	-21.76	-23.20	-23.75	0.150	0.145	0.144
30.0%	-18.85	-20.37	-20.74	0.161	0.155	0.154
40.0%	-16.13	-18.07	-18.43	0.171	0.164	0.163
50.0%	-14.66	-16.13	-16.40	0.182	0.173	0.172
60.0%	-12.92	-14.33	-14.64	0.194	0.184	0.182
70.0%	-11.28	-12.44	-12.90	0.207	0.196	0.193
80.0%	-9.46	-10.60	-11.04	0.224	0.212	0.208
90.0%	-7.46	-8.45	-8.84	0.250	0.236	0.231
95.0%	-5.96	-6.96	-7.30	0.275	0.258	0.252
97.5%	-4.89	-5.83	-6.14	0.299	0.279	0.273
99.0%	-3.82	-4.76	-4.89	0.334	0.301	0.297

	MZ_t^{GLS}			P_T^{GLS}		
	$T=\infty$	$T=100$		$T=\infty$	$T=100$	
		Model 1	Model 2		Model 1	Model 2
1.0%	-4.50	-4.53	-4.59	6.80	6.62	6.47
2.5%	-4.17	-4.22	-4.25	7.86	7.59	7.44
5.0%	-3.89	-3.99	-4.02	8.93	8.50	8.44
10.0%	-3.61	-3.71	-3.73	10.34	9.80	9.75
20.0%	-3.27	-3.38	-3.42	12.56	11.80	11.61
30.0%	-3.04	-3.16	-3.19	14.44	13.52	13.32
40.0%	-2.85	-2.97	-3.00	16.37	15.22	14.98
50.0%	-2.67	-2.81	-2.82	18.47	16.98	16.86
60.0%	-2.50	-2.64	-2.67	20.93	19.08	18.82
70.0%	-2.33	-2.45	-2.50	23.80	21.92	21.27
80.0%	-2.13	-2.26	-2.30	28.10	25.61	24.79
90.0%	-1.87	-1.99	-2.05	34.97	31.97	30.72
95.0%	-1.64	-1.79	-1.84	42.67	37.93	36.88
97.5%	-1.44	-1.62	-1.67	50.47	45.36	43.76
99.0%	-1.24	-1.40	-1.43	62.11	53.33	51.77

Table 2: Exact Size of the tests at selected values of k in s^2
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS} ; MA case

T	θ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	1.000	0.798	0.372	0.249	0.199	0.191	0.195	0.251	0.250	0.291	0.336
	-0.40	0.588	0.081	0.037	0.075	0.119	0.189	0.247	0.324	0.397	0.459	0.529
	0.00	0.021	0.030	0.050	0.126	0.172	0.254	0.326	0.403	0.445	0.538	0.588
	0.40	0.000	0.112	0.034	0.153	0.174	0.277	0.333	0.424	0.471	0.546	0.587
	0.80	0.000	0.269	0.007	0.274	0.099	0.372	0.245	0.471	0.412	0.599	0.546
200	-0.80	1.000	0.993	0.748	0.376	0.189	0.125	0.093	0.084	0.092	0.094	0.107
	-0.40	0.725	0.143	0.043	0.044	0.054	0.068	0.104	0.137	0.175	0.210	0.245
	0.00	0.029	0.029	0.043	0.073	0.093	0.120	0.156	0.194	0.215	0.254	0.299
	0.40	0.001	0.117	0.030	0.080	0.092	0.140	0.158	0.196	0.240	0.273	0.311
	0.80	0.000	0.238	0.008	0.178	0.038	0.213	0.099	0.252	0.194	0.316	0.284

b) P_T^{GLS} ; MA case

T	θ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	1.000	0.793	0.396	0.258	0.202	0.204	0.204	0.255	0.255	0.296	0.341
	-0.40	0.538	0.083	0.042	0.080	0.124	0.192	0.249	0.329	0.400	0.460	0.523
	0.00	0.018	0.033	0.053	0.133	0.179	0.254	0.325	0.399	0.446	0.534	0.590
	0.40	0.000	0.112	0.036	0.153	0.178	0.278	0.328	0.417	0.465	0.545	0.581
	0.80	0.000	0.241	0.009	0.270	0.103	0.371	0.250	0.469	0.407	0.597	0.541
200	-0.80	1.000	0.991	0.743	0.385	0.195	0.132	0.098	0.088	0.093	0.101	0.112
	-0.40	0.691	0.134	0.044	0.048	0.053	0.073	0.106	0.137	0.175	0.208	0.245
	0.00	0.024	0.028	0.044	0.071	0.090	0.116	0.155	0.198	0.218	0.259	0.296
	0.40	0.001	0.105	0.029	0.082	0.092	0.140	0.159	0.203	0.238	0.278	0.312
	0.80	0.000	0.224	0.009	0.171	0.041	0.210	0.101	0.252	0.193	0.311	0.279

c) MZ_{α}^{GLS} ; AR case

T	ρ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	0.979	0.000	0.003	0.015	0.047	0.101	0.192	0.270	0.343	0.410	0.472
	-0.40	0.415	0.008	0.020	0.085	0.141	0.223	0.275	0.371	0.438	0.494	0.554
	0.00	0.021	0.030	0.050	0.126	0.172	0.254	0.326	0.403	0.445	0.538	0.588
	0.40	0.000	0.046	0.090	0.146	0.205	0.288	0.339	0.420	0.488	0.565	0.597
	0.80	0.000	0.105	0.151	0.219	0.271	0.354	0.413	0.481	0.545	0.620	0.638
200	-0.80	0.992	0.000	0.000	0.004	0.016	0.035	0.068	0.087	0.117	0.145	0.183
	-0.40	0.506	0.019	0.025	0.043	0.065	0.093	0.125	0.163	0.198	0.242	0.276
	0.00	0.029	0.029	0.043	0.073	0.093	0.120	0.156	0.194	0.215	0.254	0.299
	0.40	0.000	0.038	0.055	0.080	0.111	0.151	0.178	0.207	0.240	0.281	0.333
	0.80	0.000	0.066	0.093	0.105	0.139	0.170	0.211	0.242	0.268	0.334	0.351

d) P_T^{GLS} ; AR case

T	ρ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	0.971	0.000	0.003	0.017	0.056	0.104	0.194	0.273	0.347	0.413	0.474
	-0.40	0.383	0.008	0.026	0.094	0.149	0.226	0.282	0.366	0.439	0.493	0.557
	0.00	0.018	0.033	0.053	0.133	0.179	0.254	0.325	0.399	0.446	0.534	0.590
	0.40	0.000	0.043	0.088	0.151	0.200	0.290	0.336	0.423	0.479	0.561	0.587
	0.80	0.000	0.096	0.130	0.205	0.257	0.348	0.402	0.475	0.532	0.605	0.628
200	-0.80	0.987	0.000	0.001	0.004	0.019	0.036	0.069	0.091	0.124	0.148	0.187
	-0.40	0.470	0.020	0.026	0.046	0.065	0.089	0.127	0.163	0.198	0.237	0.274
	0.00	0.024	0.028	0.044	0.071	0.090	0.116	0.155	0.198	0.218	0.259	0.296
	0.40	0.000	0.035	0.054	0.080	0.104	0.144	0.174	0.208	0.242	0.278	0.329
	0.80	0.000	0.058	0.082	0.096	0.137	0.0170	0.212	0.239	0.261	0.323	0.340

Table 3. Selected values of k using IC and MIC (MZ_α^{GLS})
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

	MA Case						AR Case					
T	θ	k^*	AIC	BIC	MAIC	MBIC	ρ	k^*	AIC	BIC	MAIC	MBIC
100	-0.8	5	2	0	4	3	-0.8	4	1	1	1	1
	-0.4	2	1	0	2	1	-0.4	2	1	1	1	1
	0.0	1	0	0	0	0	0.0	1	0	0	0	0
	0.4	2	2	1	2	1	0.4	1	1	1	1	1
	0.8	4	5	3	5	3	0.8	0	2	1	1	1
200	-0.8	7	4	2	6	4	-0.8	5	1	1	1	1
	-0.4	2	2	1	2	1	-0.4	3	1	1	1	1
	0.0	0	0	0	0	0	0.0	0	0	0	0	0
	0.4	2	2	1	2	1	0.4	1	1	1	1	1
	0.8	4	7	4	6	4	0.8	1	1	1	1	1

Table 4: Finite sample critical values; Model I, choosing T_B minimizing the test statistic ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

a) MZ_α^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-7608.23	-515.94	-113.94	-38.39	-9.21	-7.59	-6.72	-6.19
	BIC	-37.74	-33.49	-28.88	-24.45	-8.37	-7.07	-6.25	-5.02
	MAIC	-31.06	-27.87	-25.17	-22.79	-8.45	-7.20	-6.38	-5.70
	MBIC	-31.54	-28.22	-24.99	-22.86	-8.02	-6.84	-6.13	-4.91
	t-sig	-84.47	-58.31	-40.30	-31.49	-9.21	-7.64	-6.76	-5.73
T=150	AIC	-176.15	-73.06	-44.79	-33.36	-8.75	-7.60	-6.78	-5.96
	BIC	-37.89	-33.36	-28.99	-25.24	-8.43	-7.15	-6.41	-5.18
	MAIC	-34.38	-30.01	-27.72	-24.30	-8.38	-7.26	-6.68	-5.69
	MBIC	-34.30	-30.38	-27.25	-24.26	-8.27	-7.02	-6.01	-4.82
	t-sig	-53.46	-39.59	-35.28	-29.70	-8.94	-8.05	-6.85	-5.87
T=200	AIC	-133.76	-67.31	-42.94	-30.82	-8.47	-7.14	-6.22	-5.01
	BIC	-36.94	-31.99	-29.22	-25.63	-8.18	-6.69	-5.18	-4.50
	MAIC	-36.63	-31.33	-27.35	-24.28	-8.15	-6.88	-5.83	-4.75
	MBIC	-35.13	-31.39	-28.08	-24.23	-7.99	-6.59	-5.05	-4.43
	t-sig	-60.01	-43.31	-35.37	-30.49	-8.52	-7.16	-5.82	-4.94
T=∞		-40.88	-35.48	-31.63	-27.46	-7.46	-6.00	-4.97	-4.09

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0081	0.0311	0.0662	0.1140	0.2269	0.2455	0.2608	0.2768
	BIC	0.1145	0.1214	0.1306	0.1412	0.2400	0.2583	0.2713	0.2954
	MAIC	0.1263	0.1324	0.1407	0.1468	0.2380	0.2549	0.2675	0.2819
	MBIC	0.1258	0.1321	0.1405	0.1466	0.2424	0.2616	0.2761	0.3041
	t-sig	0.0767	0.0926	0.1110	0.1252	0.2264	0.2439	0.2630	0.2822
T=150	AIC	0.0533	0.0826	0.1057	0.1220	0.2312	0.2453	0.2585	0.2713
	BIC	0.1148	0.1206	0.1294	0.1393	0.2365	0.2529	0.2686	0.2914
	MAIC	0.1197	0.1285	0.1339	0.1423	0.2387	0.2533	0.2665	0.2847
	MBIC	0.1196	0.1283	0.1349	0.1428	0.2385	0.2554	0.2742	0.2931
	t-sig	0.0966	0.1124	0.1186	0.1284	0.2280	0.2413	0.2570	0.2815
T=200	AIC	0.0611	0.0860	0.1077	0.1272	0.2357	0.2588	0.2708	0.2891
	BIC	0.1163	0.1238	0.1296	0.1391	0.2407	0.2606	0.2934	0.3234
	MAIC	0.1168	0.1259	0.1342	0.1422	0.2395	0.2603	0.2783	0.2973
	MBIC	0.1188	0.1255	0.1323	0.1420	0.2434	0.2621	0.2973	0.3240
	t-sig	0.0913	0.1074	0.1184	0.1275	0.2354	0.2564	0.2682	0.3048
T=∞		0.1096	0.1179	0.1250	0.1338	0.2495	0.2754	0.2985	0.3239

Table 4 (cont'd): Finite sample critical values; Model I, choosing T_B minimizing the test statistic ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-61.67	-16.05	-7.54	-4.37	-2.09	-1.91	-1.77	-1.64
	BIC	-4.33	-4.07	-3.79	-3.48	-1.99	-1.81	-1.70	-1.52
	MAIC	-3.92	-3.70	-3.50	-3.34	-2.01	-1.84	-1.73	-1.62
	MBIC	-3.94	-3.75	-3.51	-3.34	-1.92	-1.80	-1.66	-1.47
	t-sig	-6.48	-5.39	-4.47	-3.95	-2.11	-1.89	-1.79	-1.62
T=150	AIC	-9.38	-6.03	-4.73	-4.05	-2.05	-1.87	-1.79	-1.66
	BIC	-4.33	-4.00	-3.79	-3.53	-1.99	-1.82	-1.72	-1.54
	MAIC	-4.11	-3.87	-3.70	-3.47	-1.99	-1.84	-1.76	-1.59
	MBIC	-4.10	-3.87	-3.65	-3.47	-1.96	-1.79	-1.70	-1.50
	t-sig	-5.16	-4.44	-4.19	-3.84	-2.07	-1.91	-1.80	-1.66
T=200	AIC	-8.17	-5.79	-4.62	-3.89	-1.97	-1.84	-1.73	-1.48
	BIC	-4.29	-3.98	-3.78	-3.55	-1.94	-1.76	-1.57	-1.44
	MAIC	-4.25	-3.94	-3.68	-3.45	-1.94	-1.80	-1.66	-1.47
	MBIC	-4.18	-3.93	-3.71	-3.45	-1.93	-1.74	-1.56	-1.42
	t-sig	-5.47	-4.65	-4.19	-3.87	-2.00	-1.84	-1.66	-1.53
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.03	0.53	2.36	6.92	29.44	35.17	39.32	44.20
	BIC	7.20	8.19	9.54	11.22	33.65	38.41	41.19	50.77
	MAIC	8.77	9.89	11.04	11.96	32.73	37.35	40.59	45.68
	MBIC	8.71	9.73	10.95	12.04	34.73	39.32	44.20	54.49
	t-sig	3.19	4.61	6.72	8.51	29.44	35.27	40.21	45.47
T=150	AIC	1.51	3.61	5.91	8.10	30.56	35.74	38.98	43.96
	BIC	7.21	8.50	9.45	10.82	31.69	37.99	41.33	51.18
	MAIC	7.93	8.91	9.92	11.02	32.14	37.95	40.40	46.65
	MBIC	8.26	8.99	10.12	11.18	32.35	38.70	42.83	53.02
	t-sig	4.96	6.72	7.57	9.15	30.41	34.52	38.90	46.47
T=200	AIC	2.06	3.94	6.28	8.92	32.08	38.05	43.74	53.01
	BIC	7.19	8.58	9.41	10.73	33.52	40.38	51.91	61.75
	MAIC	7.55	8.78	9.93	11.34	33.49	38.75	46.11	54.40
	MBIC	7.78	8.80	9.90	11.44	34.23	40.49	53.01	61.75
	t-sig	4.37	6.25	7.58	9.02	31.33	38.16	44.11	54.71
T=∞		6.59	7.69	8.52	9.82	34.65	42.57	49.76	58.73

Table 4 (cont'd): Finite sample critical values; Model I, choosing T_B minimizing the test statistic ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-5.04	-4.79	-4.57	-4.26	-2.31	-2.11	-1.99	-1.87
	BIC	-5.00	-4.67	-4.34	-3.94	-2.16	-1.98	-1.87	-1.69
	MAIC	-4.62	-4.26	-3.91	-3.67	-2.06	-1.92	-1.82	-1.67
	MBIC	-4.62	-4.25	-3.92	-3.68	-2.04	-1.89	-1.73	-1.57
	t-sig	-5.00	-4.69	-4.43	-4.12	-2.27	-2.05	-1.93	-1.85
T=150	AIC	-4.95	-4.64	-4.37	-4.04	-2.22	-2.06	-1.90	-1.80
	BIC	-4.72	-4.38	-4.18	-3.83	-2.09	-1.93	-1.83	-1.70
	MAIC	-4.52	-4.21	-3.94	-3.64	-2.05	-1.89	-1.80	-1.65
	MBIC	-4.52	-4.21	-3.94	-3.66	-2.02	-1.85	-1.76	-1.53
	t-sig	-4.80	-4.61	-4.28	-3.95	-2.16	-2.05	-1.89	-1.80
T=200	AIC	-4.93	-4.46	-4.21	-3.92	-2.08	-1.90	-1.81	-1.67
	BIC	-4.72	-4.26	-4.06	-3.74	-2.00	-1.83	-1.67	-1.50
	MAIC	-4.47	-4.06	-3.77	-3.50	-1.99	-1.83	-1.73	-1.52
	MBIC	-4.47	-4.17	-3.81	-3.52	-1.97	-1.79	-1.63	-1.48
	t-sig	-4.90	-4.46	-4.19	-3.91	-2.08	-1.89	-1.76	-1.62
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

Table 5. Finite sample critical values; Model I, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-83.53	-38.75	-29.23	-23.88	-7.29	-6.13	-4.93	-3.95
	BIC	-34.40	-29.93	-26.96	-23.31	-7.74	-6.56	-5.70	-4.20
	MAIC	-27.16	-24.14	-22.79	-20.21	-6.69	-5.52	-4.42	-3.47
	MBIC	-29.58	-26.96	-23.82	-21.58	-7.31	-6.34	-5.41	-4.12
	t-sig	-40.72	-34.53	-29.42	-24.43	-7.49	-6.44	-5.52	-4.27
T=150	AIC	-65.38	-36.68	-30.42	-25.65	-7.76	-6.69	-5.86	-4.82
	BIC	-35.17	-31.12	-27.44	-24.24	-8.09	-6.81	-5.86	-4.82
	MAIC	-29.23	-26.63	-24.47	-21.73	-7.36	-6.37	-5.32	-4.56
	MBIC	-32.27	-27.55	-25.71	-23.06	-7.91	-6.65	-5.62	-4.62
	t-sig	-43.21	-33.98	-29.08	-24.59	-8.24	-6.86	-5.82	-4.79
T=200	AIC	-55.64	-35.10	-29.97	-25.11	-7.51	-6.24	-5.21	-4.28
	BIC	-35.10	-31.11	-28.28	-23.97	-7.84	-6.27	-4.92	-4.05
	MAIC	-31.11	-28.26	-25.11	-21.88	-7.18	-5.99	-5.12	-4.05
	MBIC	-32.97	-30.64	-26.70	-22.88	-7.64	-6.27	-4.92	-4.05
	t-sig	-39.66	-33.50	-29.00	-25.17	-7.64	-6.27	-5.18	-4.17
T=∞		-41.00	-34.96	-30.74	-26.40	-7.45	-5.95	-4.88	-3.81

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0774	0.1133	0.1300	0.1434	0.2527	0.2693	0.2957	0.3323
	BIC	0.1197	0.1285	0.1345	0.1455	0.2463	0.2622	0.2771	0.3063
	MAIC	0.1344	0.1424	0.1474	0.1555	0.2639	0.2848	0.3098	0.3514
	MBIC	0.1289	0.1349	0.1437	0.1502	0.2521	0.2683	0.2848	0.3098
	t-sig	0.1106	0.1197	0.1303	0.1419	0.2492	0.2644	0.2807	0.3081
T=150	AIC	0.0874	0.1165	0.1277	0.1382	0.2459	0.2614	0.2730	0.2919
	BIC	0.1188	0.1249	0.1342	0.1427	0.2412	0.2570	0.2754	0.2956
	MAIC	0.1307	0.1367	0.1420	0.1496	0.2521	0.2687	0.2818	0.3075
	MBIC	0.1233	0.1326	0.1387	0.1450	0.2439	0.2617	0.2808	0.3012
	t-sig	0.1071	0.1204	0.1308	0.1416	0.2400	0.2560	0.2754	0.2938
T=200	AIC	0.0948	0.1192	0.1283	0.1402	0.2500	0.2706	0.2915	0.3060
	BIC	0.1192	0.1259	0.1323	0.1429	0.2458	0.2675	0.2997	0.3246
	MAIC	0.1262	0.1326	0.1401	0.1502	0.2566	0.2757	0.2974	0.3112
	MBIC	0.1225	0.1276	0.1359	0.1465	0.2471	0.2683	0.2997	0.3246
	t-sig	0.1122	0.1220	0.1306	0.1399	0.2464	0.2676	0.2899	0.3171
T=∞		0.1098	0.1187	0.1266	0.1365	0.2498	0.2751	0.2991	0.3339

Table 5 (cont'd). Finite sample critical values; Model I, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-6.46	-4.38	-3.82	-3.43	-1.85	-1.66	-1.45	-1.21
	BIC	-4.11	-3.84	-3.65	-3.38	-1.90	-1.74	-1.53	-1.32
	MAIC	-3.68	-3.45	-3.34	-3.14	-1.75	-1.55	-1.37	-1.17
	MBIC	-3.83	-3.65	-3.41	-3.26	-1.87	-1.70	-1.51	-1.29
	t-sig	-4.49	-4.13	-3.82	-3.47	-1.88	-1.72	-1.53	-1.32
T=150	AIC	-5.71	-4.27	-3.89	-3.56	-1.91	-1.76	-1.56	-1.36
	BIC	-4.17	-3.95	-3.68	-3.47	-1.95	-1.76	-1.56	-1.34
	MAIC	-3.82	-3.62	-3.49	-3.26	-1.85	-1.70	-1.51	-1.34
	MBIC	-3.98	-3.69	-3.58	-3.37	-1.92	-1.73	-1.53	-1.34
	t-sig	-4.62	-4.09	-3.76	-3.49	-1.95	-1.77	-1.57	-1.36
T=200	AIC	-5.27	-4.15	-3.85	-3.52	-1.85	-1.66	-1.49	-1.28
	BIC	-4.18	-3.92	-3.73	-3.43	-1.89	-1.67	-1.45	-1.18
	MAIC	-3.92	-3.73	-3.50	-3.29	-1.82	-1.65	-1.47	-1.24
	MBIC	-4.05	-3.89	-3.62	-3.36	-1.87	-1.67	-1.45	-1.18
	t-sig	-4.45	-4.08	-3.78	-3.52	-1.86	-1.67	-1.52	-1.31
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	3.10	6.85	9.13	11.37	36.35	42.43	50.46	68.14
	BIC	8.02	9.03	10.18	11.82	34.78	39.98	44.34	56.24
	MAIC	10.12	11.25	12.15	13.60	39.61	47.67	58.28	73.74
	MBIC	9.13	10.71	11.43	12.64	36.22	40.59	46.30	57.67
	t-sig	6.50	7.69	9.22	11.14	35.28	40.37	47.54	57.98
T=150	AIC	4.02	7.25	8.82	10.73	35.03	40.24	44.36	53.21
	BIC	7.78	8.76	9.85	11.24	32.84	39.13	43.20	53.21
	MAIC	9.32	10.25	11.10	12.77	36.65	41.15	47.61	54.93
	MBIC	8.64	9.75	10.63	11.81	34.31	39.48	45.26	53.25
	t-sig	6.19	8.10	9.52	10.92	32.84	39.03	43.75	52.96
T=200	AIC	4.90	7.78	9.27	10.88	36.14	42.20	50.31	56.81
	BIC	7.64	8.80	9.74	11.42	34.90	41.11	53.40	62.41
	MAIC	8.72	9.65	10.98	12.52	37.53	43.70	51.36	58.84
	MBIC	8.19	9.11	10.32	11.97	35.26	41.31	53.40	62.41
	t-sig	6.73	8.07	9.36	10.98	35.01	42.20	48.33	58.87
T=∞		6.79	7.85	8.92	10.34	34.97	42.67	50.46	62.11

Table 5 (cont'd). Finite sample critical values; Model I, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-4.97	-4.64	-4.37	-4.04	-2.08	-1.86	-1.73	-1.49
	BIC	-4.95	-4.57	-4.19	-3.85	-2.04	-1.82	-1.60	-1.39
	MAIC	-4.54	-4.19	-3.82	-3.56	-1.88	-1.71	-1.49	-1.28
	MBIC	-4.54	-4.19	-3.84	-3.60	-1.92	-1.73	-1.54	-1.29
	t-sig	-4.91	-4.60	-4.33	-3.99	-2.10	-1.87	-1.72	-1.46
T=150	AIC	-4.78	-4.53	-4.22	-3.89	-2.06	-1.83	-1.64	-1.41
	BIC	-4.67	-4.34	-4.05	-3.74	-2.02	-1.82	-1.60	-1.37
	MAIC	-4.43	-4.07	-3.85	-3.55	-1.88	-1.72	-1.54	-1.39
	MBIC	-4.42	-4.08	-3.87	-3.60	-1.94	-1.75	-1.54	-1.36
	t-sig	-4.70	-4.48	-4.22	-3.87	-2.06	-1.87	-1.70	-1.48
T=200	AIC	-4.82	-4.38	-4.12	-3.80	-1.95	-1.76	-1.58	-1.35
	BIC	-4.67	-4.22	-4.00	-3.65	-1.95	-1.73	-1.50	-1.25
	MAIC	-4.39	-4.00	-3.65	-3.44	-1.86	-1.67	-1.49	-1.25
	MBIC	-4.41	-4.05	-3.68	-3.47	-1.90	-1.69	-1.46	-1.25
	t-sig	-4.79	-4.36	-4.06	-3.73	-1.97	-1.77	-1.55	-1.35
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

Table 6. Finite sample critical values; Model II, choosing T_B minimizing the test statistic ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

a) MZ_α^{GLS}

T	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
100	AIC	-59308.98	-1952.48	-217.19	-55.98	-11.38	-10.04	-9.06	-7.91
	BIC	-39.68	-34.48	-29.93	-26.72	-9.77	-8.40	-7.26	-6.43
	MAIC	-34.04	-30.66	-28.24	-24.48	-10.11	-8.96	-8.28	-6.58
	MBIC	-31.24	-28.80	-27.04	-23.90	-9.53	-8.15	-6.92	-6.04
	t-sig	-104.55	-66.24	-46.00	-36.03	-11.10	-9.70	-8.63	-6.93
150	AIC	-656.63	-107.04	-58.51	-37.62	-10.25	-8.74	-7.82	-6.60
	BIC	-39.80	-34.12	-30.37	-26.73	-9.11	-7.98	-6.74	-5.21
	MAIC	-35.36	-32.25	-28.89	-25.82	-9.34	-8.08	-7.44	-6.21
	MBIC	-34.28	-30.38	-27.99	-25.39	-8.90	-7.88	-6.48	-5.19
	t-sig	-59.30	-45.13	-37.90	-32.02	-10.27	-8.81	-7.83	-7.00
200	AIC	-235.34	-77.82	-51.48	-33.92	-9.52	-7.75	-6.84	-5.54
	BIC	-39.99	-33.51	-30.53	-26.60	-8.49	-7.14	-5.65	-4.73
	MAIC	-37.61	-32.42	-29.01	-25.79	-8.91	-7.62	-6.68	-5.54
	MBIC	-36.71	-32.52	-29.20	-25.76	-8.48	-7.06	-5.54	-4.62
	t-sig	-61.78	-48.12	-40.62	-33.09	-9.71	-8.11	-6.75	-5.13
∞		-40.88	-35.48	-31.63	-27.46	-7.46	-6.00	-4.97	-4.09

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
100	AIC	0.0029	0.0160	0.0480	0.0945	0.2050	0.2168	0.2270	0.2423
	BIC	0.1122	0.1193	0.1287	0.1363	0.2198	0.2380	0.2512	0.2680
	MAIC	0.1196	0.1265	0.1322	0.1419	0.2161	0.2309	0.2417	0.2563
	MBIC	0.1244	0.1301	0.1353	0.1434	0.2238	0.2416	0.2456	0.2752
	t-sig	0.0692	0.0869	0.1034	0.1172	0.2065	0.2240	0.2346	0.2518
150	AIC	0.0276	0.0681	0.0924	0.1147	0.2142	0.2312	0.2426	0.2622
	BIC	0.1120	0.1199	0.1277	0.1362	0.2259	0.2402	0.2556	0.2847
	MAIC	0.1173	0.1231	0.1304	0.1381	0.2250	0.2392	0.2503	0.2689
	MBIC	0.1187	0.1275	0.1328	0.1393	0.2272	0.2425	0.2666	0.2915
	t-sig	0.0918	0.1052	0.1148	0.1244	0.2150	0.2294	0.2417	0.2571
200	AIC	0.0461	0.0801	0.0985	0.1209	0.2221	0.2467	0.2591	0.2789
	BIC	0.1116	0.1213	0.1276	0.1363	0.2348	0.2538	0.2814	0.3081
	MAIC	0.1153	0.1230	0.1299	0.1385	0.2294	0.2495	0.2628	0.2789
	MBIC	0.1164	0.1230	0.1287	0.1385	0.2349	0.2552	0.2871	0.3117
	t-sig	0.0900	0.1017	0.1107	0.1223	0.2209	0.2400	0.2570	0.2891
∞		0.1096	0.1179	0.1250	0.1338	0.2495	0.2754	0.2985	0.3239

Table 6 (cont'd). Finite sample critical values; Model II, choosing T_B minimizing the test statistic ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-172.20	-31.24	-10.42	-5.29	-2.33	-2.19	-2.06	-1.94
	BIC	-4.45	-4.11	-3.85	-3.61	-2.15	-2.00	-1.81	-1.69
	MAIC	-4.12	-3.90	-3.75	-3.46	-2.19	-2.06	-1.95	-1.74
	MBIC	-3.93	-3.78	-3.60	-3.42	-2.12	-1.96	-1.79	-1.68
	t-sig	-7.23	-5.75	-4.79	-4.22	-2.28	-2.12	-2.01	-1.80
T=150	AIC	-18.11	-7.29	-5.40	-4.31	-2.19	-2.03	-1.90	-1.78
	BIC	-4.46	-4.13	-3.88	-3.63	-2.07	-1.91	-1.76	-1.54
	MAIC	-4.19	-3.98	-3.78	-3.58	-2.10	-1.95	-1.85	-1.74
	MBIC	-4.13	-3.88	-3.72	-3.53	-2.05	-1.90	-1.75	-1.53
	t-sig	-5.44	-4.74	-4.34	-3.96	-2.20	-2.05	-1.88	-1.79
T=200	AIC	-10.84	-6.23	-5.06	-4.10	-2.12	-1.92	-1.79	-1.60
	BIC	-4.46	-4.04	-3.90	-3.62	-1.99	-1.83	-1.62	-1.52
	MAIC	-4.33	-3.98	-3.76	-3.57	-2.06	-1.89	-1.76	-1.58
	MBIC	-4.27	-4.02	-3.79	-3.56	-1.98	-1.79	-1.61	-1.48
	t-sig	-5.55	-4.89	-4.50	-4.05	-2.14	-1.93	-1.73	-1.57
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.00	0.14	1.23	4.90	25.13	28.49	33.01	37.90
	BIC	6.73	7.93	9.13	10.40	30.03	34.93	39.33	44.15
	MAIC	8.20	9.14	9.90	11.25	28.51	32.76	37.60	41.88
	MBIC	8.86	9.47	10.55	11.62	30.73	36.41	39.60	47.77
	t-sig	2.56	3.92	5.79	7.63	25.79	30.51	35.43	40.08
T=150	AIC	0.41	2.70	4.62	7.31	27.15	32.15	36.19	40.28
	BIC	6.87	8.00	9.13	10.41	30.19	35.15	40.28	49.64
	MAIC	7.56	8.56	9.39	10.65	29.86	34.68	37.81	43.94
	MBIC	8.01	8.99	9.88	10.84	31.01	35.99	41.35	51.22
	t-sig	4.44	5.99	7.23	8.53	27.25	32.07	34.92	40.61
T=200	AIC	1.12	3.41	5.20	7.99	29.38	36.10	40.64	51.32
	BIC	6.88	8.11	8.95	10.29	31.78	38.41	48.44	58.57
	MAIC	7.11	8.55	9.66	10.65	31.11	36.55	40.74	51.32
	MBIC	7.48	8.48	9.57	10.73	32.19	38.64	49.74	60.45
	t-sig	4.32	5.60	6.56	8.37	28.86	33.38	40.57	52.61
T=∞		6.59	7.69	8.52	9.82	34.65	42.57	49.76	58.73

Table 6 (cont'd). Finite sample critical values; Model II, choosing T_B minimizing the test statistic ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-5.30	-4.98	-4.74	-4.43	-2.60	-2.47	-2.35	-2.17
	BIC	-5.07	-4.80	-4.49	-4.23	-2.38	-2.18	-2.03	-1.86
	MAIC	-4.69	-4.43	-4.16	-3.85	-2.27	-2.12	-1.98	-1.85
	MBIC	-4.73	-4.42	-4.13	-3.86	-2.24	-2.07	-1.90	-1.74
	t-sig	-5.18	-4.89	-4.64	-4.36	-2.54	-2.39	-2.22	-2.05
T=150	AIC	-5.03	-4.82	-4.56	-4.23	-2.38	-2.20	-2.08	-1.93
	BIC	-4.82	-4.48	-4.27	-3.99	-2.19	-2.06	-1.91	-1.73
	MAIC	-4.62	-4.28	-4.02	-3.77	-2.16	-2.00	-1.89	-1.76
	MBIC	-4.62	-4.27	-4.03	-3.80	-2.13	-1.96	-1.83	-1.60
	t-sig	-4.98	-4.69	-4.46	-4.14	-2.32	-2.18	-2.02	-1.86
T=200	AIC	-5.03	-4.73	-4.39	-4.10	-2.20	-2.01	-1.88	-1.75
	BIC	-4.89	-4.44	-4.20	-3.85	-2.08	-1.88	-1.69	-1.56
	MAIC	-4.48	-4.22	-3.89	-3.60	-2.06	-1.89	-1.79	-1.65
	MBIC	-4.50	-4.28	-3.96	-3.65	-2.04	-1.86	-1.67	-1.51
	t-sig	-5.00	-4.51	-4.32	-4.09	-2.23	-2.05	-1.87	-1.68
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

Table 7. Finite sample critical values; Model II, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

a) MZ_α^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-210.88	-40.93	-29.41	-24.86	-7.28	-6.15	-5.07	-4.05
	BIC	-33.32	-29.05	-26.95	-13.19	-7.85	-6.64	-5.94	-4.82
	MAIC	-26.83	-24.32	-22.53	-19.89	-6.82	-5.91	-4.58	-3.87
	MBIC	-27.34	-25.56	-23.32	-21.16	-7.39	-6.45	-5.80	-4.90
	t-sig	-53.91	-35.70	-29.30	-25.41	-7.81	-6.71	-5.91	-4.59
T=150	AIC	-67.55	-38.00	-31.38	-25.84	-7.96	-6.82	-5.80	-4.96
	BIC	-34.49	-28.97	-27.17	-23.75	-8.06	-6.82	-5.89	-4.86
	MAIC	-28.71	-26.46	-24.12	-21.45	-7.36	-6.23	-5.47	-4.51
	MBIC	-30.21	-27.89	-24.86	-22.26	-7.84	-6.54	-5.66	-4.77
	t-sig	-39.38	-34.44	-29.00	-25.28	-8.15	-6.91	-5.92	-4.81
T=200	AIC	-56.97	-37.78	-29.24	-24.98	-7.53	-6.39	-5.30	-4.35
	BIC	-33.15	-30.16	-27.38	-23.51	-7.74	-6.33	-4.94	-4.17
	MAIC	-31.57	-27.62	-24.27	-21.72	-7.09	-5.97	-5.08	-4.17
	MBIC	-32.76	-29.47	-25.24	-22.10	-7.63	-6.06	-4.94	-4.14
	t-sig	-41.20	-35.10	-29.58	-25.16	-8.03	-6.49	-5.13	-4.25
T=∞		-41.00	-34.96	-30.74	-26.40	-7.45	-5.95	-4.88	-3.81

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0487	0.1105	0.1290	0.1407	0.2535	0.2694	0.2940	0.3197
	BIC	0.1218	0.1305	0.1348	0.1451	0.2454	0.2627	0.2758	0.2996
	MAIC	0.1356	0.1422	0.1468	0.1568	0.2617	0.2777	0.3009	0.3402
	MBIC	0.1331	0.1389	0.1449	0.1527	0.2521	0.2660	0.2817	0.3071
	t-sig	0.0963	0.1182	0.1302	0.1393	0.2460	0.2622	0.2772	0.2959
T=150	AIC	0.0860	0.1145	0.1257	0.1383	0.2424	0.2614	0.2771	0.2931
	BIC	0.1198	0.1300	0.1345	0.1439	0.2415	0.2585	0.2745	0.2931
	MAIC	0.1318	0.1367	0.1433	0.1512	0.2521	0.2710	0.2865	0.3033
	MBIC	0.1279	0.1332	0.1406	0.1481	0.2463	0.2630	0.2825	0.3001
	t-sig	0.1127	0.1198	0.1303	0.1394	0.2394	0.2561	0.2752	0.2993
T=200	AIC	0.0936	0.1150	0.1299	0.1403	0.2509	0.2684	0.2908	0.3044
	BIC	0.1218	0.1282	0.1338	0.1456	0.2471	0.2675	0.2985	0.3269
	MAIC	0.1249	0.1328	0.1415	0.1504	0.2570	0.2748	0.2935	0.3126
	MBIC	0.1226	0.1292	0.1393	0.1490	0.2504	0.2702	0.3005	0.3269
	t-sig	0.1096	0.1192	0.1295	0.1401	0.2423	0.2670	0.2908	0.3173
T=∞		0.1098	0.1187	0.1266	0.1365	0.2498	0.2751	0.2991	0.3339

Table 7 (cont'd). Finite sample critical values; Model II, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-10.26	-4.52	-3.81	-3.50	-1.85	-1.66	-1.44	-1.25
	BIC	-4.05	-3.79	-3.63	-3.37	-1.92	-1.75	-1.59	-1.41
	MAIC	-3.64	-3.46	-3.33	-3.14	-1.78	-1.59	-1.40	-1.18
	MBIC	-3.67	-3.52	-3.37	-3.20	-1.87	-1.71	-1.56	-1.32
	t-sig	-5.19	-4.22	-3.82	-3.54	-1.92	-1.76	-1.59	-1.34
T=150	AIC	-5.78	-4.35	-3.94	-3.55	-1.93	-1.76	-1.60	-1.34
	BIC	-4.13	-3.78	-3.66	-3.42	-1.95	-1.74	-1.56	-1.34
	MAIC	-3.77	-3.61	-3.43	-3.25	-1.83	-1.67	-1.54	-1.33
	MBIC	-3.88	-3.72	-3.51	-3.31	-1.91	-1.73	-1.54	-1.34
	t-sig	-4.43	-4.13	-3.78	-3.53	-1.94	-1.74	-1.62	-1.35
T=200	AIC	-5.33	-4.33	-3.82	-3.50	-1.86	-1.69	-1.56	-1.33
	BIC	-4.06	-3.87	-3.67	-3.39	-1.87	-1.69	-1.46	-1.19
	MAIC	-3.95	-3.67	-3.47	-3.27	-1.84	-1.67	-1.50	-1.25
	MBIC	-4.03	-3.83	-3.53	-3.30	-1.87	-1.66	-1.46	-1.19
	t-sig	-4.52	-4.11	-3.83	-3.52	-1.90	-1.71	-1.46	-1.33
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	1.25	6.43	9.03	10.90	36.81	41.58	51.42	66.17
	BIC	8.31	9.25	10.31	11.72	35.33	39.40	43.79	53.96
	MAIC	10.46	11.15	12.26	13.71	38.92	45.63	54.98	69.13
	MBIC	10.15	10.86	11.72	13.21	36.81	40.67	46.33	56.51
	t-sig	5.08	7.35	9.13	10.80	35.32	39.72	45.69	55.00
T=150	AIC	4.02	7.05	8.76	10.67	33.82	39.54	45.12	52.33
	BIC	7.96	9.39	10.14	11.58	33.34	38.98	44.16	52.33
	MAIC	9.39	10.59	11.40	12.81	36.29	42.29	48.16	54.90
	MBIC	9.10	9.87	10.94	12.35	34.68	40.09	46.79	54.23
	t-sig	6.92	7.96	9.35	10.75	33.05	38.83	45.85	52.33
T=200	AIC	4.71	7.55	9.30	10.98	36.11	41.45	49.42	56.46
	BIC	8.08	9.01	9.94	11.74	35.20	41.34	54.00	62.47
	MAIC	8.73	9.93	11.26	12.57	37.65	44.13	51.13	56.97
	MBIC	8.14	9.27	10.93	12.38	35.66	41.45	54.15	62.47
	t-sig	6.60	7.88	9.12	10.85	34.47	41.34	48.78	61.64
T=∞		6.79	7.85	8.92	10.34	34.97	42.67	50.46	62.11

Table 7 (cont'd). Finite sample critical values; Model II, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-4.89	-4.49	-4.31	-4.06	-2.08	-1.91	-1.75	-1.45
	BIC	-4.89	-4.42	-4.17	-3.84	-2.04	-1.84	-1.71	-1.49
	MAIC	-4.35	-4.12	-3.83	-3.57	-1.90	-1.72	-1.56	-1.39
	MBIC	-4.35	-4.12	-3.83	-3.58	-1.94	-1.77	-1.60	-1.37
	t-sig	-4.90	-4.49	-4.26	-4.02	-2.08	-1.91	-1.78	-1.54
T=150	AIC	-4.66	-4.41	-4.12	-3.85	-2.07	-1.87	-1.69	-1.38
	BIC	-4.59	-4.23	-4.03	-3.71	-2.02	-1.82	-1.59	-1.38
	MAIC	-4.31	-4.07	-3.77	-3.49	-1.86	-1.72	-1.55	-1.38
	MBIC	-4.31	-4.08	-3.79	-3.53	-1.91	-1.76	-1.55	-1.34
	t-sig	-4.59	-4.43	-4.10	-3.83	-2.05	-1.84	-1.69	-1.46
T=200	AIC	-4.73	-4.33	-4.10	-3.72	-1.96	-1.76	-1.59	-1.38
	BIC	-4.46	-4.19	-3.93	-3.61	-1.94	-1.75	-1.55	-1.26
	MAIC	-4.32	-3.87	-3.61	-3.28	-1.87	-1.69	-1.50	-1.24
	MBIC	-4.37	-3.94	-3.65	-3.41	-1.90	-1.69	-1.50	-1.26
	t-sig	-4.69	-4.24	-4.00	-3.71	-1.99	-1.78	-1.56	-1.38
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

Table 8: Size and Power; choosing T_B minimizing the tests; Model I; T=100
($\bar{c}=0$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.051	0.050	0.051	0.050	0.050	0.068	0.068	0.068	0.069	0.384
	BIC	0.050	0.050	0.051	0.051	0.051	0.500	0.495	0.495	0.521	0.497
	MAIC	0.050	0.051	0.051	0.050	0.050	0.409	0.424	0.422	0.441	0.487
	MBIC	0.050	0.050	0.051	0.050	0.050	0.553	0.553	0.551	0.563	0.503
	t-sig	0.050	0.051	0.050	0.050	0.051	0.212	0.213	0.213	0.214	0.464
	AIC	0.037	0.037	0.037	0.037	0.771	0.100	0.100	0.100	0.100	0.960
$\theta=-0.8$	BIC	0.771	0.769	0.766	0.776	0.973	0.919	0.919	0.919	0.918	1.000
	MAIC	0.117	0.117	0.121	0.131	0.353	0.363	0.363	0.365	0.370	0.871
	MBIC	0.132	0.132	0.135	0.142	0.353	0.373	0.370	0.374	0.376	0.874
	t-sig	0.324	0.323	0.324	0.324	0.749	0.670	0.670	0.670	0.671	0.973
	AIC	0.042	0.042	0.042	0.042	0.208	0.066	0.066	0.066	0.067	0.700
$\theta=-0.4$	BIC	0.299	0.296	0.297	0.306	0.429	0.751	0.753	0.752	0.768	0.904
	MAIC	0.060	0.064	0.069	0.076	0.120	0.277	0.279	0.291	0.319	0.475
	MBIC	0.084	0.084	0.088	0.095	0.118	0.343	0.342	0.339	0.367	0.491
	t-sig	0.056	0.055	0.058	0.060	0.227	0.396	0.396	0.397	0.407	0.729
	AIC	0.057	0.057	0.057	0.057	0.060	0.100	0.100	0.100	0.099	0.326
$\theta=0.4$	BIC	0.152	0.153	0.151	0.159	0.077	0.626	0.620	0.621	0.652	0.506
	MAIC	0.118	0.127	0.125	0.128	0.018	0.307	0.314	0.326	0.349	0.107
	MBIC	0.061	0.066	0.059	0.056	0.003	0.241	0.245	0.247	0.252	0.100
	t-sig	0.074	0.074	0.073	0.078	0.059	0.338	0.335	0.343	0.360	0.398
	AIC	0.146	0.145	0.146	0.148	0.057	0.296	0.296	0.296	0.298	0.168
$\theta=0.8$	BIC	0.289	0.287	0.286	0.289	0.109	0.585	0.581	0.583	0.600	0.383
	MAIC	0.206	0.213	0.217	0.222	0.012	0.298	0.301	0.324	0.345	0.103
	MBIC	0.053	0.056	0.057	0.056	0.001	0.118	0.113	0.124	0.151	0.031
	t-sig	0.101	0.101	0.102	0.105	0.050	0.326	0.320	0.327	0.332	0.227
	AIC	0.043	0.043	0.043	0.043	0.048	0.053	0.053	0.053	0.054	0.361
$\rho=-0.8$	BIC	0.006	0.006	0.006	0.007	0.044	0.022	0.021	0.022	0.025	0.462
	MAIC	0.006	0.007	0.006	0.006	0.047	0.006	0.006	0.006	0.007	0.410
	MBIC	0.000	0.000	0.000	0.001	0.047	0.003	0.003	0.003	0.004	0.406
	t-sig	0.012	0.011	0.012	0.014	0.044	0.062	0.062	0.063	0.067	0.432
	AIC	0.052	0.052	0.052	0.052	0.070	0.079	0.079	0.079	0.080	0.389
$\rho=-0.4$	BIC	0.094	0.096	0.096	0.103	0.140	0.365	0.360	0.364	0.394	0.567
	MAIC	0.032	0.034	0.037	0.039	0.062	0.188	0.194	0.212	0.238	0.417
	MBIC	0.026	0.024	0.027	0.032	0.059	0.240	0.231	0.251	0.283	0.421
	t-sig	0.040	0.040	0.041	0.042	0.071	0.212	0.208	0.215	0.218	0.459
	AIC	0.051	0.051	0.050	0.053	0.047	0.089	0.089	0.089	0.090	0.221
$\rho=0.4$	BIC	0.095	0.092	0.094	0.106	0.042	0.460	0.456	0.460	0.498	0.291
	MAIC	0.120	0.127	0.124	0.126	0.021	0.301	0.308	0.321	0.331	0.050
	MBIC	0.067	0.072	0.065	0.069	0.001	0.250	0.248	0.251	0.240	0.036
	t-sig	0.065	0.065	0.067	0.069	0.045	0.239	0.236	0.240	0.251	0.292
	AIC	0.067	0.068	0.067	0.067	0.083	0.083	0.083	0.083	0.083	0.114
$\rho=0.8$	BIC	0.177	0.183	0.169	0.170	0.067	0.286	0.286	0.280	0.301	0.118
	MAIC	0.247	0.269	0.250	0.240	0.074	0.298	0.307	0.320	0.330	0.145
	MBIC	0.202	0.213	0.197	0.186	0.064	0.337	0.341	0.340	0.350	0.031
	t-sig	0.131	0.133	0.128	0.123	0.072	0.185	0.186	0.185	0.188	0.134

Table 9: Size and Power; choosing T_B minimizing the tests; Model I; T=200
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.050	0.050	0.050	0.051	0.051	0.201	0.201	0.198	0.198	0.460
	BIC	0.051	0.052	0.051	0.050	0.051	0.538	0.532	0.555	0.538	0.535
	MAIC	0.051	0.051	0.050	0.051	0.050	0.466	0.463	0.461	0.468	0.509
	MBIC	0.051	0.050	0.050	0.050	0.051	0.532	0.521	0.541	0.553	0.531
	t-sig	0.050	0.050	0.050	0.050	0.051	0.373	0.371	0.367	0.359	0.468
$\theta = -0.8$	AIC	0.187	0.187	0.188	0.197	0.539	0.562	0.563	0.562	0.569	0.942
	BIC	0.617	0.616	0.626	0.628	0.873	0.907	0.907	0.909	0.911	0.999
	MAIC	0.035	0.034	0.036	0.039	0.206	0.200	0.193	0.202	0.209	0.668
	MBIC	0.042	0.041	0.045	0.046	0.197	0.209	0.205	0.216	0.226	0.670
	t-sig	0.233	0.232	0.233	0.237	0.566	0.573	0.571	0.574	0.579	0.975
$\theta = -0.4$	AIC	0.043	0.043	0.043	0.046	0.119	0.252	0.251	0.252	0.262	0.641
	BIC	0.177	0.174	0.185	0.186	0.228	0.778	0.772	0.789	0.800	0.848
	MAIC	0.064	0.063	0.063	0.066	0.074	0.349	0.342	0.347	0.364	0.447
	MBIC	0.078	0.077	0.078	0.078	0.072	0.460	0.461	0.479	0.485	0.464
	t-sig	0.046	0.047	0.045	0.047	0.119	0.422	0.415	0.423	0.432	0.647
$\theta = 0.4$	AIC	0.074	0.074	0.073	0.076	0.058	0.287	0.286	0.288	0.304	0.409
	BIC	0.101	0.102	0.105	0.108	0.096	0.656	0.643	0.669	0.677	0.590
	MAIC	0.080	0.080	0.078	0.080	0.048	0.450	0.444	0.454	0.470	0.367
	MBIC	0.054	0.052	0.056	0.059	0.008	0.441	0.437	0.455	0.474	0.090
	t-sig	0.074	0.072	0.072	0.067	0.052	0.433	0.427	0.432	0.437	0.426
$\theta = 0.8$	AIC	0.161	0.162	0.162	0.162	0.059	0.457	0.457	0.459	0.463	0.251
	BIC	0.203	0.197	0.208	0.202	0.083	0.629	0.621	0.640	0.643	0.424
	MAIC	0.147	0.150	0.145	0.147	0.024	0.425	0.417	0.422	0.439	0.185
	MBIC	0.052	0.052	0.052	0.054	0.017	0.253	0.247	0.266	0.293	0.167
	t-sig	0.157	0.159	0.155	0.155	0.066	0.567	0.567	0.567	0.569	0.335
$\rho = -0.8$	AIC	0.034	0.034	0.034	0.034	0.056	0.067	0.066	0.067	0.068	0.440
	BIC	0.003	0.003	0.003	0.003	0.047	0.032	0.029	0.036	0.042	0.516
	MAIC	0.002	0.002	0.002	0.002	0.047	0.015	0.013	0.016	0.024	0.452
	MBIC	0.000	0.000	0.000	0.000	0.044	0.009	0.008	0.010	0.020	0.448
	t-sig	0.015	0.013	0.015	0.014	0.047	0.072	0.071	0.073	0.080	0.466
$\rho = -0.4$	AIC	0.051	0.051	0.051	0.051	0.053	0.172	0.172	0.171	0.176	0.441
	BIC	0.037	0.034	0.041	0.043	0.050	0.409	0.404	0.433	0.443	0.515
	MAIC	0.036	0.035	0.039	0.041	0.048	0.339	0.325	0.333	0.346	0.476
	MBIC	0.029	0.026	0.028	0.034	0.048	0.367	0.358	0.396	0.421	0.470
	t-sig	0.042	0.043	0.042	0.042	0.043	0.268	0.264	0.269	0.285	0.456
$\rho = 0.4$	AIC	0.065	0.066	0.064	0.063	0.050	0.239	0.240	0.243	0.249	0.381
	BIC	0.072	0.069	0.073	0.077	0.044	0.538	0.527	0.569	0.573	0.454
	MAIC	0.080	0.080	0.079	0.078	0.056	0.476	0.470	0.472	0.483	0.381
	MBIC	0.064	0.065	0.065	0.063	0.005	0.481	0.473	0.491	0.501	0.018
	t-sig	0.063	0.062	0.061	0.063	0.046	0.404	0.403	0.405	0.401	0.396
$\rho = 0.8$	AIC	0.076	0.077	0.076	0.076	0.068	0.184	0.185	0.182	0.181	0.223
	BIC	0.103	0.106	0.109	0.105	0.064	0.351	0.341	0.358	0.369	0.248
	MAIC	0.120	0.126	0.120	0.117	0.072	0.342	0.337	0.342	0.348	0.247
	MBIC	0.097	0.102	0.101	0.106	0.070	0.336	0.329	0.347	0.370	0.245
	t-sig	0.091	0.088	0.090	0.086	0.061	0.282	0.281	0.279	0.280	0.232

Table 10: Size and Power; choosing T_B maximizing $|t_{\widehat{\beta}_2}|$; Model I; T=100
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.051	0.050	0.051	0.050	0.050	0.240	0.232	0.234	0.231	0.388
	BIC	0.050	0.050	0.050	0.050	0.050	0.457	0.429	0.467	0.488	0.495
	MAIC	0.050	0.050	0.050	0.051	0.050	0.253	0.253	0.262	0.275	0.414
	MBIC	0.050	0.052	0.051	0.051	0.050	0.353	0.345	0.366	0.359	0.417
	t-sig	0.051	0.051	0.051	0.051	0.050	0.212	0.216	0.213	0.219	0.428
$\theta = -0.8$	AIC	0.104	0.104	0.104	0.107	0.628	0.175	0.174	0.175	0.174	0.874
	BIC	0.373	0.363	0.373	0.374	0.892	0.446	0.445	0.448	0.443	0.998
	MAIC	0.005	0.005	0.006	0.006	0.164	0.028	0.028	0.028	0.028	0.554
	MBIC	0.008	0.008	0.008	0.007	0.160	0.026	0.027	0.027	0.027	0.553
	t-sig	0.097	0.097	0.098	0.102	0.575	0.183	0.183	0.183	0.184	0.913
$\theta = -0.4$	AIC	0.071	0.069	0.071	0.072	0.195	0.208	0.204	0.208	0.204	0.609
	BIC	0.187	0.181	0.186	0.189	0.411	0.477	0.462	0.483	0.490	0.860
	MAIC	0.027	0.028	0.027	0.030	0.091	0.080	0.081	0.081	0.086	0.310
	MBIC	0.029	0.027	0.033	0.035	0.090	0.098	0.093	0.108	0.102	0.313
	t-sig	0.051	0.051	0.052	0.054	0.194	0.187	0.190	0.191	0.189	0.614
$\theta = 0.4$	AIC	0.082	0.083	0.081	0.078	0.070	0.220	0.214	0.220	0.217	0.307
	BIC	0.122	0.109	0.125	0.124	0.087	0.389	0.377	0.395	0.410	0.486
	MAIC	0.063	0.066	0.067	0.069	0.012	0.119	0.118	0.124	0.137	0.091
	MBIC	0.025	0.027	0.024	0.022	0.004	0.103	0.099	0.114	0.114	0.091
	t-sig	0.083	0.085	0.083	0.084	0.052	0.235	0.240	0.239	0.247	0.359
$\theta = 0.8$	AIC	0.148	0.151	0.146	0.146	0.050	0.137	0.136	0.136	0.139	0.145
	BIC	0.146	0.144	0.145	0.150	0.088	0.205	0.202	0.209	0.212	0.328
	MAIC	0.065	0.071	0.064	0.069	0.009	0.049	0.048	0.050	0.055	0.071
	MBIC	0.017	0.018	0.015	0.017	0.001	0.022	0.020	0.023	0.023	0.028
	t-sig	0.078	0.080	0.078	0.079	0.044	0.122	0.122	0.123	0.126	0.183
$\rho = -0.8$	AIC	0.014	0.014	0.014	0.014	0.036	0.014	0.014	0.014	0.014	0.254
	BIC	0.000	0.000	0.000	0.000	0.031	0.004	0.003	0.004	0.004	0.364
	MAIC	0.000	0.000	0.000	0.000	0.030	0.003	0.003	0.004	0.004	0.232
	MBIC	0.000	0.000	0.000	0.000	0.031	0.003	0.002	0.003	0.003	0.231
	t-sig	0.003	0.003	0.003	0.003	0.032	0.013	0.013	0.013	0.013	0.304
$\rho = -0.4$	AIC	0.051	0.050	0.050	0.050	0.064	0.110	0.106	0.110	0.112	0.333
	BIC	0.051	0.046	0.050	0.050	0.135	0.215	0.207	0.218	0.228	0.513
	MAIC	0.015	0.013	0.018	0.024	0.047	0.073	0.072	0.081	0.091	0.297
	MBIC	0.013	0.013	0.015	0.013	0.043	0.083	0.078	0.094	0.087	0.295
	t-sig	0.038	0.039	0.039	0.041	0.060	0.111	0.112	0.113	0.118	0.376
$\rho = 0.4$	AIC	0.089	0.092	0.086	0.089	0.043	0.218	0.212	0.217	0.215	0.219
	BIC	0.082	0.078	0.083	0.084	0.041	0.329	0.314	0.338	0.352	0.304
	MAIC	0.081	0.088	0.081	0.083	0.014	0.156	0.157	0.158	0.168	0.042
	MBIC	0.035	0.038	0.034	0.029	0.001	0.101	0.103	0.107	0.100	0.034
	t-sig	0.077	0.080	0.076	0.084	0.038	0.225	0.225	0.226	0.231	0.251
$\rho = 0.8$	AIC	0.145	0.149	0.142	0.131	0.045	0.160	0.162	0.153	0.150	0.099
	BIC	0.137	0.134	0.131	0.133	0.038	0.224	0.210	0.225	0.227	0.116
	MAIC	0.140	0.157	0.130	0.138	0.051	0.193	0.199	0.195	0.210	0.127
	MBIC	0.136	0.145	0.137	0.129	0.038	0.212	0.215	0.222	0.217	0.017
	t-sig	0.133	0.139	0.130	0.124	0.041	0.170	0.178	0.168	0.169	0.109

Table 11: Size and Power; choosing T_B maximizing $|t_{\widehat{\beta}_2}|$; Model I; T=200
($\bar{c} = 0$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.051	0.050	0.050	0.050	0.051	0.307	0.299	0.312	0.327	0.395
	BIC	0.050	0.050	0.051	0.050	0.050	0.465	0.463	0.474	0.482	0.480
	MAIC	0.051	0.049	0.051	0.050	0.051	0.365	0.352	0.376	0.367	0.460
	MBIC	0.050	0.050	0.051	0.051	0.051	0.424	0.416	0.433	0.436	0.496
	t-sig	0.051	0.051	0.050	0.051	0.050	0.320	0.317	0.328	0.310	0.428
$\theta = -0.8$	AIC	0.042	0.040	0.042	0.043	0.302	0.145	0.143	0.145	0.147	0.741
	BIC	0.238	0.238	0.241	0.239	0.675	0.427	0.430	0.428	0.422	0.977
	MAIC	0.001	0.001	0.002	0.002	0.068	0.020	0.018	0.021	0.021	0.287
	MBIC	0.004	0.004	0.004	0.004	0.071	0.020	0.019	0.018	0.019	0.286
	t-sig	0.050	0.049	0.050	0.050	0.355	0.171	0.170	0.173	0.174	0.865
$\theta = -0.4$	AIC	0.042	0.040	0.041	0.050	0.102	0.261	0.262	0.269	0.279	0.500
	BIC	0.113	0.113	0.115	0.121	0.212	0.589	0.586	0.591	0.585	0.757
	MAIC	0.032	0.033	0.037	0.036	0.066	0.160	0.155	0.170	0.159	0.322
	MBIC	0.039	0.038	0.041	0.039	0.064	0.194	0.193	0.197	0.199	0.336
	t-sig	0.049	0.047	0.050	0.046	0.109	0.290	0.293	0.295	0.291	0.538
$\theta = 0.4$	AIC	0.065	0.064	0.064	0.067	0.049	0.288	0.282	0.293	0.308	0.345
	BIC	0.082	0.085	0.082	0.084	0.084	0.495	0.487	0.501	0.494	0.539
	MAIC	0.056	0.055	0.061	0.061	0.057	0.288	0.283	0.305	0.305	0.313
	MBIC	0.030	0.030	0.033	0.034	0.006	0.217	0.216	0.223	0.228	0.092
	t-sig	0.065	0.067	0.064	0.064	0.054	0.341	0.335	0.344	0.329	0.405
$\theta = 0.8$	AIC	0.093	0.092	0.093	0.098	0.047	0.215	0.207	0.216	0.236	0.189
	BIC	0.098	0.100	0.098	0.101	0.072	0.326	0.325	0.332	0.337	0.356
	MAIC	0.057	0.057	0.060	0.067	0.021	0.142	0.133	0.156	0.160	0.149
	MBIC	0.018	0.016	0.020	0.021	0.017	0.087	0.083	0.093	0.093	0.153
	t-sig	0.113	0.118	0.117	0.113	0.067	0.392	0.390	0.397	0.382	0.329
$\rho = -0.8$	AIC	0.007	0.007	0.007	0.008	0.040	0.032	0.032	0.033	0.040	0.280
	BIC	0.001	0.001	0.001	0.001	0.036	0.014	0.012	0.013	0.016	0.345
	MAIC	0.000	0.000	0.000	0.000	0.035	0.012	0.010	0.014	0.013	0.289
	MBIC	0.000	0.000	0.000	0.000	0.035	0.006	0.006	0.007	0.012	0.281
	t-sig	0.005	0.004	0.005	0.005	0.041	0.035	0.032	0.037	0.039	0.325
$\rho = -0.4$	AIC	0.026	0.025	0.028	0.031	0.048	0.187	0.177	0.190	0.193	0.343
	BIC	0.027	0.027	0.027	0.027	0.044	0.300	0.297	0.306	0.301	0.410
	MAIC	0.022	0.023	0.026	0.026	0.047	0.195	0.192	0.214	0.212	0.379
	MBIC	0.015	0.016	0.017	0.017	0.049	0.220	0.215	0.228	0.224	0.383
	t-sig	0.026	0.026	0.027	0.027	0.049	0.226	0.229	0.235	0.219	0.371
$\rho = 0.4$	AIC	0.065	0.065	0.063	0.072	0.040	0.329	0.321	0.333	0.350	0.345
	BIC	0.059	0.062	0.057	0.063	0.042	0.461	0.456	0.471	0.472	0.403
	MAIC	0.066	0.066	0.069	0.066	0.058	0.344	0.333	0.354	0.356	0.336
	MBIC	0.044	0.046	0.042	0.044	0.005	0.262	0.257	0.265	0.260	0.018
	t-sig	0.083	0.087	0.081	0.082	0.044	0.375	0.371	0.382	0.373	0.382
$\rho = 0.8$	AIC	0.079	0.082	0.079	0.084	0.047	0.228	0.221	0.231	0.245	0.187
	BIC	0.071	0.077	0.070	0.075	0.045	0.296	0.292	0.298	0.294	0.219
	MAIC	0.090	0.096	0.089	0.092	0.057	0.262	0.258	0.278	0.272	0.246
	MBIC	0.080	0.080	0.074	0.073	0.054	0.284	0.280	0.294	0.294	0.247
	t-sig	0.074	0.074	0.074	0.068	0.046	0.261	0.255	0.262	0.260	0.239

Tables 12.a: Results for the Real Wages and Stock Prices Series
Choosing the break point minimizing the tests ($\bar{c} = 0$ to construct s^2)

Serie	T	Criteria	MZ_t	k	T_B	P_T	k	T_B	ADF	k	T_B	α
Real Wages	71	BIC	-3.85^d	1	1940	9.49^d	1	1938	-4.63^d	1	1938	0.62
		MAIC	-3.85^c	1	1940	9.49^c	1	1938	-4.63^b	1	1938	0.62
		MBIC	-3.85^a	1	1940	9.49^b	1	1938	-4.63^b	1	1938	0.62
Stock Prices	100	BIC	-4.69^a	1	1945	6.24^a	1	1945	-5.12^a	1	1937	0.67
		MAIC	-4.69^a	1	1945	6.24^a	1	1945	-5.12^a	1	1937	0.67
		MBIC	-4.63^a	1	1937	6.45^a	1	1937	-5.12^a	1	1937	0.67

Notes: 1) For the applications, we impose a minimal value $k = 1$; 2) the superscipts a, b,c and d denote significance levels at the 1.0%, 2.5%, 5.0%, and 10.0%, respectively.

Tables 12.b: Results for the Real Wages and Stock Prices Series
Choosing the break point maximizing $|t_{\hat{\beta}_2}|$ ($\bar{c} = 0$ to construct s^2)

Series	T	T_B	Criteria	MZ_t	k	P_T	k	ADF	k	α
Real Wages	71	1933	BIC	-3.37^d	1	11.46^d	1	-3.83	1	0.70
			MAIC	-3.37^c	1	11.46^b	1	-3.83^d	1	0.70
			MBIC	-3.37^c	1	11.46^b	1	-3.83^d	1	0.70
Stock Prices	100	1931	BIC	-3.87^b	1	9.14^b	1	-4.16^c	1	0.75
			MAIC	-3.04	2	14.66	2	-3.25	2	0.79
			MBIC	-3.04	2	14.66	2	-3.25	2	0.79

Notes: 1) For the applications, we impose a minimal value $k = 1$; 2) the superscipts a, b,c and d denote significance levels at the 1.0%, 2.5%, 5.0%, and 10.0%, respectively.

Annex 2

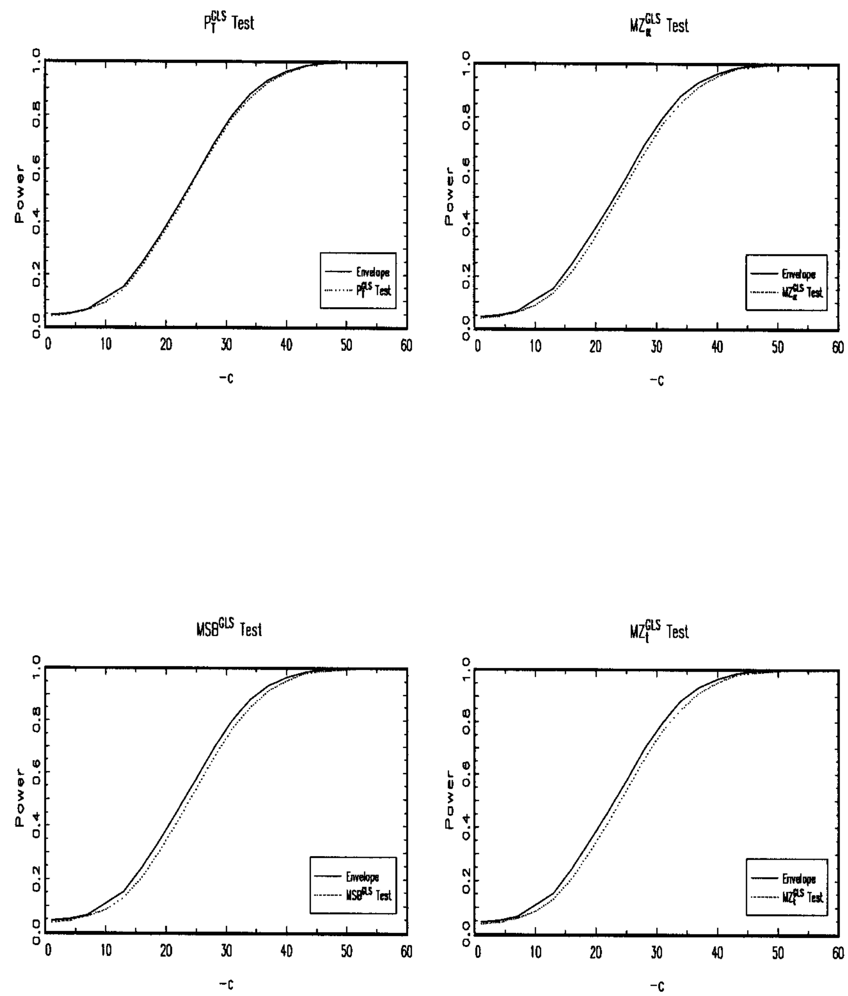

Figure 1. Envelope and asymptotic power functions choosing T_B minimizing the tests

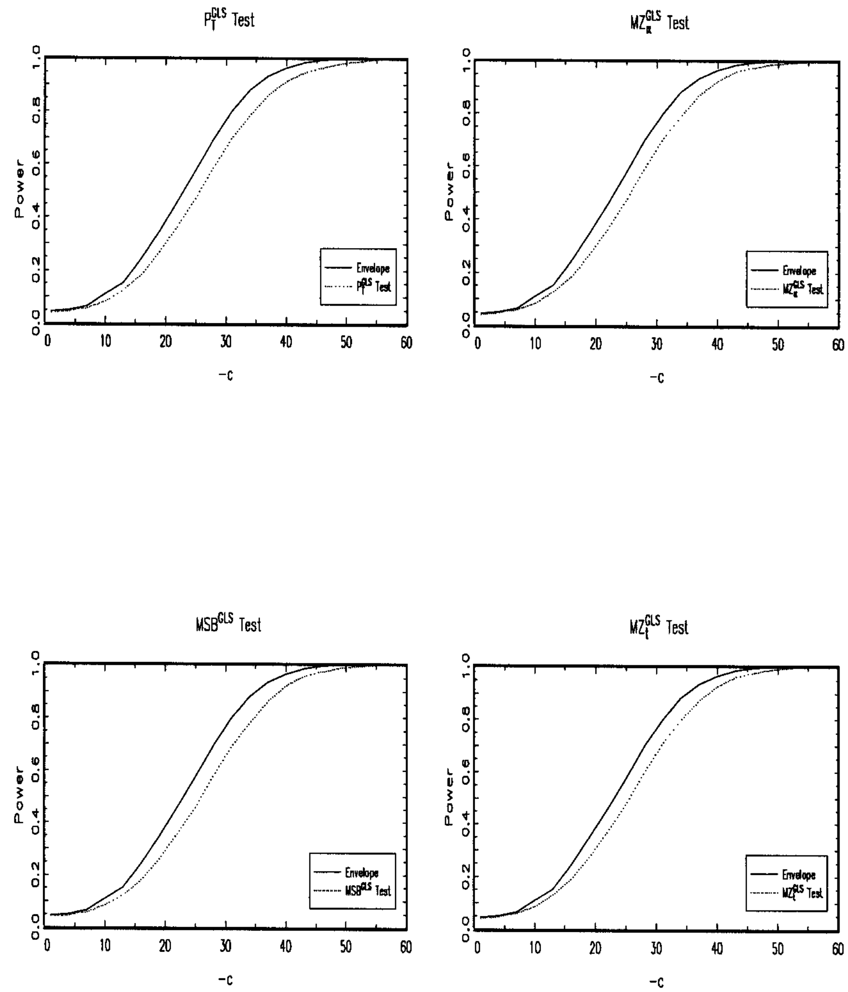

Figure 2. Envelope and asymptotic power functions choosing T_B maximizing $| t_{\widehat{\beta}_2} |$

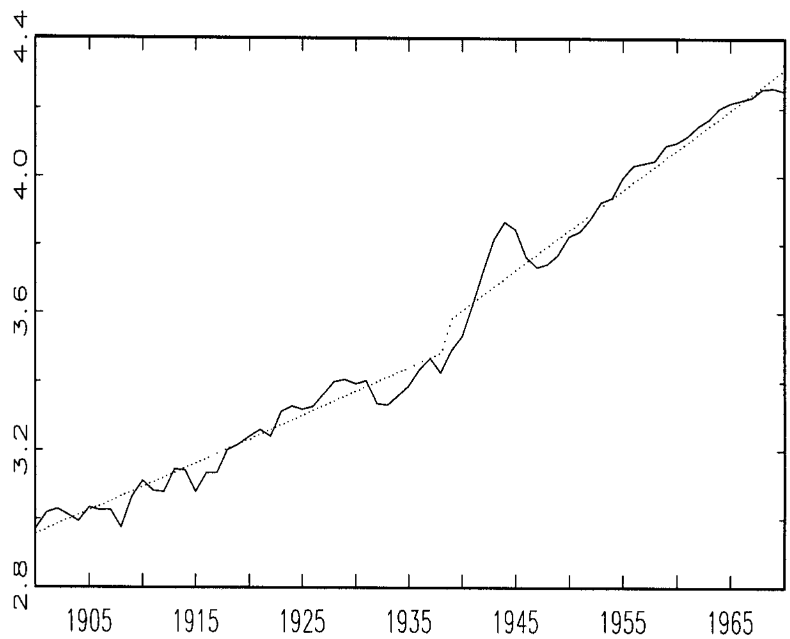

Figure 3. Logarithm of Real wages (1900-1970) and estimated broken trend ($T_B = 1938$)

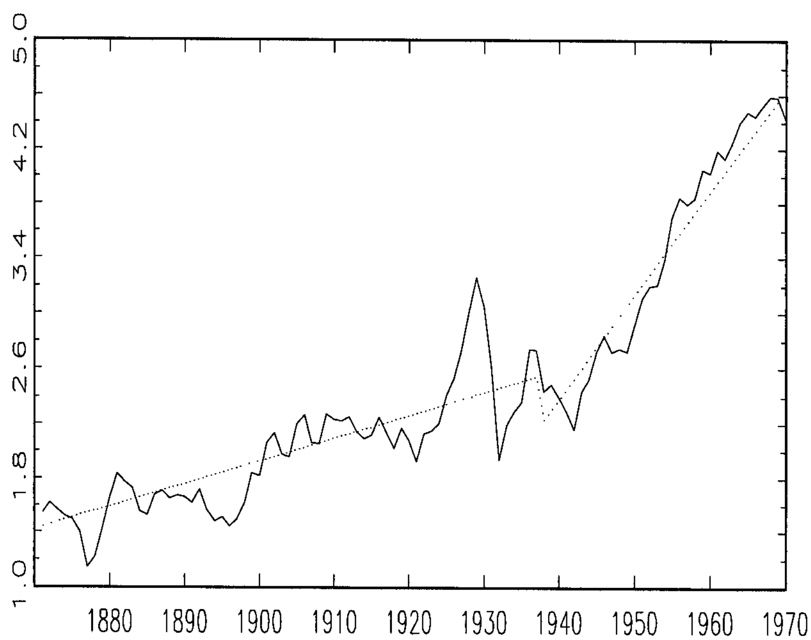

Figure 4. Logarithm of Common stock prices (1871-1970) and estimated broken trend ($T_B = 1937$)

Annex 3

Table 13: Exact Size of the tests at selected values of k in s^2
($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS} ; MA case

T	θ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	1.000	0.880	0.602	0.499	0.442	0.473	0.497	0.538	0.557	0.604	0.656
	-0.40	0.432	0.106	0.090	0.168	0.244	0.350	0.479	0.592	0.663	0.748	0.806
	0.00	0.010	0.044	0.110	0.241	0.338	0.462	0.561	0.679	0.731	0.822	0.849
	0.40	0.000	0.150	0.083	0.285	0.340	0.505	0.592	0.695	0.761	0.827	0.850
	0.80	0.000	0.314	0.016	0.457	0.220	0.627	0.504	0.754	0.715	0.853	0.829
200	-0.80	1.000	0.994	0.842	0.539	0.340	0.266	0.202	0.197	0.197	0.224	0.245
	-0.40	0.685	0.169	0.070	0.078	0.083	0.131	0.179	0.245	0.293	0.350	0.421
	0.00	0.020	0.038	0.056	0.099	0.129	0.194	0.250	0.316	0.364	0.416	0.477
	0.40	0.001	0.130	0.043	0.121	0.136	0.208	0.253	0.328	0.372	0.438	0.492
	0.80	0.000	0.257	0.009	0.246	0.076	0.315	0.180	0.403	0.318	0.482	0.456

b) P_T^{GLS} ; MA case

T	θ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	1.000	0.872	0.602	0.503	0.446	0.478	0.495	0.539	0.559	0.604	0.658
	-0.40	0.418	0.107	0.087	0.169	0.238	0.349	0.474	0.584	0.652	0.739	0.800
	0.00	0.008	0.044	0.099	0.231	0.315	0.451	0.555	0.666	0.724	0.816	0.839
	0.40	0.000	0.136	0.075	0.265	0.324	0.487	0.573	0.678	0.750	0.822	0.844
	0.80	0.000	0.291	0.016	0.432	0.212	0.607	0.485	0.737	0.704	0.838	0.822
200	-0.80	1.000	0.993	0.833	0.529	0.338	0.267	0.203	0.194	0.196	0.220	0.245
	-0.40	0.665	0.157	0.067	0.077	0.080	0.126	0.170	0.239	0.288	0.340	0.409
	0.00	0.021	0.034	0.054	0.093	0.128	0.184	0.241	0.303	0.357	0.406	0.469
	0.40	0.001	0.119	0.039	0.112	0.129	0.200	0.241	0.312	0.365	0.421	0.486
	0.80	0.000	0.241	0.009	0.226	0.071	0.304	0.173	0.388	0.302	0.472	0.440

c) MZ_{α}^{GLS} ; AR case

T	ρ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	0.960	0.000	0.004	0.040	0.138	0.255	0.386	0.512	0.618	0.700	0.767
	-0.40	0.287	0.014	0.053	0.179	0.279	0.398	0.518	0.636	0.708	0.792	0.832
	0.00	0.010	0.044	0.110	0.241	0.338	0.462	0.561	0.679	0.731	0.822	0.849
	0.40	0.000	0.067	0.162	0.277	0.377	0.524	0.618	0.710	0.773	0.840	0.864
	0.80	0.000	0.183	0.308	0.462	0.588	0.713	0.777	0.834	0.875	0.909	0.932
200	-0.80	0.990	0.002	0.001	0.007	0.035	0.059	0.119	0.169	0.232	0.290	0.345
	-0.40	0.460	0.020	0.030	0.078	0.102	0.155	0.208	0.277	0.342	0.389	0.453
	0.00	0.020	0.038	0.056	0.099	0.129	0.194	0.250	0.316	0.364	0.416	0.477
	0.40	0.000	0.045	0.078	0.118	0.159	0.221	0.272	0.335	0.381	0.458	0.501
	0.80	0.000	0.088	0.135	0.179	0.239	0.317	0.365	0.437	0.514	0.565	0.604

d) P_T^{GLS} ; AR case

T	ρ	k=0	1	2	3	4	5	6	7	8	9	10
100	-0.80	0.952	0.000	0.004	0.044	0.142	0.254	0.388	0.510	0.618	0.694	0.768
	-0.40	0.281	0.011	0.053	0.173	0.268	0.391	0.507	0.629	0.696	0.785	0.826
	0.00	0.008	0.044	0.099	0.231	0.315	0.451	0.555	0.666	0.724	0.816	0.839
	0.40	0.000	0.058	0.138	0.259	0.357	0.503	0.606	0.696	0.762	0.831	0.856
	0.80	0.000	0.147	0.278	0.426	0.560	0.682	0.763	0.826	0.871	0.908	0.924
200	-0.80	0.984	0.002	0.001	0.010	0.034	0.062	0.118	0.169	0.225	0.284	0.340
	-0.40	0.441	0.021	0.029	0.077	0.098	0.149	0.210	0.275	0.333	0.383	0.441
	0.00	0.021	0.034	0.054	0.093	0.128	0.184	0.241	0.303	0.357	0.406	0.469
	0.40	0.000	0.041	0.074	0.108	0.144	0.209	0.257	0.320	0.372	0.445	0.487
	0.80	0.000	0.076	0.117	0.165	0.216	0.294	0.351	0.421	0.494	0.551	0.588

Table 14. Selected values of k using IC and MIC (MZ_{α}^{GLS})
($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

	MA Case						AR Case					
T	θ	k^*	AIC	BIC	MAIC	MBIC	ρ	k^*	AIC	BIC	MAIC	MBIC
100	-0.8	4	1	0	5	4	-0.8	3	1	1	1	1
	-0.4	2	1	0	2	2	-0.4	2	1	1	1	1
	0.0	1	0	0	0	0	0.0	1	0	0	0	0
	0.4	2	2	1	2	0	0.4	1	1	1	1	0
	0.8	2	5	3	4	2	0.8	0	1	1	1	1
200	-0.8	7	3	1	8	8	-0.8	4	1	1	1	1
	-0.4	2	2	1	3	2	-0.4	2	1	1	1	1
	0.0	1	0	0	0	0	0.0	1	0	0	0	0
	0.4	2	2	1	2	2	0.4	1	1	1	1	1
	0.8	4	7	4	6	2	0.8	1	1	1	1	1

Table 15. Finite sample critical values; Model I, choosing T_B minimizing the test statistic ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-2980689.66	-206553.61	-2349.24	-175.28	-10.09	-8.54	-7.30	-6.51
	BIC	-43.37	-34.62	-29.47	-24.96	-8.41	-7.13	-6.32	-5.26
	MAIC	-29.47	-26.39	-23.84	-21.63	-8.14	-7.04	-6.33	-5.47
	MBIC	-29.36	-26.14	-23.32	-21.32	-7.62	-6.61	-5.76	-4.77
	t-sig	-189.00	-117.00	-73.61	-44.30	-9.79	-8.44	-7.29	-6.50
T=150	AIC	-26042.22	-445.23	-130.48	-54.00	-9.74	-8.30	-7.19	-6.60
	BIC	-37.23	-32.19	-28.95	-24.78	-8.53	-7.20	-6.40	-5.53
	MAIC	-32.50	-29.52	-25.67	-23.10	-8.31	-7.09	-6.42	-5.50
	MBIC	-31.69	-28.84	-25.67	-23.24	-8.07	-6.80	-5.94	-4.85
	t-sig	-64.53	-52.85	-41.95	-34.38	-9.49	-8.50	-7.51	-6.60
T=200	AIC	-759.07	-169.02	-66.72	-38.90	-9.05	-7.40	-6.65	-5.64
	BIC	-38.32	-30.85	-28.29	-24.97	-8.13	-6.68	-5.20	-4.64
	MAIC	-33.33	-29.54	-26.04	-22.37	-8.10	-6.93	-5.87	-4.84
	MBIC	-33.20	-29.55	-25.43	-22.50	-7.88	-6.57	-5.02	-4.29
	t-sig	-70.04	-53.61	-43.56	-35.05	-9.09	-7.62	-6.26	-5.02
T=∞		-40.88	-35.48	-31.63	-27.46	-7.46	-6.00	-4.97	-4.09

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0004	0.0016	0.0146	0.0553	0.2190	0.2352	0.2514	0.2643
	BIC	0.1067	0.1195	0.1296	0.1390	0.2363	0.2554	0.2690	0.2838
	MAIC	0.1299	0.1367	0.1441	0.1505	0.2419	0.2557	0.2706	0.2852
	MBIC	0.1305	0.1374	0.1444	0.1512	0.2459	0.2644	0.2797	0.3080
	t-sig	0.0512	0.0649	0.0822	0.1059	0.2193	0.2362	0.2543	0.2659
T=150	AIC	0.0044	0.0335	0.0619	0.0958	0.2206	0.2393	0.2528	0.2690
	BIC	0.1158	0.1233	0.1308	0.1406	0.2341	0.2528	0.2673	0.2878
	MAIC	0.1229	0.1295	0.1393	0.1460	0.2392	0.2549	0.2699	0.2908
	MBIC	0.1236	0.1308	0.1394	0.1460	0.2417	0.2597	0.2758	0.2946
	t-sig	0.0878	0.0967	0.1090	0.1199	0.2215	0.2358	0.2489	0.2670
T=200	AIC	0.0257	0.0542	0.0857	0.1119	0.2272	0.2536	0.2660	0.2881
	BIC	0.1139	0.1259	0.1313	0.1404	0.2396	0.2606	0.2925	0.3181
	MAIC	0.1218	0.1298	0.1380	0.1474	0.2402	0.2588	0.2793	0.2995
	MBIC	0.1218	0.1297	0.1385	0.1472	0.2453	0.2631	0.2974	0.3250
	t-sig	0.0845	0.0964	0.1069	0.1193	0.2277	0.2474	0.2681	0.2944
T=∞		0.1096	0.1179	0.1250	0.1338	0.2495	0.2754	0.2985	0.3239

Table 15 (cont'd). Finite sample critical values; Model I, choosing T_B minimizing the test statistic ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-1220.79	-321.36	-34.23	-9.35	-2.20	-2.01	-1.86	-1.77
	BIC	-4.63	-4.15	-3.81	-3.50	-2.01	-1.82	-1.73	-1.58
	MAIC	-3.82	-3.62	-3.41	-3.25	-1.96	-1.82	-1.72	-1.57
	MBIC	-3.81	-3.58	-3.40	-3.23	-1.89	-1.77	-1.62	-1.48
	t-sig	-9.67	-7.59	-6.04	-4.70	-2.16	-1.98	-1.84	-1.75
T=150	AIC	-114.11	-14.91	-8.07	-5.17	-2.15	-1.99	-1.85	-1.75
	BIC	-4.31	-4.00	-3.76	-3.50	-2.00	-1.85	-1.74	-1.60
	MAIC	-4.00	-3.82	-3.58	-3.35	-1.98	-1.82	-1.72	-1.56
	MBIC	-3.96	-3.78	-3.57	-3.36	-1.93	-1.78	-1.66	-1.48
	t-sig	-5.66	-5.13	-4.57	-4.12	-2.13	-2.00	-1.85	-1.72
T=200	AIC	-19.48	-9.16	-5.75	-4.39	-2.06	-1.86	-1.77	-1.61
	BIC	-4.36	-3.89	-3.74	-3.52	-1.94	-1.76	-1.57	-1.45
	MAIC	-4.06	-3.79	-3.58	-3.33	-1.93	-1.79	-1.67	-1.48
	MBIC	-4.05	-3.83	-3.55	-3.33	-1.92	-1.73	-1.56	-1.41
	t-sig	-5.91	-5.08	-4.66	-4.16	-2.05	-1.88	-1.73	-1.56
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.00	0.00	0.12	1.59	27.74	32.76	35.53	39.81
	BIC	6.21	7.76	9.16	11.04	32.76	37.41	41.29	47.93
	MAIC	9.09	10.20	11.37	12.83	34.31	37.45	41.29	47.52
	MBIC	9.11	10.29	11.40	12.83	35.61	39.81	45.44	55.52
	t-sig	1.43	2.42	3.75	6.04	27.83	32.93	36.54	41.47
T=150	AIC	0.01	0.59	2.14	5.03	28.43	32.67	37.75	40.69
	BIC	7.09	8.31	9.45	10.90	31.19	38.03	41.35	50.07
	MAIC	8.27	9.24	10.37	11.74	32.48	38.04	41.93	50.62
	MBIC	8.65	9.42	10.45	11.74	33.41	39.35	44.10	52.86
	t-sig	4.16	5.13	6.51	7.89	28.70	32.19	36.72	41.28
T=200	AIC	0.37	1.70	4.01	6.93	30.27	36.67	41.10	49.37
	BIC	7.07	8.72	9.57	10.88	33.31	39.94	50.23	60.68
	MAIC	8.04	9.35	10.52	12.04	33.63	38.42	46.38	54.22
	MBIC	8.05	9.24	10.56	12.07	34.47	40.57	52.71	61.97
	t-sig	3.81	5.12	6.25	7.71	29.63	35.68	41.88	53.93
T=∞		6.59	7.69	8.52	9.82	34.65	42.57	49.76	58.73

Table 15 (cont'd). Finite sample critical values; Model I, choosing T_B minimizing the test statistic ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-5.04	-4.79	-4..57	-4.26	-2.31	-2.11	-1.99	-1.87
	BIC	-5.00	-4.67	-4.34	-3.94	-2.16	-1.98	-1.87	-1.69
	MAIC	-4.62	-4.26	-3.91	-3.67	-2.06	-1.92	-1.82	-1.67
	MBIC	-4.62	-4.25	-3.92	-3.68	-2.04	-1.89	-1.73	-1.57
	t-sig	-5.00	-4.69	-4.43	-4.12	-2.27	-2.05	-1.93	-1.85
T=150	AIC	-4.95	-4.64	-4.37	-4.04	-2.22	-2.06	-1.90	-1.80
	BIC	-4.72	-4.38	-4.18	-3.83	-2.09	-1.93	-1.83	-1.70
	MAIC	-4.52	-4.21	-3.94	-3.64	-2.05	-1.89	-1.80	-1.65
	MBIC	-4.52	-4.21	-3.94	-3.66	-2.02	-1.85	-1.76	-1.53
	t-sig	-4.80	-4.61	-4.28	-3.95	-2.16	-2.05	-1.89	-1.80
T=200	AIC	-4.93	-4.46	-4.21	-3.92	-2.08	-1.90	-1.81	-1.67
	BIC	-4.72	-4.26	-4.06	-3.74	-2.00	-1.83	-1.67	-1.50
	MAIC	-4.47	-4.06	-3.77	-3.50	-1.99	-1.83	-1.73	-1.52
	MBIC	-4.47	-4.17	-3.81	-3.52	-1.97	-1.79	-1.63	-1.48
	t-sig	-4.90	-4.46	-4.19	-3.91	-2.08	-1.89	-1.76	-1.62
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

Table 16. Finite sample critical values; Model I, choosing T_B maximizing $|t_{\widehat{\beta_2}}|$ ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-2010.00	-299.38	-112.86	-43.39	-8.31	-7.02	-6.10	-4.89
	BIC	-40.39	-31.48	-27.13	-23.31	-7.82	-6.68	-5.78	-4.29
	MAIC	-28.49	-25.84	-22.43	-20.74	-7.08	-6.10	-5.05	-4.27
	MBIC	-28.39	-25.65	-22.51	-20.83	-7.13	-6.26	-5.18	-4.11
	t-sig	-141.06	-87.62	-58.32	-36.58	-8.70	-7.13	-6.10	-4.50
T=150	AIC	-499.25	-105.80	-56.82	-34.83	-8.36	-7.10	-6.20	-4.98
	BIC	-35.58	-31.42	-27.44	-24.01	-8.15	-6.81	-5.79	-4.93
	MAIC	-31.34	-27.86	-24.65	-22.08	-7.48	-6.36	-5.42	-4.68
	MBIC	-31.34	-27.93	-24.89	-22.21	-7.55	-6.39	-5.46	-4.66
	t-sig	-63.56	-45.83	-38.73	-30.74	-8.71	-7.50	-6.45	-5.04
T=200	AIC	-417.79	-83.67	-48.19	-29.93	-7.49	-6.71	-5.56	-4.41
	BIC	-36.33	-30.22	-27.56	-24.01	-7.86	-6.34	-4.90	-4.05
	MAIC	-32.47	-27.49	-24.48	-21.48	-7.24	-6.20	-5.11	-3.60
	MBIC	-32.47	-28.29	-24.50	-21.59	-7.37	-6.05	-4.66	-3.63
	t-sig	-64.86	-47.89	-39.66	-30.50	-8.14	-6.88	-5.55	-4.26
T=∞		-41.00	-34.96	-30.74	-26.40	-7.45	-5.95	-4.88	-3.81

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0158	0.0409	0.0663	0.1072	0.2392	0.2564	0.2695	0.2839
	BIC	0.1112	0.1255	0.1352	0.1452	0.2462	0.2633	0.2790	0.3032
	MAIC	0.1324	0.1384	0.1472	0.1539	0.2564	0.2711	0.2880	0.3230
	MBIC	0.1325	0.1387	0.1470	0.1536	0.2560	0.2704	0.2880	0.3118
	t-sig	0.0593	0.0755	0.0922	0.1166	0.2339	0.2546	0.2695	0.2992
T=150	AIC	0.0316	0.0684	0.0936	0.1196	0.2381	0.2535	0.2680	0.2902
	BIC	0.1185	0.1248	0.1337	0.1435	0.2415	0.2569	0.2749	0.2975
	MAIC	0.1248	0.1335	0.1418	0.1492	0.2509	0.2687	0.2889	0.3028
	MBIC	0.1249	0.1334	0.1411	0.1485	0.2502	0.2667	0.2828	0.3053
	t-sig	0.0884	0.1040	0.1136	0.1260	0.2332	0.2503	0.2667	0.2894
T=200	AIC	0.0346	0.0773	0.1015	0.1291	0.2439	0.2634	0.2845	0.3059
	BIC	0.1165	0.1276	0.1336	0.1433	0.2460	0.2680	0.2999	0.3254
	MAIC	0.1240	0.1325	0.1422	0.1515	0.2536	0.2706	0.2951	0.3123
	MBIC	0.1240	0.1321	0.1410	0.1513	0.2508	0.2729	0.3005	0.3275
	t-sig	0.0878	0.1019	0.1116	0.1279	0.2396	0.2600	0.2832	0.3097
T=∞		0.1098	0.1187	0.1266	0.1365	0.2498	0.2751	0.2991	0.3339

Table 16 (cont'd). Finite sample critical values; Model I, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-31.70	-12.23	-7.48	-4.64	-1.99	-1.79	-1.67	-1.45
	BIC	-4.49	-3.94	-3.66	-3.39	-1.89	-1.74	-1.55	-1.30
	MAIC	-3.76	-3.58	-3.32	-3.17	-1.81	-1.63	-1.45	-1.27
	MBIC	-3.76	-3.58	-3.33	-3.17	-1.83	-1.66	-1.49	-1.26
	t-sig	-8.37	-6.60	-5.38	-4.25	-2.01	-1.81	-1.66	-1.42
T=150	AIC	-15.79	-7.21	-5.31	-4.16	-1.99	-1.80	-1.60	-1.39
	BIC	-4.21	-3.95	-3.69	-3.44	-1.95	-1.76	-1.57	-1.34
	MAIC	-3.90	-3.71	-3.50	-3.29	-1.85	-1.68	-1.51	-1.36
	MBIC	-3.90	-3.72	-3.51	-3.30	-1.86	-1.68	-1.51	-1.33
	t-sig	-5.59	-4.76	-4.37	-3.87	-2.01	-1.84	-1.69	-1.48
T=200	AIC	-14.45	-6.46	-4.90	-3.84	-1.91	-1.74	-1.54	-1.32
	BIC	-4.23	-3.87	-3.71	-3.43	-1.89	-1.70	-1.47	-1.18
	MAIC	-4.00	-3.70	-3.45	-3.26	-1.84	-1.67	-1.46	-1.22
	MBIC	-4.00	-3.74	-3.48	-3.27	-1.85	-1.65	-1.43	-1.18
	t-sig	-5.69	-4.87	-4.43	-3.86	-1.92	-1.77	-1.54	-1.34
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.13	0.88	2.62	6.25	32.81	37.66	41.81	48.70
	BIC	6.94	8.66	9.90	11.81	35.04	40.12	44.29	56.04
	MAIC	9.54	10.60	12.09	13.28	37.71	43.26	49.41	60.08
	MBIC	9.66	10.74	12.01	13.28	37.60	42.12	49.08	58.76
	t-sig	1.89	3.14	4.71	7.42	30.88	37.44	41.31	52.97
T=150	AIC	0.55	2.62	4.84	7.79	31.76	37.68	41.25	50.56
	BIC	7.50	8.65	10.09	11.32	32.85	39.22	43.83	53.07
	MAIC	8.91	9.81	11.08	12.42	36.76	40.88	47.85	53.96
	MBIC	8.91	9.63	11.01	12.21	35.92	40.88	47.10	54.59
	t-sig	4.27	5.89	7.20	8.93	31.11	36.41	41.25	47.75
T=200	AIC	0.63	3.28	5.74	9.20	34.10	39.71	47.70	55.12
	BIC	7.66	9.02	9.81	11.56	34.80	41.13	53.54	62.71
	MAIC	8.33	9.81	11.32	12.72	37.20	41.83	51.94	62.19
	MBIC	8.33	9.72	11.06	12.58	36.57	42.80	53.72	62.78
	t-sig	4.13	5.74	6.95	8.96	32.92	39.75	45.96	56.94
T=∞		6.79	7.85	8.92	10.34	34.97	42.67	50.46	62.11

Table 16 (cont'd). Finite sample critical values; Model I, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-4.97	-4.64	-4.37	-4.04	-2.08	-1.86	-1.73	-1.49
	BIC	-4.95	-4.57	-4.19	-3.85	-2.04	-1.82	-1.60	-1.39
	MAIC	-4.54	-4.19	-3.82	-3.56	-1.88	-1.71	-1.49	-1.28
	MBIC	-4.54	-4.19	-3.84	-3.60	-1.92	-1.73	-1.54	-1.29
	t-sig	-4.91	-4.60	-4.33	-3.99	-2.10	-1.87	-1.72	-1.46
T=150	AIC	-4.78	-4.53	-4.22	-3.89	-2.06	-1.83	-1.64	-1.41
	BIC	-4.67	-4.34	-4.05	-3.74	-2.02	-1.82	-1.60	-1.37
	MAIC	-4.43	-4.07	-3.85	-3.55	-1.88	-1.72	-1.54	-1.39
	MBIC	-4.42	-4.08	-3.87	-3.60	-1.94	-1.75	-1.54	-1.36
	t-sig	-4.70	-4.48	-4.22	-3.87	-2.06	-1.87	-1.70	-1.48
T=200	AIC	-4.82	-4.38	-4.12	-3.80	-1.95	-1.76	-1.58	-1.35
	BIC	-4.67	-4.22	-4.00	-3.65	-1.95	-1.73	-1.50	-1.25
	MAIC	-4.39	-4.00	-3.65	-3.44	-1.86	-1.67	-1.49	-1.25
	MBIC	-4.41	-4.05	-3.68	-3.47	-1.90	-1.69	-1.46	-1.25
	t-sig	-4.79	-4.36	-4.06	-3.73	-1.97	-1.77	-1.55	-1.35
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

Table 17. Finite sample critical values; Model II, choosing T_B minimizing the test statistic ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-9780888.73	-302874.69	-16476.12	-579.76	-13.06	-11.48	-10.31	-8.61
	BIC	-99.97	-49.17	-35.75	-28.47	-10.12	-8.79	-7.63	-6.78
	MAIC	-30.37	-28.38	-26.00	-22.91	-9.76	-8.69	-7.83	-6.87
	MBIC	-29.76	-27.59	-25.71	-22.66	-9.24	-7.95	-7.09	-6.15
	t-sig	-300.91	-166.18	-99.97	-62.22	-12.82	-11.44	-10.05	-8.88
T=150	AIC	-334485.43	-5867.90	-334.99	-104.73	-11.41	-9.70	-8.55	-7.48
	BIC	-41.14	-35.62	-31.07	-26.64	-9.25	-8.24	-7.25	-5.80
	MAIC	-33.81	-30.09	-27.11	-24.28	-9.21	-8.23	-7.38	-5.97
	MBIC	-32.95	-29.20	-27.00	-24.23	-8.84	-7.72	-6.41	-5.28
	t-sig	-81.19	-63.34	-50.53	-40.53	-10.89	-9.64	-8.34	-7.29
T=200	AIC	-2597.71	-463.64	-120.98	-55.65	-10.12	-8.56	-7.48	-5.90
	BIC	-40.85	-33.13	-30.13	-26.71	-8.66	-7.35	-5.67	-4.90
	MAIC	-35.28	-30.50	-27.01	-23.90	-8.75	-7.50	-6.39	-5.66
	MBIC	-33.46	-31.20	-27.54	-24.13	-8.37	-6.95	-5.64	-4.43
	t-sig	-83.71	-63.09	-50.28	-40.12	-10.35	-9.04	-7.45	-5.64
T=∞		-40.88	-35.48	-31.63	-27.46	-7.46	-6.00	-4.97	-4.09

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0002	0.0013	0.0055	0.0294	0.1916	0.2052	0.2145	0.2331
	BIC	0.0707	0.1007	0.1175	0.1310	0.2174	0.2331	0.2463	0.2598
	MAIC	0.1266	0.1321	0.1375	0.1460	0.2208	0.2329	0.2427	0.2552
	MBIC	0.1289	0.1330	0.1387	0.1467	0.2266	0.2422	0.2540	0.2687
	t-sig	0.0406	0.0549	0.0707	0.0894	0.1925	0.2067	0.2202	0.2340
T=150	AIC	0.0012	0.0092	0.0386	0.0691	0.2036	0.2178	0.2328	0.2484
	BIC	0.1102	0.1180	0.1260	0.1358	0.2232	0.2380	0.2512	0.2712
	MAIC	0.1203	0.1277	0.1349	0.1426	0.2244	0.2405	0.2513	0.2712
	MBIC	0.1221	0.1302	0.1356	0.1433	0.2298	0.2460	0.2665	0.2932
	t-sig	0.0785	0.0888	0.0990	0.1103	0.2077	0.2215	0.2354	0.2503
T=200	AIC	0.0139	0.0328	0.0643	0.0940	0.2139	0.2370	0.2531	0.2697
	BIC	0.1102	0.1221	0.1286	0.1359	0.2334	0.2519	0.2762	0.3023
	MAIC	0.1187	0.1272	0.1346	0.1433	0.2334	0.2495	0.2612	0.2890
	MBIC	0.1215	0.1259	0.1333	0.1427	0.2361	0.2564	0.2870	0.3116
	t-sig	0.0773	0.0888	0.0995	0.1109	0.2132	0.2307	0.2497	0.2890
T=∞		0.1096	0.1179	0.1250	0.1338	0.2495	0.2754	0.2985	0.3239

Table 17 (cont'd). Finite sample critical values; Model II, choosing T_B minimizing the test statistic ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-2211.43	-389.14	-90.76	-17.02	-2.52	-2.37	-2.21	-2.03
	BIC	-7.06	-4.95	-4.18	-3.75	-2.20	-2.03	-1.88	-1.73
	MAIC	-3.87	-3.75	-3.57	-3.36	-2.16	-2.03	-1.90	-1.74
	MBIC	-3.85	-3.70	-3.54	-3.34	-2.09	-1.95	-1.79	-1.68
	t-sig	-12.23	-9.11	-7.03	-5.57	-2.49	-2.34	-2.21	-2.06
T=150	AIC	-408.95	-54.16	-12.94	-7.22	-2.35	-2.14	-1.99	-1.87
	BIC	-4.53	-4.18	-3.88	-3.62	-2.10	-1.95	-1.84	-1.65
	MAIC	-1.09	-3.83	-3.67	-3.47	-2.10	-1.94	-1.84	-1.71
	MBIC	-4.03	-3.79	-3.63	-3.46	-2.04	-1.88	-1.74	-1.54
	t-sig	-6.37	-5.57	-5.00	-4.47	-2.28	-2.13	-1.95	-1.85
T=200	AIC	-36.03	-15.22	-7.76	-5.18	-2.17	-2.01	-1.88	-1.67
	BIC	-4.50	-4.06	-3.85	-3.63	-2.01	-1.83	-1.64	-1.52
	MAIC	-4.18	-3.88	-3.64	-3.42	-2.04	-1.88	-1.73	-1.63
	MBIC	-4.08	-3.92	-3.66	-3.44	-1.97	-1.80	-1.62	-1.47
	t-sig	-6.46	-5.60	-5.00	-4.45	-2.23	-2.05	-1.85	-1.62
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.00	0.00	0.01	0.47	22.27	25.41	28.53	34.88
	BIC	2.62	5.43	7.68	9.61	29.05	33.65	38.02	42.96
	MAIC	8.95	9.53	10.55	12.00	29.39	33.41	36.80	42.54
	MBIC	9.06	9.91	10.86	12.11	32.03	36.65	41.08	46.60
	t-sig	0.91	1.69	2.84	4.45	22.81	25.72	29.46	34.43
T=150	AIC	0.00	0.04	0.79	2.73	24.67	28.57	33.12	37.28
	BIC	6.36	7.62	8.95	10.31	29.48	34.42	38.53	45.39
	MAIC	8.04	9.06	10.04	11.18	30.11	34.70	39.61	44.64
	MBIC	8.24	9.29	10.11	11.25	31.89	36.41	41.52	50.95
	t-sig	3.35	4.50	5.50	6.79	25.87	29.19	33.44	38.39
T=200	AIC	0.10	0.58	2.32	4.97	27.72	32.99	38.06	48.06
	BIC	6.63	8.12	9.08	10.26	31.69	38.02	48.06	57.38
	MAIC	7.76	8.87	10.19	11.48	31.51	36.73	41.86	51.61
	MBIC	8.00	8.83	10.06	11.38	32.67	38.90	48.91	59.19
	t-sig	3.27	4.38	5.42	6.90	26.79	31.23	37.87	48.91
T=∞		6.59	7.69	8.52	9.82	34.65	42.57	49.76	58.73

Table 17 (cont'd). Finite sample critical values; Model II, choosing T_B minimizing the test statistic ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-5.30	-4.98	-4.74	-4.43	-2.60	-2.47	-2.35	-2.17
	BIC	-5.07	-4.80	-4.49	-4.23	-2.38	-2.18	-2.03	-1.86
	MAIC	-4.69	-4.43	-4.16	-3.85	-2.27	-2.12	-1.98	-1.85
	MBIC	-4.73	-4.42	-4.13	-3.86	-2.24	-2.07	-1.90	-1.74
	t-sig	-5.18	-4.89	-4.64	-4.36	-2.54	-2.39	-2.22	-2.05
T=150	AIC	-5.03	-4.82	-4.56	-4.23	-2.38	-2.20	-2.08	-1.93
	BIC	-4.82	-4.48	-4.27	-3.99	-2.19	-2.06	-1.91	-1.73
	MAIC	-4.62	-4.28	-4.02	-3.77	-2.16	-2.00	-1.89	-1.76
	MBIC	-4.62	-4.27	-4.03	-3.80	-2.13	-1.96	-1.83	-1.60
	t-sig	-4.98	-4.69	-4.46	-4.14	-2.32	-2.18	-2.02	-1.86
T=200	AIC	-5.03	-4.73	-4.39	-4.10	-2.20	-2.01	-1.88	-1.75
	BIC	-4.89	-4.44	-4.20	-3.85	-2.08	-1.88	-1.69	-1.56
	MAIC	-4.48	-4.22	-3.89	-3.60	-2.06	-1.89	-1.79	-1.65
	MBIC	-4.50	-4.28	-3.96	-3.65	-2.04	-1.86	-1.67	-1.51
	t-sig	-5.00	-4.51	-4.32	-4.09	-2.23	-2.05	-1.87	-1.68
T=∞		-4.49	-4.18	-3.96	-3.68	-1.87	-1.66	-1.51	-1.35

Table 18. Finite sample critical values; Model II, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

a) MZ_{α}^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-2082.70	-432.85	-115.26	-45.03	-8.50	-7.17	-6.41	-5.17
	BIC	-37.23	-31.46	-27.21	-23.93	-7.92	-6.71	-5.99	-5.02
	MAIC	-27.59	-25.04	-22.53	-20.48	-7.14	-6.24	-5.17	-4.54
	MBIC	-27.05	-24.81	-22.53	-20.48	-7.14	-6.38	-5.30	-4.54
	t-sig	-131.79	-87.40	-57.17	-36.56	-8.64	-7.21	-6.64	-5.63
T=150	AIC	-302.58	-103.84	-63.95	-34.53	-8.50	-7.39	-6.19	-5.11
	BIC	-34.19	-30.14	-26.96	-23.40	-8.12	-6.73	-5.85	-4.91
	MAIC	-30.14	-27.10	-24.19	-21.27	-7.46	-6.22	-5.45	-4.40
	MBIC	-30.14	-27.11	-24.19	-21.51	-7.61	-6.24	-5.45	-4.40
	t-sig	-58.00	-44.04	-38.32	-29.80	-8.58	-7.46	-6.41	-5.11
T=200	AIC	-296.21	-75.99	-44.67	-28.72	-7.96	-6.72	-5.58	-4.52
	BIC	-33.23	-29.88	-26.95	-23.24	-7.75	-6.40	-4.93	-4.15
	MAIC	-31.82	-26.54	-23.98	-21.14	-7.30	-6.06	-5.00	-3.66
	MBIC	-31.82	-27.00	-24.08	-21.23	-7.38	-5.92	-4.93	-3.66
	t-sig	-59.96	-49.03	-39.19	-29.65	-8.26	-6.78	-5.58	-4.26
T=∞		-41.00	-34.96	-30.74	-26.40	-7.45	-5.95	-4.88	-3.81

b) MSB^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.0155	0.0340	0.0656	0.1053	0.2363	0.2552	0.2656	0.2755
	BIC	0.1154	0.1257	0.1343	0.1437	0.2455	0.2627	0.2739	0.2973
	MAIC	0.1339	0.1405	0.1464	0.1548	0.2560	0.2707	0.2866	0.3085
	MBIC	0.1346	0.1405	0.1479	0.1548	0.2554	0.2699	0.2851	0.3085
	t-sig	0.0608	0.0756	0.0933	0.1167	0.2332	0.2533	0.2666	0.2810
T=150	AIC	0.0406	0.0693	0.0882	0.1201	0.2345	0.2511	0.2709	0.2868
	BIC	0.1208	0.1282	0.1356	0.1451	0.2420	0.2592	0.2763	0.2952
	MAIC	0.1285	0.1356	0.1431	0.1512	0.2514	0.2694	0.2840	0.3025
	MBIC	0.1285	0.1354	0.1431	0.1505	0.2492	0.2683	0.2866	0.3025
	t-sig	0.0926	0.1064	0.1141	0.1287	0.2322	0.2511	0.2664	0.2866
T=200	AIC	0.0411	0.0811	0.1042	0.1309	0.2431	0.2644	0.2849	0.3039
	BIC	0.1209	0.1288	0.1348	0.1450	0.2471	0.2679	0.2983	0.3262
	MAIC	0.1253	0.1361	0.1436	0.1526	0.2560	0.2716	0.2940	0.3169
	MBIC	0.1253	0.1334	0.1436	0.1519	0.2540	0.2716	0.3015	0.3287
	t-sig	0.0907	0.1006	0.1123	0.1295	0.2387	0.2604	0.2855	0.3084
T=∞		0.1098	0.1187	0.1266	0.1365	0.2498	0.2751	0.2991	0.3339

Table 18 (cont'd). Finite sample critical values; Model II, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

c) MZ_t^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-32.26	-14.70	-7.56	-4.74	-1.99	-1.83	-1.68	-1.37
	BIC	-4.30	-3.94	-3.67	-3.40	-1.92	-1.76	-1.62	-1.43
	MAIC	-3.71	-3.51	-3.34	-3.17	-1.83	-1.66	-1.50	-1.31
	MBIC	-3.67	-3.51	-3.33	-3.17	-1.84	-1.68	-1.50	-1.31
	t-sig	-8.01	-6.60	-5.25	-4.24	-2.00	-1.85	-1.72	-1.55
T=150	AIC	-12.29	-7.20	-5.63	-4.12	-2.01	-1.81	-1.65	-1.35
	BIC	-4.12	-3.82	-3.65	-3.40	-1.95	-1.74	-1.56	-1.35
	MAIC	-3.86	-3.67	-3.45	-3.23	-1.83	-1.67	-1.51	-1.35
	MBIC	-3.86	-3.67	-3.46	-3.25	-1.85	-1.69	-1.52	-1.28
	t-sig	-5.38	-4.68	-4.36	-3.83	-2.01	-1.82	-1.67	-1.43
T=200	AIC	-12.16	-6.16	-4.70	-3.77	-1.91	-1.74	-1.58	-1.34
	BIC	-4.05	-3.85	-3.64	-3.38	-1.89	-1.68	-1.47	-1.19
	MAIC	-3.98	-3.62	-3.43	-3.24	-1.84	-1.68	-1.47	-1.23
	MBIC	-3.98	-3.64	-3.45	-3.24	-1.85	-1.65	-1.45	-1.19
	t-sig	-5.47	-4.94	-4.41	-3.83	-1.94	-1.74	-1.54	-1.38
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

d) P_T^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	0.13	0.65	2.47	6.09	32.40	37.78	41.26	46.76
	BIC	7.23	8.61	9.91	11.78	35.07	38.83	42.94	51.90
	MAIC	9.88	10.91	12.17	13.39	37.67	42.94	49.52	57.80
	MBIC	9.91	10.94	12.13	13.39	37.67	42.50	48.55	57.88
	t-sig	2.09	3.16	4.89	7.52	31.49	37.07	40.85	48.71
T=150	AIC	0.89	2.59	4.55	7.87	31.43	37.50	41.77	50.91
	BIC	7.89	9.32	10.31	11.73	33.47	39.14	44.29	52.97
	MAIC	9.26	10.10	11.11	13.00	36.24	41.96	47.62	52.97
	MBIC	9.26	10.04	11.11	12.68	35.46	41.65	47.31	53.88
	t-sig	4.78	6.36	7.20	9.22	31.02	36.85	41.26	47.62
T=200	AIC	0.90	3.49	6.13	9.46	34.18	39.55	47.56	55.69
	BIC	8.18	9.23	10.11	11.92	35.15	41.46	54.12	62.59
	MAIC	8.27	10.27	11.38	12.93	37.49	41.64	52.13	62.59
	MBIC	8.27	10.21	11.36	12.87	36.75	42.31	54.25	63.31
	t-sig	4.64	5.59	7.05	9.31	32.72	39.82	46.46	57.37
T=∞		6.79	7.85	8.92	10.34	34.97	42.67	50.46	62.11

Table 18 (cont'd). Finite sample critical values; Model II, choosing T_B maximizing $|t_{\widehat{\beta}_2}|$ ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

e) ADF^{GLS}

	Criteria	1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
T=100	AIC	-4.89	-4.49	-4.31	-4.06	-2.08	-1.91	-1.75	-1.45
	BIC	-4.89	-4.42	-4.17	-3.84	-2.04	-1.84	-1.71	-1.49
	MAIC	-4.35	-4.12	-3.83	-3.57	-1.90	-1.72	-1.56	-1.39
	MBIC	-4.35	-4.12	-3.83	-3.58	-1.94	-1.77	-1.60	-1.37
	t-sig	-4.90	-4.49	-4.26	-4.02	-2.08	-1.91	-1.78	-1.54
T=150	AIC	-4.66	-4.41	-4.12	-3.85	-2.07	-1.87	-1.69	-1.38
	BIC	-4.59	-4.23	-4.03	-3.71	-2.02	-1.82	-1.59	-1.38
	MAIC	-4.31	-4.07	-3.77	-3.49	-1.86	-1.72	-1.55	-1.38
	MBIC	-4.31	-4.08	-3.79	-3.53	-1.91	-1.76	-1.55	-1.34
	t-sig	-4.59	-4.43	-4.10	-3.83	-2.05	-1.84	-1.69	-1.46
T=200	AIC	-4.73	-4.33	-4.10	-3.72	-1.96	-1.76	-1.59	-1.38
	BIC	-4.46	-4.19	-3.93	-3.61	-1.94	-1.75	-1.55	-1.26
	MAIC	-4.32	-3.87	-3.61	-3.28	-1.87	-1.69	-1.50	-1.24
	MBIC	-4.37	-3.94	-3.65	-3.41	-1.90	-1.69	-1.50	-1.26
	t-sig	-4.69	-4.24	-4.00	-3.71	-1.99	-1.78	-1.56	-1.38
T=∞		-4.50	-4.16	-3.89	-3.61	-1.86	-1.64	-1.44	-1.23

Table 19: Size and Power; choosing T_B minimizing the tests; Model I; T=100
($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.051	0.051	0.051	0.051	0.050	0.055	0.055	0.055	0.055	0.384
	BIC	0.050	0.050	0.051	0.051	0.051	0.342	0.344	0.342	0.360	0.497
	MAIC	0.051	0.050	0.051	0.051	0.050	0.483	0.492	0.500	0.497	0.487
	MBIC	0.051	0.050	0.050	0.051	0.050	0.531	0.517	0.526	0.517	0.503
	t-sig	0.051	0.051	0.051	0.050	0.051	0.134	0.135	0.133	0.138	0.464
$\theta = -0.8$	AIC	0.044	0.044	0.044	0.044	0.771	0.123	0.123	0.123	0.124	0.960
	BIC	0.922	0.923	0.922	0.921	0.973	1.000	1.000	1.000	1.000	1.000
	MAIC	0.337	0.338	0.343	0.342	0.353	0.841	0.843	0.843	0.841	0.871
	MBIC	0.353	0.345	0.348	0.345	0.353	0.848	0.848	0.848	0.845	0.874
	t-sig	0.078	0.079	0.078	0.081	0.749	0.368	0.369	0.374	0.392	0.973
$\theta = -0.4$	AIC	0.045	0.045	0.045	0.045	0.208	0.055	0.055	0.055	0.056	0.700
	BIC	0.352	0.360	0.352	0.337	0.429	0.863	0.861	0.862	0.866	0.904
	MAIC	0.136	0.139	0.143	0.142	0.120	0.515	0.520	0.524	0.525	0.475
	MBIC	0.148	0.140	0.145	0.139	0.118	0.550	0.535	0.544	0.542	0.491
	t-sig	0.027	0.027	0.027	0.027	0.227	0.109	0.110	0.109	0.110	0.729
$\theta = 0.4$	AIC	0.061	0.061	0.061	0.062	0.060	0.062	0.062	0.062	0.063	0.326
	BIC	0.219	0.224	0.222	0.205	0.077	0.703	0.706	0.703	0.706	0.506
	MAIC	0.118	0.119	0.125	0.114	0.018	0.188	0.192	0.206	0.203	0.107
	MBIC	0.003	0.003	0.003	0.003	0.003	0.089	0.080	0.094	0.094	0.100
	t-sig	0.050	0.051	0.050	0.050	0.059	0.159	0.159	0.159	0.162	0.398
$\theta = 0.8$	AIC	0.155	0.155	0.155	0.158	0.057	0.189	0.189	0.189	0.192	0.168
	BIC	0.507	0.511	0.503	0.488	0.109	0.715	0.716	0.713	0.716	0.383
	MAIC	0.221	0.226	0.227	0.207	0.012	0.381	0.383	0.395	0.398	0.103
	MBIC	0.013	0.012	0.012	0.010	0.001	0.049	0.040	0.047	0.048	0.031
	t-sig	0.072	0.071	0.070	0.068	0.050	0.231	0.232	0.230	0.234	0.227
$\rho = -0.8$	AIC	0.053	0.053	0.053	0.054	0.048	0.076	0.076	0.076	0.078	0.361
	BIC	0.015	0.015	0.016	0.016	0.044	0.048	0.047	0.046	0.047	0.462
	MAIC	0.001	0.001	0.002	0.002	0.047	0.013	0.014	0.016	0.019	0.410
	MBIC	0.001	0.000	0.001	0.001	0.047	0.017	0.015	0.016	0.018	0.406
	t-sig	0.016	0.017	0.017	0.017	0.044	0.064	0.063	0.064	0.068	0.432
$\rho = -0.4$	AIC	0.049	0.049	0.049	0.050	0.070	0.062	0.062	0.062	0.063	0.389
	BIC	0.130	0.131	0.127	0.118	0.140	0.476	0.477	0.477	0.481	0.567
	MAIC	0.061	0.060	0.067	0.063	0.062	0.367	0.369	0.391	0.386	0.417
	MBIC	0.068	0.062	0.067	0.064	0.059	0.405	0.382	0.401	0.399	0.421
	t-sig	0.034	0.034	0.034	0.033	0.071	0.128	0.127	0.129	0.130	0.459
$\rho = 0.4$	AIC	0.065	0.065	0.065	0.065	0.047	0.061	0.061	0.061	0.063	0.221
	BIC	0.157	0.159	0.157	0.144	0.042	0.502	0.507	0.506	0.502	0.291
	MAIC	0.129	0.142	0.129	0.120	0.021	0.112	0.111	0.119	0.118	0.050
	MBIC	0.003	0.003	0.003	0.002	0.001	0.030	0.026	0.032	0.028	0.036
	t-sig	0.056	0.056	0.057	0.055	0.045	0.152	0.152	0.152	0.150	0.292
$\rho = 0.8$	AIC	0.161	0.161	0.161	0.163	0.083	0.070	0.070	0.070	0.070	0.114
	BIC	0.365	0.381	0.353	0.334	0.067	0.354	0.364	0.350	0.340	0.118
	MAIC	0.432	0.455	0.429	0.400	0.074	0.409	0.421	0.414	0.396	0.145
	MBIC	0.356	0.362	0.343	0.305	0.064	0.222	0.220	0.218	0.196	0.031
	t-sig	0.158	0.159	0.159	0.152	0.072	0.111	0.111	0.111	0.114	0.134

Table 20. Size and Power; choosing T_B minimizing the tests; Model 1; T=200
($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.051	0.050	0.050	0.051	0.051	0.107	0.107	0.109	0.107	0.460
	BIC	0.051	0.050	0.050	0.051	0.051	0.537	0.522	0.543	0.548	0.535
	MAIC	0.050	0.051	0.051	0.050	0.050	0.500	0.507	0.501	0.539	0.509
	MBIC	0.051	0.050	0.050	0.050	0.051	0.561	0.535	0.557	0.565	0.531
	t-sig	0.050	0.050	0.050	0.050	0.051	0.254	0.255	0.251	0.265	0.468
$\theta = -0.8$	AIC	0.204	0.200	0.205	0.202	0.539	0.742	0.738	0.745	0.739	0.942
	BIC	0.809	0.804	0.811	0.806	0.873	0.993	0.993	0.993	0.993	0.999
	MAIC	0.112	0.111	0.113	0.116	0.206	0.505	0.507	0.506	0.516	0.668
	MBIC	0.117	0.113	0.116	0.116	0.197	0.521	0.514	0.519	0.527	0.670
	t-sig	0.299	0.301	0.296	0.305	0.566	0.789	0.789	0.787	0.789	0.975
$\theta = -0.4$	AIC	0.042	0.043	0.042	0.038	0.119	0.129	0.125	0.131	0.130	0.641
	BIC	0.239	0.229	0.238	0.242	0.228	0.842	0.831	0.842	0.849	0.848
	MAIC	0.083	0.081	0.083	0.085	0.074	0.475	0.480	0.480	0.499	0.447
	MBIC	0.089	0.086	0.085	0.085	0.072	0.523	0.512	0.520	0.519	0.464
	t-sig	0.035	0.034	0.035	0.037	0.119	0.312	0.319	0.309	0.328	0.647
$\theta = 0.4$	AIC	0.068	0.066	0.067	0.064	0.058	0.164	0.158	0.164	0.158	0.409
	BIC	0.152	0.151	0.153	0.151	0.096	0.723	0.709	0.722	0.727	0.590
	MAIC	0.100	0.108	0.095	0.096	0.048	0.516	0.520	0.518	0.535	0.367
	MBIC	0.012	0.012	0.011	0.011	0.008	0.114	0.106	0.113	0.116	0.090
	t-sig	0.070	0.070	0.069	0.066	0.052	0.346	0.347	0.344	0.357	0.426
$\theta = 0.8$	AIC	0.175	0.174	0.173	0.164	0.059	0.395	0.385	0.397	0.392	0.251
	BIC	0.317	0.314	0.314	0.305	0.083	0.727	0.718	0.732	0.731	0.424
	MAIC	0.199	0.205	0.194	0.193	0.024	0.481	0.484	0.482	0.499	0.185
	MBIC	0.038	0.040	0.036	0.038	0.017	0.316	0.306	0.311	0.314	0.167
	t-sig	0.156	0.160	0.154	0.147	0.066	0.533	0.534	0.534	0.538	0.335
$\rho = -0.8$	AIC	0.049	0.049	0.049	0.046	0.056	0.083	0.082	0.083	0.083	0.440
	BIC	0.006	0.007	0.007	0.007	0.047	0.058	0.053	0.059	0.065	0.516
	MAIC	0.003	0.003	0.003	0.002	0.047	0.039	0.040	0.040	0.048	0.452
	MBIC	0.001	0.001	0.001	0.001	0.044	0.046	0.042	0.048	0.048	0.448
	t-sig	0.016	0.016	0.015	0.014	0.047	0.069	0.069	0.067	0.073	0.466
$\rho = -0.4$	AIC	0.052	0.049	0.052	0.050	0.053	0.122	0.122	0.123	0.125	0.441
	BIC	0.057	0.056	0.056	0.057	0.050	0.491	0.473	0.500	0.512	0.515
	MAIC	0.043	0.046	0.041	0.042	0.048	0.428	0.429	0.429	0.460	0.476
	MBIC	0.049	0.046	0.044	0.042	0.048	0.469	0.448	0.461	0.476	0.470
	t-sig	0.037	0.037	0.035	0.037	0.043	0.217	0.217	0.216	0.223	0.456
$\rho = 0.4$	AIC	0.068	0.066	0.069	0.065	0.050	0.129	0.129	0.129	0.126	0.381
	BIC	0.106	0.108	0.106	0.101	0.044	0.614	0.588	0.617	0.613	0.454
	MAIC	0.097	0.099	0.096	0.097	0.056	0.506	0.506	0.506	0.527	0.381
	MBIC	0.027	0.024	0.023	0.022	0.005	0.024	0.021	0.021	0.023	0.018
	t-sig	0.071	0.073	0.070	0.066	0.046	0.282	0.285	0.282	0.292	0.396
$\rho = 0.8$	AIC	0.128	0.127	0.130	0.119	0.068	0.137	0.135	0.137	0.131	0.223
	BIC	0.171	0.168	0.169	0.150	0.064	0.428	0.418	0.427	0.426	0.248
	MAIC	0.200	0.215	0.195	0.189	0.072	0.410	0.416	0.410	0.419	0.247
	MBIC	0.177	0.185	0.166	0.153	0.070	0.460	0.440	0.446	0.449	0.245
	t-sig	0.097	0.101	0.095	0.093	0.061	0.228	0.232	0.227	0.227	0.232

Table 21. Size and Power; choosing T_B maximizing $|t_{\widehat{\beta}_2}|$; Model 1; T=100 ($\bar{c} = -23$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.051	0.050	0.050	0.050	0.050	0.072	0.072	0.074	0.079	0.388
	BIC	0.051	0.051	0.050	0.051	0.050	0.397	0.396	0.403	0.389	0.495
	MAIC	0.047	0.050	0.051	0.050	0.050	0.436	0.422	0.460	0.454	0.414
	MBIC	0.051	0.050	0.050	0.050	0.050	0.456	0.432	0.469	0.460	0.417
	t-sig	0.051	0.049	0.051	0.051	0.050	0.129	0.128	0.130	0.131	0.428
$\theta = -0.8$	AIC	0.024	0.024	0.024	0.026	0.628	0.075	0.075	0.075	0.078	0.874
	BIC	0.842	0.843	0.839	0.830	0.892	0.996	0.996	0.996	0.994	0.998
	MAIC	0.162	0.161	0.167	0.166	0.164	0.542	0.542	0.542	0.541	0.554
	MBIC	0.164	0.161	0.167	0.164	0.160	0.540	0.540	0.539	0.538	0.553
	t-sig	0.050	0.049	0.050	0.050	0.575	0.251	0.251	0.251	0.252	0.913
$\theta = -0.4$	AIC	0.040	0.040	0.039	0.044	0.195	0.055	0.055	0.055	0.056	0.609
	BIC	0.372	0.376	0.373	0.347	0.411	0.814	0.816	0.816	0.808	0.860
	MAIC	0.108	0.108	0.119	0.123	0.091	0.353	0.349	0.363	0.359	0.310
	MBIC	0.120	0.112	0.124	0.126	0.090	0.377	0.363	0.377	0.374	0.313
	t-sig	0.024	0.024	0.024	0.025	0.194	0.078	0.077	0.078	0.077	0.614
$\theta = 0.4$	AIC	0.063	0.062	0.063	0.067	0.057	0.101	0.099	0.101	0.107	0.307
	BIC	0.213	0.221	0.209	0.199	0.087	0.666	0.662	0.673	0.660	0.486
	MAIC	0.059	0.063	0.063	0.065	0.012	0.125	0.121	0.142	0.154	0.091
	MBIC	0.004	0.003	0.004	0.004	0.004	0.091	0.078	0.095	0.103	0.091
	t-sig	0.051	0.050	0.052	0.052	0.052	0.170	0.169	0.169	0.175	0.359
$\theta = 0.8$	AIC	0.200	0.199	0.200	0.208	0.050	0.299	0.298	0.299	0.324	0.145
	BIC	0.385	0.392	0.383	0.368	0.088	0.638	0.636	0.639	0.637	0.328
	MAIC	0.099	0.100	0.103	0.104	0.009	0.260	0.252	0.283	0.281	0.071
	MBIC	0.005	0.003	0.006	0.006	0.001	0.032	0.025	0.035	0.029	0.028
	t-sig	0.077	0.076	0.076	0.077	0.044	0.225	0.224	0.228	0.231	0.183
$\rho = -0.8$	AIC	0.046	0.046	0.047	0.053	0.036	0.039	0.038	0.039	0.039	0.254
	BIC	0.006	0.006	0.006	0.006	0.031	0.024	0.025	0.024	0.023	0.364
	MAIC	0.001	0.001	0.001	0.001	0.030	0.010	0.009	0.012	0.013	0.232
	MBIC	0.001	0.001	0.001	0.001	0.031	0.011	0.009	0.012	0.013	0.231
	t-sig	0.015	0.015	0.015	0.015	0.032	0.040	0.039	0.040	0.043	0.304
$\rho = -0.4$	AIC	0.056	0.056	0.056	0.059	0.064	0.076	0.076	0.076	0.078	0.333
	BIC	0.125	0.122	0.127	0.119	0.135	0.427	0.426	0.431	0.431	0.513
	MAIC	0.054	0.054	0.057	0.054	0.047	0.280	0.269	0.296	0.296	0.297
	MBIC	0.058	0.052	0.058	0.053	0.043	0.294	0.272	0.301	0.301	0.295
	t-sig	0.036	0.036	0.034	0.036	0.060	0.096	0.094	0.097	0.097	0.376
$\rho = 0.4$	AIC	0.068	0.068	0.068	0.069	0.043	0.103	0.103	0.103	0.106	0.219
	BIC	0.144	0.147	0.141	0.130	0.041	0.508	0.506	0.516	0.495	0.304
	MAIC	0.077	0.082	0.082	0.081	0.014	0.073	0.072	0.077	0.079	0.042
	MBIC	0.002	0.003	0.002	0.001	0.001	0.034	0.030	0.034	0.033	0.034
	t-sig	0.052	0.053	0.051	0.053	0.038	0.137	0.137	0.136	0.136	0.251
$\rho = 0.8$	AIC	0.073	0.072	0.074	0.078	0.045	0.083	0.084	0.083	0.089	0.099
	BIC	0.204	0.213	0.197	0.187	0.038	0.331	0.336	0.331	0.315	0.116
	MAIC	0.213	0.243	0.210	0.207	0.051	0.330	0.334	0.339	0.341	0.127
	MBIC	0.209	0.223	0.196	0.186	0.038	0.115	0.106	0.114	0.109	0.017
	t-sig	0.092	0.092	0.092	0.089	0.041	0.106	0.105	0.107	0.108	0.109

Table 22. Size and Power; choosing T_B maximizing $|t_{\widehat{\beta}_2}|$; Model 1; T=200 ($\overline{c} = -23$ when constructing s^2 for M^{GLS} tests)

		Size					Power				
	Criteria	MZ_α	MSB	MZ_t	P_T	ADF	MZ_α	MSB	MZ_t	P_T	ADF
i.i.d.	AIC	0.050	0.050	0.051	0.050	0.051	0.097	0.097	0.097	0.099	0.395
	BIC	0.051	0.050	0.050	0.051	0.050	0.481	0.455	0.470	0.458	0.480
	MAIC	0.050	0.050	0.050	0.051	0.051	0.467	0.467	0.490	0.481	0.460
	MBIC	0.051	0.051	0.050	0.050	0.051	0.498	0.477	0.507	0.490	0.496
	t-sig	0.051	0.051	0.050	0.050	0.050	0.220	0.212	0.219	0.218	0.428
$\theta = -0.8$	AIC	0.154	0.154	0.152	0.159	0.302	0.563	0.563	0.560	0.560	0.741
	BIC	0.610	0.607	0.606	0.591	0.675	0.952	0.953	0.951	0.937	0.977
	MAIC	0.036	0.036	0.038	0.039	0.068	0.196	0.195	0.198	0.200	0.287
	MBIC	0.037	0.035	0.036	0.037	0.071	0.198	0.188	0.196	0.193	0.286
	t-sig	0.172	0.172	0.172	0.171	0.355	0.579	0.579	0.579	0.563	0.865
$\theta = -0.4$	AIC	0.034	0.035	0.034	0.034	0.102	0.154	0.153	0.151	0.153	0.500
	BIC	0.217	0.213	0.214	0.218	0.212	0.757	0.754	0.752	0.734	0.757
	MAIC	0.074	0.074	0.081	0.077	0.066	0.349	0.349	0.360	0.363	0.322
	MBIC	0.076	0.069	0.074	0.073	0.064	0.369	0.352	0.373	0.361	0.336
	t-sig	0.034	0.033	0.035	0.036	0.109	0.256	0.254	0.262	0.259	0.538
$\theta = 0.4$	AIC	0.066	0.067	0.064	0.065	0.049	0.211	0.207	0.208	0.215	0.345
	BIC	0.134	0.136	0.132	0.132	0.084	0.664	0.653	0.658	0.653	0.539
	MAIC	0.093	0.096	0.091	0.093	0.057	0.447	0.446	0.463	0.457	0.313
	MBIC	0.012	0.010	0.012	0.010	0.006	0.096	0.080	0.096	0.092	0.092
	t-sig	0.063	0.065	0.063	0.062	0.054	0.325	0.320	0.327	0.326	0.405
$\theta = 0.8$	AIC	0.157	0.156	0.153	0.158	0.047	0.401	0.402	0.401	0.410	0.189
	BIC	0.252	0.257	0.245	0.245	0.072	0.653	0.643	0.650	0.639	0.356
	MAIC	0.131	0.136	0.134	0.136	0.021	0.382	0.380	0.405	0.409	0.149
	MBIC	0.034	0.032	0.033	0.031	0.017	0.254	0.236	0.258	0.251	0.153
	t-sig	0.145	0.146	0.145	0.144	0.067	0.491	0.484	0.495	0.494	0.329
$\rho = -0.8$	AIC	0.028	0.028	0.028	0.028	0.040	0.055	0.055	0.055	0.053	0.280
	BIC	0.004	0.005	0.004	0.004	0.036	0.037	0.036	0.036	0.035	0.345
	MAIC	0.000	0.000	0.000	0.001	0.035	0.028	0.028	0.032	0.029	0.289
	MBIC	0.000	0.000	0.000	0.001	0.035	0.028	0.028	0.028	0.027	0.281
	t-sig	0.012	0.013	0.012	0.011	0.041	0.046	0.046	0.047	0.047	0.325
$\rho = -0.4$	AIC	0.051	0.051	0.051	0.051	0.048	0.103	0.102	0.102	0.104	0.343
	BIC	0.046	0.045	0.042	0.041	0.044	0.398	0.391	0.389	0.383	0.410
	MAIC	0.044	0.044	0.049	0.047	0.047	0.339	0.347	0.356	0.375	0.379
	MBIC	0.046	0.041	0.045	0.043	0.049	0.361	0.336	0.362	0.349	0.383
	t-sig	0.030	0.029	0.030	0.032	0.049	0.168	0.164	0.169	0.166	0.371
$\rho = 0.4$	AIC	0.060	0.060	0.060	0.060	0.040	0.166	0.165	0.164	0.171	0.345
	BIC	0.094	0.099	0.089	0.089	0.042	0.571	0.557	0.562	0.561	0.403
	MAIC	0.092	0.096	0.096	0.095	0.058	0.448	0.451	0.460	0.466	0.336
	MBIC	0.013	0.012	0.012	0.012	0.005	0.019	0.014	0.020	0.020	0.018
	t-sig	0.060	0.059	0.059	0.058	0.044	0.268	0.265	0.272	0.277	0.382
$\rho = 0.8$	AIC	0.066	0.065	0.066	0.068	0.047	0.123	0.124	0.123	0.125	0.187
	BIC	0.111	0.118	0.105	0.104	0.045	0.376	0.369	0.366	0.368	0.219
	MAIC	0.122	0.135	0.121	0.127	0.057	0.380	0.373	0.388	0.389	0.246
	MBIC	0.130	0.131	0.122	0.120	0.054	0.409	0.384	0.409	0.396	0.247
	t-sig	0.069	0.069	0.065	0.063	0.046	0.198	0.196	0.201	0.203	0.239

Residual Based Tests for Cointegration with GLS Detrended Data

Abstract

We propose residual based tests for cointegration using local *GLS* detrending (Elliott, Rothemberg and Stock (1996), ERS) to eliminate separately the deterministic components in the series. We consider two cases, one where only a constant is included and one with a constant and a time trend are included. The limiting distributions of various residuals based tests are derived for a general quasi-differencing parameter $\overline{c}$ and critical values are tabulated for values of $\overline{c} = 0$ irrespective of the nature of the deterministic components and the values suggested by ERS. Simulations show that using *GLS* detrending allows tests with higher power and that using $\overline{c} = -7.0$ or $\overline{c} = -13.5$, as the quasi-differencing parameter, according to the two cases analyzed, is preferable.

1 Introduction

Even though they are applicable only under some specific conditions, residuals based tests for cointegration, developed by Phillips and Ouliaris (1990), have been quite popular in applied work mostly because of their computational simplicity. The statistics introduced are designed to test the null hypothesis of no cointegration in a single equation setting assuming that the variables introduced as regressors are not cointegrated. These tests also have some appeal because they follow quite intuitively from the basic definition of cointegration as laid out in Engle and Granger (1987). If the system of variables is cointegrated, then there exists a linear combination (given by the cointegrating vector) that is stationary. In this case, the residuals from a simple static regression are stationary and, as shown by Stock (1987), this regression estimated by *OLS* will provide a consistent estimate of the cointegrating vector. In the absence of cointegration, the residuals from the static regression are nonstationary for any choice of the parameter vector and we have what has been labelled, following Granger and Newbold (1974) and later Phillips (1986), a spurious regression. Hence, an obvious testing strategy is to test the null hypothesis of no cointegration using some unit root test on the estimated residuals from the simple static regression.

Of course, there are many alternative approaches available, some applicable under less restrictive conditions; for example, the system based tests of Johansen (1991) and Stock and Watson (1988). The reader is referred to one of the many available surveys; for example, Watson (1994), Perron and Campbell (1992) and Banerjee, Dolado, Galbraith and Hendry (1993).

In an important paper, Elliott, Rothemberg and Stock (1996, hereafter ERS) show that several unit root tests constructed using *GLS* or quasi-differenced data have asymptotic power functions close to the Gaussian local asymptotic power envelope. Hence, they enjoy some optimal properties over tests constructed using *OLS* detrended data and the simulations in ERS showed substantial power gains in finite samples. If such a detrending device is beneficial for unit root tests, it is natural to think that it would also be for cointegration tests.

Our aim, accordingly, is to analyze residual based tests for cointegration when they are constructed using *GLS* detrended or quasi-differenced data. We consider the standard *ADF* test and the class of modified unit root tests analyzed in Stock (1990) and Ng and Perron (1999). We derive their asymptotic distribution assuming a general quasi-differencing parameter $\bar{c}$ and tabulate critical values for two choices: a) $\bar{c} = 0$ irrespective of the nature of the deterministic components; b) $\bar{c} = -7.0$ for the constant only case and $\bar{c} = -13.5$ for the linear trend case, as suggested by ERS in the context of unit root tests. Our simulation results about power reveal that important power gains can indeed be achieved by using *GLS* detrended data, especially if the quasi-difference parameter is set as suggested by ERS.

This paper is organized as follows. Section 2 presents the data-generating process considered and offers preliminary results concerning the limit of the estimate of the coefficients of the trend function and the estimate from the static regression with *GLS* detrended data. Section 3 presents the tests considered and derives their asymptotic distributions. Tabulated asymptotic and finite sample critical values are presented in Section 4. Section 5 assesses the size and power of the tests in a simple bivariate setting.

Section 6 offers brief concluding remarks and an appendix contains technical derivations.

2 The Model and Preliminary Results

We consider the following Data Generating Process:

$$\begin{aligned} z_t &= d_t + u_t \\ u_t &= Au_{t-1} + v_t \end{aligned} \tag{1}$$

where $z_t = (y_t, x_t')'$ and $u_t = (u_{1t}, u_{2t}')'$ are n-vectors, x_t and u_{2t} are m-vectors ($n = m+1$), y_t is a scalar, A is a diagonal matrix with typical element given by $\alpha = 1+c/T$. Then u_t is a near-integrated process whith c as a non-centrality parameter, and $\{v_t\}$ is an unobserved stationary mean-zero process with finite variance and spectral density matrix $f_{vv}(\lambda)$. We use the assumption that $u_0 = 0$ throughout, though the results hold for the weaker requirement that $E(u_0^2) < \infty$. The noise function is $v_t = \Phi(L)\tau_t = \sum_{i=0}^{\infty} \phi_i \tau_{t-i}$ with $\sum_{i=0}^{\infty} i \det |\phi_i| < \infty$ and $\tau_t \sim i.i.d.(0, \sigma_\tau^2)$ [1]. Under these conditions, the partial sum process constructed from $\{v_t\}$ satisfies a multivariate invariance principle. More specifically, for $r \in [0, 1]$, and as $T \to \infty$, we have

$$X_T(r) = T^{-1/2} \sum_{t=1}^{[Tr]} v_t \Rightarrow B_c(r),$$

[1]This assumption is made for simplicity of exposition. All results holds under more general conditions, in particular allowing τ_t to be a martingale difference sequence.

where $B_c(r)$ is a n-vector Brownian motion with covariance matrix given by

$$\begin{aligned}\Omega &= \lim_{T\to\infty} T^{-1}E\{[\sum_{t=1}^{T} v_t][\sum_{t=1}^{T} v_t']\} \\ &= 2\pi f_{vv}(0) \\ &= 2\pi\sigma_\tau^2\Phi(1)\Phi(1)'.\end{aligned}$$

We define the following partition of Ω and $B_c(r)$:

$$\Omega = \begin{bmatrix} w_{11} & w_{21}' \\ w_{21} & \Omega_{22} \end{bmatrix}$$

$$B_c(r) = \begin{bmatrix} B_{1c}(r) \\ B_{2c}(r) \end{bmatrix}$$

where w_{11} and $B_{1c}(r)$ are scalars, w_{21} and $B_{2c}(r)$ are vectors of dimension m and Ω_{22} is a positive definite matrix of dimension m by m. We also define the block triangular decomposition $\Omega = L'L$, with

$$L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & L_{22} \end{bmatrix}$$

where

$$\begin{aligned}l_{11} &= (w_{11} - w_{21}'\Omega_{22}^{-1}w_{21})^{1/2}, \\ l_{21} &= \Omega_{22}^{-1/2}w_{21}, \\ L_{22} &= \Omega_{22}^{-1/2}.\end{aligned}$$

Given that Ω_{22} is a positive definite matrix, the null hypothesis of no cointegration implies that $l_{11} \neq 0$ and, accordingly that Ω is of full rank. This assumption is maintained throughout.

In (1), the deterministic component d_t is specified as $d_t = \psi' m_t$, where m_t is a set of deterministic terms. In this paper, we only consider d_t to be a polynomial in time, i.e. $d_t = \sum_{i=0}^{p} \psi_i t^i$, with special focus on $p = 0, 1$.

2.1 Estimates of the Trend Function under GLS Detrending

To apply the *GLS* detrending procedure suggested by ERS, we start by defining the transformed data $z_t^{\bar{\alpha}}$ and $m_t^{\bar{\alpha}}$ as:

$$z_t^{\bar{\alpha}} = (1 - \bar{\alpha}L) z_t,$$

$$m_t^{\bar{\alpha}} = (1 - \bar{\alpha}L) m_t,$$

for $t = 2, ..., T$, and $z_1^{\bar{\alpha}} = z_1$, $m_1^{\bar{\alpha}} = m_1$. We let $\hat{\psi}$ be the estimator that minimizes:

$$S(\psi) = \left(z_t^{\bar{\alpha}} - \psi' m_t^{\bar{\alpha}}\right)' \left(z_t^{\bar{\alpha}} - \psi' m_t^{\bar{\alpha}}\right). \tag{2}$$

As a preliminary result, we consider the limiting distribution of the estimates $\hat{\psi}$ obtained from local to unity *GLS* detrending. We consider detrending each series separately assuming the quasi-difference parameter $\bar{\alpha} = 1 + \bar{c}/T$. The following lemma gives the asymptotic properties of the estimates of the coefficients of the trend function in this case.

Lemma 1 *Suppose that z_t is generated by (1) with A a diagonal matrix with typical elements given by $\alpha = 1 + c/T$, and $d_t = \psi' m_t$. Suppose that each variable in the $n-$vector z_t is detrended separately. Let $\hat{\psi}$ be the GLS estimates, from minimizing (2), of the coefficients of the trend function obtained using $\bar{\alpha} = 1 + \bar{c}/T$.*

a. *If $p = 0$, with $m_t = 1$ for all t, then:*

$$\Upsilon_T(\hat{\psi} - \psi)' \Rightarrow 0'$$

where $\Upsilon_T = diag(T^{-1/2}, ..., T^{-1/2})$, a n by n matrix.

b. *If $p = 1$, with $m_t' = (1, t)$ for all t, then:*

$$\Upsilon_T vec[(\hat{\psi} - \psi)'] \Rightarrow \begin{bmatrix} 0' \\ \Lambda B_c(1) + 3(I_n - \Lambda) \int sB_c(r)ds \end{bmatrix} \equiv \begin{bmatrix} 0' \\ D \end{bmatrix}$$

where $\Upsilon_T = [diag(T^{-1/2}, ..., T^{-1/2}), diag(T^{1/2}, ..., T^{1/2})]$, a $2n$ by $2n$ matrix and Λ is a n by n diagonal matrix with typical elements given by $\lambda = (1 - \bar{c})/(1 - \bar{c} + \bar{c}^2/3)$.

2.2 Limit Distributions for the Static Regression

We now consider residuals based tests for cointegration in the spirit of Phillips and Ouliaris (1990) but using GLS detrended variables defined by $(y_t^a, x_t^{a\prime})' = (y_t, x_t')' - \hat{\psi} m_t$ where $\hat{\psi}$ is obtained by minimizing (2). The relevant regression estimated by OLS is

$$y_t^a = \gamma' x_t^a + e_t. \tag{3}$$

The following theorem gives the limiting behavior of the estimate $\hat{\gamma}$ under the null hypothesis of no cointegration.

Theorem 1 *Suppose that z_t is generated by (1) with A a diagonal matrix with typical elements given by $\alpha = 1 + c/T$, and $d_t = \psi' m_t$. Let y_t^a and x_t^a be GLS detrended variables with non-centrality parameter $\bar{\alpha} = 1 + \bar{c}/T$. Let $\hat{g} = (1, -\widehat{\gamma})'$ be the OLS estimates, from (3), of the cointegrating vector.*

1. *If $p = 0$, with $m_t = 1$ for all t, then:*

$$\hat{g} = (1, -\widehat{\gamma})' \Rightarrow \eta' = (1, -g_{21}' G_{22})$$

where $g_{21} = \int_0^1 B_{2c}(r) B_{1c}(r) dr$ and $G_{22} = \int_0^1 B_{2c}(r) B_{2c}'(r) dr$.

2. *If $p = 1$, with $m_t' = (1, t)$ for all t, then:*

$$\hat{g} = (1, -\widehat{\gamma})' \Rightarrow \overline{\eta}' = (1, -\overline{g}_{21}' \overline{G}_{22})$$

where $\overline{g}_{21} = \int_0^1 \overline{B}_{2c}(r) \overline{B}_{1c}(r) dr$, $\overline{G}_{22} = \int_0^1 \overline{B}_{2c}(r) \overline{B}_{2c}'(r) dr$ and $\overline{B}_c(r) = B_c(r) - rD'$.

3 The Tests and their Asymptotic Distributions

3.1 The Tests

The M-tests, originally proposed by Stock (1990), and further analyzed by Perron and Ng (1996) and Ng and Perron (1999), exploit the feature that a series converges with different rates of normalization under the null and the alternative hypothesis. Constructed using the residuals from the cointegrating regression, they are defined by:

$$MZ_{\alpha}^{GLS} = (T^{-1}\hat{e}_T^2 - s^2)\left(2T^{-2}\sum_{t=1}^{T}\hat{e}_t^2\right)^{-1} \quad (4)$$

$$MSB^{GLS} = \left(T^{-2}\sum_{t=1}^{T}\hat{e}_t^2/s^2\right)^{1/2} \quad (5)$$

$$MZ_t^{GLS} = \left(T^{-1}\hat{e}_T^2 - s^2\right)\left(4s^2T^{-2}\sum_{t=1}^{T}\hat{e}_t^2\right)^{-1/2} \tag{6}$$

where $\hat{e}_t$ is the residual vector obtained from the cointegration regression (3), i.e.,

$$\hat{e}_t = y_t^a - \hat{\gamma}' x_t^a. \tag{7}$$

The term s^2 is an autoregressive estimate of the spectral density at frequency zero of v_t, defined as:

$$s^2 = s_{\eta k}^2 / \left(1 - \hat{b}(1)\right)^2, \tag{8}$$

where $s_{\eta k}^2 = T^{-1}\sum_{t=k+1}^{T}\hat{\eta}_{tk}^2$, $\hat{b}(1) = \sum_{j=1}^{k}\hat{b}_j$, with $\hat{b}_j$ and $\{\hat{\eta}_{tk}\}$ obtained from the autoregression [2]:

$$\Delta\hat{e}_t = b_0\hat{e}_{t-1} + \sum_{j=1}^{k} b_j \Delta\hat{e}_{t-j} + \eta_{tk}. \tag{9}$$

The first statistic is a modified version of the Phillips and Perron (1988) Z_α test originally developed by Phillips (1987). The second statistic is a modified version of Bhargava's (1986) R_1 statistic which builds upon the work of Sargan and Bhargava (1983). The third statistic is a modified version of the Phillips and Perron (1988) Z_t test. As Perron and Ng (1996) showed, the MSB and Z_α tests are approximately related by:

$$Z_t \approx MSB \cdot Z_\alpha.$$

[2] The advantages of using this autoregressive-based spectral density estimator over the more traditional kernel-based methods are discussed in Perron and Ng (1998).

This relation suggests the MZ_t^{GLS} test defined by (6) since it satisfies the relation

$$MZ_t^{GLS} = MSB^{GLS} \cdot MZ_\alpha^{GLS}.$$

Another test of interest is the so-called ADF test (Said and Dickey, 1984) which is the t-statistic for testing $b_0 = 0$ in regression (9). We denote this test by ADF^{GLS}.

3.2 The Asymptotic Distributions of the Tests

The following theorem gives the asymptotic distribution of the residual based tests for cointegration under the null hypothesis of no-cointegration.

Theorem 2 *Suppose that z_t is generated by (1) with A a diagonal matrix with typical elements given by $\alpha = 1 + c/T$, and $d_t = \psi' m_t$. Let $MZ_\alpha^{GLS}, MSB^{GLS}$, and MZ_t^{GLS} be defined as in (4), (5) and (6) with residuals obtained from (7). Let ADF^{GLS} be the t-statistic for testing b_0 in regression (9).*

a. *If $p = 0$, with $m_t = 1$ for all t, then:*

$$\begin{aligned} MZ_\alpha^{GLS} &\Rightarrow \frac{0.5\ K_1(c)}{K_2(c)} \equiv H^{MZ_\alpha}(c) \\ MSB^{GLS} &\Rightarrow \left(\frac{K_2(c)}{\kappa'\kappa}\right)^{1/2} \equiv H^{MSB}(c) \\ MZ_t^{GLS} &\Rightarrow \frac{0.5\ K_1(c)}{(K_2(c)\ \kappa'\kappa)^{1/2}} \equiv H^{MZ_t}(c) \\ ADF^{GLS} &\Rightarrow \frac{0.5\ K_1(c)}{(K_2(c)\ \kappa'\kappa)^{1/2}} \equiv H^{ADF}(c) \end{aligned} \tag{10}$$

where

$$\begin{aligned} K_1(c) &= Q_c(1)^2 - k'k, \\ K_2(c) &= \int_0^1 Q_c(r)^2 dr, \end{aligned}$$

$$
\begin{aligned}
Q_c(r) &= W_{1c}(r) - (\int_0^1 W_{1c}(s)W'_{2c}(s)ds)(\int_0^1 W_{2c}(s)W'_{2c}(s)ds)^{-1}W_{2c}(r), \\
k' &= (1, -f_{21}F_{22}^{-1}),
\end{aligned}
$$

and $W_c(r)$ is a multivariate Ornstein-Uhlenbeck process, i.e. the solution to the stochastic differential equation $dW_c(r) = CW_c(r) + dW(r)$ with $W_c(0) = 0$, where $C = diag(c, ..., c)$ an n by n diagonal matrix and $W(r)$ is an n vector of independent Wiener processes.

b. *If $p = 1$, with $z'_t = (1, t)$ for all t, then:*

$$
\begin{aligned}
MZ_\alpha^{GLS} &\Rightarrow \frac{0.5\,\overline{K}_1(c)}{\overline{K}_2(c)} \equiv H^{MZ_\alpha}(c) \\
MSB^{GLS} &\Rightarrow \left(\frac{\overline{K}_2(c)}{\overline{\kappa}'\overline{\kappa}}\right)^{1/2} \equiv H^{MSB}(c) \\
MZ_t^{GLS} &\Rightarrow \frac{0.5\,\overline{K}_1(c)}{(\overline{K}_2(c)\ \overline{\kappa}'\overline{\kappa})^{1/2}} \equiv H^{MZ_t}(c) \\
ADF^{GLS} &\Rightarrow \frac{0.5\,\overline{K}_1(c)}{(\overline{K}_2(c)\ \overline{\kappa}'\overline{\kappa})^{1/2}} \equiv H^{ADF}(c)
\end{aligned}
\qquad (11)
$$

where

$$
\begin{aligned}
\overline{K}_1(c) &= \overline{Q}_c(1)^2 - \overline{k}'\overline{k}, \\
\overline{K}_2(c) &= \int_0^1 \overline{Q}_c(r)^2 dr, \\
\overline{Q}_c(r) &= \overline{W}_{1c}(r) - (\int_0^1 \overline{W}_{1c}(s)\overline{W}'_{2c}(s)ds)(\int_0^1 \overline{W}_{2c}(s)\overline{W}'_{2c}(s)ds)^{-1}\overline{W}_{2c}(r), \\
\overline{W}_c &= W_c(r) - rE', \\
E' &= (\Lambda W_c(1) + 3(I_n - \Lambda)\int sW_c(s)ds)', \\
\overline{k}' &= (1, -\overline{f}_{21}\overline{F}_{22}^{-1}).
\end{aligned}
$$

Note that when $c = 0$ the limiting distributions for the tests MZ_α^{GLS}, MZ_t^{GLS} and ADF^{GLS} shown in (10) for the case $p = 0$ are those derived by Phillips and Ouliaris (1990) for the Z_α, Z_t and ADF tests when no deterministic component is present. This is a simple consequence of the use of local *GLS* detrending because the intercept is bounded in probability.

4 Asymptotic Critical Values

As seen in Theorem 2, the limiting distributions of the test statistics depend on both non-centrality parameters c and $\bar{c}$. Since the concept of cointegration usually involves variables with a unit root and that residual based tests are constructed from a regression with $I(1)$ variables (most often as a result of some unit root pre-test), it is natural to consider tabulating critical values setting $c = 0$. Accordingly, a sensible choice of the non-centrality parameter used to detrend the data by *GLS* is $\bar{c} = 0$. Nevertheless, we shall assess the extent to which the quality of the inference is dependent on that particular choice by also considering the size and power of the tests for the values suggested by ERS, namely $\bar{c} = -7.0$ when $p = 0$ and $\bar{c} = -13.5$ when $p = 1$.

We simulate directly the asymptotic distributions using 1,000 steps to approximate the Wiener process $W(r)$ as the partial sums of a vector of n independent *i.i.d.* $N(0,1)$ random variables. In all cases, 10,000 replications are used and we present critical values for $m = 1, 2, 3, 4$, and 5, where m is the number of right-hand side variables in the cointegrating regression. To assess the adequacy of the asymptotic distribution, we also present critical values for the finite sample distributions of the statistics for a sample of size $T = 100$ with data generated by independent random walks with zero initial condition and *i.i.d.* $N(0,1)$ errors. Here, k is set to unity to construct s^2.

The results are presented in Table 1 for the case where $p = 0$ and $\bar{c} = 0$; in Table 2 for $p = 0$ and $\bar{c} = -7.0$; in Table 3 for $p = 1$ and $\bar{c} = 0$, and in Table 4 for $p = 1$

and $\overline{c} = -13.5$. In general, the asymptotic distributions provide good approximations to the finite sample distributions. Note that when the model includes a time trend and GLS detrending is performed at $\overline{c} = 0$ the asymptotic critical values are similar to those derived by Phillips and Ouliaris (1990) for the case where only an intercept is included and removed by OLS, even though the theoretical limit distributions are different.

5 Size and Power of the Tests in Finite Samples

In this section, we evaluate the size and power properties of the tests considered in section 3. For comparison, we also include tests based on OLS detrended series, namely the Z_α, Z_t and ADF tests of Phillips and Ouliaris (1990) and the MZ_α test proposed by Stock (1990) (we do not report results for the MZ_t and MSB as they are very similar to those of MZ_α). This will allow us to assess the advantages of using GLS detrended series to construct the tests. Again, we only report results for MZ_α^{GLS} as they are similar to those for MZ_t^{GLS} and MSB^{GLS}.

In the tables, we use the following notation. The tests that use GLS detrended data with the non-centrality parameter $\overline{c} = 0$ are denoted as $\widetilde{Z}_\alpha$, $\widetilde{Z}_t$, $\widetilde{A}DF$ and $\widetilde{M}Z_\alpha$. The corresponding tests that use GLS detrended data with $\overline{c} = -7.0$ or $\overline{c} = -13.5$, according to whether a constant or a constant and a time trend is included are denoted $\overline{Z}_\alpha$, $\overline{Z}_t$, $\overline{A}DF$ and $\overline{M}Z_\alpha$.

5.1 The Monte Carlo Design

The data-generating process considered is given by:

$$
\begin{aligned}
y_t - \beta x_t &= u_{1t} \qquad (12)\\
a_1 y_t - a_2 x_t &= u_{2t}\\
u_{1t} &= \alpha u_{1t-1} + e_{1t}\\
u_{2t} &= u_{2t-1} + e_{2t}\\
\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} &= i.i.d. \quad N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \eta\sigma \\ \eta\sigma & \sigma^2 \end{pmatrix}\right]
\end{aligned}
$$

This data-generating process is similar to that used by Banerjee et al. (1986), Gonzalo (1994) and Haug (1996). Here, y_t and x_t are scalars. For each Monte-Carlo experiment, we generate 2,000 series of length T and we start at $u_{20} = 0$ and $u_{10} = 0$. We use the *rndns* function in Gauss-Unix with a *seed* = 123456 in order to generate the pseudo normal variates e_{1t} and e_{2t}.

The values of the various parameters considered are similar to those used in Gonzalo (1994) and Haug (1996), among others. The parameter β is assumed unity and a_2 is assumed minus unity. For the parameter a_1, two cases are considered. The first sets $a_1 = 0$, which is equivalent to assuming that the variable x_t is exogenous with respect to the parameter β. We also consider $a_1 = 1$, in which case x_t is endogenous. The parameter η measures the presence of contemporaneous correlation across the errors. Here, we consider three possible values (-0.5, 0.0 and 0.5). Notice that the case where $\eta = 0$ in conjunction with $a_1 = 0$ implies strong exogeneity. The parameter σ is known as the ratio signal-noise and it measures how big is the random walk component in the series. We consider three possible values for this parameter (0.25, 1.0 and 4.0). Two sample sizes are used, $T = 100$ and 200.

Note that we consider only the case with *i.i.d.* errors and, hence, the presence of correlation is not analyzed. Our principal goal is to appreciate gains in power from tests constructed using GLS detrended data compared to the case where OLS detrended series are used[3]. Finally, the BIC is used to select lag length k when estimating the autoregression (9) to construct s^2 and the ADF tests.

The size and power are evaluated using asymptotic critical values. For the Z_α, MZ_α, Z_t and ADF tests, we use those from Phillips and Ouliaris (1990). For the GLS-based tests, we use the asymptotic critical values tabulated in Tables 1 to 4 for the case where $m = 1$.

5.2 Size of the Tests

Size is evaluated at $\alpha = 1$. The results are reported in Tables 5 and 6 for the case where the model includes only a constant and in Tables 7 and 8 for the case where both a constant and a time trend are included. Tables 5 and 7 consider the case where x_t is endogenous and Tables 6 and 8 the case where x_t is weakly exogenous. Given that we use $2,000$ replications an approximate 95% confidence interval for the p-value .05 is $(.04, .06)$. What transpires from the results is that the exact sizes of all tests, except MZ_α, are within this confidence interval (or very close to it) in all cases considered.

[3]We also conducted experiments for MA(1) and AR(1) process in the errors. The relative ranking and the general conclusions remain the same unless these is strong correlation in which case the class of M tests suffer much less from size distortions. The results are available upon request.

The tests MZ_α have exact sizes slightly below the nominal 5% size when $T = 100$. These distortions are reduced when $T = 200$ for the case $p = 0$ but remain somewhat, though to a lesser extent, when $p = 1$. The conservative nature of the M tests was also noted by Perron and Ng (1996) in the case of unit root tests.

5.3 Power of the Tests

Power is evaluated at $\alpha = 0.85$ when $T = 100$ and at $\alpha = 0.90$ when $T = 200$. The results are reported in Tables 9 and 10 for the case where the model includes only a constant and in Tables 11 and 12 for the case where both a constant and a time trend are included. Tables 9 and 11 consider the case where x_t is endogenous and Tables 10 and 12 the case where x_t is weakly exogenous.

The qualitative results are consistent across the various configurations considered. The most striking and important feature is the fact that substantial power gains can be achieved using *GLS* detrended data instead of *OLS* detrended data. These gains are more important when the quasi-differencing parameter is set to $\bar{c} = -7.0$ (if $p = 0$) or $\bar{c} = -13.5$ (if $p = 1$), particularly when $\sigma \geqslant 1$. For example, consider the *ADF* test in the case $p = 0$, $\sigma = 4$, $a_1 = 1$ and $\eta = 0.0$ with $T = 100$ (whose results are presented in the second to last column of Table 9). The power of the test using the standard test of Phillips and Ouliaris (1990) based on *OLS* detrended data is .449. With *GLS* detrended data this power increases to .611 when using a quasi-differencing parameter $\bar{c} = 0$. When using a the non-centrality parameter $\bar{c} = -7.0$ (as suggested by ERS for unit root tests), the power is further increased to .668. This increase in power holds in all configurations considered.

Concerning the relative ranking of the various tests considered, the most powerful are the Z_t and ADF followed by the Z_α. The test MZ_α appears significantly less powerful, no doubt due to the fact that its exact size is conservative.

6 Conclusions

We have considered residuals-based tests for cointegration using GLS detrended data and provided asymptotic critical values for use in applied work. Our simulation results show that substantial gains in power can be obtained relative to the OLS residuals-based tests of Phillips and Ouliaris (1990) especially when using a quasi-differencing parameter $\bar{c} = -7.0$ or $\bar{c} = -7.0$ according to the nature of the deterministic components included in the regression. Given that our tests are basically as easy to construct, they provide a useful alternative for empirical applications.

While we have shown that substantial power gains can be achieved, our approach nevertheless lacks any optimality criterion. Unlike Elliott, Rothemberg and Stock (1996) for unit root tests, we have not shown that the local asymptotic power functions of the tests lie close to some Gaussian power envelope. Deriving such power envelope is more delicate in the context of tests for cointegration. This is the subject of ongoing research by us.

7 Appendix

Throughout, we use the following lemmas which are by now standard (see Lemma 3.1 of Phillips (1988) and Lemma 2.2 of Phillips and Ouliaris (1990)).

Lemma A.1: Let $\{u_t\}$ be a near-integrated series generated by (1), then: a) $T^{-1/2}u_{[Tr]} \Rightarrow B_c(r)$; b) $T^{-3/2}\sum_{t=1}^{T} u_t \Rightarrow \int_0^1 B_c(r)dr$; c) $T^{-2}\sum_{t=1}^{T} u_t u_t' \Rightarrow \int_0^1 B_c(r)B_c'(r)dr$; d) $T^{-1}\sum_{t=1}^{T} u_{t-1}v_t \Rightarrow \{\int_0^1 B_c(r)dB_c(r)' + \Omega_1\}$ with $\Omega_1 = \sum_{k=1}^{\infty} E(v_0 v_k')$.

Lemma A.2: Using the notation defined in Section 2, we have: a) $B_c(r) \equiv L'W_c(r)$; b) $L\eta \equiv l_{11}k, \quad \eta'\Omega\eta \equiv w_{11.2}k'k$; c) $\eta' B_c(r) \equiv l_{11}Q_c(r)$; d) $\eta' \int_0^1 B_c(r)dB_c'(r)\eta \equiv w_{11.2}\int_0^1 Q_c(r)dQ_c(r)$; e) $\eta' A\eta \equiv w_{11.2}\int_0^1 Q_c(r)^2 dr$; with $w_{11.2} = w_{11} - w_{21}'\Omega_{22}^{-1}w_{21} = l_{11}^2$.

Proof of Lemma 1: In both cases, the proof is a multivariate extension of arguments given in ERS (1996) for the univariate case and is, hence, omitted.

Proof of Theorem 1: Consider first the case where $m_t = \{1\}$. Let $z_t^a = (y_t^a, x_t^{a\prime})'$ be a n-vector, $M_t = (u_t', m_t)'$ be a $n+1$-vector and let $\Upsilon_T = diag(I_n, T^{-1/2})$ be a diagonal matrix. We can write:

$$T^{-2}\sum_{t=1}^{T} z_t^a z_t^{a\prime} = \begin{bmatrix} I_n \\ -(\hat{\psi}-\psi) \end{bmatrix}' \Upsilon_T \begin{bmatrix} T^{-2}\Upsilon_T^{-1}\sum_{t=1}^{T} M_t M_t' \Upsilon_T^{-1} \end{bmatrix} \Upsilon_T \begin{bmatrix} I_n \\ -(\hat{\psi}-\psi) \end{bmatrix} \qquad \text{(A.1)}$$

Using the notation $\psi_0 = (\psi_{0,y}, \psi_{0,x})$ where $\psi_{0,y}$ corresponds to the mean of the dependent variable and $\psi_{0,x}$ is an m-vector of means associated with the regressors x_t (a similar partition is used for $\hat{\psi}_0$). The first element in (A.1) can be written as

$$\begin{aligned}
\begin{bmatrix} I_n & -(\hat{\psi}-\psi)' \end{bmatrix} \Upsilon_T &= \begin{bmatrix} 1 & \mathbf{0} & -(\hat{\psi}_{0,y}-\psi_{0,y}) \\ \mathbf{0}' & I_m & -(\hat{\psi}_{0,x}-\psi_{0,x}) \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}' & 0 \\ \mathbf{0} & I_m & \mathbf{0} \\ 0 & 0 & T^{-1/2} \end{bmatrix} \\
&= \begin{bmatrix} 1 & \mathbf{0}' & -T^{-1/2}(\hat{\psi}_{0,y}-\psi_{0,y}) \\ \mathbf{0} & I_m & -T^{-1/2}(\hat{\psi}_{0,x}-\psi_{0,x}) \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} 1 & \mathbf{0}' & 0 \\ \mathbf{0} & I_m & \mathbf{0} \end{bmatrix}
\end{aligned}$$

The limit of the second element is:

$$\begin{aligned}
\left[T^{-2}\Upsilon_T^{-1}\sum_{t=1}^{T} M_t M_t' \Upsilon_T^{-1}\right] &= T^{-2}\begin{bmatrix} 1 & \mathbf{0}' & 0 \\ \mathbf{0} & I_m & \mathbf{0} \\ 0 & 0 & T^{1/2} \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{T} u_{1t}u_{1t} & \sum_{t=1}^{T} u_{1t}u_{2t}' & \sum_{t=1}^{T} u_{1t} \\ \sum_{t=1}^{T} u_{2t}u_{1t} & \sum_{t=1}^{T} u_{2t}u_{2t}' & \sum_{t=1}^{T} u_{2t}' \\ \sum_{t=1}^{T} u_{1t} & \sum_{t=1}^{T} u_{2t} & T \end{bmatrix} \\
&\quad \begin{bmatrix} 1 & \mathbf{0}' & 0 \\ \mathbf{0} & I_m & \mathbf{0} \\ 0 & 0 & T^{1/2} \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} T^{-2}\sum_{t=1}^{T} u_{1t}u_{1t} & T^{-2}\sum_{t=1}^{T} u_{1t}u_{2t}' & T^{-3/2}\sum_{t=1}^{T} u_{1t} \\ T^{-2}\sum_{t=1}^{T} u_{2t}u_{1t} & T^{-2}\sum_{t=1}^{T} u_{2t}u_{2t}' & \sum_{t=1}^{T} u_{2t}' \\ T^{-3/2}\sum_{t=1}^{T} u_{1t} & T^{-3/2}\sum_{t=1}^{T} u_{2t} & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \int_0^1 B_{1c}(r)B_{1c}(r)dr & \int_0^1 B_{1c}(r)B_{2c}'(r)dr & \int_0^1 B_{1c}(r)dr \\ \int_0^1 B_{2c}(r)B_{1c}(r)dr & \int_0^1 B_{2c}(r)B_{2c}'(r)dr & \int_0^1 B_{2c}'(r)dr \\ \int_0^1 B_{1c}(r)dr & \int_0^1 B_{2c}(r)dr & 1 \end{bmatrix}$$

By symmetry, the third element converges to the transpose of the limit of the first and we have

$$\begin{aligned} T^{-2}\sum_{t=1}^{T} z_t^a z_t^{a\prime} &\Rightarrow \begin{bmatrix} \int_0^1 B_{1c}(r)B_{1c}(r)dr & \int_0^1 B_{1c}(r)B_{2c}'(r)dr \\ \int_0^1 B_{2c}(r)B_{1c}(r)dr & \int_0^1 B_{2c}(r)B_{2c}'(r)dr \end{bmatrix} \\ &\equiv \begin{bmatrix} g_{11} & g_{21}' \\ g_{21} & G_{22} \end{bmatrix} \equiv G \end{aligned} \tag{A.2}$$

The limit of $\hat{g}$ follows directly. Now, consider the case where $m_t = \{1, t\}$. Let $\Upsilon_T = diag(I_n, T^{-1/2}, T^{1/2})$ be a diagonal matrix and using a similar partition for $(\hat{\psi} - \psi)$.

$$\begin{aligned} \begin{bmatrix} I_n & -(\hat{\psi}-\psi)' \end{bmatrix} \Upsilon_T &= \begin{bmatrix} 1 & \mathbf{0}' & -(\hat{\psi}_{0,y}-\psi_{0,y}) & -(\hat{\psi}_{1,y}-\psi_{1,y}) \\ \mathbf{0} & I_m & -(\hat{\psi}_{0,x}-\psi_{0,x}) & -(\hat{\psi}_{1,x}-\psi_{1,x}) \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}' & 0 & 0 \\ \mathbf{0} & I_m & 0 & 0 \\ 0 & 0 & T^{-1/2} & 0 \\ 0 & 0 & 0 & T^{1/2} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \mathbf{0}' & -T^{-1/2}(\hat{\psi}_{0,y}-\psi_{0,y}) & -T^{-1/2}(\hat{\psi}_{1,y}-\psi_{1,y}) \\ \mathbf{0} & I_m & -T^{-1/2}(\hat{\psi}_{0,x}-\psi_{0,x}) & -T^{-1/2}(\hat{\psi}_{1,x}-\psi_{1,x}) \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} 1 & \mathbf{0}' & 0 & -D_y \\ \mathbf{0} & I_m & 0 & -D_x \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} D_y &= \lambda B_{1c}(1) + 3(1-\lambda)\int_0^1 sB_{1c}(r)dr \\ D_x &= \Lambda_x B_{2c}(1) + 3(I_m - \Lambda_x)\int_0^1 sB_{2c}(r)dr \end{aligned}$$

and $\Lambda_x = diag(\lambda, ..., \lambda)$ a m by m diagonal matrix. Also,

$$\left[T^{-2}\Upsilon_T^{-1}\sum_{t=1}^{T} M_t M_t' \Upsilon_T^{-1}\right] = \begin{bmatrix} T^{-2}\sum_{t=1}^{T} u_{1t}u_{1t} & T^{-2}\sum_{t=1}^{T} u_{1t}u_{2t}' & T^{-3/2}\sum_{t=1}^{T} u_{1t} & T^{-5/2}\sum_{t=1}^{T} tu_{1t} \\ T^{-2}\sum_{t=1}^{T} u_{2t}u_{1t} & T^{-2}\sum_{t=1}^{T} u_{2t}u_{2t}' & T^{-3/2}\sum_{t=1}^{T} u_{2t}' & T^{-5/2}\sum_{t=1}^{T} tu_{2t}' \\ T^{-3/2}\sum_{t=1}^{T} u_{1t} & T^{-3/2}\sum_{t=1}^{T} u_{2t} & 1 & T^{-2}\sum_{t=1}^{T} t \\ T^{-5/2}\sum_{t=1}^{T} tu_{1t} & T^{-5/2}\sum_{t=1}^{T} tu_{2t} & T^{-2}\sum_{t=1}^{T} t & T^{-3}\sum_{t=1}^{T} t^2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \int_0^1 B_{1c}(r)B_{1c}(r)dr & \int_0^1 B_{1c}(r)B_{2c}'(r)dr & \int_0^1 B_{1c}(r)dr & \int_0^1 rB_{1c}(r)dr \\ \int_0^1 B_{2c}(r)B_{1c}(r)dr & \int_0^1 B_{2c}(r)B_{2c}'(r)dr & \int_0^1 B_{2c}'(r)dr & \int_0^1 rB_{2c}'(r)dr \\ \int_0^1 B_{1c}(r)dr & \int_0^1 B_{2c}(r)dr & 1 & \int_0^1 rdr \\ \int_0^1 rB_{1c}(r)dr & \int_0^1 rB_{2c}(r)dr & \int_0^1 rdr & \int_0^1 r^2dr \end{bmatrix}$$

Using these results, we have

$$\begin{aligned} T^{-2}\sum_{t=1}^{T} z_t^a z_t^{a\prime} &\Rightarrow \begin{bmatrix} \int_0^1 \overline{B}_{1c}(r)\overline{B}_{1c}(r)dr & \int_0^1 \overline{B}_{1c}(r)\overline{B}_{2c}'(r)dr \\ \int_0^1 \overline{B}_{2c}(r)\overline{B}_{1c}(r)dr & \int_0^1 \overline{B}_{2c}(r)\overline{B}_{2c}'(r)dr \end{bmatrix} \\ &\equiv \begin{bmatrix} \overline{g}_{11} & \overline{g}_{21}' \\ \overline{g}_{21} & \overline{G}_{22} \end{bmatrix} \equiv \overline{G} \end{aligned} \tag{A.3}$$

where $\overline{B}_c = B_c(r) - rD$, and the final expression for the limit of $\hat{g}$ follows directly.

Proof of Theorem 2: Consider first the case where $m_t = \{1\}$. We have following results:

a. $2T^{-2}\sum_{t=1}^{T}\hat{e}_t^2 = 2T^{-2}\sum_{t=1}^{T}\{y_t^a - \hat{\gamma}'x_t^a\}^2 = 2\hat{g}'\{T^{-2}\sum_{t=1}^{T} z_t^a z_t^{\prime a}\}\hat{g} \Rightarrow 2\eta' G\eta$

From expression (A.2), we have $G = \int_0^1 B_c(r)B_c(r)'dr$, and, from Lemma A.2, $B_c(r) = L'W_c(r)$ and $L\eta = l_{11}k$. Furthermore, using the results $k'W_c(r) = Q_c(r)$ and $w_{11.2} = w_{11} - w_{21}'\Omega_{22}^{-1}w_{21} = l_{11}^2$, we have:

$$\begin{aligned} 2\eta' G\eta &= 2\int_0^1 [\eta' B_c(r)][\eta' B_c(r)]'dr \\ &= 2\int_0^1 [\eta' L'W_c(r)][\eta' L'W_c(r)]'dr \\ &= 2l_{11}^2\int_0^1 [k'W_c(r)][k'W_c(r)]'dr \\ &= 2l_{11}^2\int_0^1 Q_c(r)^2dr \\ &= 2w_{11.2}\int_0^1 Q_c(r)^2dr. \end{aligned}$$

b. $T^{-1}\hat{e}_T^2 = T^{-1}\{y_T^a - \hat{\gamma}'x_T^a\}^2 = \hat{g}'\{T^{-1}z_T^a z_T^{\prime a}\}\hat{g} \Rightarrow \eta' B_c(1)B_c(1)'\eta$

Using similar arguments as in (a), we have:

$$\begin{aligned} [\eta' B_c(1)][\eta' B_c(1)]' &= [\eta' L'W_c(1)][\eta' L'W_c(1)]' \\ &= l_{11}^2[k'W_c(1)][k'W_c(1)]' \\ &= l_{11}^2 Q_c(1)^2 \\ &= w_{11.2}Q_c(1)^2. \end{aligned}$$

c. $s^2 \Rightarrow \eta'\Omega\eta$.

We have $s^2 \Rightarrow \lim_{T\to\infty} Var(\sum_{t=1}^{T} \Delta\hat{e}_t)$. Now $\Delta\hat{e}_t = \hat{g}'\Delta z_t = \hat{g}'u_t$. Using the fact that $\hat{g} \Rightarrow \eta$, we have

$$\begin{aligned} s^2 &\Rightarrow \eta' \left[\lim_{T\to\infty} Var(\sum_{t=1}^{T} u_t) \right] \eta \\ &= \eta'\Omega\eta. \end{aligned}$$

Using Lemma A.2, $\eta'\Omega\eta = w_{11.2}\kappa'\kappa$, where $k' = (1, -f_{21}'F_{22}^{-1})$ and f_{21} and F_{22} are elements of the n by n matrix F defined as

$$\begin{aligned} F &= \begin{bmatrix} f_{11} & f_{21}' \\ f_{21} & F_{22} \end{bmatrix} \\ &= \begin{bmatrix} \int_0^1 W_{1c}(r)W_{1c}(r)dr & \int_0^1 W_{1c}(r)W_{2c}'(r)dr \\ \int_0^1 W_{2c}(r)W_{1c}(r)dr & \int_0^1 W_{2c}(r)W_{2c}'(r)dr \end{bmatrix} \end{aligned} \quad (A.4)$$

Using these results in expression (4), the proof is complete for the MZ_α^{GLS} test. For the MSB^{GLS} test, it is sufficient to use (a) and (c) in expression (5). Given the equivalence $MZ_t^{GLS} = MZ_\alpha^{GLS} * MSB^{GLS}$, the proof for the MZ_t^{GLS} test follows.

Consider now the case where $m_t = \{1, t\}$. We have following results:

a. $2T^{-2}\sum_{t=1}^{T} \hat{e}_t^2 = 2T^{-2}\sum_{t=1}^{T}\{y_t^a - \hat{\gamma}'x_t^a\}^2 = 2\hat{g}'\{T^{-2}\sum_{t=1}^{T} z_t^a z_t'^a\}\hat{g} \Rightarrow 2\overline{\eta}'\overline{G}\overline{\eta}$

From expression (A.3), we have $\overline{G} = \int_0^1 \overline{B}_c(r)\overline{B}_c(r)'dr$, and, from Lemma A.2, $\overline{B}_c(r) = L'\bar{W}_c(r)$ and $L\bar{\eta} = l_{11}\bar{k}$. Furthermore, using the results $\bar{k}'\bar{W}_c(r) = \bar{Q}_c(r)$ and $w_{11.2} = w_{11} - w_{21}'\Omega_{22}^{-1}w_{21} = l_{11}^2$, we have:

$$\begin{aligned} 2\overline{\eta}'\overline{G}\overline{\eta} &= 2\int_0^1 [\overline{\eta}'\overline{B}_c(r)][\overline{\eta}'\overline{B}_c(r)]'dr \\ &= 2\int_0^1 [\overline{\eta}'L'\overline{W}_c(r)][\overline{\eta}'L'\overline{W}_c(r)]'dr \\ &= 2l_{11}^2\int_0^1 [\overline{k}'\overline{W}_c(r)][\overline{k}'\overline{W}_c(r)]'dr \\ &= 2l_{11}^2\int_0^1 \bar{Q}_c(r)^2 dr \\ &= 2w_{11.2}\int_0^1 \bar{Q}_c(r)^2 dr \end{aligned}$$

b. $T^{-1}\hat{e}_T^2 = T^{-1}\{y_T^a - \hat{\gamma}'x_T^a\}^2 = \hat{g}'\{T^{-1}z_T^a z_T'^a\}\hat{g} \Rightarrow \overline{\eta}'\overline{B}_c(1)\overline{B}_c(1)'\overline{\eta}$

Using similar arguments as in (a), we have:

$$\begin{aligned} [\overline{\eta}'\overline{B}_c(1)][\overline{\eta}'\overline{B}_c(1)]' &= [\overline{\eta}'L'\overline{W}_c(1)][\overline{\eta}'L'\overline{W}_c(1)]' \\ &= l_{11}^2[\overline{k}'\overline{W}_c(1)][\overline{k}'\overline{W}_c(1)]' \end{aligned}$$

$$= l_{11}^2 \bar{Q}_c(1)^2$$
$$= w_{11.2} \bar{Q}_c(1)^2$$

c. $s^2 \Rightarrow \bar{\eta}'\Omega\bar{\eta}$, using arguments similar to the case with $m_t = \{1\}$.

Using Lemma A.2, $\overline{\eta}'\Omega\overline{\eta} = w_{11.2}\overline{\kappa}'\overline{\kappa}$, where $\overline{k}' = (1, -\overline{f}'_{21}\overline{F}_{22}^{-1})$ and $\overline{f}_{21}$ and $\overline{F}_{22}$ are elements from a n by n matrix $\bar{F}$ given by (A.4) with $\bar{W}_{1c}(r)$ and $\bar{W}'_{2c}(r)$ instead of $W_{1c}(r)$ and $W'_{2c}(r)$. Using these results in expression (4), the proof is complete for the MZ_α^{GLS} test. For the MSB^{GLS} test, it is sufficient to use (a) and (c) in expression (5). Given the equivalence $MZ_t^{GLS} = MZ_\alpha^{GLS} * MSB^{GLS}$, the proof for the MZ_t^{GLS} test follows.

References

[1] Banerjee, A., J. J. Dolado, J. W. Galbraith and D. H. Hendry (1993), *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press.

[2] Banerjee, A., J. J. Dolado, D. F. Hendry and G. W. Smith (1986), "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence," *Oxford Bulletin of Economics and Statistics* **48**, 253-277.

[3] Bhargava, A. (1986), "On the Theory of Testing for Unit Root in Observed Time Series," *Review of Economic Studies* **53**, 369-384.

[4] Elliott, G., T. Rothemberg and J. H. Stock (1996), "Efficient Tests for an Autoregessive Unit Roots," *Econometrica* **64**, 813-836.

[5] Engle, R. F. and C. W .J. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica* **55**, 251-276.

[6] Gonzalo, J. (1994), "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," *Journal of Econometrics* **60**, 203-233.

[7] Granger, C. W. J. and P. Newbold (1974), "Spurious Regression in Econometrics," *Journal of Econometrics* **2**, 111-120.

[8] Haug, A. A. (1996), "Tests for Cointegration. A Monte Carlo Comparison," *Journal of Econometrics* **71**, 89-115.

[9] Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* **59**, 1551-1580.

[10] Ng, S. and P. Perron (1999), "Lag Length Selection and the Construction of Unit Root Tests With Good Size and Power," manuscript, Department of Economics, Boston College.

[11] Perron, P. and J. Y. Campbell (1992), "Racines Unitaires en Macroeconomie: Le Cas Multidimensionnel," *Annales D'Economie et de Statistique* **27**, 1-50.

[12] Perron, P. and S. Ng (1996), "Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties," *Review of Economic Studies* **63**, 435-463.

[13] Perron, P. and S. Ng (1998), "An Autoregressive Spectral Density Estimator at Frequency Zero for Nonstationarity Tests," *Econometric Theory* **14**, 560-603.

[14] Phillips, P. C. B. (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics* **33**, 311-340.

[15] Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root," *Econometrica* **55**, 277-302.

[16] Phillips, P. C. B. (1988), "Regression Theory for Near-Integrated Time Series," *Econometrica* **56**, 1021-1043.

[17] Phillips, P. C. B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica* **58**, 165-193.

[18] Phillips, P. C. B. and P. Perron (1988), "Testing for Unit Roots in Time Series Regression," *Biometrika* **75**, 335-346.

[19] Said, S. E. and D. A. Dickey (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika* 71, 599-608.

[20] Sargan, J. D. and A. Bhargava (1983), "Testing Residuals from Least Squares Regression for being generated by the Gaussian Random Walk," *Econometrica* **51**, 153-174.

[21] Stock, J. H. (1987), "Asymptotic Properties of Least-Squares Estimators of Cointegrating Vectors," *Econometrica* **55**, 1035-1056.

[22] Stock, J. H. (1990), "A Class of Tests for Integration and Cointegration," Unpublished manuscript, Kennedy School of Government, Harvard University.

[23] Stock, J. H. and M. W. Watson (1988), "Testing for Common Trends," *Journal of the American Statistical Association* **83**, 1097-1107.

[24] Watson, M. W. (1994), "Vector Autoregressions and Cointegration,"in *Handbook of Econometrics*, vol. IV, Engle, R. F. and D. L. McFadden (eds.), chapter 47.

Annex

Table 1. Percentage points of the tests under the null hypothesis ($c = 0$)
Case with $p = 0$ and $\overline{c} = 0$

a) MZ_{α}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-22.84	-30.72	-38.03	-43.72	-50.16	-22.84	-29.48	-34.96	-40.33	-45.07
2.5%	-18.96	-26.23	-32.09	-38.17	-44.46	-18.92	-25.02	-30.15	-35.62	-40.07
5.0%	-15.84	-22.63	-28.34	-33.90	-39.87	-15.89	-21.87	-26.49	-31.49	-35.70
7.5%	-14.07	-20.42	-25.62	-31.34	-36.97	-14.01	-19.91	-24.57	-28.97	-33.14
10.0%	-12.80	-18.85	-24.05	-29.60	-34.98	-12.69	-18.47	-22.94	-27.30	-31.46
15.0%	-10.93	-16.48	-21.63	-26.87	-31.80	-10.70	-16.22	-20.46	-24.78	-28.90
20.0%	-9.59	-14.82	-19.75	-24.63	-29.32	-9.43	-14.58	-18.73	-22.89	-26.74

b) MSB^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	0.146	0.126	0.113	0.106	0.099	0.145	0.129	0.118	0.110	0.105
2.5%	0.159	0.136	0.123	0.113	0.105	0.160	0.139	0.127	0.117	0.110
5.0%	0.173	0.146	0.131	0.119	0.111	0.174	0.148	0.135	0.124	0.117
7.5%	0.183	0.153	0.137	0.124	0.115	0.183	0.155	0.140	0.129	0.121
10.0%	0.191	0.159	0.142	0.128	0.118	0.192	0.161	0.145	0.133	0.124
15.0%	0.206	0.169	0.149	0.134	0.124	0.207	0.171	0.153	0.140	0.130
20.0%	0.218	0.178	0.155	0.140	0.128	0.220	0.180	0.159	0.145	0.134

c) MZ_{t}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.34	-3.89	-4.33	-4.64	-4.97	-3.31	-3.80	-4.12	-4.45	-4.72
2.5%	-3.03	-3.57	-3.97	-4.33	-4.68	-3.00	-3.48	-3.84	-4.17	-4.42
5.0%	-2.76	-3.32	-3.72	-4.08	-4.42	-2.76	-2.25	-3.60	-3.92	-4.19
7.5%	-2.59	-3.14	-3.54	-3.92	-4.26	-2.59	-3.10	-3.46	-3.76	-4.03
10.0%	-2.48	-3.02	-3.42	-3.81	-4.13	-2.46	-2.97	-3.34	-3.65	-3.92
15.0%	-2.27	-2.82	-3.24	-3.62	-3.94	-2.25	-2.79	-3.15	-3.47	-3.75
20.0%	-2.11	-2.65	-3.08	-3.46	-3.78	-2.10	-2.65	-3.01	-3.33	-3.61

d) ADF^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.34	-3.89	-4.33	-4.64	-4.97	-3.41	-3.95	-4.35	-4.75	-5.09
2.5%	-3.03	-3.57	-3.97	-4.33	-4.68	-3.06	-3.60	-4.05	-4.43	-4.75
5.0%	-2.76	-3.32	-3.72	-4.08	-4.42	-2.80	-3.33	-3.74	-4.13	-4.46
7.5%	-2.59	-3.14	-3.54	-3.92	-4.26	-2.59	-3.16	-3.58	-3.95	-4.26
10.0%	-2.48	-3.02	-3.42	-3.81	-4.13	-2.47	-3.02	-3.44	-3.81	-4.13
15.0%	-2.27	-2.82	-3.24	-3.62	-3.94	-2.25	-2.82	-3.24	-3.60	-3.94
20.0%	-2.11	-2.65	-3.08	-3.46	-3.78	-2.09	-2.67	-3.08	-3.44	-3.77

Table 2. Percentage points of the tests under the null hypothesis ($c = 0$)
Case with $p = 0$ and $\bar{c} = -7.0$

a) MZ_{α}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-22.84	-30.72	-38.03	-43.72	-50.16	-23.03	-29.54	-34.82	-40.36	-44.97
2.5%	-18.96	-26.23	-32.09	-38.17	-44.46	-19.32	-24.89	-30.09	-35.58	-39.81
5.0%	-15.84	-22.63	-28.34	-33.90	-39.87	-16.08	-21.85	-26.50	-31.41	-35.72
7.5%	-14.07	-20.42	-25.62	-31.34	-36.97	-14.30	-19.53	-24.65	-28.99	-33.08
10.0%	-12.80	-18.84	-24.05	-29.60	-34.98	-12.96	-18.47	-22.93	-27.18	-31.39
15.0%	-10.93	-16.48	-21.63	-26.87	-31.80	-11.07	-16.30	-20.50	-24.75	-28.88
20.0%	-9.59	-14.82	-19.75	-24.63	-29.32	-9.72	-14.76	-18.74	-22.86	-26.68

b) MSB^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	0.145	0.126	0.113	0.106	0.099	0.144	0.129	0.119	0.110	0.105
2.5%	0.159	0.136	0.123	0.113	0.105	0.158	0.139	0.127	0.117	0.111
5.0%	0.173	0.146	0.131	0.119	0.111	0.172	0.148	0.135	0.124	0.117
7.5%	0.183	0.153	0.137	0.124	0.114	0.181	0.155	0.140	0.129	0.121
10.0%	0.191	0.159	0.1421	0.128	0.118	0.190	0.160	0.145	0.133	0.124
15.0%	0.206	0.169	0.149	0.134	0.124	0.205	0.170	0.153	0.140	0.130
20.0%	0.218	0.178	0.155	0.140	0.128	0.217	0.179	0.160	0.145	0.134

c) MZ_{t}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.34	-3.89	-4.33	-4.64	-4.97	-3.34	-3.80	-4.12	-4.45	-4.71
2.5%	-3.03	-3.57	-3.97	-4.33	-4.68	-3.03	-3.48	-3.84	-4.17	-4.41
5.0%	-2.76	-3.32	-3.72	-4.08	-4.42	-2.79	-3.25	-3.60	-3.93	-4.18
7.5%	-2.59	-3.14	-3.54	-3.92	−4.26	-2.61	-3.10	-3.46	-3.76	-4.02
10.0%	-2.48	-3.02	-3.42	-3.81	-4.13	-2.48	-2.98	-3.34	-3.65	-3.92
15.0%	-2.27	-2.82	-3.24	-3.62	-3.94	-2.28	-2.80	-3.15	-3.47	-3.75
20.0%	-2.11	-2.65	-3.08	-3.46	-3.78	-2.14	-2.65	-3.01	-3.33	-3.60

d) ADF^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.33	-3.88	-4.32	-4.64	-4.97	-3.45	-3.97	-4.35	-4.76	-5.09
2.5%	-3.03	-3.57	-3.97	-4.32	-4.68	-3.09	-3.60	-4.04	-4.41	-4.74
5.0%	-2.76	-3.32	-3.71	-4.07	-4.42	-2.82	-3.33	-3.74	-4.13	-4.46
7.5%	-2.59	-3.14	-3.53	-3.91	−4.25	-2.63	-3.16	-3.58	-3.94	-4.26
10.0%	-2.47	-3.02	-3.42	-3.80	-4.13	-2.50	-3.04	-3.44	-3.81	-4.13
15.0%	-2.27	-2.81	-3.24	-3.61	-3.94	-2.29	-2.84	-3.24	-3.60	-3.94
20.0%	-2.11	-2.65	-3.08	-3.45	-3.78	-2.13	-2.68	-3.08	-3.44	-3.77

Table 3. Percentage points of the tests under the null hypothesis ($c = 0$)
Case with $p = 1$ and $\bar{c} = 0$

a) MZ_{α}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-28.40	-37.16	-43.89	-52.83	-60.92	-26.17	-31.69	-37.06	-42.29	-46.14
2.5%	-23.49	-31.76	-37.58	-44.59	-51.81	-22.24	-27.81	-32.63	-37.47	-41.34
5.0%	-20.08	-27.56	-33.20	-40.13	-46.16	-18.87	-24.37	-28.96	-33.21	-37.23
7.5%	-18.23	-25.09	-30.79	-37.17	-42.61	-17.21	-22.42	-26.78	-30.97	-34.88
10.0%	-16.90	-23.18	-28.91	-34.92	-40.41	-15.91	-20.99	-25.34	-29.33	-33.27
15.0%	-14.78	-20.81	-25.96	-31.78	-37.12	-14.03	-18.82	-23.08	-26.76	-30.77
20.0%	-13.22	-18.89	-23.83	-29.23	-34.44	-12.56	-17.09	-21.40	-24.86	-28.70

b) MSB^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	0.133	0.116	0.107	0.097	0.091	0.138	0.126	0.116	0.109	0.104
2.5%	0.146	0.126	0.115	0.106	0.098	0.150	0.134	0.124	0.116	0.110
5.0%	0.158	0.135	0.123	0.111	0.104	0.163	0.143	0.131	0.123	0.116
7.5%	0.166	0.141	0.127	0.116	0.108	0.170	0.149	0.137	0.127	0.120
10.0%	0.172	0.147	0.132	0.119	0.111	0.177	0.154	0.140	0.131	0.123
15.0%	0.184	0.155	0.139	0.125	0.116	0.189	0.163	0.147	0.137	0.127
20.0%	0.194	0.163	0.145	0.131	0.121	0.200	0.171	0.153	0.142	0.132

c) MZ_{t}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.77	-4.31	-4.68	-5.14	-5.52	-3.62	-3.98	-4.31	-4.60	-4.80
2.5%	-3.43	-3.99	-4.33	-4.72	-5.09	-3.34	-3.73	-4.04	-4.33	-4.55
5.0%	-3.17	-3.71	-4.07	-4.48	-4.80	-3.07	-3.49	-3.81	-4.08	-4.31
7.5%	-3.02	-3.54	-3.92	-4.31	-4.62	-2.93	-3.35	-3.66	-3.94	-4.18
10.0%	-2.91	-3.40	-3.80	-4.18	-4.49	-2.82	-3.24	-3.56	-3.83	-4.08
15.0%	-2.72	-3.23	-3.60	-3.99	-4.31	-2.65	-3.07	-3.40	-3.66	-3.92
20.0%	-2.57	-3.07	-3.45	-3.82	-4.15	-2.51	-2.92	-3.27	-3.53	-3.79

d) ADF^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.77	-4.31	-4.68	-5.14	-5.52	-3.78	-4.23	-4.63	-5.00	-5.26
2.5%	-3.43	-3.99	-4.33	-4.72	-5.09	-3.46	-3.90	-4.28	-4.64	-4.93
5.0%	-3.17	-3.71	-4.07	-4.48	-4.80	-3.17	-3.63	-4.01	-4.36	-4.66
7.5%	-3.02	-3.54	-3.92	-4.31	-4.62	-3.00	-3.47	-3.84	-4.19	-4.50
10.0%	-2.91	-3.40	-3.80	-4.18	-4.49	-2.88	-3.34	-3.72	-4.05	-4.36
15.0%	-2.72	-3.23	-3.60	-3.99	-4.31	-2.69	-3.16	-3.53	-3.84	-4.17
20.0%	-2.57	-3.07	-3.45	-3.82	-4.15	-2.53	-2.99	-3.39	-3.70	-4.01

Table 4. Percentage points of the tests under the null hypothesis ($c = 0$)
Case with $p = 1$ and $\bar{c} = -13.5$

a) MZ_{α}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-31.81	-38.85	-46.38	-53.92	-63.83	-30.12	-34.63	-39.91	-44.65	-48.23
2.5%	-26.79	-33.93	-40.38	-47.04	-55.17	-26.12	-30.61	-35.00	-39.38	-43.70
5.0%	-22.82	-29.69	-35.67	-41.81	-49.29	-22.78	-26.98	-31.22	-35.37	-38.94
7.5%	-20.76	-27.19	-32.74	-38.48	-45.28	-20.56	-24.90	-29.06	-32.92	-36.51
10.0%	-19.29	-25.29	-30.90	-36.19	-42.64	-19.21	-23.47	-27.39	-31.12	-34.76
15.0%	-16.97	-22.58	-27.86	-33.13	-38.84	-17.18	-21.25	-25.15	-28.49	-32.18
20.0%	-15.48	-20.63	-25.64	-30.58	-36.04	-15.58	-19.63	-23.26	-26.60	-30.29

b) MSB^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	0.125	0.113	0.103	0.096	0.088	0.128	0.120	0.111	0.105	0.102
2.5%	0.136	0.121	0.111	0.103	0.095	0.137	0.127	0.119	0.112	0.107
5.0%	0.146	0128	0.117	0.109	0.100	0.146	0.135	0.126	0.118	0.113
7.5%	0.153	0.134	0.123	0.113	0.104	0.154	0.140	0.130	0.122	0.116
10.0%	0.159	0.139	0.126	0.117	0.108	0.159	0.144	0.134	0.126	0.119
15.0%	0.169	0.147	0.133	0.122	0.113	0.168	0.152	0.140	0.131	0.124
20.0%	0.177	0.154	0.138	0.127	0.117	0.176	0.158	0.145	0.136	0.128

c) MZ_{t}^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.97	-4.39	-4.79	-5.18	-5.63	-3.86	-4.14	-4.45	-4.71	-4.89
2.5%	-3.64	-4.09	-4.46	-4.83	-5.24	-3.58	-3.88	-4.16	-4.42	-4.66
5.0%	-3.35	-3.83	-4.20	-4.55	-4.94	-3.33	-3.65	-3.92	-4.18	-4.39
7.5%	-3.19	-3.66	-4.03	-4.36	-4.74	-3.18	-3.50	-3.78	-4.03	-4.25
10.0%	-3.07	-3.53	-3.90	-4.23	-4.59	-3.07	-3.40	-3.68	-3.92	-4.15
15.0%	-2.88	-3.33	-3.70	-4.04	-4.38	-2.90	-3.23	-3.51	-3.75	-3.99
20.0%	-2.74	-3.18	-3.55	-3.88	-4.22	-2.75	-3.10	-3.38	-3.62	-3.87

d) ADF^{GLS}

	$T=\infty$					$T=100$				
	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1.0%	-3.97	-4.39	-4.79	-5.18	-5.63	-4.03	-4.41	-4.77	-5.10	-5.32
2.5%	-3.64	-4.09	-4.46	-4.83	-5.24	-3.75	-4.08	-4.42	-4.76	-5.02
5.0%	-3.35	-3.83	-4.20	-4.55	-4.94	-3.44	-3.81	-4.14	-4.46	-4.74
7.5%	-3.19	-3.66	-4.03	-4.36	-4.74	-3.27	-3.64	-3.97	-4.29	-4.58
10.0%	-3.07	-3.53	-3.90	-4.23	-4.59	-3.13	-3.52	-3.85	-4.16	-4.45
15.0%	-2.88	-3.33	-3.70	-4.04	-4.38	-2.95	-3.33	-3.66	-3.94	-4.24
20.0%	-2.74	-3.18	-3.55	-3.88	-4.22	-2.80	-3.17	-3.50	-3.79	-4.09

Table 5. Size of residual-based tests for cointegration ($a_1 = 1$)
Model includes only a constant ($p = 0$)

	Test	$\sigma = 0.25$			$\sigma = 1$			$\sigma = 4$		
		$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.040	0.040	0.040	0.041	0.039	0.037	0.039	0.039	0.041
	Z_t	0.050	0.053	0.053	0.051	0.055	0.050	0.053	0.052	0.053
	ADF	0.052	0.054	0.054	0.051	0.055	0.053	0.053	0.052	0.053
	MZ_α	0.020	0.020	0.020	0.021	0.021	0.019	0.019	0.021	0.019
	$\widetilde{Z}_\alpha$	0.046	0.044	0.044	0.041	0.040	0.042	0.041	0.044	0.043
	$\widetilde{Z}_t$	0.055	0.052	0.052	0.053	0.051	0.052	0.050	0.053	0.051
	$\widetilde{ADF}$	0.055	0.052	0.052	0.054	0.051	0.052	0.051	0.053	0.051
	$\widetilde{M}Z_\alpha$	0.032	0.031	0.031	0.029	0.029	0.030	0.030	0.032	0.031
	$\overline{Z}_\alpha$	0.049	0.049	0.049	0.048	0.046	0.049	0.046	0.050	0.048
	$\overline{Z}_t$	0.063	0.062	0.062	0.057	0.060	0.062	0.060	0.060	0.057
	$\overline{ADF}$	0.061	0.061	0.061	0.058	0.058	0.060	0.058	0.059	0.056
	$\overline{M}Z_\alpha$	0.034	0.034	0.034	0.033	0.035	0.037	0.037	0.036	0.036
T=200	Z_α	0.049	0.048	0.048	0.049	0.050	0.052	0.051	0.047	0.043
	Z_t	0.056	0.053	0.053	0.056	0.057	0.055	0.058	0.052	0.048
	ADF	0.058	0.055	0.055	0.057	0.058	0.056	0.059	0.053	0.049
	MZ_α	0.035	0.033	0.034	0.033	0.035	0.033	0.033	0.033	0.027
	$\widetilde{Z}_\alpha$	0.046	0.048	0.048	0.051	0.050	0.050	0.050	0.048	0.045
	$\widetilde{Z}_t$	0.053	0.054	0.055	0.058	0.058	0.054	0.058	0.055	0.051
	$\widetilde{ADF}$	0.053	0.054	0.055	0.058	0.058	0.053	0.058	0.054	0.051
	$\widetilde{M}Z_\alpha$	0.042	0.044	0.043	0.046	0.043	0.043	0.043	0.043	0.039
	$\overline{Z}_\alpha$	0.048	0.046	0.046	0.053	0.052	0.051	0.052	0.047	0.046
	$\overline{Z}_t$	0.056	0.057	0.056	0.059	0.059	0.056	0.059	0.055	0.054
	$\overline{ADF}$	0.056	0.056	0.056	0.058	0.059	0.056	0.059	0.055	0.054
	$\overline{M}Z_\alpha$	0.044	0.043	0.043	0.047	0.047	0.044	0.046	0.042	0.040

Table 6. Size of residual-based tests for cointegration ($a_1 = 0$)
Model includes only a constant ($p = 0$)

	Test	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.037	0.036	0.043
	Z_t	0.050	0.046	0.054
	ADF	0.053	0.046	0.056
	MZ_α	0.019	0.022	0.018
	$\widetilde{Z}_\alpha$	0.042	0.044	0.043
	$\widetilde{Z}_t$	0.052	0.052	0.056
	$\widetilde{A}DF$	0.052	0.052	0.055
	$\widetilde{M}Z_\alpha$	0.030	0.030	0.030
	$\overline{Z}_\alpha$	0.049	0.048	0.045
	$\overline{Z}_t$	0.062	0.058	0.058
	$\overline{A}DF$	0.060	0.057	0.058
	$\overline{M}Z_\alpha$	0.037	0.035	0.036
T=200	Z_α	0.052	0.044	0.049
	Z_t	0.055	0.049	0.053
	ADF	0.056	0.050	0.053
	MZ_α	0.033	0.029	0.031
	$\widetilde{Z}_\alpha$	0.050	0.049	0.046
	$\widetilde{Z}_t$	0.054	0.056	0.054
	$\widetilde{A}DF$	0.053	0.056	0.053
	$\widetilde{M}Z_\alpha$	0.043	0.042	0.040
	$\overline{Z}_\alpha$	0.051	0.048	0.048
	$\overline{Z}_t$	0.056	0.056	0.056
	$\overline{A}DF$	0.056	0.056	0.055
	$\overline{M}Z_\alpha$	0.044	0.041	0.040

Table 7. Size of residual-based tests for cointegration ($a_1 = 1$)
Model includes a constant and a time trend ($p = 1$)

	Test	$\sigma = 0.25$			$\sigma = 1$			$\sigma = 4$		
		$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.040	0.040	0.040	0.037	0.036	0.038	0.037	0.043	0.039
	Z_t	0.060	0.059	0.059	0.059	0.056	0.062	0.057	0.062	0.059
	ADF	0.058	0.056	0.056	0.057	0.052	0.060	0.055	0.060	0.058
	MZ_α	0.014	0.013	0.013	0.014	0.013	0.017	0.015	0.020	0.017
	$\widetilde{Z}_\alpha$	0.036	0.036	0.036	0.037	0.036	0.040	0.038	0.039	0.042
	$\widetilde{Z}_t$	0.056	0.056	0.056	0.053	0.055	0.059	0.055	0.058	0.057
	$\widetilde{ADF}$	0.055	0.055	0.055	0.053	0.055	0.058	0.055	0.058	0.058
	$\widetilde{MZ}_\alpha$	0.026	0.026	0.026	0.024	0.024	0.024	0.023	0.024	0.028
	$\overline{Z}_\alpha$	0.045	0.048	0.048	0.045	0.043	0.041	0.041	0.041	0.043
	$\overline{Z}_t$	0.071	0.070	0.070	0.068	0.066	0.064	0.064	0.068	0.068
	$\overline{ADF}$	0.068	0.068	0.068	0.066	0.064	0.063	0.063	0.067	0.069
	$\overline{MZ}_\alpha$	0.023	0.023	0.023	0.023	0.022	0.023	0.024	0.025	0.027
T=200	Z_α	0.047	0.047	0.047	0.044	0.045	0.045	0.044	0.044	0.043
	Z_t	0.061	0.062	0.062	0.059	0.056	0.056	0.055	0.052	0.055
	ADF	0.061	0.062	0.062	0.060	0.056	0.056	0.056	0.051	0.053
	MZ_α	0.027	0.027	0.027	0.027	0.026	0.026	0.026	0.026	0.021
	$\widetilde{Z}_\alpha$	0.045	0.046	0.045	0.044	0.045	0.046	0.046	0.045	0.041
	$\widetilde{Z}_t$	0.055	0.054	0.054	0.052	0.053	0.055	0.055	0.052	0.051
	$\widetilde{ADF}$	0.055	0.054	0.054	0.052	0.053	0.055	0.055	0.052	0.052
	$\widetilde{MZ}_\alpha$	0.033	0.034	0.034	0.035	0.037	0.037	0.039	0.036	0.034
	$\overline{Z}_\alpha$	0.043	0.041	0.041	0.044	0.044	0.047	0.045	0.045	0.044
	$\overline{Z}_t$	0.053	0.051	0.051	0.054	0.054	0.058	0.054	0.059	0.053
	$\overline{ADF}$	0.053	0.051	0.051	0.054	0.053	0.057	0.054	0.058	0.053
	$\overline{MZ}_\alpha$	0.033	0.033	0.032	0.034	0.036	0.037	0.036	0.036	0.034

Table 8. Size of residual-based tests for cointegration ($a_1 = 0$)
Model includes a constant and a time trend ($p = 1$)

	Test	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.038	0.040	0.038
	Z_t	0.062	0.060	0.059
	ADF	0.060	0.060	0.060
	MZ_α	0.017	0.018	0.016
	$\widetilde{Z}_\alpha$	0.040	0.041	0.037
	$\widetilde{Z}_t$	0.059	0.060	0.057
	$\widetilde{ADF}$	0.058	0.060	0.057
	$\widetilde{MZ}_\alpha$	0.024	0.024	0.025
	$\overline{Z}_\alpha$	0.041	0.040	0.040
	$\overline{Z}_t$	0.064	0.067	0.064
	$\overline{ADF}$	0.063	0.066	0.064
	$\overline{MZ}_\alpha$	0.023	0.025	0.025
T=200	Z_α	0.045	0.042	0.041
	Z_t	0.056	0.056	0.051
	ADF	0.056	0.054	0.051
	MZ_α	0.026	0.025	0.022
	$\widetilde{Z}_\alpha$	0.046	0.042	0.044
	$\widetilde{Z}_t$	0.055	0.049	0.051
	$\widetilde{ADF}$	0.055	0.049	0.051
	$\widetilde{MZ}_\alpha$	0.037	0.036	0.037
	$\overline{Z}_\alpha$	0.047	0.045	0.041
	$\overline{Z}_t$	0.058	0.058	0.050
	$\overline{ADF}$	0.057	0.058	0.049
	$\overline{MZ}_\alpha$	0.037	0.034	0.033

Table 9. Power of residual-based tests for cointegration ($a_1 = 1$)
Model includes only a constant ($p = 0$)

	Test	$\sigma = 0.25$			$\sigma = 1$			$\sigma = 4$		
		$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.037	0.066	0.063	0.198	0.337	0.412	0.362	0.470	0.453
	Z_t	0.042	0.083	0.082	0.198	0.331	0.399	0.350	0.452	0.436
	ADF	0.043	0.084	0.082	0.199	0.331	0.401	0.350	0.449	0.434
	MZ_α	0.018	0.026	0.026	0.126	0.205	0.264	0.235	0.320	0.312
	$\widetilde{Z}_\alpha$	0.108	0.134	0.124	0.338	0.445	0.483	0.460	0.542	0.520
	$\widetilde{Z}_t$	0.128	0.162	0.154	0.392	0.509	0.565	0.530	0.613	0.592
	$\widetilde{ADF}$	0.127	0.161	0.154	0.390	0.509	0.566	0.530	0.611	0.591
	$\widetilde{MZ}_\alpha$	0.084	0.107	0.100	0.276	0.354	0.406	0.383	0.457	0.447
	$\overline{Z}_\alpha$	0.083	0.104	0.099	0.338	0.472	0.533	0.506	0.597	0.580
	$\overline{Z}_t$	0.094	0.136	0.124	0.393	0.534	0.614	0.571	0.670	0.661
	$\overline{ADF}$	0.093	0.135	0.124	0.391	0.533	0.612	0.570	0.668	0.658
	$\overline{MZ}_\alpha$	0.063	0.077	0.076	0.269	0.383	0.453	0.419	0.515	0.496
T=200	Z_α	0.069	0.129	0.118	0.411	0.593	0.682	0.627	0.742	0.721
	Z_t	0.066	0.121	0.118	0.349	0.519	0.611	0.554	0.675	0.654
	ADF	0.067	0.120	0.119	0.347	0.519	0.611	0.553	0.675	0.655
	MZ_α	0.051	0.094	0.078	0.347	0.512	0.608	0.554	0.676	0.660
	$\widetilde{Z}_\alpha$	0.228	0.277	0.271	0.591	0.721	0.774	0.750	0.813	0.797
	$\widetilde{Z}_t$	0.244	0.289	0.291	0.629	0.750	0.797	0.779	0.834	0.820
	$\widetilde{ADF}$	0.244	0.287	0.290	0.628	0.750	0.796	0.778	0.834	0.819
	$\widetilde{MZ}_\alpha$	0.210	0.253	0.248	0.557	0.687	0.739	0.710	0.781	0.769
	$\overline{Z}_\alpha$	0.192	0.243	0.233	0.598	0.735	0.804	0.779	0.840	0.830
	$\overline{Z}_t$	0.201	0.259	0.255	0.637	0.772	0.826	0.803	0.858	0.849
	$\overline{ADF}$	0.199	0.258	0.254	0.637	0.772	0.825	0.802	0.858	0.848
	$\overline{MZ}_\alpha$	0.177	0.220	0.213	0.568	0.703	0.771	0.737	0.811	0.804

Table 10. Power of residual-based tests for cointegration ($a_1 = 0$)
Model includes only a constant ($p = 0$)

	Test	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.417	0.477	0.420
	Z_t	0.402	0.465	0.398
	ADF	0.400	0.462	0.398
	MZ_α	0.274	0.331	0.280
	$\widetilde{Z}_\alpha$	0.501	0.549	0.496
	$\widetilde{Z}_t$	0.576	0.623	0.571
	$\widetilde{ADF}$	0.574	0.620	0.569
	$\widetilde{M}Z_\alpha$	0.415	0.466	0.423
	$\overline{Z}_\alpha$	0.556	0.608	0.546
	$\overline{Z}_t$	0.628	0.680	0.629
	$\overline{ADF}$	0.626	0.677	0.626
	$\overline{M}Z_\alpha$	0.467	0.523	0.464
T=200	Z_α	0.687	0.751	0.684
	Z_t	0.613	0.683	0.607
	ADF	0.612	0.683	0.606
	MZ_α	0.615	0.687	0.615
	$\widetilde{Z}_\alpha$	0.784	0.816	0.777
	$\widetilde{Z}_t$	0.807	0.837	0.799
	$\widetilde{ADF}$	0.806	0.837	0.799
	$\widetilde{M}Z_\alpha$	0.751	0.785	0.748
	$\overline{Z}_\alpha$	0.813	0.844	0.807
	$\overline{Z}_t$	0.840	0.861	0.828
	$\overline{ADF}$	0.839	0.861	0.828
	$\overline{M}Z_\alpha$	0.785	0.817	0.778

Table 11. Power of residual-based tests for cointegration ($a_1 = 1$)
Model includes a constant and a time trend ($p = 1$)

	Test	$\sigma = 0.25$			$\sigma = 1$			$\sigma = 4$		
		$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.027	0.043	0.026	0.078	0.157	0.208	0.176	0.267	0.263
	Z_t	0.041	0.064	0.049	0.110	0.213	0.270	0.225	0.322	0.327
	ADF	0.039	0.063	0.046	0.109	0.212	0.267	0.223	0.321	0.322
	MZ_α	0.012	0.010	0.009	0.028	0.060	0.083	0.069	0.113	0.113
	$\widetilde{Z}_\alpha$	0.041	0.050	0.042	0.142	0.209	0.260	0.242	0.316	0.301
	$\widetilde{Z}_t$	0.057	0.072	0.059	0.187	0.279	0.328	0.311	0.397	0.385
	$\widetilde{ADF}$	0.056	0.071	0.059	0.188	0.278	0.327	0.311	0.397	0.383
	$\widetilde{MZ}_\alpha$	0.027	0.028	0.025	0.090	0.137	0.178	0.159	0.213	0.214
	$\overline{Z}_\alpha$	0.032	0.044	0.030	0.125	0.211	0.275	0.251	0.353	0.354
	$\overline{Z}_t$	0.047	0.072	0.048	0.178	0.303	0.379	0.348	0.487	0.466
	$\overline{ADF}$	0.046	0.070	0.046	0.176	0.298	0.378	0.345	0.484	0.464
	$\overline{MZ}_\alpha$	0.016	0.021	0.019	0.064	0.124	0.171	0.148	0.223	0.226
T=200	Z_α	0.039	0.073	0.048	0.175	0.345	0.428	0.374	0.501	0.492
	Z_t	0.048	0.084	0.057	0.177	0.347	0.425	0.365	0.500	0.474
	ADF	0.048	0.085	0.059	0.177	0.347	0.426	0.366	0.499	0.474
	MZ_α	0.021	0.044	0.023	0.111	0.247	0.326	0.274	0.411	0.379
	$\widetilde{Z}_\alpha$	0.085	0.108	0.070	0.284	0.395	0.469	0.437	0.540	0.530
	$\widetilde{Z}_t$	0.104	0.126	0.085	0.321	0.443	0.512	0.487	0.581	0.568
	$\widetilde{ADF}$	0.104	0.126	0.085	0.320	0.441	0.511	0.487	0.581	0.568
	$\widetilde{MZ}_\alpha$	0.071	0.084	0.058	0.249	0.351	0.419	0.390	0.483	0.474
	$\overline{Z}_\alpha$	0.048	0.082	0.050	0.257	0.418	0.523	0.472	0.617	0.593
	$\overline{Z}_t$	0.061	0.102	0.063	0.296	0.472	0.592	0.522	0.676	0.656
	$\overline{ADF}$	0.061	0.100	0.063	0.295	0.471	0.591	0.520	0.676	0.655
	$\overline{MZ}_\alpha$	0.038	0.063	0.034	0.209	0.361	0.452	0.394	0.537	0.515

Table 12. Power of residual-based tests for cointegration ($a_1 = 0$)
Model includes a constant and a time trend ($p = 1$)

	Test	$\eta = -.5$	$\eta = 0$	$\eta = .5$
T=100	Z_α	0.220	0.282	0.228
	Z_t	0.274	0.346	0.284
	ADF	0.271	0.342	0.279
	MZ_α	0.092	0.120	0.091
	$\widetilde{Z}_\alpha$	0.277	0.318	0.274
	$\widetilde{Z}_t$	0.359	0.412	0.352
	$\widetilde{ADF}$	0.359	0.410	0.351
	$\widetilde{MZ}_\alpha$	0.184	0.220	0.190
	$\overline{Z}_\alpha$	0.309	0.373	0.307
	$\overline{Z}_t$	0.425	0.502	0.408
	$\overline{ADF}$	0.422	0.500	0.407
	$\overline{MZ}_\alpha$	0.186	0.233	0.193
T=200	Z_α	0.431	0.524	0.431
	Z_t	0.424	0.511	0.419
	ADF	0.424	0.511	0.418
	MZ_α	0.346	0.424	0.329
	$\widetilde{Z}_\alpha$	0.487	0.555	0.485
	$\widetilde{Z}_t$	0.529	0.592	0.537
	$\widetilde{ADF}$	0.529	0.592	0.537
	$\widetilde{MZ}_\alpha$	0.441	0.500	0.434
	$\overline{Z}_\alpha$	0.536	0.632	0.537
	$\overline{Z}_t$	0.605	0.687	0.608
	$\overline{ADF}$	0.604	0.687	0.607
	$\overline{MZ}_\alpha$	0.473	0.555	0.473

Searching for Additive Outliers In Nonstationary Time Series

Abstract

Recently, Vogelsang (1999) proposed a method to detect outliers which explicitly imposes the null hypothesis of a unit root. It works in an iterative fashion to select multiple outliers in a given series. We show, via simulations, that under the null hypothesis of no outliers, it has the right size in finite samples to detect a single outlier but when applied in an iterative fashion to select multiple outliers, it exhibits severe size distortions towards finding an excessive number of outliers. We show that his iterative method is incorrect and derive the appropriate limiting distribution of the test at each step of the search. Whether corrected or not, we also show that the outliers need to be very large for the method to have any decent power. We propose an alternative method based on first-differenced data that has considerably more power. The issues are illustrated using two US/Finland real-exchange rate series.

1 Introduction

From Fox (1972), who introduced the notion of additive and innovational outliers, issues related to this type of atypical observations in time series have received considerable attention in the statistics and econometric literature. The outlier detection issue, itself, has received particular attention [1]. Another topic of interest in the research has been the estimation of *ARMA* models in the presence of outliers. In this case, as mentioned by Chen and Liu (1993), a common approach is to identify the locations and the types of outliers and then to accommodate the effects of outliers using intervention models as proposed by Box and Tiao (1975). This approach requires iterations between stages of outlier detection and estimation of the model [2].

In the context of integrated data (processes with an autoregressive unit root), the effects of additive outliers have recently been the object of sustained research. It is by now well recognized that outliers affect the properties of unit root tests (e.g., Franses and Haldrup (1994)). They do so by inducing a negative moving average component in the noise function which causes most unit root tests to exhibit substantial size distortions towards rejecting the null hypothesis too often. Franses and Haldrup (1994) suggested applying Dickey-Fuller (1979) unit root tests by incorporating dummy variables in the autoregression chosen on the basis of the outlier detection procedure proposed by Chen

[1] See, e.g., Hawkins (1980) who presents a set of methods proposed before 1980 and Hawkins (1973) who proposed one of the most used methods, based on order statistics, to detect for outliers.

[2] Some references are Chang, Tiao and Chen (1988) and Tsay (1986). Chen and Liu (1993) also followed this way and they proposed another method to detect the locations of the outliers and the joint estimation of the parameters of the model. Their point of view was the fact that even if the model is well specified, outliers may still produce biased estimates of the parameters and, hence, may affect the outlier detection procedure. This is because atypical observations, in general, affect the variance of the estimates (e.g., Peña (1990)).

and Liu (1993). This procedure has been implemented in the computer program *TRAM* (Time Series Regression with *ARIMA* Noise and Missing Values) written by Gómez and Maravall (1992b), which lets us to estimate *ARIMA* models where missing observations may be treated as additive outliers.

In an interesting recent paper, Vogelsang (1999) makes two contributions to the issue about the effects of additive outliers on unit root tests. First, recognizing that outliers induce a negative moving average component, he suggests using unit root test developed by Stock (1990) and Perron and Ng (1996) that are robust, in terms of achieving exact size close to nominal size in small samples, even in the presence of a substantial negative moving average component. He shows via simulations that these unit root tests are little affected by systematic outliers. Secondly, he recognized that one can take advantage of the null hypothesis of a unit root in devising an outlier detection procedure. This allows the derivation of a non-degenerate limiting distribution for the t-statistic on the relevant one-time dummy.

In this paper, we make further contributions along the second suggestion of Vogelsang (1999). We show, via simulations, that Vogelsang's (1999) procedure, under the null hypothesis of no outliers, has the right size in finite samples to detect a single outlier but, when applied in an iterative fashion to select multiple outliers, it exhibits severe size distortions towards finding an excessive number of outliers. We show that an alternative estimate of the variance of the errors alleviates this problem but that there is a basic flaw in the iterative method suggested by Vogelsang (1999). In effect, contrary to what he implicitly assumes, the limiting distribution of the test used is different at each iteration of the outlier detection procedure. We derive the appropriate limiting distribution and tabulate some critical values. When so corrected, his method is shown to have very low power to detect outliers (even a single one without the correction made) unless the magnitude of the outlier is very large. As an alternative, we propose a method based on first-differenced data which has considerably more power. All of the methods considered are illustrated using two US/Finland real-exchange rate series.

This rest of the paper is organized as follows. Section 2 deals with the model and the issue of outlier detection. It reviews the procedure suggested by Vogelsang (1999) and suggests two simple modifications to the estimator of the variance of the residual function. Simulations about their size and power are also presented. Section 3 derives the correct limiting distribution of the test suggested by Vogelsang (1999) for each iteration of the outlier detection procedure. Section 4 presents the procedure based on first-differenced data and compares its size and power to methods based on levels of the data using simulations. An empirical illustration using two US/Finland real-exchange rate series is presented in Section 5. Section 6 presents brief concluding remarks and some details about the data used and technical derivations are put in appendices.

2 The model and the issue of outlier detection

There is a large literature in statistics and econometrics on the subject of outlier detection in *ARMA* models. The standard approach is to estimate a fully parameterized *ARMA* model and construct a t-statistic for the presence of an outlier. Such a t-statistic is constructed at all possible dates and the supremun is taken. The value of the supremun is then compared to a critical value to decide if an outlier is present. Some references are Tsay (1986), Chang, Tiao and Chen (1988), Shin, Sharkar and Lee (1996) and Chen and

Liu (1993). Using a time series with an $ARIMA$ noise function, Gómez and Maravall (1992a) proposed to analyze missing observations as additive outliers. This paper was the basis for the computer program TRAM written by Gómez and Maravall (1992b) to estimate $ARIMA$ models with missing observations, which was used by Franses and Haldrup (1994) in the context of outlier detection in time series with unit roots.

The issue of outlier detection in the unit root framework offers a distinct advantage, namely that one can work under the null hypothesis that a unit root is present. This is the approach taken by Vogelsang (1999) whose procedure has two useful features. First, it does not require a fully parametric model of the noise function and is valid for a wide class of processes. Second, an asymptotic distribution can be obtained and critical values tabulated even without having to make specific distributional and parametric assumptions about the data-generating process.

The data-generating process entertained is of the following general form:

$$y_t = d_t + \sum_{j=1}^{m} \delta_j D(T_{ao,j})_t + u_t \tag{1}$$

where $D(T_{ao,j})_t = 1$ if $t = T_{ao,j}$ and 0 otherwise. This permits the presence of m additive outliers occurring at dates $T_{ao,j}$ $(j = 1, ..., m)$. The term d_t specifies the deterministic components. In most cases, $d_t = \mu$ if the series is non-trending or $d_t = \mu + \beta t$ if the series is trending (of course, other specifications are possible). The noise function is integrated of order one, i.e.

$$u_t = u_{t-1} + v_t \tag{2}$$

where v_t can be, for example, a linear process of the form $v_t = \varphi(L)e_t$ with $\varphi(L) =$

$\sum_{i=0}^{\infty} \varphi_i L^i$ ($\sum_{i=0}^{\infty} i^2 \varphi_i^2 < \infty$) and e_t is a martingale difference sequence with mean 0 and $\sigma_e^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E(e_t^2)$ is finite. What is important is that the sequence v_t satisfies the condition for the application of a functional central limit theorem such that $T^{-1/2} \sum_{t=1}^{[Tr]} v_t \Rightarrow \sigma W(r)$ where $W(r)$ is the unit Wiener process, $\Rightarrow$ denotes weak convergence in distribution and $\sigma^2 = \lim_{T\to\infty} T^{-1} E(\sum_{t=1}^{T} v_t)^2$ with $0 < \sigma^2 < \infty$.

The detection procedure, suggested by Vogelsang (1999), starts with the following regression estimated by OLS (if necessary, a time trend can also be included),

$$y_t = \widehat{\mu} + \widehat{\delta} D(T_{ao})_t + \widehat{u}_t \tag{3}$$

where $D(T_{ao})_t = 1$ if $t = T_{ao}$ and 0 otherwise. Let $t_{\widehat{\delta}}(T_{ao})$ denote the t-statistic for testing $\delta = 0$ in (3). Following Chen and Liu (1993), the presence of an additive outlier can be tested using

$$\tau = \sup_{T_{ao}} \mid t_{\widehat{\delta}}(T_{ao}) \mid .$$

Assuming that $\lambda = T_{ao}/T$ remains fixed as T grows, Vogelsang (1999) showed that as $T \to \infty$,

$$t_{\widehat{\delta}}(T_{ao}) \Rightarrow H(\lambda) = \frac{W^*(\lambda)}{(\int_0^1 W^*(r)^2 dr)^{1/2}} \tag{4}$$

where $W^*(\lambda)$ denotes a demeaned standard Wiener process (i.e. $W^*(\lambda) = W(\lambda) - \int_0^1 W(s)ds$). If (3) also includes a time trend, $W^*(\lambda)$ will denote a detrended Wiener process. Furthermore, from the continuous mapping theorem it follows that,

$$\tau \Rightarrow \sup_{\lambda} |H(\lambda)| \equiv H^*. \tag{5}$$

The distribution given in (5) is non-standard but is invariant with respect to any nuisance parameters, including the correlation structure of the noise function. The asymptotic critical values for τ were tabulated using simulations. The Wiener processes were approximated by normalized sums of *i.i.d.* $N(0,1)$ random deviates using 1000 steps and $50,000$ replications. Two cases were considered according to the deterministic components included in (3). When there is an intercept in (3) the critical values are 3.5325, 3.1143 and 2.9151 at the 1, 5 and 10% significance levels, respectively. Finally, if a time trend is also included in (3) the corresponding critical values are 3.7252, 3.3088 and 3.1154.[3]

The outlier detection procedure recommended by Vogelsang (1999) is implemented as follows[4]. First, compute the τ statistic for the entire series and compare τ to the appropriate critical value. If τ exceeds the critical value, then an outlier is detected at date $\widehat{T}_{ao} = \arg\max_{T_{ao}} |t_{\widehat{\delta}}(T_{ao})|$. The outlier and the corresponding row of the regression is dropped and (3) is again estimated and tested for the presence of another outlier. This continues until the test shows a non-rejection.

An important element in the procedure is that, when applied in an iterative fashion, the sample used to calculate the variance of the residuals changes at each step. In

[3] Critical values were also tabulated for the case where no deterministic components are included in (3). The critical values at 1%, 5% and 10% significance levels are 3.2170, 2.8360 and 2.6527, respectively.

[4] This is equivalent to the *stepwise* procedures to select for multiple outliers. See Hawkins (1980).

particular, as more observations are labelled as outliers, the samples used keep only the observations with the least variance. This may be a desired feature to maximize power when there actually are outliers but can potentially lead to severe problems under the null hypothesis of no outlier. The reason is the following. Suppose that no outlier is present but that the procedure finds one at the first step, the variance of the residuals is computed without this observation. Since it is the largest, the estimated variance is biased downward which increases the probability of finding an outlier again. The downward bias in the estimate of the variance gets more severe as the procedure iterates.

To alleviate this potential problem, we consider a simple modification of the procedure suggested by Vogelsang (1999). It amounts to estimating the variance of the residuals using the estimated errors from the regression

$$y_t = \widehat{\mu} + \widehat{u}_t$$

or

$$y_t = \widehat{\mu} + \widehat{\beta} t + \widehat{u}_t$$

if a trend is present, at each step in the iterative procedure (including the first one). This should permit a better estimate of the variance of the residuals under the null hypothesis of no outlier and avoid finding an excessively number when using the full iterative procedure. This alternative procedure will be denoted as τ_1.

We also will consider another modification to estimate the variance of the residuals. It is to use the estimate of the variance of the residuals from the first step of Vogelsang's

(1999) procedure for all subsequent steps in the iterative procedure. This alternative procedure will be denoted τ_2.

2.1 Simulation experiments for size

To asses the properties of the various methods in finite samples, we performed simulation experiments under the hypothesis that the series contain no outlier. We consider a simple data-generating process with an autoregressive unit root, i.e.

$$y_t = y_{t-1} + u_t$$

Two cases are considered for the errors u_t; namely $MA(1)$ processes of the form $u_t = v_t + \theta v_{t-1}$ and $AR(1)$ processes of the form $u_t = \rho u_{t-1} + v_t$. In all cases, $v_t \sim$ *i.i.d.* $N(0,1)$. We consider values of θ and ρ in the range $[-0.8, 0.8]$ with a step size of 0.2. Two sample sizes are used, $T = 100$ and $T = 200$. The number of replications used was 1,000 and tests at the 5% and 10% significance levels were performed[5].

We first consider the size of the procedures in what we label the "one pass" case. That is the number of times an observation is categorized as an outlier when searching for a single outlier (without iterating any further for a given sample). The results are presented in Table 1. [6]

For the *i.i.d.* case, Vogelsang's method has an exact size close to nominal size and the procedure based on τ_1 shows some size distortions with the test being conservative,

[5] Results using 1% significance level were also tabulated and they are available upon request.
[6] Of course, the results for the procedure using τ_2 are identical to those in Vogelsang (1998) and, hence, omitted.

although these distortions are alleviated when the sample size increases. On the hand, when more deterministic components are included, the procedures are more conservative. For the case with negative moving average errors, both methods present size distortions (the tests being liberal) although these are smaller for the procedure using τ_1. Again, these distortions are higher when more deterministic components are included in the models. For positive moving average errors and particularly for the model that includes a time trend, both procedures are undersized. A similar result is observed when there are negative autoregressive errors.

The next experiments consider the properties of the methods when applied in a full iterative fashion, i.e. continuing to search for additional outliers when one is found. Here, we record the total number of observations categorized as outliers divided by the number of replications. These values can be labelled as the expected number of outliers found. If the tests have the correct size α, say, at each steps of the iterations, this number should be close to $\alpha/(1-\alpha)$, that is .111 for a significance level 10% and .053 for a significance level 5%.

The results for all three procedures are presented in Table 2. The main thing to note is that Vogelsang's procedure (τ) finds many more outliers than would be expected if the test had the correct size at each steps. For example, for the model with only a constant with *i.i.d.* errors, $T = 100$, and a significance level of 10%, the number is .354 instead of .111, i.e. an average of 3.54 outliers for each replication which contains at least one outlier. These distortions increase when T increases with a value of .552 (instead of .111) which corresponds to approximately 5.5 outliers per replications which have at least one outlier. When $T = 100$, these distortions are substantially reduced using the procedure based on τ_1. Here, the number is .100 (even lower than the expected value of .111). Nevertheless, when T increases to 200, it too shows some distortions (a value of .201 instead of .111) though much lower than for Vogelsang's procedure since now this represents approximately 2 outliers found per replications which have at least one outlier.

2.2 Simulation experiments for power

We now report results concerning the power of the procedures based on τ and τ_1 to detect outliers. The data-generating process considered is one with a single outlier affecting an integrated process, i.e.

$$
\begin{aligned}
y_t &= \delta D(T_{ao})_t + u_t, \\
u_t &= u_{t-1} + v_t.
\end{aligned}
$$

Again, two cases are considered for the errors v_t; namely $MA(1)$ processes of the form $v_t = e_t + \theta e_{t-1}$ and $AR(1)$ processes of the form $v_t = \rho v_{t-1} + e_t$. In all cases, $e_t \sim i.i.d.$ $N(0,1)$. We consider values of θ and ρ in the range $[-0.5, 0.5]$ with a step size of 0.5. The size, δ, considered for the magnitude of the outlier are 1, 5 and 10. The sample size is $T = 100$ and the outlier is located at mid-sample, i.e. $T_{ao} = 50$. The number of replications used was $1,000$ and tests at the 1%, 5% and 10% significance levels were performed. We tabulate the probability of finding at least one outlier, that is the power of the application of a single iteration for each procedure.

The results are presented in Table 3. There are two main features to note from these results. First, even though the size of the modified procedure τ_1 is below nominal size, its power is basically the same as that of Vogelsang's procedure (only slightly below). The second feature is that both procedures show decent power only when the outlier is very large, i.e. around 10 standard deviations (relative to the variance of the innovations). Thus, it appears that detecting outliers in series with an autoregressive unit root may be difficult [7]. A last feature of the results is that the power is higher when there are more deterministic components included in the estimated model.

3 The distribution of the test τ at different iterations

In the last section, we showed that the original procedure of Vogelsang (1999) has severe size distortions when applied in an iterative fashion to search for outliers. The reason for this is that the limiting distribution of the τ test given by (5) is only valid in the first step of the iteration. In subsequent steps, the asymptotic critical values used need to be modified. The correct limiting distribution at each step is given in the following Theorem.

Theorem 1 *Suppose that y_t is generated by (1) with $\delta_i = 0$ $(i = 1, ..., m)$ and let $\tau^{(i)}$ be the statistic τ obtained at step i of the iterative search for outliers, then*

$$\lim_{T\to\infty} \Pr[\tau^{(i)} > x] = \Pr[H^* > x]/\alpha^{i-1}$$

where α is the significance level of the test. Hence, the correct α-percentage point of the limiting distribution of $\tau^{(i)}$ is the α^i percentage point of the distribution of H^*.

[7] This low power may be related to the *"masking effect"* (see, for example, Hawkins (1980)).

P roof. The basic reason for this result is that at different steps the tests are not independent; indeed they are asymptotically equivalent because of the fact that the series is integrated. Hence, at each step $\tau^{(i)} \Rightarrow H^*$ unconditionally on what happened in the previous steps. But subsequent steps are applied only if the previous one showed a rejection, hence one must consider the limiting distribution conditional upon a rejection at the previous step. For simplicity, consider this limiting distribution for the second step. It is given by, where x_α is the α-percentage point of the distribution of H^*,

$$\lim_{T\to\infty} \Pr[\tau^{(2)} > x|\tau^{(1)} > x_\alpha] = \frac{\lim_{T\to\infty} \Pr[(\tau^{(2)} > x) \cap (\tau^{(1)} > x_\alpha)]}{\lim_{T\to\infty} \Pr[\tau^{(1)} > x_\alpha]}$$
$$= \frac{\lim_{T\to\infty} \Pr[(\tau^{(2)} > x) \cap (\tau^{(1)} > x_\alpha)]}{\alpha}$$

since $\tau^{(1)} \Rightarrow H^*$. Now, since we also have $\tau^{(2)} \Rightarrow H^*$,

$$\lim_{T\to\infty} \Pr[\tau^{(2)} > x|\tau^{(1)} > x_\alpha] = \frac{\Pr[(H^* > x) \cap (H^* > x_\alpha)]}{\alpha}$$
$$= \frac{\Pr[H^* > x]}{\alpha}$$

provided $x \geq x_\alpha$, which we shall need to have tests with correct sizes. The result stated in the theorem follows using further iterations of the same arguments. ■

We shall denote by τ_3 the iterative outlier detection procedure that uses the correct (and different) asymptotic critical values at different steps. We have simulated some asymptotic critical values. We approximate the Wiener process by normalized sums of *i.i.d.* $N(0,1)$ random variables. To obtain a fair range of critical values, we used 2 million replications. Nevertheless, even with such a large number of replications, the critical values for only a few cases can be obtained. This is because as we get further in the iterations of the outlier detection, we need percentage points of the distribution of H^* that are very far in the tail. For example, if the significance level is $\alpha = .05$, the percentage point needed at the 4th iteration is approximately .00001. Hence, even with 2 million replications we can only present critical values up to $i = 4$ for $\alpha = .05$, $i = 5$ for $\alpha = .10$, and $i = 7$ for $\alpha = .20$. These are presented in Table 4 [8].

4 A test using the first differences of the data

As discussed in Section 2, Vogelsang's procedure is not powerful unless the size of the outlier is very large. One can infer from this that the full corrected iterative procedure will be even less powerful since the critical values to be used at each iteration increase. Simulation evidence to that effect will presented in the next section. Hence, it is desirable to entertain an alternative outlier detecting procedure that is less likely to suffer from this low power problem.

We propose an iterative strategy using tests based on the first-differences of the data. Consider data generated by (1) with $d_t = \mu$, and a single outlier occurring at date T_{ao} with magnitude δ. Then,

$$\Delta y_t = \delta[D(T_{ao})_t - D(T_{ao})_{t-1}] + v_t, \tag{6}$$

where $D(T_{ao})_t = 1$, if $t = T_{ao}$ (0 otherwise) and $D(T_{ao})_{t-1} = 1$, if $t = T_{ao} - 1$ (0 otherwise).

[8]Note that the critical values with $i = 1$ are not quite identical to those presented in Section 2 of this paper or in Vogelsang (1998) since 200 instead of 1000 steps where used to approximate the Weiner process. The differences, however, are minor and do not affect subsequent results.

This reflects the fact that a unit root process with an outlier is characterized in first-differences by two successive outliers of equal magnitude but with opposite signs. Let $t_{\hat{\delta}}(T_{ao})$ denote the t-statistic for testing $\delta = 0$ in (6) estimated by OLS (if the data are trending a constant should be included). Then, the presence of an additive outlier can be tested using the statistic

$$\tau_4 = \sup_{T_{ao}} |t_{\hat{\delta}}(T_{ao})|. \tag{7}$$

To detect multiple outliers, we can follow a strategy similar to that suggested by Vogelsang (1999), by dropping the observation labelled as an outlier before proceeding to the next step. The disadvantage of this procedure, compared to that based on the level of the data, is that the limiting distribution of the test τ_4 not only depends on the presence of serial correlation and heteroskedasticity in the errors v_t but also on its specific distribution. This problem is exactly the same as that for finding outliers in stationary time series since by differencing we effectively work with a stationary series. Nevertheless, following standard practice in the literature, we shall simulate critical values assuming *i.i.d.* normal errors and assess the extent to which inference is affected when the data deviates from these specifications. So the data generating process is again

$$y_t = y_{t-1} + u_t \tag{8}$$

where $u_t \sim N(0,1)$. Two samples sizes are considered, namely $T = 100$ and $T = 200$. The number of replications used was 5,000. The percentage points of the test τ_4 are presented in Table 5.

4.1 Simulations for size and power

In this section, we present results about the size and, especially, the power of the various procedures when multiple outliers are present. The Data Generating Process considered is

$$y_t = \sum_{j=1}^{m} \delta_j D(T_{ao,j})_t + u_t, \tag{9}$$

$$u_t = u_{t-1} + v_t, \tag{10}$$

where $v_t \sim N(0,1)$ and $D(T_{ao,j}) = 1$ if $t = T_{ao,j}$ and 0 otherwise. This permits the presence of m additive outliers occurring at dates $T_{ao,j}$ $(j = 1, ..., m)$ and we consider up to $m = 4$ outliers. All simulations are based on a sample size $T = 100$ and $2,000$ replications were performed. We present results only for the case where a constant is included in the set of deterministic components. The significance level of the test is set to 5%. For the procedures based on τ, τ_1 and τ_2, the asymptotic critical value stated in Section 2 was used (this value is 3.1143). Note that these use the same critical value at each iteration of the outlier detection procedure and are, therefore, not valid procedures given the results of Section 3. We, nevertheless, present results for these procedures to asses how the need to use different critical values at different iterations has an impact on power. For the procedures based on τ_3 and τ_4, we used the asymptotic critical values presented in Tables 4 and 5, respectively.

The locations of the outliers are specified as follows. When there is one outlier, it is located at date 50; when there are two, they are located at dates 25 and 75; with 3 outliers, the dates are $25, 50$, and 75; and with four the dates are $20, 40, 60$, and 80. The magnitudes of the outliers considered are $\delta = 0, 2, 3, 5$. When $\delta = 0$, the data-generating process is equivalent to a data-generating process without additive outliers, which is useful to see if the procedures have the correct size. The results are presented in Table 6.

Consider first the behavior of the test when there is no outlier. All procedures considered have the correct size (except for τ_1 which has an exact size slightly below nominal size) at the first step of the outlier detection procedure. Consistent with the results of the previous section, Vogelsang's (1999) method based of the τ test shows a probability of finding an excessive number of outliers in subsequent steps, even as far as the 4th iteration. When the correct critical values are used in the different steps (method τ_3), these distortions disappear. The procedure based on first-differenced data (τ_4) show no size distortion and the probability of finding more than one outlier is basically null.

Consider now the power properties. The results for all the configurations considered all point to the same conclusion. Any method based on the level of the data has basically no power while the method based on first-differenced data has excellent power even for outliers of moderate size. Consider for example, the case with a single outlier of size 3. The power of Vogelsang's (1999) method (appropriately corrected, τ_3) is .08 while the method based on first differenced data (τ_4) has power .75. When the magnitude of the outlier increases to 5, the power of Vogelsang's (1999) method increases only to .15 while that based on first-differenced data is 1.00. The same important power differences remain for other cases with multiple outliers.

5 Empirical applications

The procedures analyzed in the last sections were applied to two series of real-exchange rates for US/Finland. The first series covers the period 1900-1988 and it is constructed using the Consumption Price Index (*CPI*). The other series spans the years 1900-1987 and is constructed using the Gross Domestic Product (*GDP*) deflator. The series are shown in Figures 1 and 2, respectively. These are the same series used by Vogelsang (1999), Franses and Haldrup (1994) and Perron and Vogelsang (1992) and are described in more details in Appendix A.

Franses and Haldrup (1994) used the *TRAM* program (Time Series Regression With *ARIMA* Noise and Missing Values) written by Gómez and Maravall (1992b) to search for outliers in these two real-exchange rate series. They considered two types of outliers, additive outliers and outliers that produce temporary changes, denoted *AO* and *TC* outliers, respectively. For the US/Finland real-exchange rate series based on the *CPI* index, they found four additive outliers at dates 1918, 1922, 1945 and 1948. The observations associated with the years 1917, 1932 and 1949 were found to be outliers that produce temporary changes (*TC* outliers). For the US/Finland real-exchange rate series based on the *GDP* deflator, an additive outlier was found only at date 1918, whereas outliers that produce temporary changes were found at dates 1917, 1932, 1949 and 1957.

Table 7 reports the empirical results from applying the procedures discussed in this paper using 5% and 10% significance levels. Vogelsang (1999) presents results for additive outliers only for the US/Finland real-exchange rate series based on the *CPI* index. The dates he found (using the procedure τ) were 1917-1919, 1921 and 1932 (these are reproduced in Table 7). Consistent with our simulation experiments, the procedures based on τ_1 and τ_2 find less outliers. Of course, all three methods are not theoretically valid in the context of multiple outliers detection. When appropriately corrected, Vogelsang's (1999) method find outliers only for the year 1918 at the 5% level and for 1918 and 1919 at the 10% level, illustrating the fact that when it is not corrected it tends to select more outliers than warranted. The procedure based on first-differenced data (τ_4) finds outliers at dates 1917, 1918, 1919, 1932 and 1948 at both the 5% and 10% significance levels. This illustrates how this latter method is more powerful.

For the US/Finland real-exchange rate series based on the *GDP* deflator, all methods based on the level of the data find no outlier. As mentioned by Vogelsang (1999), this may be due to the presence of a shift in the mean of the series as documented by Perron and Vogelsang (1992). The procedure based on first-differenced data (τ_4) is, nevertheless, able to identify the years 1918 and 1948 as outliers. These two dates are not associated with the change in mean identified by Perron and Vogelsang (1992) as occurring in 1937. The fact that our procedure identifies the year 1918 as an outlier is comforting since visual inspection clearly points in that direction.

6 Conclusions

We analyzed in this paper the size and power properties of some test procedures for multiple outliers in series with an autoregressive unit root. We showed, via simulations, that the procedure suggested by Vogelsang has indeed the right size when applied to detect a single outlier but that it finds an excessive number of outliers when applied in an iterative fashion. We showed that a simple modification based on a different estimate of the variances of the residuals can alleviate this problem in finite sample. However, both methods were also shown to be theoretically incorrect and we derived the appropriate limiting distribution for each step of the iterations. We also showed that, whether corrected or not such outlier detection methods based on the level of the data have very low power unless the magnitude of the outliers is unrealistically large. Our suggestion was to use a procedure based on first-differenced data which was shown to have considerably more power.

Of course, the method based on first-differenced data has a drawback common to most outlier detection methods for stationary time series, namely the fact that the asymptotic distribution of the test depends on the nature of the serial correlation and the distribution of the data. Hence, one needs to resort to some ad hoc method to select critical values. We have done so using the base case of independent and identically normally distributed errors. This drawback is not shared by Vogelsang's method based on the level of the data, but the price in terms of power loss seems to big to argue in its favor.

7 Appendix A: The Data

The US/Finland real-exchange rate series based on the *CPI* index and the *GDP* deflator were kindly provided by Tim Vogelsang. They are the same series used in Vogelsang (1999), Franses and Haldrup (1994) and Perron and Vogelsang (1992). The US/Finland real-exchange rate series based on the *CPI* index is annual from 1900 to 1988, whereas that based on the *GDP* deflator is from 1900 to 1987. The details of the sources is as follows (see appendix A of Perron and Vogelsang (1992)): Nominal exchange rate series —1900-1988 from the Bank of Finland; *CPI* —1900-1985 from the Bank of Finland, 1986-1988 from the *IMF* (1988); *GDP* deflator —1900-1985 from the Bank of Finland, 1986-1987 from *IMF* (1988). The sources of the U.S. data are: for the *GNP* deflator —1869-1975 from Friedman and Schwartz (1982), 1976-1988 from *IMF* (1988); for the *CPI* —1860-1970 from the U.S. Bureau of the Census (1976) and 1971-1988 from *IMF* (1988).

8 Appendix B: Mathematical Derivations

Vogelsang (1999) showed the proof of the limiting expression found in (4) for the general case. Here, for convenience of the reader, we include the proofs related to the two particular cases analyzed in this paper according to the set of deterministic components included in the regression to detect for additive outliers.

Consider first the case where only an intercept is included in the regression used to detect for additive outliers. The regression estimated by *OLS* is

$$y_t = \widehat{\mu} + \widehat{\delta} D(T_{ao})_t + \widehat{u}_t. \tag{A.1}$$

The t-statistic for testing the null hypothesis that $\delta = 0$ can be expressed as

$$t_{\widehat{\delta}} = \frac{\widehat{\delta}}{(s^2[x'x]_{22}^{-1})^{1/2}} \tag{A.2}$$

where $(x'x)_{22}^{-1}$ is the element $(2,2)$ of the inverse matrix of $(x'x)$. Here $x = [x_1, ..., x_T]'$ with $x_t = (1, D(T_{ao})_t)$. From standard *OLS* formulae we have:

$$\begin{aligned}
\begin{bmatrix} \widehat{\mu} \\ \widehat{\delta} \end{bmatrix} &= \begin{bmatrix} T & \sum_{t=1}^{T} D(T_{ao})_t \\ \sum_{t=1}^{T} D(T_{ao})_t & \sum_{t=1}^{T} D(T_{ao})_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} D(T_{ao})_t y_t \end{bmatrix} \\
&= \begin{bmatrix} T & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=T_{ao}}^{T} y_t \end{bmatrix}
\end{aligned} \tag{A.3}$$

Let $\Psi_1 = diag(T^{-1/2}, T^{-1/2})$ and $\Psi_2 = diag(T^{-3/2}, T^{-1/2})$ be two standardizing matrices and

$\widehat{\psi} = (\widehat{\mu}, \widehat{\delta})'$. We can then write expression (A.3) as

$$\Psi_1 \widehat{\psi} = \Psi_1 \begin{bmatrix} T & 1 \\ 1 & 1 \end{bmatrix}^{-1} \Psi_2^{-1} \Psi_2 \begin{bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=[T_{ao}]}^{T} y_t \end{bmatrix} \tag{A.4}$$

Assuming that $\lambda = T_{ao}/T$, we have, after some algebra,

$$\begin{bmatrix} T^{-1/2}\widehat{\mu} \\ T^{-1/2}\widehat{\delta} \end{bmatrix} \Rightarrow \begin{bmatrix} \sigma \int_0^1 W(r)dr \\ \sigma\{W(\lambda) - \int_0^1 W(r)dr\} \end{bmatrix} \tag{A.5}$$

From expression (A.3), we have that $(x'x)_{22}^{-1} = 1$. We need to find the limit of:

$$T^{-1}s^2 = T^{-1} \sum_{t=1}^{T} [y_t - \widehat{\mu} - \widehat{\delta} D(T_{ao})_t]^2 \tag{A.6}$$

After some algebra, we have:

$$\begin{aligned} T^{-1}s^2 &\Rightarrow \sigma^2 \int_0^1 [W(r) - \int_0^1 W(r)dr]^2 \\ &= \sigma^2 \int_0^1 [W^{\mu}(r)]^2 dr \end{aligned} \tag{A.7}$$

The proof follows substituting expressions (A.7) and (A.5) in (A.2).

Consider now the case where a time trend is also included in the regression used to detect for additive outliers. The estimated *OLS* regression is

$$y_t = \widehat{\mu} + \widehat{\beta} t + \widehat{\delta} D(T_{ao})_t + \widehat{u}_t. \tag{A.8}$$

The t-statistic for testing the null hypothesis that $\delta = 0$ can be expressed as

$$t_{\widehat{\delta}} = \frac{\widehat{\delta}}{(s^2[x'x]_{33}^{-1})^{1/2}} \tag{A.9}$$

where $(x'x)_{33}^{-1}$ is the element $(3,3)$ of the inverse of $(x'x)$ where now $x_t = (1, t, D(T_{ao})_t)$. From standard *OLS* formulae we have:

$$\begin{bmatrix} \widehat{\mu} \\ \widehat{\beta} \\ \widehat{\delta} \end{bmatrix} = \begin{bmatrix} T & \sum_{t=1}^{T} t & \sum_{t=1}^{T} D(T_{ao})_t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 & \sum_{t=1}^{T} tD(T_{ao})_t \\ \sum_{t=1}^{T} D(T_{ao})_t & \sum_{t=1}^{T} tD(T_{ao})_t & \sum_{t=1}^{T} D(T_{ao})_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} ty_t \\ \sum_{t=1}^{T} D(T_{ao})_t y_t \end{bmatrix}$$

$$= \begin{bmatrix} T & \sum_{t=1}^{T} t & 1 \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 & \sum_{t=T_{ao}}^{T} t \\ 1 & \sum_{t=T_{ao}}^{T} t & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} t y_t \\ \sum_{t=T_{ao}}^{T} y_t \end{bmatrix} \tag{A.10}$$

Let $\Psi_1 = diag(T^{-1/2}, T^{1/2}, T^{-1/2})$ and $\Psi_2 = diag(T^{-3/2}, T^{-5/2}, T^{-1/2})$ be two standardizing matrices and $\widehat{\psi} = (\widehat{\mu}, \widehat{\beta}, \widehat{\delta})'$. We can then write expression (A.10) as

$$\Psi_1 \widehat{\psi} = \Psi_1 \begin{bmatrix} T & \sum_{t=1}^{T} t & 1 \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 & \sum_{t=T_{ao}}^{T} t \\ 1 & \sum_{t=T_{ao}}^{T} t & 1 \end{bmatrix}^{-1} \Psi_2^{-1} \Psi_2 \begin{bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} t y_t \\ \sum_{t=T_{ao}}^{T} y_t \end{bmatrix} \tag{A.11}$$

Assuming that $\lambda = T_{ao}/T$, we have, after some algebra,

$$\begin{bmatrix} T^{-1/2}\widehat{\mu} \\ T^{1/2}\widehat{\beta} \\ T^{-1/2}\widehat{\delta} \end{bmatrix} \Rightarrow \begin{bmatrix} \sigma\{4\int_0^1 W(r)dr - 6\int_0^1 rW(r)dr\} \\ \sigma\{-6\int_0^1 W(r)dr + 12\int_0^1 rW(r)dr\} \\ \sigma\{(6\lambda - 4)\int_0^1 W(r)dr + (6 - 12\lambda)\int_0^1 rW(r)dr + W(\lambda)\} \end{bmatrix} \tag{A.12}$$

From expression (A.10), we have $(x'x)_{33}^{-1} = 1$. We need to find the limit of

$$T^{-1}s^2 = T^{-1}\sum_{t=1}^{T}[y_t - \widehat{\mu} - \widehat{\beta}t - \widehat{\delta}D(T_{ao})_t]^2 \tag{A.13}$$

After some algebra, we have:

$$\begin{aligned} T^{-1}s^2 &\Rightarrow \sigma^2 \int_0^1 \{W(r) - (4 - 6s)\int_0^1 W(r)dr - (12s - 6)\int_0^1 rW(r)dr\}^2 \\ &= \sigma^2 \int_0^1 [W^{\tau}(r)]^2 dr \end{aligned} \tag{A.14}$$

The proof follows substituting expressions (A.14) and (A.12) in (A.9).

References

[1] Box, G.E.P. and Tiao, G. C. (1975), "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of the American Statistical Association* **70**, 70-79.

[2] Chang, I., Tiao, G. C. and Chen, C. (1988), "Estimation of Time Series Parameters in the Presence of Outliers," *Technometrics* **30**, 193-204.

[3] Chen, C. and L. Liu (1993), "Joint Estimation of Model Parameters and Outlier Effects in Time Series," *Journal of the American Statistical Association* **74**, 427-431.

[4] 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* **12**, 471-478.

[5] Fox, A. J. (1972), "Outliers in Time Series," *Journal of the Royal Statistical Association* **74**, 140-146.

[6] Franses, P. H. and N. Haldrup (1994), "The Effects of Additive Outliers on Tests for Unit Roots and Cointegration," *Journal of Business & Economic Statistics* **12**, 471-478.

[7] Friedman, M. and Schwartz, A. J. (1982), *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices and Interest Rates, 1867-1975*, Chicago: The University of Chicago Press.

[8] Gómez, V. and A. Maravall (1992a), "Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter," European University Institute, Working Paper ECO 92/80.

[9] Gómez, V. and A. Maravall (1992b), "Time Series Regression with ARIMA Noise and Missing Observations. Program TRAM," European University Institute, Working Paper ECO 92/81.

[10] Hawkins, D. M. (1973), "Repeated Testing for Outliers," *Statistica Neerlandica*, **27**, 1-10.

[11] Hawkins, D. M. (1980), *Identification of Outliers*, New York: Chapman and Hall.

[12] International Monetary Fund (1988), *International Financial Statistics: Yearbook*, Washington, DC.

[13] Peña, D. (1990), "Influential Observations in Time Series," *Journal of Business & Economic Statistics* **8**, 235-241.

[14] Perron, P. and S. Ng (1996), "Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties," *Review of Economic Studies* **63**, 435-463.

[15] Perron, P. and T. J. Vogelsang (1992), "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity," *Journal of Business and Economic Statistics* **10**, 301-320.

[16] Shin, D. W., S. Sarkar and J. H. Lee (1996), "Unit Root Tests for Time Series with Outliers," *Statistics and Probability Letters* **30**, 189-197.

[17] Stock, J. H. (1990), "A Class of Tests for Integration and Cointegration", unpublished manuscript, Kennedy School of Government, Harvard University.

[18] Tsay, R. S. (1986), "Time Series Model Specification in the Presence of Outliers," *Journal of the American Statistical Association* **81**, 132-141.

[19] U.S. Bureau of the Census (1976), *The Statistical History of the United States, From Colonial Times to the Present*, New York: Basic Books.

[20] Vogelsang, T. J. (1999), "Two Simple Procedures for Testing for a Unit Root when there are Additive Outliers," *Journal of Time Series Analysis* **20 (2)**, 237-252.

Annex 1

Table 1. Exact size of single outlier detection methods

a) MA case

		constant only				constant and time trend			
		5.0%		10.0%		5.0%		10.0%	
T	θ	Test τ_1	Test τ	Test τ_1	Test τ	Test τ_1	Test τ	Test τ_1	Test τ
100	-0.80	0.104	0.200	0.240	0.352	0.084	0.157	0.178	0.278
	-0.60	0.059	0.103	0.127	0.197	0.056	0.114	0.128	0.202
	-0.40	0.041	0.066	0.095	0.145	0.033	0.077	0.087	0.145
	-0.20	0.028	0.055	0.075	0.125	0.025	0.058	0.061	0.112
	0.00	0.026	0.049	0.064	0.104	0.017	0.047	0.057	0.093
	0.20	0.024	0.045	0.056	0.091	0.020	0.036	0.045	0.082
	0.40	0.025	0.044	0.053	0.088	0.017	0.036	0.040	0.073
	0.60	0.022	0.042	0.052	0.088	0.017	0.036	0.037	0.068
	0.80	0.019	0.041	0.053	0.087	0.018	0.035	0.041	0.067
200	-0.80	0.173	0.232	0.356	0.411	0.125	0.183	0.276	0.348
	-0.60	0.071	0.097	0.149	0.186	0.063	0.095	0.146	0.187
	-0.40	0.049	0.061	0.104	0.123	0.042	0.062	0.092	0.112
	-0.20	0.039	0.053	0.082	0.102	0.025	0.035	0.060	0.086
	0.00	0.037	0.046	0.076	0.091	0.018	0.028	0.049	0.075
	0.20	0.037	0.046	0.067	0.087	0.014	0.023	0.049	0.069
	0.40	0.037	0.046	0.067	0.083	0.014	0.023	0.045	0.067
	0.60	0.037	0.045	0.067	0.080	0.014	0.020	0.043	0.063
	0.80	0.037	0.045	0.067	0.080	0.013	0.020	0.041	0.061

Table 1 (cont'd). Exact size of single outlier detection methods

b) AR case

		Constant only				Constant and time trend			
		5.0%		10.0%		5.0%		10.0%	
T	ρ	Test τ_1	Test τ	Test τ_1	Test τ	Test τ_1	Test τ	Test τ_1	Test τ
100	-0.80	0.039	0.071	0.095	0.138	0.033	0.070	0.075	0.132
	-0.60	0.037	0.068	0.088	0.129	0.026	0.067	0.076	0.130
	-0.40	0.035	0.054	0.079	0.126	0.027	0.060	0.069	0.121
	-0.20	0.027	0.055	0.069	0.119	0.023	0.056	0.058	0.109
	0.20	0.024	0.046	0.054	0.088	0.016	0.036	0.046	0.078
	0.40	0.017	0.040	0.046	0.074	0.015	0.036	0.040	0.065
	0.60	0.014	0.031	0.039	0.059	0.011	0.029	0.032	0.053
	0.80	0.014	0.021	0.027	0.051	0.015	0.027	0.028	0.039
200	-0.80	0.057	0.074	0.120	0.153	0.037	0.067	0.107	0.142
	-0.60	0.052	0.063	0.098	0.120	0.035	0.051	0.079	0.115
	-0.40	0.044	0.054	0.089	0.111	0.030	0.045	0.066	0.098
	-0.20	0.039	0.049	0.081	0.097	0.021	0.034	0.056	0.086
	0.20	0.036	0.045	0.068	0.083	0.014	0.023	0.047	0.066
	0.40	0.033	0.041	0.060	0.078	0.012	0.017	0.036	0.053
	0.60	0.027	0.038	0.050	0.064	0.010	0.016	0.032	0.045
	0.80	0.021	0.029	0.041	0.047	0.008	0.016	0.027	0.039

Table 2. Expected number of outliers found using multiple outliers detection methods

a) MA case

		Constant only					
		5.0%			10.0%		
T	θ	Test τ_1	Test τ_2	Test τ	Test τ_1	Test τ_2	Test τ
100	-0.80	0.108	0.213	0.232	0.257	0.405	0.474
	-0.60	0.060	0.118	0.144	0.148	0.257	0.317
	-0.40	0.046	0.089	0.112	0.127	0.206	0.332
	-0.20	0.034	0.082	0.125	0.108	0.198	0.342
	0.00	0.036	0.080	0.128	0.100	0.184	0.354
	0.20	0.034	0.083	0.132	0.097	0.175	0.354
	0.40	0.035	0.081	0.144	0.099	0.169	0.338
	0.60	0.033	0.082	0.142	0.100	0.169	0.361
	0.80	0.029	0.079	0.139	0.099	0.174	0.359
200	-0.80	0.189	0.266	0.286	0.440	0.550	0.652
	-0.60	0.090	0.139	0.204	0.228	0.311	0.454
	-0.40	0.081	0.113	0.219	0.201	0.269	0.474
	-0.20	0.082	0.124	0.280	0.202	0.262	0.498
	0.00	0.086	0.118	0.286	0.201	0.265	0.552
	0.20	0.089	0.122	0.286	0.201	0.266	0.544
	0.40	0.089	0.124	0.289	0.204	0.265	0.541
	0.60	0.089	0.123	0.291	0.208	0.262	0.544
	0.80	0.089	0.126	0.293	0.205	0.266	0.557

Table 2 (cont'd). Expected number of outliers found using multiple outliers detection methods

a) MA case

		Constant and time trend					
		5.0%			10.0%		
T	θ	Test τ_1	Test τ_2	Test τ	Test τ_1	Test τ_2	Test τ
100	-0.80	0.084	0.165	0.177	0.185	0.299	0.335
	-0.60	0.056	0.116	0.123	0.131	0.217	0.240
	-0.40	0.034	0.081	0.092	0.091	0.165	0.195
	-0.20	0.026	0.071	0.092	0.076	0.148	0.211
	0.00	0.020	0.062	0.098	0.076	0.146	0.226
	0.20	0.028	0.060	0.087	0.072	0.145	0.229
	0.40	0.023	0.060	0.098	0.065	0.129	0.225
	0.60	0.023	0.063	0.096	0.063	0.128	0.212
	0.80	0.023	0.061	0.104	0.069	0.124	0.212
200	-0.80	0.131	0.200	0.214	0.312	0.413	0.464
	-0.60	0.068	0.108	0.121	0.173	0.228	0.277
	-0.40	0.051	0.078	0.105	0.128	0.162	0.246
	-0.20	0.035	0.050	0.089	0.096	0.155	0.273
	0.00	0.030	0.045	0.075	0.089	0.152	0.301
	0.20	0.024	0.040	0.079	0.094	0.156	0.324
	0.40	0.023	0.043	0.085	0.091	0.160	0.324
	0.60	0.022	0.040	0.083	0.092	0.156	0.314
	0.80	0.020	0.039	0.071	0.085	0.159	0.323

Table 2 (cont'd). **Expected number of outliers found using multiple outliers detection methods**

b) AR case

		Constant only					
		5.0%			10.0%		
T	ρ	Test τ_1	Test τ_2	Test τ	Test τ_1	Test τ_2	Test τ
100	-0.80	0.045	0.088	0.115	0.115	0.195	0.289
	-0.60	0.043	0.084	0.132	0.111	0.186	0.286
	-0.40	0.043	0.075	0.113	0.108	0.191	0.352
	-0.20	0.035	0.082	0.125	0.100	0.194	0.349
	0.20	0.034	0.084	0.136	0.096	0.170	0.340
	0.40	0.028	0.076	0.142	0.094	0.156	0.335
	0.60	0.018	0.062	0.145	0.083	0.151	0.322
	0.80	0.020	0.044	0.117	0.056	0.140	0.308
200	-0.80	0.083	0.122	0.196	0.203	0.289	0.471
	-0.60	0.091	0.116	0.223	0.198	0.261	0.523
	-0.40	0.081	0.114	0.247	0.207	0.271	0.499
	-0.20	0.081	0.121	0.293	0.203	0.265	0.506
	0.20	0.084	0.119	0.287	0.204	0.268	0.530
	0.40	0.084	0.122	0.294	0.196	0.270	0.549
	0.60	0.089	0.121	0.286	0.184	0.258	0.556
	0.80	0.071	0.109	0.272	0.182	0.229	0.531

Table 2 (cont'd). **Expected number of outliers found using multiple outliers detection methods**

b) AR case

		Constant and time trend					
		5.0%			10.0%		
T	ρ	Test τ_1	Test τ_2	Test τ	Test τ_1	Test τ_2	Test τ
100	-0.80	0.038	0.082	0.095	0.087	0.158	0.183
	-0.60	0.028	0.074	0.083	0.084	0.148	0.189
	-0.40	0.029	0.065	0.092	0.074	0.148	0.212
	-0.20	0.025	0.069	0.094	0.072	0.145	0.220
	0.20	0.022	0.062	0.094	0.073	0.142	0.230
	0.40	0.020	0.062	0.115	0.068	0.131	0.218
	0.60	0.017	0.053	0.102	0.058	0.115	0.231
	0.80	0.020	0.060	0.133	0.063	0.103	0.255
200	-0.80	0.050	0.084	0.098	0.138	0.196	0.251
	-0.60	0.043	0.069	0.079	0.111	0.174	0.235
	-0.40	0.039	0.061	0.088	0.099	0.161	0.283
	-0.20	0.031	0.049	0.078	0.092	0.155	0.279
	0.20	0.024	0.039	0.077	0.092	0.152	0.327
	0.40	0.020	0.036	0.067	0.081	0.147	0.295
	0.60	0.020	0.038	0.065	0.080	0.128	0.302
	0.80	0.018	0.036	0.104	0.073	0.117	0.297

Table 3. Probability to find at least one outlier

	Significance Level	Test	Constant only: magnitude of single outlier (δ) 1	5	10	Constant and time trend: magnitude of single outlier (δ) 1	5	10
i.i.d	1.0%	τ_1	0.001	0.019	0.286	0.001	0.060	0.590
		τ	0.008	0.036	0.341	0.005	0.099	0.664
	5.0%	τ_1	0.027	0.073	0.404	0.018	0.151	0.716
		τ	0.049	0.112	0.447	0.048	0.204	0.752
	10.0%	τ_1	0.061	0.127	0.465	0.057	0.214	0.758
		τ	0.106	0.177	0.497	0.095	0.278	0.789
MA, $\theta = -0.5$	1.0%	τ_1	0.007	0.198	0.767	0.003	0.384	0.953
		τ	0.019	0.250	0.809	0.017	0.486	0.969
	5.0%	τ_1	0.044	0.340	0.854	0.043	0.569	0.975
		τ	0.086	0.382	0.876	0.096	0.635	0.980
	10.0%	τ_1	0.109	0.406	0.891	0.101	0.650	0.980
		τ	0.170	0.466	0.908	0.178	0.711	0.988
MA, $\theta = 0.5$	1.0%	τ_1	0.002	0.008	0.079	0.001	0.016	0.218
		τ	0.005	0.014	0.116	0.007	0.026	0.275
	5.0%	τ_1	0.025	0.032	0.168	0.018	0.044	0.342
		τ	0.046	0.056	0.209	0.037	0.087	0.422
	10.0%	τ_1	0.054	0.066	0.224	0.039	0.089	0.436
		τ	0.087	0.105	0.264	0.069	0.123	0.493
AR, $\rho = -0.5$	1.0%	τ_1	0.005	0.098	0.583	0.003	0.222	0.872
		τ	0.011	0.131	0.635	0.006	0.312	0.897
	5.0%	τ_1	0.037	0.192	0.682	0.029	0.373	0.917
		τ	0.068	0.242	0.712	0.065	0.443	0.932
	10.0%	τ_1	0.084	0.271	0.727	0.074	0.459	0.933
		τ	0.130	0.321	0.756	0.131	0.535	0.942
AR, $\rho = 0.5$	1.0%	τ_1	0.001	0.003	0.029	0.000	0.004	0.096
		τ	0.005	0.012	0.046	0.006	0.014	0.144
	5.0%	τ_1	0.015	0.023	0.077	0.015	0.031	0.196
		τ	0.035	0.041	0.108	0.033	0.054	0.242
	10.0%	τ_1	0.044	0.049	0.125	0.037	0.059	0.249
		τ	0.070	0.077	0.169	0.062	0.085	0.299

Table 4. Asymptotic critical values of the test τ_3

α	i	Model 1 $z_t = \{1\}$	Model 2 $z_t = \{1, t\}$
0.05	1	2.9894	3.3327
	2	3.6893	4.8641
	3	4.2938	13.1591
	4	4.4250	18.1991
0.10	1	2.8068	3.1050
	2	3.3790	3.9415
	3	3.8810	6.0802
	4	4.3335	14.4341
	5	4.7833	36.4413
0.20	1	2.6063	2.8741
	2	3.0456	3.4077
	3	3.4291	4.0501
	4	3.7858	5.3950
	5	4.1168	8.8770
	6	4.4185	18.0403
	7	4.7319	33.4132

Table 5. Finite sample critical values of the test τ_4

Level of significance	Model 1 $z_t = \{1\}$		Model 2 $z_t = \{1, t\}$	
	$T = 100$	$T = 200$	$T = 100$	$T = 200$
1.0%	4.0498	4.1118	4.0842	4.1171
2.5%	3.7799	3.8875	3.7943	3.9039
5.0%	3.5731	3.7137	3.5911	3.7274
10.0%	3.3517	3.5170	3.3656	3.5234
20.0%	3.1062	3.2933	3.1186	3.3012
30.0%	2.9480	3.1612	2.9689	3.1675
40.0%	2.8423	3.0463	2.8584	3.0531
50.0%	2.7419	2.9599	2.7532	2.9665
60.0%	2.6468	2.8655	2.6612	2.8720
70.0%	2.5585	2.7756	2.5738	2.7851
80.0%	2.4587	2.6766	2.4695	2.6836
90.0%	2.3235	2.5557	2.3374	2.5611
95.0%	2.2273	2.4726	2.2418	2.4797
97.5%	2.1445	2.3997	2.1602	2.4057
99.0%	2.0588	2.3198	2.0670	2.3312

Table 6. Size and power of the tests to detect outliers

Size of outliers	Probability to find	Test τ	Test τ_1	Test τ_2	Test τ_3	Test τ_4
no outlier	first outlier	0.0405	0.0220	0.0405	0.0680	0.0500
	second outlier	0.0220	0.0070	0.0155	0.0035	0.0020
	third outlier	0.0155	0.0015	0.0055	0.0000	0.0000
	fourth outlier	0.0085	0.0005	0.0025	0.0000	0.0000
$\delta_1 = 3, \delta_2 = 0, \delta_3 = 0, \delta_4 = 0$	first outlier	0.0485	0.0270	0.0485	0.0810	0.7465
	second outlier	0.0205	0.0060	0.0130	0.0035	0.0415
	third outlier	0.0130	0.0020	0.0055	0.0000	0.0010
	fourth outlier	0.0075	0.0005	0.0020	0.0000	0.0000
$\delta_1 = 5, \delta_2 = 0, \delta_3 = 0, \delta_4 = 0$	first outlier	0.1115	0.0795	0.1115	0.1510	1.0000
	second outlier	0.0210	0.0065	0.0125	0.0025	0.0555
	third outlier	0.0130	0.0020	0.0050	0.0000	0.0015
	fourth outlier	0.0085	0.0000	0.0015	0.0000	0.0000
$\delta_1 = 5, \delta_2 = 3, \delta_3 = 0, \delta_4 = 0$	first outlier	0.1170	0.0830	0.1170	0.1505	1.0000
	second outlier	0.0205	0.0055	0.0120	0.0025	0.7395
	third outlier	0.0135	0.0015	0.0040	0.0000	0.0400
	fourth outlier	0.0080	0.0000	0.0015	0.0000	0.0005
$\delta_1 = 3, \delta_2 = 3, \delta_3 = 0, \delta_4 = 0$	first outlier	0.0560	0.0315	0.0560	0.0850	0.8915
	second outlier	0.0190	0.0060	0.0130	0.0030	0.5495
	third outlier	0.0130	0.0010	0.0055	0.0000	0.0315
	fourth outlier	0.0075	0.0000	0.0015	0.0000	0.0005
$\delta_1 = 3, \delta_2 = 3, \delta_3 = 2, \delta_4 = 0$	first outlier	0.0575	0.0310	0.0575	0.0855	0.8790
	second outlier	0.0190	0.0060	0.0115	0.0030	0.5625
	third outlier	0.0130	0.0010	0.0040	0.0000	0.1560
	fourth outlier	0.0080	0.0000	0.0015	0.0000	0.0160
$\delta_1 = 5, \delta_2 = 3, \delta_3 = 2, \delta_4 = 0$	first outlier	0.1130	0.0805	0.1130	0.1510	0.9995
	second outlier	0.0185	0.0060	0.0110	0.0030	0.7605
	third outlier	0.0125	0.0005	0.0035	0.0000	0.2085
	fourth outlier	0.0080	0.0000	0.0015	0.0000	0.0175
$\delta_1 = 3, \delta_2 = 3, \delta_3 = 3, \delta_4 = 3$	first outlier	0.0645	0.0400	0.0645	0.0980	0.9375
	second outlier	0.0170	0.0040	0.0115	0.0020	0.8035
	third outlier	0.0120	0.0005	0.0040	0.0000	0.6020
	fourth outlier	0.0080	0.0000	0.0015	0.0000	0.3060
$\delta_1 = 5, \delta_2 = 3, \delta_3 = 2, \delta_4 = 2$	first outlier	0.1095	0.0775	0.1095	0.1505	0.9990
	second outlier	0.0175	0.0050	0.0100	0.0030	0.7680
	third outlier	0.0105	0.0005	0.0030	0.0000	0.2840
	fourth outlier	0.0070	0.0000	0.0015	0.0000	0.0515

Table 7. Empirical results: Logarithm of the US/Finland real exchange rate

Significance level	Test	CPI-based series 1900-1988	GDP-based series 1900-1987
5.0%	τ	1917,1918,1919,1932	no outliers
	τ_1	1918	no outliers
	τ_2	1918,1919	no outliers
	τ_3	1918	no outliers
	τ_4	1917,1918,1919,1932,1948	1918, 1948
10.0%	τ	1917,1918,1919,1921,1932	no outliers
	τ_1	1918	no outliers
	τ_2	1917,1918,1919	no outliers
	τ_3	1918,1919	no outliers
	τ_4	1917,1918,1919,1932,1948	1918,1948

Annex 2

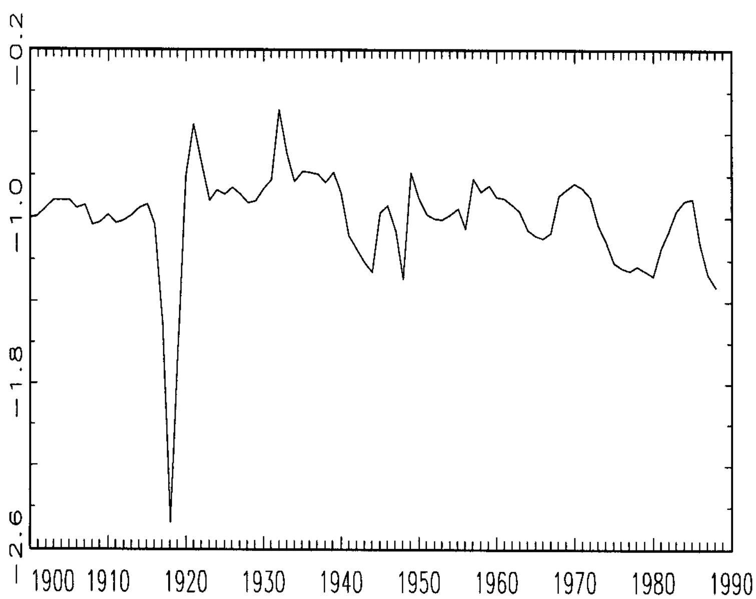

Figure 1. Logarithm of the US/Finland Real Exchange Rate Based on the Consumer Price Indexes (CPI); 1900-1988, Annual

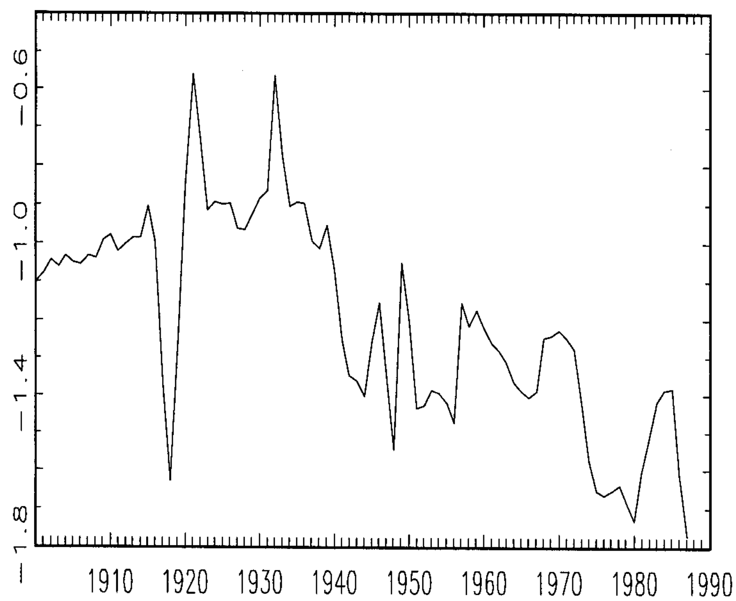

Figure 2. Logarithm of the US/Finland Real Exchange Rate Based on the GNP (GDP) deflators as the Price Indexes; 1900-1987, Annual