Essays in Unit Root Testing

by

Lynette Vanessa Smith, B.Sc., M.Sc.

A Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, May 2003

To my parents

Acknowledgments

I am indebted to my supervisors Professor Steve Leybourne and Professor Paul Newbold for their support, guidance, tolerance and recognition of my curiosities throughout the writing of this thesis. I also owe gratitude to Dr. Tae-Hwan Kim for his valuable help and discussions.

My family and friends are deeply appreciated for their generous understanding of the selfishness I have manifested during my graduate studies. In particular, I would like to thank my parents and my sister for their continuous support and encouragement. My aunties Soula and Mary for their love and appreciation. My office mates Cihan and Thanaset for some relaxing conversations. Fred and Ortenca and Jesus and Auxi, for some delightful evenings. Special thanks also goes to the Warden Dr Brian Atkin, his wife Jan and the staff at Ancaster Hall as well as the resident tutors for providing a supportive and enjoyable environment throughout the writing of this thesis. I would specifically like to thank Dorit and Carole for their friendship and for a number of memorable evenings. Last but not least I would like to thank a very special person, Nikos, for his continuous support and understanding, for lifting the spirits when they were low and for always making me look at the bright side of life.

Thanks is also due to the University of Nottingham and NATO for funding this thesis.

Abstract

This thesis is a collection of four essays with main focus on testing for a unit root under structural change, and on the behaviour of power-enhancing unit root tests that have recently emerged as a solution to the well-known power deficiency of traditional such tests. New tests and variants of commonly applied ones are introduced in response to the need for reliable statistical techniques in modelling economic series over time.

The first essay explores the possibility that a time series may change structure from trend-stationarity to difference-stationarity, or vice versa as has been recognised by economists for several years. Taking difference-stationarity as the null hypothesis, tests are developed for this possibility, where neither the location nor direction of any possible change under the alternative hypothesis need be specified. Application of these tests to series on consumer price inflation in the G7 countries reveals evidence of a change from trend-stationarity to difference-stationarity in the majority of these countries.

In the second essay we apply two elaboration principles of standard unit root tests in the more flexible setting of testing for a unit root against the alternative of stationarity around a smooth transition in linear trend. In comparison to the standard case, the modified tests within this context generate only moderate additional power, a phenomenon which appears to be related to the elaborate nature of the trend function under the alternative. An empirical application of the modified smooth transition tests to common macroeconomic time series in the US economy leads to stronger evidence in favour of the smooth transition alternative than do the unmodified tests.

In the third essay we show that more powerful variants of commonly applied unit root tests to panel data, seeking mean or trend reversion, are readily available. Moreover, power gains persist when the modifications are applied to bootstrap procedures that may be employed when cross-correlation of a rather general sort among individual panel members is suspected. That such an approach can strongly influence inference is demonstrated through an application to a panel of real exchange rates against the US dollar.

The final essay explores the behaviour of the power-enhancing unit root test most widely applied in the empirical literature. The principle issue is that such a test can have very low power for certain parameter configurations and sample sizes relative to conventional unit root tests. A theoretical attempt is made to identify these unsatisfactory cases relying on local to unity asymptotics, through investigation of the relative efficiencies in the case of an unknown mean. Extensive Monte Carlo results highlight the shortcomings of such a test under higher order autoregressive processes and indicate preference for its existing rivals.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The content of this thesis draws upon two important areas in econometrics, which have jointly and independently played an increasingly important role in the search of an optimal way to characterise the behaviour of economic time series, namely unit root testing and structural change. How macroeconomic series evolve over time has profound implications for economic theories purporting to explain economic events, for econometric modelling strategies and for forecasting accuracy. Thus, the need for reliable statistical techniques to characterise their behaviour over time becomes imperative.

Unit root tests involve statistical techniques designed to distinguish between deterministic and stochastic trends in observed economic time series. Traditionally, empirical researchers have treated observed trends in macroeconomic and financial time series as deterministic functions of time. Under this view, current stochastic shocks have only a temporary effect on the long-run movement of a series. Consequently, long-run forecasts from such a model may be expected to be fairly precise as long as the trend is consistently estimated. Following the work of Nelson and Plosser (1982), there was a general acceptance that most macroeconomic variables exhibited some form of stochastic nonstationarity - e.g. could be modelled as unit root processes. In this case, current shocks have an enduring effect on the evolution of the series; hence, long-run forecasts are expected to be quite poor. Different conclusions are thus derived based on the interpretation of observed trend behaviour of economic series, with a subsequent influence on the way in which macroeconomic theorists view the working of an economy. Significant contributions in this area include among others the work of Dickey and

Fuller (1979, 1981) and Phillips and Perron (1988), the well-known Dickey-Fuller (DF) and Phillips-Perron (PP) unit root tests, respectively. The former are with no doubt the most widely employed in empirical applications and they are the type of tests we will concentrate on throughout this thesis.[1]

Structural change, although not given an exact definition in the literature, is usually interpreted as changes of regression parameters in the underlying econometric model. It is a phenomenon pervasive in economic time series relationships particularly when investigated over a long time span, and it can be quite perilous to ignore. Events like the great depression, oil price shocks, abrupt policy changes could be responsible for any parameter shifts observed in the underlying series. A structural change may affect any or all of the model parameters, and these cases have different implications. A large number of statistical tests have been developed to detect such changes dating back to the traditional Chow (1960) test, which involves testing the null hypothesis of parameter constancy against the alternative of a known break point a priori, under the assumption of constant variances. The earliest tests for structural breaks in the economic literature are for stationary variables and a single break. A number of later developments in this area involve tests that allow for more general alternatives of structural change occurring at some unknown point in time, allowing also for non-stationary regressors and even multiple breaks.

While the aforementioned areas have to a large extent evolved independently, arguments related to the restrictive nature of the linear time trend under the trend-stationary alternative and the adverse effect of changes in regression parameters on traditional unit root tests, have sparked intense and active research over the last decade in their joint investigation. Alternative, more flexible specifications of the trend function have been proposed including structural breaks, Markov regime-switching and smooth transitions which have resulted in unit root tests under various forms of structural change, the majority linked to shifts in the trend function. We contribute to this strand of literature by proposing

[1]The Dickey-Fuller tests involve augmenting the underlying model with lag differences to account for any correlation in the residuals, while the Phillips-Perron tests modify the statistics using a nonparametric approach to obtain consistent estimators. Several simulation studies (Schwert, 1989; DeJong *et al.* , 1992) have shown that the latter tests have serious size distortions in finite samples when the data generating process has a predominance of negative autocorrelations in first-differences.

unit root tests against the alternative of structural change associated with a shift in the (dominant) autoregressive parameter of the underlying time series, namely a change in persistence. Although a number of studies are available that suggest the possibility of a change of this form, not much attention has been given to designing official tests for this purpose.

A separate important issue regarding traditional unit roots is their well-known power deficiency. Research efforts in this direction have proved rewarding by giving rise to unit root tests characterised by improved power, the majority of which are modifications of Dickey-Fuller type tests. A pervasive concern in this thesis is how these power-enhancing unit root tests behave when adopted in a number of alternative settings. Specifically, we seek to investigate how two such modifications perform when adopted in the context of testing for a unit root under the flexible alternative of stationarity around a smooth transition in linear trend. Subsequently we explore whether these modifications in the basic time series case, when extended to the panel unit root testing setting, can result in more powerful panel data unit root tests. We also aim to highlight certain shortcomings related to the power-enhancing unit root test most widely employed in empirical applications, the GLS test proposed by Elliott *et al.* (1996), which in any case proves less powerful than the two elaborated tests we consider.

Thus, the original contribution of this thesis lies in the development of unit root tests under a relatively well-documented, though under-explored in the testing literature, type of structural change related to a change in persistence. In addition to this, more powerful variants of commonly applied panel data unit root tests are introduced. Such developments are oriented towards the need for reliable statistical techniques in modelling economic series over time, in response to which we also explore the behaviour of recently emerged unit root tests in more flexible settings.

The outline of this thesis is as follows. In the second chapter unit root tests are considered against the alternative hypothesis of structural change characterised by "a change in persistence". Such a term describes a change in structure of a time series from difference stationarity, $I(1)$, to trend stationarity, $I(0)$, or from $I(0)$ to $I(1)$. A number of studies have argued the likelihood of a switch in persistence for particular time series. Taking the null hypothesis to be that of $I(1)$ throughout, Dickey-Fuller type tests are developed that are designed to have power against the alternative hypothesis of a switch at an unknown point in time, either from $I(0)$ to $I(1)$, or vice versa. Sequential estimation procedures are considered for this purpose. In the most general case, it is not necessary a priori to specify the direction of any possible switch under the alternative hypothesis. In all cases, a consistent estimator of any break fraction results as a by-product of the tests. Asymptotic properties of the test statistics are investigated both under the null and alternative hypotheses. We report simulation evidence on the performance of the tests in finite samples in terms of size robustness and power. An application of the tests to consumer price inflation in the G-7 countries reveals strong and consistent evidence of a change from $I(0)$ to $I(1)$ behaviour in the majority of countries.

In the third chapter we explore and assess the performance of the more powerful elaborations of unit root tests when the alternative hypothesis is that of stationarity around smooth transition in linear trend. Smooth transition regression models have been employed to characterise the behaviour of a number of macroeconomic series. They allow for a continuum of states between two extremes permitting a more plausible and flexible specification of the trend function, while they include the possibility of an abrupt structural break as a special case. We demonstrate how the modified statistics are incorporated into the testing procedure, and obtain estimated percentiles of the limiting null distributions of the tests through simulation. Extensive finite sample Monte Carlo results show that, although power gains are not as substantial as in the standard case of the linear trend alternative, the use of the modified tests is worth the while as it signifies extra power at the minimum cost of a little more computational complexity. In particular, power gains are more prominent in the simpler case of a smooth transition in the constant term only. An empirical application based on US macroeconomic data suggests that the smooth transition hypothesis is an attractive characterisation for a further number of series when using the modified smooth transition tests compared to the unmodified tests.

The fourth chapter is concerned with unit root testing in the panel data context, where a topic of increasing interest is whether or not the individual series are generated by unit root processes. Such a concern has been expressed in the literature on purchasing power parity, where the issue of interest is the possible mean reversion of real exchange rates. This issue can be addressed through an extension of commonly applied unit root tests, such as the DF test or a Lagrange Multiplier test. We contribute to this strand of literature by showing that power-enhancing elaborations of unit root tests in the single time series case, maintain these power gains when applied in the panel data unit root testing context. In particular, modified panel unit root tests are introduced starting from the base case, where independence over panel data members can be assumed. However, it is well known that difficulties, particularly spurious rejections of the null hypothesis, can arise when individual panel series are generated by cross-sectionally correlated innovations. An important special case, which is readily dealt with through the subtraction of time-specific means is investigated. Reported simulation results show that, while the modified panel data unit root tests retain size reliability, they can produce appreciable gains in power. We analyse a panel of series of real exchange rates against the US dollar and employ the bootstrap to accommodate the heterogeneous nature of cross-section correlation found amongst the innovations generating the individual time series in the panel. We find through simulation that modified bootstrap tests retain the power gains noted in simpler cases, and, moreover, that the application of these tests yields appreciably stronger evidence against the unit root null hypothesis for our data than do the unmodified tests.

The fifth chapter, provides a thorough investigation into the properties of the most widely applied, modified DF-type test based on generalised least squares detrending (GLS), that has emerged as a solution to the well-known power deficiency of conventional unit root tests. While a number of studies address the behaviour of such a test only limited detailed results are available to date, mainly considering ARMA type models. Given the well-known autoregressive approximation to this class of models (Said and Dickey, 1984) and the importance of the autoregressive structure in characterising key economic series, we explore the reliability of such a test under higher order autoregressive processes and under the more natural alternative of 'strict stationarity'. A theoretical attempt is made to predict the low power that such tests appear to display relative to the DF test for certain parameter configurations and sample sizes, relying on local to unity asymptotics. This is achieved through an investigation of the relative efficiencies of the GLS and OLS estimators, in the case of an unknown mean. The limiting results of the derived approximate relative efficiencies predict to some extent this issue, through the localising parameter of the autoregressive structure. However, Monte Carlo simulations are required to uncover the significant finite sample effects associated with the higher order autoregressive parameters. Overall, the relative efficiencies of the mean estimator appear to predict power well. In comparing power across alternative power-enhancing unit root tests in this context, preference is clearly in favour of the alternative tests that appear not to be affected by the shortcomings related to the GLS test, in that they maintain correct size and superior power in all cases.

In the last chapter we conclude and provide some suggestions for future research.

References

Chow, G.C., 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28, 591-605.

DeJong, D.N., Nankervis, J.C., Savin, N.E., Whiteman, C.H., 1992. The power problems of unit root tests in time series with autoregressive errors. *Journal of Econometrics*, 53, 323-343.

Dickey, D., Fuller, W., 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.

Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49, 1057-1072.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. *Econometrica*, 64, 813-836.

Nelson, C.R., Plosser C.I., 1982. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10, 139–162.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. *Biometrika*, 75, 335-346.

Schwert, G.W., 1989. Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economic Statistics*, 7, 147-159.

Chapter 2

Tests for a Change in Persistence Against the Null of Difference - Stationarity with Application to the Behaviour of Price Inflation

2.1 Introduction

Over the years the properties of important macroeconomic and financial time series such as output, inflation rates and interest rates have been subjected to scrutiny by numerous empirical studies, in an attempt to best characterise their behaviour. To this end, a number of unit root tests have been employed, which have arisen in response to the need for discriminating between deterministic and stochastic trends in observed economic time series and thus investigating whether the effect of current shocks is temporary or permanent respectively. The empirical results have led many to accept the notion that a wide variety of economic time series contain unit roots, and therefore, stochastic trends.

While unit root tests have served as the basis for testing the effect/persistence of current shocks on the evolution of a series, the variation of the permanent impact of shocks to a series is linked to the phenomenon of structural change. Investigating the properties of economic series over a long time span it is unlikely that no change in structure will have occurred. Major events of some form have usually taken place during the period under examination, affecting the behaviour of the series in one way or another. It is this observation that has motivated testing for structural change, which constitutes another active area of research. Contributions to the problem of testing for structural change have been made by a number of authors. Brown, Durbin and Evans (1975) test for parameter constancy against general alternatives, including the case of a single break, based on recursive residuals. Ploberger, Kramer and Kontrus (1989) propose a test based on successive parameter estimates rather than on recursive residuals to test parameter constancy. Andrews (1993) develops likelihood-ratio, Wald and Lagrange multiplier type tests of parameter stability against a one-time structural change with an unknown change point, limited to nontrending regressors. Andrews and Ploberger (1994) extend the tests in Andrews (1993) to allow for stronger optimality properties. Further statistics are proposed by Hansen (1992) for testing constancy of the parameters against the alternative of sudden breaks and of a gradual change. Chu and White (1992) consider tests of a trend stationary process against the alternative of a change in the trend function at some unknown point in time, while Bai and Perron (1998) estimate and test linear models with multiple structural changes occurring at unknown dates.

Unit root tests and tests for structural change have both played an increasingly important role in the search for an optimal way to characterise the behaviour of economic series over time. Of the literature to date when combining both the aforementioned, it was Perron (1989, 1990) and Rapporort and Reichlin (1989) who discovered that the usual unit root tests behave rather poorly against time series whose mean or trend function might have undergone important structural changes. They demonstrated how exogenous breaks in the series can lead to biased results in favour of difference stationarity. Since then, a number of studies have emerged that endogenise the choice of the break date by making it data dependent, including work by Banerjee *et al.* (1992), Zivot and Andrews (1992), Perron (1997), Perron and Vogelsang (1992), Saikkonen and Lütkepohl (2002).[1] The majority of such work addresses the effects of structural breaks related to a

[1]Evidence from these studies reveals that when the date of the break is treated as unknown there is less compelling evidence against the unit root hypothesis. The critical values of the limiting test statistics are further out in the tail than those of the exogenous trend break statistics, thus being harder to reject the null of a unit root.

change in the intercept and/or slope term of the trend function, on augmented Dickey-Fuller (ADF) type tests.[2] To capture such changes various dummy variables are typically added to the regressions, while use of sequential and recursive estimation methods is made to reveal the unknown break date.[3]

Further extensions in the same direction have been considered by Leybourne *et al.* (1998) who test for unit roots against the alternative of stationarity around a smooth transition in linear trend. Under the alternative, a smooth transition function controls the transition between regimes. It incorporates no change and an instantaneous change as limiting cases. A further form of structural change popularised by Hamilton (1989) and considered widely in the literature thereafter, involves the Markov regime-switching type. The structural model in this case specifies the time series process of interest as the sum of a trend function the parameters of which change according to an unobservable state variable that is specified as a Markov chain, plus an autoregressive process with a root on the unit circle, or possibly a stationary autoregressive process as in Lam (1990). However, technical difficulties still remain in the case of testing whether the linear part of the process does have a unit root or not and thus combining unit testing with this form of structural change.

While the main interest in the literature when testing for a unit root against the alternative of structural change has centered around shifts in the underlying trend function, we concentrate in this chapter on a different type of shift, namely a change in persistence. By change in persistence we refer to the phenomenon whereby the series under investigation changes from a trend-stationary process, $I(0)$, to a difference-stationary, $I(1)$, process or vice versa.

The issue of structural change characterised by change in persistence has recently been addressed by Kim (2000). He maintains the null hypothesis of $I(0)$ throughout, and therefore adapts stationarity tests such as those of Kwiatkowski *et al.* (1992) and Leybourne and McCabe (1994). We take a different approach by adopting a null of $I(1)$ and using test statistics based on the forward and reverse

[2]An exception is the work of Saikkonen and Lütkepohl (2002), who in testing for unit roots against changes in the trend function employ generalized least squares estimation rather than augmenting the underlying model with lagged differences.

[3]The endogenous break literature has concentrated on testing the unit-root null against a one break alternative. Lumsdaine and Papell (1995) extend this methodology by allowing for a two-break alternative.

Dickey-Fuller statistics, as advocated in Leybourne (1995). The problem, therefore, is that of testing for a unit root against the alternative of structural change at some unknown point, with the selection of the break point determined from the estimation procedure designed to fit to an observed series y_t a representation in which the series changes from stationarity to non-stationarity or conversely.

There are a number of studies available relating to the issue of a change in persistence in key macroeconomic variables. Parker (1989) studies the persistence of price shocks in the pre-World War I and post-World war II eras and finds that inflation has greater persistence in the post war period. De Long and Summers (1988) demonstrate the greater persistence of output in the post-WWII period. Ball and Cecchetti (1990) find that as trend inflation has risen over the past 100 years, the persistence of changes in inflation has also increased, and they fail to reject nonsationarity for most of the countries in their sample based on post-War data. This corroborates the findings of Brasky (1987) that quarterly US inflation evolved from a white noise process in the pre-World War I years to a highly persistent nonstationary $ARIMA$ process in the post-1960 period. Brunner and Hess (1993) find evidence based on unit root tests that inflation is an $I(0)$ process for the period from 1947-1959 and an $I(1)$ process for 1960-1992, with results qualitatively similar using seasonally adjusted and non adjusted data. They, thus, model inflation as containing a unit root subsequent to 1960 and as stationary prior to this time.

The short term interest rate is another macroeconomic variable generally agreed via empirical evidence as having followed a stationary, $I(0)$, process during 1890-1910, whilst characterised as non-stationary over the period 1920-1933. Various studies have investigated this issue attempting to uncover the exact timing and speed of such a break. Different dates have been characterised as potential break dates depending on the dataset and method employed (see Fishe, 1991 and Fishe and Wohar, 1990). Among these studies, Mankiw, Miron and Weil (1987) employ switching regression techniques in the case of US and UK interest rates and find a break in early 1915. Adopting a recursive method based on Bayesian learning, Kool (1995) detects a switch to nonstationarity in late 1915 for the UK and in late 1917 for the US, relating these dates to the start of interest rate targeting. Newbold *et al.* (2001) develop a testing procedure to model the structural break from an $I(0)$ process to an $I(1)$ process based on the logistic function, with a smooth transition taking place in both the intercept and the autoregressive parameter. Using US and UK monthly and weekly nominal interest rates over the period 1890-1934 they cannot reject the null hypothesis, $I(1)$, for the UK, while for the US data they can reject the $I(1)$ hypothesis in favour of the hypothesis that a structural change occurred from $I(0)$ to $I(1)$. Their best estimate is of a rapid structural change occurring in 1917. Furthermore, Hakkio and Rush (1991) consider the possibility of a change in persistence for the government budget deficit.

Although the majority of empirical evidence indicates a change in persistence from an $I(0)$ to an $I(1)$ process, the converse is suggested in studies like that of Evans and Wachtel (1993) who employ a Markov switching model in the case of US inflation. Their estimated Markov model that allows for a stationary $AR(1)$ process in one regime and a random walk in the other suggests the possibility of structural change of the inflation series from a difference-stationary process back to level stationarity in the mid 1980's, coinciding with the collapse in oil prices.

The rest of this chapter is organised as follows. In Section 2.2, we begin by analysing the case where the direction, *but not the location*, of any change in persistence, under the alternative hypothesis, is taken as given. Initially we consider the situation where we permit in the $I(0)$ phase stationarity around an unknown mean, while in the $I(1)$ phase first differences have zero mean. In detecting a change from either $I(0)$ to $I(1)$, or $I(1)$ to $I(0)$ we use test statistics based on the minima of sequential unit root tests. In doing this we are prompted by Banerjee *et al.* (1992) - hereafter BLS - who advocated such procedures in the context of testing for $I(1)$ against the alternative of $I(0)$ with a change in the deterministic trend at an unknown time. Section 2.3 presents the limiting null distributions of the test statistics along with the asymptotic properties of the tests under the alternative hypotheses drawing on results derived in Leybourne *et al.* (2000). It also follows that, when a break does occur, the point identified by the test statistic (i.e. that at which a minimum occurs) is a consistent estimator of the true break fraction. Based on these results the development of tests against the most general and practically important alternative, where the direction of any possible change is not specified a priori is taken up in Section 2.4. In particular, we show how the tests of Section 2.2 can be combined to generate what is, in effect, a single test against a two-sided alternative of a change in persistence in an unknown direction at an unknown time. In Section 2.5, Monte Carlo simulations are undertaken to investigate the finite sample performance of the tests. Finite sample critical values of the tests are given, while robustness of the tests to residual serial correlation and error processes that depart from normality are explored. In the case of the latter, the errors are assumed to follow Student-t and chi-square distributions with five and three degrees of freedom, respectively. In Section 2.5.4 power estimates are obtained and compared across the tests, while results are also reported relating to the accuracy of the break point estimation. Extensions of the tests involving further deterministics are briefly presented in Section 2.6, where a drift is permitted under the null hypothesis, and correspondingly stationarity around a linear trend is allowed in the $I(0)$ phase of the series under the alternative. Section 2.7 reports an empirical application to quarterly CPI inflation rate series of the G-7 countries. Section 2.8 ends the chapter with some concluding remarks.

2.2 Tests For a Change in Persistence: Direction Specified

Initially, we concentrate on the case where, under the null hypothesis, the data generating process is a driftless random walk. Under the alternative hypothesis, a transition is permitted, either from a stationary first order autoregression (with unknown mean) to a random walk, or from a random walk to a stationary first order autoregression. The transition point is not assumed to be known.

Consider the following data generating process for a series of T observations on y_t. The null is that y_t is $I(1)$ throughout

$$y_t = y_{t-1} + \varepsilon_t, \qquad t = 1, 2, ..., T \tag{2.1}$$

where ε_t is taken to satisfy the following assumption.

Assumption 2.1. *ε_t is a martingale difference sequence and satisfies* $E(\varepsilon_t^2| \varepsilon_{t-1},...) = \sigma^2, E(|\varepsilon_t|^i|\varepsilon_{t-1},...) = \kappa_i$ $(i = 3,4)$, *and* $\sup_t E(|\varepsilon_t|^{4+\gamma}|\varepsilon_{t-1},...) = \kappa < \infty$ *for some* $\gamma > 0$.

This assumption is standard in the unit root literature, see for instance BLS and Stock (1994). It allows for lagged dependent regressors (e.g., autoregressive models) and it implies that the functional central limit theorem applies to the partial sums of ε_t, *i.e.* $T^{-1/2}\sum_{t=1}^{[\tau T]}\varepsilon_t \Longrightarrow \sigma W(\tau)$, uniformly for $\tau \in [0,1]$.

There are two alternatives. One is that y_t is $I(0)$ changing to $I(1)$ at time τ^*T

$$\begin{aligned} y_t &= \alpha + \rho y_{t-1} + \varepsilon_t, & t \leq \tau^*T, \quad \rho < 1, \\ y_t &= y_{t-1} + \varepsilon_t, & t > \tau^*T, \end{aligned} \tag{2.2}$$

and the other is that y_t is $I(1)$ changing to $I(0)$ at time τ^*T

$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_t, & t \leq \tau^*T \\ y_t &= \alpha_{\tau^*T} + \rho y_{t-1} + \varepsilon_t, & t > \tau^*T, \quad \rho < 1, \\ \alpha_{\tau^*T} &= (1-\rho) y_{\tau^*T}. \end{aligned}$$

We denote the null H_0^{11} and the two alternatives H_1^{01} and H_1^{10}, respectively. This definition of H_1^{10} ensures a "joining up" of the $I(1)$ and $I(0)$ parts of the series, just as H_1^{01} does. This is precisely what would be achieved in a series that is a time-reversed variant of (2.2) with $(1-\tau^*)$ in place of τ^*. This seems to be a reasonable practical requirement. Also, if it is not imposed, properties of tests will obviously depend on the mean of the $I(0)$ part of the series.[4]

2.2.1 Tests Based on the Forward Series

We now consider the alternative hypothesis of a switch from $I(0)$ to $I(1)$; that is, we require that, if H_0^{11} does not hold, then H_1^{01} must be true. Let τ denote a

[4]A similar definition is used in Banerjee *et al.* (1992). See the footnote to Table 3 of that paper.

possible break fraction. Define

$$d_{10t}(\tau) \equiv 1[t \leq \tau T]$$

where $1(\cdot)$ is the indicator function. We estimate the regression using the full sample

$$\begin{aligned} \Delta y_t &= \hat{\alpha}(\tau) d_{10t}(\tau) + \hat{\phi}(\tau) d_{10t}(\tau) y_{t-1} + \hat{\varepsilon}_t(\tau) \\ t &= 1, 2, ..., T \end{aligned} \tag{2.3}$$

and construct the corresponding t-ratio associated with $\hat{\phi}(\tau)$. We denote this by $DF_{01}^f(\tau)$.[5] Since the break fraction τ is not assumed known, Dickey-Fuller t-ratios are computed for all possible τ, and we take as a test statistic that t-ratio least favourable to the null hypothesis that the series is $I(1)$ throughout; that is, the infimum over τ of the Dickey-Fuller t-ratios. The test statistic is then

$$DF_{01}^{f\,\inf} \equiv \inf_{\tau \in \Lambda} DF_{01}^f(\tau)$$

where Λ is a closed subset of (0,1).[6] Tests of this sort that are based on the full sample are termed "sequential" by BLS.

Since the last $(1-\tau)T$ observations are taken under both the null and alternative hypotheses to be generated by an $I(1)$ process, another possibility is to apply Dickey-Fuller type tests to the first τT observations, allowing τ to vary. Tests of this sort are termed "recursive" by BLS, in which case (2.3) is estimated based on the first τ fraction of the sample, that is $t = 1, 2, ..., \tau T$. The resultant t–statistic $DF_{01}^f(\tau)$ in this case is the same as the t–statistic, denoted $\hat{t}_{DF}(\tau)$, in BLS when there is neither a time trend nor lagged Δy_t in their model, which is based on the same regression as in (2.3).

[5]The subscript 01 in the statistic indicates that the test is for $I(0)$ changing to $I(1)$, and the superscript f indicates that the test is based on the forward series y_t.

[6]This type of trimming is standard in the literature when testing for breaks in order to obtain well-defined limiting distributions.

For the generating model as specified in (2.1), the recursive procedure is inefficient, since in the ADF regression of (2.3) we use only the first τ fraction of the sample. It should be possible to estimate the variance of the innovations ε_t more efficiently using the full sample and in doing so to improve the finite sample power of the test. For this purpose, the analysis of this chapter restricts its attention to the case of sequential testing.[7]

The estimated variance for the sequential test is given by

$$\hat{\sigma}^2 \equiv \tau\tilde{\sigma}^2 + (1-\tau)\bar{\sigma}^2$$

where

$$\tilde{\sigma}^2 \equiv (\tau T)^{-1} \sum_{t=1}^{\tau T} \hat{\varepsilon}_t^2$$

which is the estimated variance corresponding to the recursive procedure, and

$$\bar{\sigma}^2 \equiv ((1-\tau)T)^{-1} \sum_{t=\tau T+1}^{T} \Delta y_t^2.$$

If instead the alternative is H_1^{10}, we define

$$d_{01t}(\tau) \equiv 1[t > \tau T]$$

and estimate the regression

$$\begin{aligned} \Delta y_t &= \hat{\alpha}(\tau)d_{01t}(\tau) + \hat{\phi}(\tau)d_{01t}(\tau)y_{t-1} + \hat{\varepsilon}_t(\tau) \\ t &= 1, 2, ..., T \end{aligned} \tag{2.4}$$

over all τ and construct the corresponding $t-$ratio associated with $\hat{\phi}(\tau)$. We

[7]It should be noted here that neither the recursive nor the sequential procedures are constructed so as to permit alternative parameters that influence persistence other than ρ, to change at time τ^*T. We would suppose that in the presence of such a change, recursive tests would be less affected of the two, simply because they impose fewer restrictions. We leave this issue to be investigated in future work.

denote this by $DF^f_{10}(\tau)$. The test statistic is then

$$DF^{f\,\inf}_{10} \equiv \inf_{\tau\in\Lambda} DF^f_{10}(\tau).$$

2.2.2 Tests Based on the Time-Reversed Series

As in Leybourne (1995), we define $z_t = y_{T-t+1}$ which is the time-reversed y_t series. Then, under H^{11}_0, z_t is still $I(1)$ throughout and under H^{10}_1, z_t is $I(0)$ changing to $I(1)$ at time $(1-\tau^*)T$, *measured in reversed time.* Similarly, under H^{01}_1, z_t is $I(1)$ changing to $I(0)$ at time $(1-\tau^*)T$. This suggests we have another pair of tests, based on the reversed series z_t.

To test H^{11}_0 against H^{10}_1 we estimate the regression

$$\begin{aligned} \Delta z_t &= \hat{\alpha}(\tau)d_{10t}(1-\tau) + \hat{\phi}(\tau)d_{10t}(1-\tau)z_{t-1} + \hat{\eta}_t(\tau) \\ t &= 1,2,...,T \end{aligned} \tag{2.5}$$

over all τ and construct the corresponding $t-$ratio associated with $\hat{\phi}(\tau)$. Denote this by $DF^r_{10}(\tau)$.[8] The test statistic is then

$$DF^{r\,\inf}_{10} \equiv \inf_{\tau\in\Lambda} DF^r_{10}(\tau).$$

Similarly, to test H^{11}_0 against H^{01}_1estimate the regression

$$\begin{aligned} \Delta z_t &= \hat{\alpha}(\tau)d_{01t}(1-\tau) + \hat{\phi}(\tau)d_{01t}(1-\tau)z_{t-1} + \hat{\eta}_t(\tau) \\ t &= 1,2,...,T \end{aligned} \tag{2.6}$$

over all τ and construct the corresponding $t-$ratio associated with $\hat{\phi}(\tau)$. We

[8]Here and throughout we use the convention that subscripts on DF relate to the alternative hypothesis for the original series. Thus, for example, DF^r_{10} is designed to test for a change from $I(1)$ to $I(0)$ in y_t.

denote this by $DF_{01}^{r}(\tau)$. The test statistic is then

$$DF_{01}^{r\,\inf} \equiv \inf_{\tau \in \Lambda} DF_{01}^{r}(\tau).$$

Our main contribution is to combine the reversed and forward tests, as we will do in Section 2.4, in order to obtain consistent tests for a break in persistence, when neither of the two possible alternatives is excluded.

It will also be shown in the sequel that under the alternatives H_1^{01} and H_1^{10} the value of τ at the infimum provides a consistent estimator of the true break fraction τ^*, and therefore the above test statistics constitute a natural choice.

2.3 Asymptotic Results

In what follows we present the asymptotic results related to the above test statistics, that is when the direction of a change in persistence is prespecified, as derived in Leybourne *et al.* (2000). These are essential, as they provide the basis for the development of tests against the more general and practical alternative of a change in persistence where the direction of the change is not specified a priori. This section is in two parts. The first part presents results related to the limiting distributions of the test statistics under the null. In the second part results are presented associated with the behaviour of the tests under the alternative hypotheses. It turns out that a particular test is consistent for the alternative hypothesis for which it was designed, but not for a change in persistence in the other direction. The theory of weak convergence and the functional central limit theorem are employed in obtaining such results, where $\Longrightarrow$ denotes weak convergence on $C[0,1]$, the set of all real continuous functions on [0,1]. Mathematical proofs are provided in Appendix 2.A.

2.3.1 Limiting Distributions of the Test Statistics Under H_0^{11}

Theorem 2.3.1 *Under H_0^{11} and Assumption 2.1,*

$$
\begin{aligned}
DF_{01}^{f}(\tau) &\Rightarrow \frac{N_1(\tau)}{D_1(\tau)^{1/2}} \\
DF_{10}^{r}(\tau) &\Rightarrow \frac{N_2(\tau)}{D_2(\tau)^{1/2}} \\
DF_{10}^{f}(\tau) &\Rightarrow \frac{N_3(\tau)}{D_3(\tau)^{1/2}} \\
DF_{01}^{r}(\tau) &\Rightarrow \frac{N_4(\tau)}{D_4(\tau)^{1/2}}
\end{aligned}
$$

where

$$
\begin{aligned}
N_1(\tau) &\equiv \frac{1}{2}\left\{W(\tau)^2-\tau\right\}-\frac{1}{\tau}W(\tau)\int_0^\tau W(r)dr \\
D_1(\tau) &\equiv \int_0^\tau W(r)^2dr-\frac{1}{\tau}\{\int_0^\tau W(r)dr\}^2 \\
N_2(\tau) &\equiv \frac{W(\tau)^2-W(1)^2-(1-\tau)}{2}-\frac{1}{1-\tau}\{W(\tau)-W(1)\}\int_\tau^1 W(r)dr \\
D_2(\tau) &\equiv \int_\tau^1 W(r)^2dr-\frac{1}{1-\tau}\{\int_\tau^1 W(r)dr\}^2 \\
N_3(\tau) &\equiv \frac{W(1)^2-W(\tau)^2-(1-\tau)}{2}-\frac{1}{1-\tau}\{W(1)-W(\tau)\}\int_\tau^1 W(r)dr \\
D_3(\tau) &\equiv \int_\tau^1 W(r)^2dr-\{\int_0^1 W(r)dr\}^2+\frac{1}{\tau}\{\int_0^\tau W(r)dr\}^2 \\
&\quad -\frac{\tau}{1-\tau}\{\int_0^1 W(r)dr-\frac{1}{\tau}\int_0^\tau W(r)dr\}^2 \\
N_4(\tau) &\equiv \frac{-W(\tau)^2-\tau}{2}+\frac{1}{\tau}W(\tau)\int_0^\tau W(r)dr \\
D_4(\tau) &\equiv \int_0^\tau W(r)^2dr-\{\int_0^1 W(r)dr\}^2+\frac{1}{1-\tau}\{\int_\tau^1 W(r)dr\}^2 \\
&\quad -\frac{1-\tau}{\tau}\{\int_0^1 W(r)dr-\frac{1}{1-\tau}\int_\tau^1 W(r)dr\}^2
\end{aligned}
$$

and

$$
\begin{aligned}
DF_{01}^{f\,\inf} &\Rightarrow \inf_{\tau} \frac{N_1(\tau)}{D_1(\tau)^{1/2}} \\
DF_{10}^{r\,\inf} &\Rightarrow \inf_{\tau} \frac{N_2(\tau)}{D_2(\tau)^{1/2}} \\
DF_{10}^{f\,\inf} &\Rightarrow \inf_{\tau} \frac{N_3(\tau)}{D_3(\tau)^{1/2}} \\
DF_{01}^{r\,\inf} &\Rightarrow \inf_{\tau} \frac{N_4(\tau)}{D_4(\tau)^{1/2}}
\end{aligned}
$$

$W(r)$ is a standard Brownian motion and $T^{-1/2}\sum_{t=2}^{[\tau T]} \epsilon_t \Longrightarrow \sigma W(\tau)$, uniformly for $\tau \in [0,1]$.

Although it is perhaps not obvious from the above results, it can be easily shown that the limiting marginal null distributions of $DF_{01}^{f}(\tau)$ and $DF_{10}^{r}(1-\tau)$ are identical. This is because $DF_{10}^{r}(1-\tau)$ is by definition the $DF_{01}^{f}(\tau)$ test with y_t replaced by z_t. Similarly, the limiting marginal null distributions of $DF_{10}^{f}(\tau)$ and $DF_{01}^{r}(1-\tau)$ are also identical.

To see this, from the results in Theorem 2.3.1 under H_0^{11} we have that

$$
DF_{10}^{r}(1-\tau) \Rightarrow \frac{N_2(1-\tau)}{D_2(1-\tau)^{1/2}}.
$$

By defining a second Brownian motion process based on the last $1-\tau$ fraction of $\{\varepsilon_t\}$, that is $\sigma^{-1}T^{-1/2}\sum_{\tau T+1}^{T} \varepsilon_t \Rightarrow B(1-\tau)$ we have $B(1-\tau) \stackrel{d}{=} W(1) - W(\tau)$ and $B(1) \stackrel{d}{=} W(1)$. It follows that

$$
\begin{aligned}
&W(1-\tau) \stackrel{d}{=} B(1) - B(\tau) \\
&\int_{1-\tau}^{1} W(r)dr \stackrel{d}{=} \tau B(1) - \int_{0}^{\tau} B(r)dr \\
&\int_{1-\tau}^{1} W(r)^2 dr \stackrel{d}{=} \tau B(1)^2 + \int_{0}^{\tau} B(r)^2 dr - 2B(1)\int_{0}^{\tau} B(r)dr
\end{aligned}
$$

and on substituting and simplifying we find

$$N_2(1-\tau) \stackrel{d}{=} \frac{1}{2}\left\{B(\tau)^2-\tau\right\} - \frac{1}{\tau}B(\tau)\int_0^\tau B(r)dr,$$
$$D_2(1-\tau) \stackrel{d}{=} \int_0^\tau B(r)^2dr - \frac{1}{\tau}\{\int_0^\tau B(r)dr\}^2. \tag{2.7}$$

Hence, comparing expressions (2.7) with $N_1(\tau)$ and $D_1(\tau)$, one can verify that the marginal limiting null distributions of $DF_{01}^f(\tau)$ and $DF_{10}^r(1-\tau)$ are the same. A similar approach establishes a parallel result for $DF_{10}^f(\tau)$ and $DF_{01}^r(1-\tau)$. It then immediately follows that the limiting marginal null distributions of $DF_{01}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$ must be identical, as are those of $DF_{10}^{f\,\inf}$ and $DF_{01}^{r\,\inf}$.

2.3.2 Behaviour of the Tests under H^{01} and H^{10}

We now present the limiting results of the tests under H_1^{01} and H_1^{10}.

Theorem 2.3.2 *Under H_1^{01} and Assumption 2.1,*

$$T^{-1/2}DF_{01}^{f\,\inf} \xrightarrow{p} k(\tau^*,\rho)$$
$$T^{-1/2}\,DF_{01}^{r\,\inf} \xrightarrow{p} k(\tau^*,\rho)$$

$$DF_{10}^{r\,\inf} = O_p(1)$$
$$DF_{10}^{f\,\inf} = O_p(1)$$

where

$$k(\tau^*,\rho) = \frac{\tau^{*1/2}(\rho-1)}{(1-\rho^2)^{1/2}} < 0$$

and, asymptotically, $DF_{01}^{f\,\inf} = DF_{01}^f(\tau^*)$ *and* $DF_{01}^{r\,\inf} = DF_{01}^r(\tau^*)$.

Hence, $DF_{01}^{f\,\inf}$ and $DF_{01}^{r\,\inf}$ will both be consistent at the rate $T^{1/2}$ under H_1^{01}, whereas $DF_{10}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$ will not be consistent. Notice also that the function $k(\tau^*,\rho)$ is monotonic decreasing in τ^*. Thus, in large samples at least, the power of the $DF_{01}^{f\,\inf}$ test should grow with increasing τ^*. This makes sense because as τ^* increases, the $I(0)$ component constitutes a larger proportion of the series. Moreover, these results suggest that each of these first two tests will provide us with a consistent estimator of the true break fraction τ^*. For instance for $DF_{01}^{f\,\inf}$, such an estimate is given by that value of τ at which $DF_{01}^{r}(\tau)$ attains its minimum.

It is not necessary to separately derive the behaviour of the tests under H_1^{10}. These can be inferred directly from the above results as H_1^{10} is simply the mirror image of H_1^{01}in terms of the reversed series z_t, with the break at time $1-\tau^*$. Thus, we have the following corollary

Corollary 2.3.1 *Under H_1^{10} and Assumption 2.1.*

$$\begin{aligned} DF_{01}^{f\,\inf} &= O_p(1) \\ DF_{01}^{r\,\inf} &= O_p(1) \end{aligned}$$

$$\begin{aligned} T^{-1/2}DF_{10}^{r\,\inf} &\xrightarrow{p} k(1-\tau^*,\rho) \\ T^{-1/2}DF_{10}^{f\,\inf} &\xrightarrow{p} k(1-\tau^*,\rho) \end{aligned}$$

and, asymptotically, $DF_{10}^{r\,\inf} = DF_{10}^{r}(\tau^*)$ *and* $DF_{10}^{f\,\inf} = DF_{10}^{f}(\tau^*)$.

Hence, $DF_{10}^{r\,\inf}$ and $DF_{10}^{f\,\inf}$ are consistent at the rate $T^{1/2}$ under H_1^{10}, whereas $DF_{01}^{f\,\inf}$ and $DF_{01}^{r\,\inf}$ will not be consistent.[9] Again each of the latter two tests yields a consistent estimate of τ^*.

In short, it appears that each test statistic is consistent for the alternative against which it is designed, but not for the alternative of a break in persistence in the opposite direction. Moreover, the use of a test against the true alternative hypothesis directly leads to consistent estimation of the true break fraction.

[9]Note the change from τ^* to $1-\tau^*$ in the limiting function k.

2.4 Pairwise Minimum Tests: Direction Not Specified

The four tests analysed so far are strictly useful only if their appropriate particular alternative is considered. For instance, $DF_{01}^{f\,\inf}$ has power against H_1^{01} in that it diverges to $-\infty$. Hence, non-rejection can be taken as evidence in favour of H_0^{11}, but only if H_1^{10} is not a possibility. Once we allow the possibility that either H_1^{01} or H_1^{10} may hold, as we often need to in practice, we cannot on the basis of a non-rejection by $DF_{01}^{f\,\inf}$, reliably distinguish between H_1^{10} and H_0^{11}, as the test is $O_p(1)$ in either case. Similarly, whilst $DF_{10}^{r\,\inf}$ has power against H_1^{10}, a non-rejection by this test does not distinguish between H_0^{11} and H_1^{01}. Furthermore, while $DF_{01}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$ are $O_p(1)$ respectively under H_1^{10} and H_1^{01}, their actual limit distributions will still be highly dependent on τ^* and ρ. For instance, if H_1^{01} is true and τ^* is large and ρ is small, y_t and z_t are then series whose behaviour, broadly speaking, has more in common with an $I(0)$ process than an $I(1)$ process. Thus, even though $DF_{10}^{r\,\inf}$ is not divergent, intuition leads us to expect that it will take more negative values than it would under H_0^{11}; that is, it will be over-sized when its null critical values are consulted. We may, therefore, quite possibly face a situation where H_1^{01} is true yet both $DF_{01}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$ reject the null hypothesis. This, of course, would be an uninformative outcome.

Thus, once more than a single alternative hypothesis is permitted we cannot conduct reliable inference based on a sequence of these individual tests. A solution is then to consider the tests jointly. Suppose we consider the behaviour of $DF_{01}^{f\,\inf}$ jointly with that of $DF_{10}^{r\,\inf}$. As we know, $DF_{01}^{f\,\inf}$ ($DF_{10}^{r\,\inf}$) has power against H_1^{01} (H_1^{10}), diverging to $-\infty$, but is $O_p(1)$ under H_0^{11} or H_1^{10} (H_1^{01}). This suggests a "two-sided" test of H_0^{11} against the union of H_1^{01} and H_1^{10} can be based on the pairwise minimum of $DF_{01}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$; that is

$$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}).$$

These same hypotheses can also be tested using

$$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf}).$$[10]

The following theorem establishes limiting null distributions for the min(.,.) statistics.

Theorem 2.4.1 *Under H_0^{11} and Assumption 2.1,*

$$\begin{aligned}
\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) &\Rightarrow \min\{\inf_{\tau}\frac{N_1(\tau)}{D_1(\tau)^{1/2}}, \inf_{\tau}\frac{N_2(\tau)}{D_2(\tau)^{1/2}}\} \\
\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf}) &\Rightarrow \min\{\inf_{\tau}\frac{N_3(\tau)}{D_3(\tau)^{1/2}}, \inf_{\tau}\frac{N_4(\tau)}{D_4(\tau)^{1/2}}\}.
\end{aligned}$$

Our interest, however, centers on their behaviour under the alternatives. This is inferred using the results reported in the previous section. Consider the $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ test. Under H_1^{01}, $DF_{01}^{f\,\inf}$ diverges to $-\infty$ at the rate $T^{1/2}$ whilst $DF_{10}^{r\,\inf}$ is $O_p(1)$. Hence, it must also be true that $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ diverges to $-\infty$ at the rate $T^{1/2}$. Asymptotically, then,

$$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) = DF_{01}^{f\,\inf} = DF_{01}^{f}(\tau^*).$$

Under H_1^{10}, $DF_{01}^{f\,\inf}$ is $O_p(1)$ and $DF_{10}^{r\,\inf}$ diverges to $-\infty$ at the rate $T^{1/2}$. Hence, $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ diverges to $-\infty$ at the rate $T^{1/2}$ and, asymptotically,

$$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) = DF_{10}^{r\,\inf} = DF_{10}^{r}(\tau^*).$$

The $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ test is therefore consistent at the rate $T^{1/2}$ in rejecting H_0^{11} against H_1^{01} or H_1^{10}. Furthermore, with the same rate of consistency we will be able to select between H_1^{01} and H_1^{10} as, if at some chosen significance level,

[10] These pairings are selected, rather than say $\min(DF_{01}^{f\,\inf}, DF_{10}^{f\,\inf})$, in order to combine statistics whose constituent elements have identical marginal limiting null distributions. It is then sensible to work with min(.,.), as equal values for the constituent statistics should constitute approximately equal evidence in favour of the corresponding alternative hypothesis.

$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ rejects H_0^{11} and $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) = DF_{01}^{f\,\inf}$ clearly the appropriate conditional decision rule is to reject H_0^{11} in favour of H_1^{01}. Conversely, if $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ rejects H_0^{11} and $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) = DF_{10}^{r\,\inf}$ we reject H_0^{11} in favour of H_1^{10}. For the reasons outlined in the previous section, this procedure also directly provides us with a consistent estimate of τ^*. These same comments apply equally to the second pairwise minimum test $\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$.

2.5 Monte Carlo Simulations

To supplement the above theoretical results on the properties of the test statistics of interest, Monte Carlo simulations are undertaken where we explore the behaviour of the asymptotically valid tests in moderate sized samples. We begin by approximating and tabulating the finite sample critical values of the tests. Since the finite sample results assume Gaussian innovations, we report next simulation evidence on the effect of non-normality on the size of the tests. Subsequently, their robustness is investigated in the presence of serial correlation in the error term, in which case augmented variants of the tests are employed. Finally, power estimates of the tests under consideration are derived and some results on the quality of the estimated breakpoint are reported.

2.5.1 Finite Sample Critical Values

The preceding analysis resulted in the following two sets of test statistics, both of which account for a known a priori as well as an unknown direction of a possible change in persistence under the alternative hypothesis

$$\begin{array}{ll} DF_{01}^{f\,\inf} & DF_{10}^{f\,\inf} \\ DF_{10}^{r\,\inf} & DF_{01}^{r\,\inf} \\ \min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) & \min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf}). \end{array} \tag{2.8}$$

We use Monte Carlo simulation to calculate the finite sample critical values of the test statistics under the null of a unit root, that is, under model (2.1). Here, and throughout, the ε_t are generated as independent standard normal variates. The variance σ_ε^2 is fixed at unity without loss of generality, since $\hat{\phi}$ is independent of σ_ε^2. To obtain samples which closely resemble stationarity under the alternative we generate $T + 200$ observations and discard the first 200 setting $y_{-201} = 0$. Tables 2.1(a) and 2.1(b) provide the 10%, 5% and 1% critical values of the test statistics for various sample sizes T. Our simulations of the null are based on 10,000 replications with the search parameter τ traversing the interval $[0.2, 0.8]$ in steps of 0.01. To reduce variability between sample size comparisons we utilised the same set of generated data for simulations with the same T.

As the sample size increases, the critical values of all tests give the appearance of converging to limits. In fact, the distributions do not change very much as T increases past the value of 150. The min(.,.) tests maintain lower critical values than their individual counterparts, while overall the first set of statistics has critical values closer to zero compared to the second set. This means that not as large values of the test statistics are required for rejection of the unit root null hypothesis in the former case, as they are in the latter.

2.5.2 Robustness of Tests to Serial Correlation

Suppose that, instead of (2.1), y_t is generated by the $ARIMA(k, 1, 0)$ process

$$\Delta y_t = \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t$$

and that under H_1^{01} the alternative generating model is

$$\begin{aligned} \Delta y_t &= \alpha + \rho y_{t-1} + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t, && t \leq \tau^* T, \quad \rho < 1, \\ \Delta y_t &= \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t, && t > \tau^* T. \end{aligned}$$

Then, as in Dickey and Fuller (1979) (2.3) is augmented to include k lagged difference terms

$$\Delta y_t = \hat{\alpha}(\tau) d_{10t}(\tau) + \hat{\phi}(\tau) d_{10t}(\tau) y_{t-1} + \sum_{i=1}^{k} \hat{\phi}_i(\tau) \Delta y_{t-i} + \hat{\varepsilon}_t(\tau).$$

The other three regressions, (2.4), (2.6) and (2.5) are similarly augmented. When k is a fixed function of the sample size such that $k \to \infty$ and $k^3/T \to 0$ as $T \to \infty$, we appeal to results of Said and Dickey (1984) to argue that the limiting distributions of the test statistics in this case remain unchanged.[11] The limiting null distributions of the tests based on these augmented regressions are then the same as those reported in Section 2.3. The behaviour of the tests under the alternative hypotheses with this sort of generalisation remains also unchanged. This approach, however, is valid only asymptotically.

To ensure that the null critical values are robust to more general $I(1)$ null DGPs when augmented variants of the tests are employed in order to eliminate possible residual autocorrelation, we examine the empirical size of the tests under such circumstances. Table 2.2 shows the results for the case in which the null is generated as an $ARIMA(1,1,0)$ model. The percentage of rejections under the null hypothesis of the lower tail tests at the nominal 5% level are reported for $T \in \{100, 200, 300\}$ and for values of the autoregressive parameter $\rho \in \{0.0, 0.3, -0.3, -0.8\}$. Results indicate that the sizes are approximately correct for all tests with small size distortions in the downward direction occurring for the extreme value of $\rho = -0.8$, in which case y_t closely resembles a stationary process. These are slightly more apparent for the second set of tests. However, the degree of under-rejection observed is only modest with the proportion of rejections increasing towards the 5% level as T increases. Such a phenomenon does not seem to occur for large and positive values of ρ, in which case the series is nearly an $I(2)$ process. Overall, the rejection probabilities under the null are very close to nominal size for all tests, indicating that the asymptotic results work well for the samples sizes reported.

2.5.3 The Case of Non-Normal Errors

The asymptotic properties of the tests in (2.8) do not depend on the normality of the disturbances, but their finite sample distributions do. Thus, while all tests are asymptotically valid under quite general conditions, Monte Carlo experiments

[11] Maddala and Kim (1998) show that the reason behind this result is that in a regression of an $I(1)$ variable on $I(1)$ and $I(0)$ variables, the asymptotic distribution of the coefficient of the $I(1)$ and $I(0)$ variables are independent.

with non-normal errors on H_0 are needed to shed some light on the relevance of asymptotic theory to finite sample behaviour. To explore what happens when the critical values are no longer obtained under normal errors, the disturbance term in (2.1) is derived from drawings from the Student-t distribution with five degrees of freedom, $t(5)$, and the chi-square distribution with three degrees of freedom, $\chi^2(3)$. The $t(5)$ distribution is symmetric but has fatter tails than the normal. The $\chi^2(3)$ distribution is fairly skewed.

The robustness of the estimated significance levels to the choice of the error distribution is illustrated in Tables 2.3(a) - 2.3(d). Replacing the normal distribution by either the $t(5)$ or the $\chi^2(3)$ distributions does not seem to have much of an effect on the size of the first set of tests. Using the critical values of Tables 2.1(a) and 2.1(b), empirical sizes for the first set of tests are very close to the conventional 10%, 5% and 1% significance levels. For the second set of tests the rejection probabilities under the null rise above nominal size, with correct size restoring rather slowly as T increases. This phenomenon is somewhat more apparent in the case of the t distribution.

Thus, robustness of the tests to the error distribution, which is known to hold asymptotically, extends reasonably well to moderate sample sizes for the first set of statistics. However, caution needs to be exercised when employing the second set of tests with data of small sample size that exhibit departures from normality.

2.5.4 Power Estimates and Comparisons

Tests should not, of course, be judged only by the degree of agreement between small sample significance levels and nominal values. While it is important to avoid tests with badly behaved significance levels, the ability of tests to detect inadequate specifications merits consideration. For this reason, power estimates are obtained and compared across the tests.

We examine finite sample power performance of the two sets of tests in (2.8) under $H_1^{01}$when the data are generated according to (2.2). As all the tests are invariant to α under H_1^{01}, we set $\alpha = 0$. Power estimates are derived based on 3000 replications, using the critical values in Tables 2.1(a) and 2.1(b). Cases in which all tests have power estimates close to unity are not very useful for making comparisons, and we therefore report results for samples of $T = 100, 200$ and

300 to illustrate how the tests compare in less extreme situations. In view of producing more precise results we conduct our power simulations for a range of parameter values for the true breakpoint τ^* and the autoregressive parameter ρ. Tables 2.4(a)-2.4(c) give the proportion of rejections under the alternative hypothesis at the nominal 0.05 level for $\rho = \{0.7, 0.8, 0.9\}$ and $\tau^* = \{0.3, 0.5, 0.7\}$. The bracketed figure beneath the rejection frequency of $\min(DF_{01}^{f\inf}, DF_{10}^{r\inf})$ is the frequency with which $\min(DF_{01}^{f\inf}, DF_{10}^{r\inf}) = DF_{01}^{f\inf}$, conditional on the fact that $\min(DF_{01}^{f\inf}, DF_{10}^{r\inf})$ rejects the null hypothesis and expressed as a proportion of the total number of replications. That is, it measures the frequency with which the correct decision H_1^{01} is made. The figure under the rejection frequency of $\min(DF_{10}^{f\inf}, DF_{01}^{r\inf})$ is analogously defined. A possible alternative test statistic to the min(.,.) for testing that either H_1^{01}or H_1^{10} holds, i.e when the direction of any possible change is not specified a priori, is that based on the absolute difference of the individual forward and reverse tests. We investigate the finite sample power behaviour of this alternative statistic, which enters the tables denoted $\left|DF_{01}^{f\inf} - DF_{10}^{r\inf}\right|$ and $\left|DF_{10}^{f\inf} - DF_{01}^{r\inf}\right|$, respectively. Moreover, we assess the effect of a change in persistence on the power of the conventional DF test, the results of which are presented in the first row of entry of the power tables.

We will initially restrict discussion to the results that correspond to the first set of statistics. From the upper section of Tables 2.4(a)-2.4(c) consistency of $DF_{01}^{f\inf}$ and $\min(DF_{01}^{f\inf}, DF_{10}^{r\inf})$ in T, for any fixed values of τ^* and ρ, is clearly evident. As we would expect, for fixed T, the powers of these tests also increase with movements away from H_0^{11} in terms of increasing $\tau^*$and/or decreasing ρ. Not surprisingly, the power of all the tests increases with τ^*, as the proportion of the series that is $I(0)$ is growing.

As for $DF_{10}^{r\inf}$, its inconsistency in T shows up reasonably well except in the case where $\tau^* = 0.7$, when convergence to a limiting distribution is rather slow. Its rejection rates for a fixed T, depend on $\tau^*$and ρ. It is slightly under-sized for small τ^* and/or large ρ , but becomes over-sized with increasing τ^* and/or decreasing ρ, as we would expect given the discussion in the previous section. Thus, just as non-rejection by $DF_{10}^{r\inf}$ does not necessarily signal H_0^{11}, a rejection need not be associated with H_1^{10}. This behaviour makes inference based on the individual tests

$DF_{01}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$ problematic once either alternative hypothesis is considered possible, as is generally the case in practice. For example, with $T = 300$, $\tau^* = 0.7$ and $\rho = 0.7$, there is a 37% chance that *both* $DF_{01}^{f\,\inf}$ and $DF_{10}^{r\,\inf}$ will reject H_0^{11} in favour of their mutually inconsistent alternatives. The pairwise mimimum test is constructed to avoid this sort of conflict. The price is, however, a loss of power. It is clear that $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ has lower power everywhere compared to $DF_{01}^{f\,\inf}$, but this is as we would expect since the latter statistic tests only for the single (correct) alternative hypothesis H_1^{01} instead of both alternatives and has null critical values that are closer to zero than the pairwise test. However, the conditional decision rules for selecting between H_1^{01} and H_1^{10} based on rejections by the pairwise test appear to work well, in that, compared to H_1^{01}, H_1^{10} is generally chosen very infrequently.

Turning to the performance of the $\left|DF_{01}^{f\,\inf} - DF_{10}^{r\,\inf}\right|$ test, results show that the power estimates for such a test statistic are overall lower than those displayed by the $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ test. When the break occurs early on in the sample, notably for $\tau^* = 0.3$, the power differences between the two tests are fairly moderate. However, for larger break fraction values these differences increase and become quite striking by the time τ^* reaches 0.7. In particular, for $T = 100$, $\tau^* = 0.7$ and $\rho = 0.7$ the power of the $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ test is 0.679 whereas the $\left|DF_{01}^{f\,\inf} - DF_{10}^{r\,\inf}\right|$ test has power equal to 0.252. The Dickey-Fuller statistic, on the other hand, demonstrates the lowest power of all consistent tests under the impact of a change in persistence, although it does seem to perform better than the inconsistent $DF_{10}^{r\,\inf}$ test.

The second set of statistics tend to exhibit on the whole higher rejection frequencies compared to the first set as illustrated in the lower part of Tables 2.4(a)-2.4(c), while qualitatively results are similar to those described above. The $\left|DF_{10}^{f\,\inf} - DF_{01}^{r\,\inf}\right|$ test has power slightly closer to the $\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$ test when the breakpoint occurs late in the sample, though its power performance continues to remain inferior.

Results in Section 2.4 showed that the two min(.,.) tests have different limiting null distributions, but the same large sample properties under H_1^{01}. Then, the choice of which test to use in practice is determined by their relative finite sample power characteristics. Tables 2.4(a)-2.4(c) imply that, on average,

$\min(DF_{10}^{f\inf}, DF_{01}^{r\inf})$ is the more powerful of the two tests, though never substantially so. This is slightly counterintuitive as $\min(DF_{01}^{f\inf}, DF_{10}^{r\inf})$ has null critical values which are closer to zero, yet the tests behave similarly under H_1^{01}, at least in large samples. Perhaps this just serves to illustrate that large sample behaviour can be a misleading indicator of finite sample behaviour.[12]

Finally, the mean and standard deviation of the associated estimates of τ^* implicit from the various tests are reported in Tables 2.5(a)-2.5(f). The estimates of τ^* based on $\min(DF_{01}^{f\inf}, DF_{10}^{r\inf})$ and $DF_{01}^{f\inf}$, and $\min(DF_{10}^{f\inf}, DF_{01}^{r\inf})$ and $DF_{01}^{r\inf}$ appear consistent, although this is most easily seen for the case where $\tau^* = 0.7$, as for smaller τ^* convergence is rather slow. As we may have anticipated, the estimates from the one-sided tests are more accurate than those from the two-sided tests. Estimates of τ^* based on $DF_{10}^{r\inf}$ and $DF_{10}^{f\inf}$ are clearly inconsistent, apart from when the breakpoint occurs at the midpoint in which case results for the latter test could be somewhat misleading. Specifically, for $\tau^* = 0.5$ the $DF_{10}^{f\inf}$ test displays estimates of τ^* closer to the true value than the $DF_{01}^{r\inf}$ test.

2.6 Extensions

A known intercept of zero under the null as considered throughout this study introduces an important element of specificity, which affects the finite sample distributions. Since economic time series may contain deterministics such as linear trends, it is required that extensions are made to take them into account. We briefly consider the case where the null generating model (2.1) incorporates a non-zero drift, and the $I(0)$ part of the alternative contains a linear trend.

As currently defined, all our test statistics are invariant under all the hypotheses considered to transformations of the form $y_t \to c_1 + y_t$ for an arbitrary constant c_1. To accomplish invariance to the transformation of the form $y_t \to c_1 + c_2 t + y_t$ we augment the regression (2.3) to include a constant and time trend dummy

[12]Note that it is not necessary to simulate under H_1^{10}, as a process that is generated under this alternative is simply a time reversal of a series generated under H_1^{01}, with $(1-\tau^*)$ in place of τ^*.

variable

$$\Delta y_t = \hat{\alpha}_1(\tau) + \hat{\alpha}_2(\tau)d_{10t}(\tau) + \hat{\beta}(\tau)d_{10t}(\tau)t + \hat{\phi}(\tau)d_{10t}(\tau)y_{t-1} + \hat{\varepsilon}_t(\tau).$$

The other three regressions (2.4), (2.6) and (2.5) are similarly augmented. The limiting null distributions of the tests based on the augmented regressions are different to those reported earlier when only a constant is included in the regression. Simulated critical values, using the same data-generating process as for the results in Tables 2.1(a) and 2.1(b), are provided in Tables 2.1(c) and 2.1(d). The limiting behaviour of the tests under the alternative hypotheses in this case remains unchanged.

2.7 Empirical Application

We apply our tests to inflation rates in the G-7 countries. The data source utilised is the seasonally adjusted quarterly Consumer Price Index from OECD Main Economic Indicators, covering the period from 1960:I-1999:IV, a total of 160 observations. The inflation rate is calculated by differencing the logarithm of the price indices. Plots of the inflation rates are given in Figure 2.1.

As a preliminary step, augmented Dickey-Fuller tests were carried out.[13] The number of lagged changes was selected through the general-to-specific data dependent method analysed in Ng and Perron (1995) where the last included lag is checked for significance using a two-tailed t-test at the 10% level. This data dependent method requires a maximal autoregressive lag length, $k\max$. Results are reported in the first column of Table 2.6, for $k\max = 12$. The Dickey-Fuller tests fail to reject the null hypothesis that each of the series is $I(1)$ throughout, at the conventional significance levels.[14]

To test the possibility of a switch from $I(0)$ to $I(1)$ or vice versa, we apply the test indicated as most powerful in the simulations of the previous section, that is $\min(DF_{10}^{f\inf}, DF_{01}^{r\inf})$. In the regressions underlying the individual tests,

[13] A time trend is not included in the Dickey-Fuller equations as this would not be consistent with the behaviour of post-war inflation.

[14] We performed a robustness check by allowing a greater value for $kmax$. The results were very similar.

the number of lagged differences used were the same as those identified in the preceding augmented Dickey-Fuller regressions, while the search procedure for the potential breakpoint was carried out over the proportion [0.2, 0.8] of the sample size. The empirical test results are shown in Table 2.7. Using the critical values of Table 2.1(b), it can be seen from column five that the $\min(DF_{10}^{f\,\mathrm{inf}}, DF_{01}^{r\,\mathrm{inf}})$ test rejects the null hypothesis of $I(1)$ throughout at the 1% critical level for two countries, namely Canada and France, and is very close to rejecting at the same level for a third country - Japan. There is also a rejection at the 10% level for Italy. In all four of these cases we find that $\min(DF_{10}^{f\,\mathrm{inf}}, DF_{01}^{r\,\mathrm{inf}}) = DF_{01}^{r\,\mathrm{inf}}$, indicating the likelihood of a change from $I(0)$ to $I(1)$.

In addition, we may wish to examine the alternative hypotheses separately. These results are reported in the first four columns of Table 2.7. If the alternative hypothesis $H_1^{01}$is considered, then $DF_{01}^{r\,\mathrm{inf}}$ obviously rejects the null for these same countries, though not for any additional ones. If H_1^{10} alone is considered, $DF_{10}^{f\,\mathrm{inf}}$ fails to reject the null for any country, which is in agreement with the outcome from the $\min(DF_{10}^{f\,\mathrm{inf}}, DF_{01}^{r\,\mathrm{inf}})$ test.

The consistency of these results indicates not only that, contrary to the information provided by the usual Dickey-Fuller tests, there is strong evidence against the hypothesis of $I(1)$ behaviour throughout, but that there is considerable evidence of a switch from $I(0)$ to $I(1)$ behaviour for at least the majority of the G-7 country inflation rates. The discontinuity manifested as a change in persistence effect occurs around 1973 for all series. This may well be associated with the oil price shock of that year.

The final two columns of Table 2.6 illustrate the results of augmented Dickey-Fuller tests applied to the two subsamples of these four series, identified according to the estimated break dates of Table 2.7. In all cases, very strong evidence of $I(0)$ behaviour is found for the pre-break series, but much less strong evidence is present in the post-break part. This outcome supports the evidence of the above test results.

2.8 Conclusions

This chapter contributes to the literature on unit root testing under the alternative of structural change. Previous studies in this area have focused mainly on a shift in the trend function. The possibility, however, of a switch in persistence of particular time series has been argued over the years by a number of economists, both on institutional and empirical grounds. Taking difference-stationarity as the null hypothesis, we have developed tests designed to have power against alternatives involving such a switch. Our alternative hypothesis is therefore of a shift either from $I(0)$ to $I(1)$ or from $I(1)$ to $I(0)$, and tests are developed for the two cases, with the location of any possible breakpoint unspecified. If, in addition, the analyst does not wish to specify the direction of change, we showed how these two tests can be combined to test against the two-sided alternative. In either case, a consistent estimator of any break fraction resulted as a by-product of the tests. Sequential testing procedures were considered.

Of course, it is quite possible that a series is $I(0)$ throughout. Although our tests will obviously have some power against this alternative, various power-enhancing modifications of the usual Dickey-Fuller test can be expected to be more powerful. Our aim here is not to replace such tests, but to supplement them with tests against a particular alternative that economists have thought to be plausible.

Simulation evidence on the behaviour of the tests in moderate sized samples was provided. The tests were found to control well for size in most cases, except in the case of non-normality of the disturbances where moderate size distortions in the upward direction were encountered for the $\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$ test. From the power comparisons conducted the $\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$ test emerged as a reliable and useful procedure as did the $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$ test. The $\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$ test, however, proved to be slightly more powerful. An empirical application of the tests to consumer price inflation in the G-7 countries, uncovered quite strong and consistent evidence of a change from $I(0)$ to $I(1)$ behaviour occurring in 1973 for the majority of these countries.

References

Amemiya, T., 1973. Regression analysis when the dependent variable is truncated normal. *Econometrica*, 41, 997-1016.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61, 821–856.

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62, 1383-1414.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66, 47-78.

Ball, L., Cecchetti, S.G., 1990. Inflation and uncertainty at short and long horizons. *Brookings Papers on Economic Activity* 1, 215-254.

Banerjee, A., Lumsdaine, R., Stock, J., 1992. Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. *Journal of Business and Economic Statistics*, 10, 271-288.

Barsky, R.B., 1987. The fisher hypothesis and the forecastabilty and persistence of inflation. *Journal of Monetary Economics*, 19, 3-24.

Brown, R.L., Durbin, J., Evans, J.M., 1975. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, Series B, 37, 149-192.

Brunner, A.D., Hess, G.D., 1993. Are higher levels of inflation less predictable? A state-dependent conditional heteroskedasticity approach. *Journal of Business and Economic Statistics*, 11, 187-197.

Chu, C-S. J., White, H., 1992. A Direct Test for Changing Trend. *Journal of Business & Economic Statistics*, 10, 289-300.

De Long, J.B., Summers, L.H., 1988. How does macroeconomic policy affect output? *Brookings Papers on Economic Activity*, 2, 433-494.

Evans, M., Wachtel, P., 1993. Inflation regimes and the sources of inflation uncertainty. *Journal of Money, Credit and Banking*, 25, 475-511.

Fishe, R.P.H., 1991. The Federal Reserve amendments of 1917: The beginning of a seasonal note issue policy. *Journal of Money, Credit and Banking*, 23, 308-326.

Fishe, R.P.H., Wohar, M.E., 1990. The adjustment of expectations to a change in regime: Comment. *American Economic Review,* 80, 968-976.

Hakkio, C.S., Rush, M., 1991. Is the budget deficit too large? *Economic Inquiry*, 29, 429-445.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the Business Cycle. *Econometrica*, 57, 357-384.

Hansen, B.E., 1992. Tests for parameter instability in regressions with I(1) processes. *Journal of Business & Economic Statistics*, 10, 321–335.

Kim, J., 2000. Detection of change in persistence of a linear time series. *Journal of Econometrics*, 95, 97-116.

Kool, C.J.M., 1995. War finance and interest rate targeting: Regime changes in 1914-1918. *Explorations in Economic History*, 32, 365-382.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159-178.

Lam, P.S., 1990. The Hamilton model with a general autoregressive component. *Journal of Monetary Economics*, 26, 409-432.

Leybourne, S.J., 1995. Testing for unit roots using forward and reverse Dickey-Fuller Regressions. *Oxford Bulletin of Economics and Statistics,* 57, 559-571.

Leybourne, S.J., McCabe, B.P.M., 1994. A consistent test for a unit root. *Journal of Business & Economic Statistics*, 12, 157-176.

Leybourne, S., Kim, T-H., Smith, L.V., Newbold, P., 2000. Tests for a change in persistence against the null of difference-stationarity with application to the behaviour of price inflation. *Working paper.* Department of Economics, University of Nottingham.

Leybourne, S., Newbold, P., Vougas, D., 1998. Unit roots and smooth transitions. *Journal of Time Series Analysis,*19, 83-97.

Lumsdaine, R.L., Papell, D.H., 1997. Multiple trend breaks and the unit-root hypothesis. *The Review of Economics and Statistics, 212-218.*

Maddala, G.S., Kim, I.M., 1998. Unit roots, cointegration and structural change. *Cambridge University Press.*

Mankiw, N.G., Miron, J.A., Weil, D.N., 1987. The adjustment of expectations to a change in regime: A study of the founding of the federal reserve. *American*

Economic Review, 77, 358-374.

Newbold, P., Leybourne, S.J., Sollis, R., Wohar, M.E., 2001. U.S and U.K interest rates 1890-1934: New evidence on structural breaks. *Journal of Money, Credit and Banking*, 33, 235-250.

Ng, S., Perron, P., 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.

Parker, R., 1989. Some evidence on the persistence of price and output shocks in the pre-World War I and post-World War II eras. *Economic Notes*, 0, 413–422.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, 57, 1361-1401.

Perron, P., 1990. Testing for a unit root in a time series with a changing mean. *Journal of Business & Economic Statistics,* 8, 153-162.

Perron, P., Vogelsang, T.J., 1992. Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business & Economic Statistics,* 10, 301-320.

Perron, P., 1997. Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics*, 80, 355-386.

Ploberger, W., Krämer, W., Kontrus, K., 1989. A new test for structural stability in the linear regression model. *Journal of Econometrics*, 40, 307–318.

Rappoport, P., Reichlin, L., 1989. Segmented trends and non-stationary time series. *The Economic journal*, 99, 168–177.

Said, E.S., Dickey, D.A , 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71, 599-607.

Saikkonen, P., Lütkepohl, H., 2002. Testing for a unit root in a time series with a level shift at unknown time. *Econometric Theory*, 18, 313-348.

Stock, J.H., 1994. Unit roots structural breaks and trends. In: Engle, R.F., McFadden, D.L. (Eds.), *Handbook of Econometrics*, Vol.4. New York: North Holland, 2740-2841.

Zivot, E., Andrews, D., 1992. Further evidence on the great crash, the oil price shock, and the unit root hypothesis. *Journal of Business & Economic Statistics,* 10, 251-270.

Appendix 2.A Mathematical Proofs

Proof of Theorem 2.3.1. First we consider the t−ratio associated with $\hat{\phi}(\tau)$ in regression (2.3)

$$DF_{01}^{f}(\tau) = \frac{\sum_2^{\tau T}(y_{t-1} - \bar{y}_1)\Delta y_t}{\hat{\sigma}\left\{\sum_2^{\tau T}(y_{t-1} - \bar{y}_1)^2\right\}} \tag{2.9}$$

where $\bar{y}_1 \equiv (\tau T)^{-1}\sum_1^{\tau T} y_t$ and $\hat{\sigma}^2 \equiv \tau\tilde{\sigma}^2 + (1-\tau)\bar{\sigma}^2$. It is straightforward to show that

$$DF_{01}^{f}(\tau) = \frac{\tilde{\sigma}}{\hat{\sigma}}\hat{t}_{DF}(\tau).$$

It has been shown in BLS that

$$\begin{aligned} \hat{t}_{DF}(\tau) &\Rightarrow \frac{N_1(\tau)}{D_1(\tau)^{1/2}} \\ \tilde{\sigma}^2 &\Rightarrow \sigma^2 \end{aligned}$$

while it can be easily derived that

$$\bar{\sigma}^2 \Rightarrow \sigma^2.$$

Therefore $DF_{01}^{f}(\tau)$ and $\hat{t}_{DF}(\tau)$ are asymptotically equivalent and hence $DF_{01}^{f}(\tau) \Rightarrow \frac{N_1(\tau)}{D_1(\tau)^{1/2}}$, which implies that $\inf_{\tau} DF_{01}^{f}(\tau) \Rightarrow \inf_{\tau} \frac{N_1(\tau)}{D_1(\tau)^{1/2}}$ by the continuous mapping theorem. The same argument can be used to establish the other claims in the Theorem. ■

Proof of Theorem 2.3.2. Let us consider the behaviour of $DF_{01}^{f\,\inf}$ under H_1^{01}. The t-ratio associated with $\widehat{\phi}(\tau)$ in regression (2.3) is given by (2.9) and $\widehat{\sigma}^2$ can be written as

$$\begin{aligned} \widehat{\sigma}^2 \;=\; & T^{-1}\sum_2^{T}\Delta y_t^2 + \widehat{a}(\tau)^2\tau + \widehat{\phi}(\tau)^2 T^{-1}\sum_2^{\tau T} y_{t-1}^2 - 2\widehat{a}(\tau)T^{-1}\sum_2^{\tau T}\Delta y_t \\ & + 2\widehat{a}(\tau)\widehat{\phi}(\tau)T^{-1}\sum_2^{\tau T} y_{t-1} - 2\widehat{\phi}(\tau)T^{-1}\sum_2^{\tau T} y_{t-1}\Delta y_t. \end{aligned} \tag{2.10}$$

Initially, the case where $\tau \le \tau^*$ is examined. Here, the components in $DF_{01}^{f}(\tau)$

apart from $\widehat{\sigma}$ involve only the y_t that are stationary $AR(1)$. So, the behaviour of $DF^f_{01}(\tau)$ is very much like that of the usual Dickey-Fuller test under the stationary alternative. Using the fact that

$$\begin{aligned}\widehat{\phi}(\tau) \xrightarrow{p} \rho - 1, \quad E(\Delta y_t^2) &= 2\sigma^2 \frac{1-\rho}{1-\rho^2}, \quad \widehat{a}(\tau) \xrightarrow{p} a, \quad E(y_t^2) = \frac{\sigma^2}{1-\rho^2} + \frac{a^2}{(1-\rho)^2} \\ E(\Delta y_t) &= 0, \quad E(y_t) = \frac{a}{1-\rho}, \quad E(y_{t-1}\Delta y_t) = \frac{\sigma^2(\rho-1)}{1-\rho^2}\end{aligned}$$

it follows that

$$\begin{aligned} T^{-1}\sum_2^T \Delta y_t^2 \xrightarrow{p} \sigma^2(1 + \tau^* \frac{1-\rho}{1+\rho}), &\quad \widehat{\phi}(\tau)^2 \tau (\tau T)^{-1} \sum_2^{\tau T} y_{t-1}^2 \xrightarrow{p} \sigma^2 \tau \frac{1-\rho}{1+\rho} + a^2 \tau \\ \widehat{a}(\tau)\widehat{\phi}(\tau) T^{-1} \sum_2^{\tau T} y_{t-1} \xrightarrow{p} -a^2\tau, &\quad \widehat{\phi}(\tau) T^{-1} \sum_2^{\tau T} y_{t-1}\Delta y_t \xrightarrow{p} \sigma^2 \tau \frac{1-\rho}{1+\rho}. \end{aligned}$$

Gathering together these results we have that

$$\widehat{\sigma}^2 \xrightarrow{p} \sigma^2\{1 + (\tau^* - \tau)\frac{1-\rho}{1+\rho}\}.$$

Then,

$$\begin{aligned} T^{-1/2} DF^f_{01}(\tau) &\xrightarrow{p} \frac{\tau^{1/2} E(y_{t-1}\Delta y_t)}{\sigma\{1 + (\tau^* - \tau)\frac{1-\rho}{1+\rho}\}^{1/2} Var(y_t)^{1/2}} \\ &= \frac{\tau^{1/2}(\rho - 1)}{\{1 + (\tau^* - \tau)\frac{1-\rho}{1+\rho}\}^{1/2}(1-\rho^2)^{1/2}}. \end{aligned} \tag{2.11}$$

Since $\rho < 1$ it follows that for $\tau \le \tau^*$ the probability limit of $T^{-1/2}DF^f_{01}(\tau)$ is negative.

Then, by the Continuous Mapping Theorem (CMT) we have

$$\inf_{\tau \in \Lambda} T^{-1/2} DF^f_{01}(\tau) \Rightarrow \inf_{\tau\in\Lambda} \frac{\tau^{1/2}(\rho-1)}{\{1 + (\tau^* - \tau)\frac{1-\rho}{1+\rho}\}^{1/2}(1-\rho^2)^{1/2}}.$$

Now because

$$\frac{\tau^{1/2}(\rho-1)}{\{1 + (\tau^* - \tau)\frac{1-\rho}{1+\rho}\}^{1/2}(1-\rho^2)^{1/2}}$$

is monotonically decreasing on the interval $(0, \tau^*]$, the minimum occurs at $\tau = \tau^*$. Thus,

$$T^{-1/2} DF_{01}^{f\,\inf} \xrightarrow{p} k(\tau^*, \rho) = \frac{\tau^{*1/2}(\rho - 1)}{(1-\rho^2)^{1/2}}$$

and so $DF_{01}^{f\,\inf}$ directly provides us with a consistent estimator of the true break fraction τ^*, where such an estimate is given by that value of τ at which $DF_{01}^{f}(\tau)$ attains its minimum.[15] Note that $\inf_{\tau} T^{-1/2} DF_{01}^{f}(\tau) = T^{-1/2} DF_{01}^{f\,\inf}$.

We consider now the case where $\tau > \tau^*$. In this case, the components in $DF_{01}^{f}(\tau)$ involve the y_t that are stationary AR(1) and then become the random walk. The random walk components dominate and so the behaviour $DF_{01}^{f}(\tau)$ is similar to that of the usual Dickey-Fuller test under the random walk null. Let us consider the numerator term in (2.9) scaled by T^{-1}

$$\begin{aligned} & T^{-1} \sum_{2}^{\tau T} (y_{t-1} - \overline{y}_1) \Delta y_t \\ = \; & T^{-1} \sum_{\tau^* T+1}^{\tau T} y_{t-1} \varepsilon_t - \tau^{-1} T^{-3/2} \sum_{\tau^* T+1}^{\tau T} y_{t-1} T^{-1/2} \sum_{\tau^* T+1}^{\tau T} \varepsilon_t + O_p(T^{-1/2}) \\ \Longrightarrow \; & \sigma^2 \frac{1}{2} \{W(\tau)^2 - W(\tau^*)^2 - (\tau - \tau^*)\} - \sigma^2 \tau^{-1} \{W(\tau) - W(\tau^*)\} \int_{\tau^*}^{\tau} W(r) dr. \end{aligned}$$

The denominator term is scaled by $T^{-1/2}$

$$\begin{aligned} T^{-2} \sum_{2}^{\tau T} (y_{t-1} - \overline{y}_1)^2 \; &= \; T^{-2} \sum_{\tau^* T+1}^{\tau T} y_{t-1}^2 - \tau^{-1} (T^{-3/2} \sum_{\tau^* T+1}^{\tau T} y_{t-1})^2 + O_p(T^{-1/2}) \\ &\Longrightarrow \; \sigma^2 \int_{\tau^*}^{\tau} W(r)^2 dr - \sigma^2 \tau^{-1} \{ \int_{\tau^*}^{\tau} W(r) dr \}^2. \end{aligned}$$

[15] A formal proof of consistency for the break fraction estimator $\widehat{\tau} = \arg\inf_{\tau \in \Lambda} DF_{01}^{f}(\tau)$, that is, $\widehat{\tau} - \tau^* = o_p(1)$ from which it follows that $\widehat{\tau} \xrightarrow{p} \tau^*$, is based on a similar argument as in Lemma 3 of Amemiya (1973). The author proves that when $Q_T(\omega, \widehat{\theta}_T(\omega)) = \sup_{\theta \in \Theta} Q_T(\omega, \theta)\ \forall\, \omega \in \Omega$, where $Q_T(\omega, \theta)$ is a measurable function on a measurable space Ω and for each $\omega \in \Omega$ a continuous function for θ in a compact set Θ, if $Q_T(\omega, \theta)$ converges to $Q(\theta)$ uniformaly for all θ in Θ, and if $Q(\theta)$ has a unique maximum at $\theta_0 \in \Theta$, then $\widehat{\theta}_T$ converges to θ_0 almost everywhere.

Hence,

$$T\widehat{\phi}(\tau) \Longrightarrow \frac{\frac{1}{2}\{W(\tau)^2 - W(\tau^*)^2 - (\tau-\tau^*)\} - \tau^{-1}\{W(\tau)-W(\tau^*)\}\int_{\tau^*}^{\tau} W(r)dr}{\int_{\tau^*}^{\tau} W(r)^2 dr - \tau^{-1}\{\int_{\tau^*}^{\tau} W(r)dr\}^2}$$

so that $\widehat{\phi}(\tau) = O_p(T^{-1})$. Although the details are omitted, a similar analysis applied to $\widehat{a}(\tau)$ also shows that $\widehat{a}(\tau) = O_p(T^{-1})$. Using these results, it then follows that in (2.14)

$$\widehat{\sigma}^2 = T^{-1}\sum_{2}^{T} \Delta y_t^2 + O_p(T^{-1}) \xrightarrow{p} \sigma^2(1+\tau^*\frac{1-\rho}{1+\rho})$$

which delivers

$$DF_{01}^{f}(\tau) \Longrightarrow (1+\tau^*\frac{1-\rho}{1+\rho})^{-1/2}[\int_{\tau^*}^{\tau} W(r)^2dr - \tau^{-1}\{\int_{\tau^*}^{\tau} W(r)dr\}^2]^{-1/2}$$

$$\times [\frac{1}{2}\{W(\tau)^2 - W(\tau^*)^2 - (\tau-\tau^*)\} - \tau^{-1}\{W(\tau)-W(\tau^*)\}\int_{\tau^*}^{\tau} W(r)dr].$$

Hence, $DF_{01}^{f}(\tau) = O_p(1)$ for any $\tau \in (0,1)$. Thus, $DF_{01}^{f\,\inf}(\tau) = O_p(1)$ also. On adopting the same standardisation as in (2.11) $T^{-1/2}DF_{01}^{f}(\tau) \xrightarrow{p} 0$.

Next, the result $DF_{10}^{r}(\tau) = O_p(1)$ under H_1^{01} is derived. Recalling the definition of $z_t \equiv y_{T+1-t}$ and further defining $\eta_t \equiv -\varepsilon_{T+2-t}$, it can be easily shown that there exist constants a^*, ρ^* with $|\rho^*| < 1$ and a zero mean whit noise $\widetilde{\eta}_t$ such that

$$\begin{aligned} z_t &= z_{t-1} + \eta_t, & t &< (1-\tau^*)T \\ z_t &= y_{\tau^* T}, & t &= (1-\tau^*)T \\ z_t &= a^* + \rho^* z_{t-1} + \widetilde{\eta}_t, & t &> (1-\tau^*)T. \end{aligned}$$

Consider an arbitrary $\tau \in (0,1)$ for which the t-ratio is given by

$$DF_{10}^{r}(\tau) = \frac{\sum_2^{\tau T}(z_{t-1} - \bar{z}_1)\Delta z_t}{\hat{\sigma}\left\{\sum_2^{\tau T}(z_{t-1} - \bar{z}_1)^2\right\}} \tag{2.12}$$

where $\bar{z}_1 \equiv (\tau T)^{-1}\sum_1^{\tau T} z_t$, $\hat{\sigma}^2 \equiv \tau\tilde{\sigma}^2 + (1-\tau)\bar{\sigma}^2$, $\tilde{\sigma}^2 \equiv (\tau T)^{-1}\sum_{t=1}^{\tau T}\hat{\eta}_t^2$ and $\bar{\sigma}^2 \equiv ((1-\tau)T)^{-1}\sum_{t=\tau T+1}^{T}\Delta z_t^2$. It is straightforward to show that $\hat{\sigma}^2 = O_p(1)$.

Firstly, the case where $\tau \leq (1-\tau^*)$ is considered. In this case, $DF_{10}^{r}(\tau)$ has the standard Dickey-Fuller limiting distribution and hence is $O_p(1)$. In the other case where $\tau > (1-\tau^*)$, the numerator scaled by T^{-1}, $T^{-1}\sum_2^{\tau T}(z_{t-1} - \bar{z}_1)\Delta z_t$, is written as follows

$$\begin{aligned}
& T^{-1}\sum_2^{(1-\tau^*)T} z_{t-1}\eta_t - \frac{1}{\tau}\{T^{-3/2}\sum_2^{(1-\tau^*)T} z_t + T^{-3/2}\sum_{(1-\tau^*)T}^{\tau T} z_t\}\{T^{-1/2}\sum_2^{(1-\tau^*)T}\eta_t \\
& +T^{-1/2}\sum_2^{(1-\tau^*)T}\Delta z_t\} + a^*T^{-1}\sum_{(1-\tau^*)T}^{\tau T} z_{t-1} + (\rho^*-1)T^{-1}\sum_{(1-\tau^*)T}^{\tau T} z_{t-1}^2 \\
& +T^{-1}\sum_{(1-\tau^*)T}^{\tau T} z_{t-1}\eta_t \\
= \; & O_p(1) + \{O_p(1) + o_p(1)\}O_p(1) + O_p(1) + O_p(1) + O_p(1) \\
= \; & O_p(1).
\end{aligned}$$

The same argument can be applied to show that $T^{-2}\left\{\sum_2^{\tau T}(z_{t-1} - \bar{z}_1)^2\right\} = O_p(1)$ which implies that $DF_{10}^{r}(\tau) = O_p(1)$ when $\tau > (1-\tau^*)$. Hence, we have $DF_{10}^{r}(\tau) = O_p(1)$ for any $\tau \in (0,1)$. Thus, $DF_{10}^{r\,\inf}(\tau) = O_p(1)$ also.

Very similar arguments to these establish the other claims in Theorem 2.3.2.

■

Proof of Theorem 2.3.2. From the Fuctional Central Limit Theorem, statistics such as $DF_{01}^{f}(\tau)$ and $DF_{10}^{r}(\tau)$ jointly converge to $\frac{N_1(\tau)}{D_1(\tau)^{1/2}}$ and $\frac{N_2(\tau)}{D_2(\tau)^{1/2}}$ uniformly in τ. This joint convergence result together with the fact that min(.,.) is continuous in both arguments implies that $\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf}) \Rightarrow \min\{\inf_\tau \frac{N_1(\tau)}{D_1(\tau)^{1/2}}, \inf_\tau \frac{N_2(\tau)}{D_2(\tau)^{1/2}}\}$ by the Continuous Mapping Theorem. A similar result holds for $\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$. ■

Appendix 2.B Tables and Figures

Table 2.1(a)

Critical values of tests with constant dummies at the 10%, 5% and 1% critical levels

	$DF_{01}^{f\,\inf}$			$DF_{10}^{r\,\inf}$			$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	-3.170	-3.429	-3.993	-3.182	-3.442	-3.989	-3.426	-3.668	-4.207
150	-3.177	-3.451	-3.995	-3.205	-3.490	-4.031	-3.459	-3.691	-4.190
200	-3.224	-3.490	-4.034	-3.199	-3.481	-3.976	-3.472	-3.718	-4.189
250	-3.215	-3.472	-3.959	-3.222	-3.493	-4.011	-3.466	-3.704	-4.194
300	-3.213	-3.491	-4.031	-3.223	-3.477	-4.044	-3.468	-3.723	-4.217
400	-3.210	-3.481	-3.998	-3.206	-3.469	-3.992	-3.462	-3.699	-4.187
500	-3.220	-3.457	-3.980	-3.209	-3.488	-4.011	-3.457	-3.714	-4.198

Table 2.1(b)

Critical values of tests with constant dummies at the 10%, 5% and 1% critical levels

	$DF_{10}^{f\,\inf}$			$DF_{01}^{r\,\inf}$			$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	-3.642	-3.922	-4.444	-3.633	-3.895	-4.434	-3.890	-4.135	-4.616
150	-3.586	-3.850	-4.385	-3.586	-3.842	-4.368	-3.835	-4.075	-4.541
200	-3.635	-3.897	-4.386	-3.618	-3.899	-4.402	-3.889	-4.125	-4.551
250	-3.663	-3.935	-4.465	-3.649	-3.914	-4.367	-3.911	-4.141	-4.550
300	-3.625	-3.897	-4.369	-3.608	-3.882	-4.373	-3.877	-4.116	-4.574
400	-3.626	-3.880	-4.359	-3.613	-3.866	-4.376	-3.865	-4.085	-4.554
500	-3.611	-3.861	-4.344	-3.633	-3.888	-4.336	-3.868	-4.089	-4.529

Table 2.1(c)

Critical values of tests with trend dummies at the 10%, 5% and 1% critical levels

	$DF_{01}^{f\,\inf}$			$DF_{10}^{r\,\inf}$			$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	-3.722	-3.984	-4.558	-3.743	-4.011	-4.586	-3.984	-4.257	-4.778
150	-3.737	-4.017	-4.513	-3.749	-3.999	-4.551	-3.990	-4.243	-4.718
200	-3.786	-4.040	-4.582	-3.758	-4.020	-4.541	-4.006	-4.264	-4.728
250	-3.776	-4.043	-4.547	-3.797	-4.061	-4.574	-4.039	-4.275	-4.758
300	-3.796	-4.058	-4.574	-3.779	-4.042	-4.549	-4.035	-4.282	-4.736
400	-3.790	-4.047	-4.570	-3.788	-4.032	-4.523	-4.023	-4.287	-4.746
500	-3.804	-4.071	-4.581	-3.793	-4.060	-4.518	-4.051	-4.276	-4.778

Table 2.1(d)

Critical values of tests with trend dummies at the 10%, 5% and 1% critical levels

	$DF_{10}^{f\,\inf}$			$DF_{01}^{r\,\inf}$			$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	-4.096	-4.357	-4.905	-4.067	-4.343	-4.865	-4.344	-4.585	-5.060
150	-4.031	-4.275	-4.778	-4.023	-4.279	-4.777	-4.270	-4.519	-4.936
200	-4.057	-4.301	-4.791	-4.092	-4.344	-4.827	-4.312	-4.548	-5.012
250	-4.110	-4.367	-4.831	-4.093	-4.345	-4.820	-4.348	-4.569	-5.005
300	-4.075	-4.318	-4.837	-4.085	-4.338	-4.830	-4.318	-4.544	-5.035
400	-4.061	-4.301	-4.827	-4.078	-4.331	-4.846	-4.308	-4.553	-4.998
500	-4.081	-4.318	-4.797	-4.087	-4.336	-4.830	-4.320	-4.539	-5.008

Table 2.2

Empirical size of tests under $\Delta y_t = \phi \Delta y_{t-1} + \varepsilon_t$

	ϕ	0.0	0.3	0.8	−0.3	−0.8
Size						
				$T = 100$		
0.05	$DF_{01}^{f\,\inf}$	0.057	0.055	0.061	0.052	0.048
	$DF_{10}^{r\,\inf}$	0.056	0.055	0.058	0.051	0.045
	$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.058	0.058	0.061	0.054	0.049
	$DF_{10}^{f\,\inf}$	0.049	0.050	0.055	0.043	0.028
	$DF_{01}^{r\,\inf}$	0.052	0.056	0.061	0.049	0.027
	$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.051	0.056	0.058	0.046	0.029
				$T = 200$		
0.05	$DF_{01}^{f\,\inf}$	0.052	0.051	0.045	0.044	0.034
	$DF_{10}^{r\,\inf}$	0.052	0.048	0.049	0.049	0.039
	$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.052	0.051	0.046	0.045	0.037
	$DF_{10}^{f\,\inf}$	0.048	0.050	0.052	0.044	0.031
	$DF_{01}^{r\,\inf}$	0.051	0.050	0.054	0.042	0.026
	$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.050	0.052	0.055	0.041	0.028
				$T = 300$		
0.05	$DF_{01}^{f\,\inf}$	0.051	0.050	0.048	0.044	0.039
	$DF_{10}^{r\,\inf}$	0.054	0.049	0.050	0.049	0.044
	$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.051	0.047	0.046	0.049	0.042
	$DF_{10}^{f\,\inf}$	0.050	0.048	0.055	0.046	0.034
	$DF_{01}^{r\,\inf}$	0.052	0.049	0.052	0.045	0.032
	$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.050	0.050	0.055	0.046	0.032

Note: Nominal size 0.05.

Table 2.3(a)

Empirical size of tests when $\varepsilon_t \tilde{} t(5)$

	$DF_{01}^{f\,\inf}$			$DF_{10}^{r\,\inf}$			$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	0.104	0.061	0.018	0.107	0.061	0.019	0.121	0.072	0.022
150	0.110	0.061	0.016	0.100	0.053	0.014	0.113	0.066	0.018
200	0.098	0.051	0.013	0.107	0.055	0.016	0.104	0.061	0.018
250	0.096	0.055	0.016	0.096	0.052	0.014	0.108	0.062	0.016
300	0.103	0.055	0.014	0.097	0.051	0.012	0.105	0.060	0.014
400	0.101	0.056	0.013	0.102	0.054	0.013	0.107	0.059	0.014
500	0.097	0.055	0.014	0.104	0.055	0.013	0.110	0.057	0.015

Table 2.3(b)

Empirical size of tests when $\varepsilon_t \tilde{} t(5)$

	$DF_{10}^{f\,\inf}$			$DF_{01}^{r\,\inf}$			$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	0.175	0.113	0.049	0.177	0.113	0.050	0.211	0.143	0.070
150	0.152	0.094	0.033	0.155	0.098	0.036	0.180	0.114	0.048
200	0.143	0.087	0.032	0.154	0.090	0.035	0.167	0.107	0.045
250	0.141	0.084	0.028	0.146	0.088	0.037	0.166	0.105	0.047
300	0.133	0.078	0.027	0.137	0.076	0.025	0.147	0.092	0.032
400	0.129	0.074	0.023	0.132	0.079	0.022	0.147	0.089	0.029
500	0.135	0.076	0.022	0.119	0.070	0.023	0.140	0.083	0.029

Table 2.3(c)

Empirical size of tests when $\varepsilon_t \tilde{} \chi^2(3)$

	$DF_{01}^{f\,\inf}$			$DF_{10}^{r\,\inf}$			$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	0.103	0.060	0.015	0.100	0.057	0.015	0.116	0.064	0.017
150	0.105	0.055	0.014	0.093	0.048	0.014	0.103	0.059	0.017
200	0.095	0.050	0.012	0.098	0.051	0.014	0.101	0.055	0.015
250	0.098	0.055	0.015	0.101	0.056	0.013	0.110	0.061	0.016
300	0.098	0.051	0.011	0.096	0.051	0.011	0.103	0.055	0.014
400	0.104	0.050	0.010	0.100	0.053	0.013	0.103	0.052	0.013
500	0.096	0.056	0.011	0.101	0.052	0.012	0.108	0.056	0.013

Table 2.3(d)

Empirical size of tests when $\varepsilon_t \tilde{} \chi^2(3)$

	$DF_{10}^{f\,\inf}$			$DF_{01}^{r\,\inf}$			$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
100	0.149	0.096	0.040	0.158	0.103	0.041	0.188	0.125	0.057
150	0.133	0.082	0.030	0.130	0.080	0.026	0.156	0.101	0.040
200	0.132	0.081	0.028	0.138	0.077	0.025	0.153	0.095	0.036
250	0.131	0.076	0.024	0.135	0.078	0.030	0.150	0.093	0.038
300	0.123	0.067	0.023	0.129	0.072	0.024	0.136	0.081	0.027
400	0.124	0.066	0.022	0.121	0.070	0.018	0.133	0.078	0.024
500	0.120	0.066	0.019	0.113	0.063	0.020	0.125	0.072	0.023

Table 2.4(a)

Simulated rejection frequencies of tests at the nominal 0.05 level under H_1^{01}

	$\tau^* = 0.3$			$\tau^* = 0.5$			$\tau^* = 0.7$		
$T = 100$	$\rho = 0.9$	0.8	0.7	$\rho = 0.9$	0.8	0.7	$\rho = 0.9$	0.8	0.7
DF	0.046	0.059	0.069	0.066	0.119	0.157	0.129	0.263	0.369
$DF_{01}^{f\,\mathrm{inf}}$	0.080	0.153	0.266	0.106	0.298	0.548	0.143	0.427	0.787
$DF_{10}^{r\,\mathrm{inf}}$	0.045	0.048	0.039	0.054	0.053	0.068	0.086	0.127	0.203
$\min(DF_{01}^{f\,\mathrm{inf}}, DF_{10}^{r\,\mathrm{inf}})$	0.061	0.112	0.195	0.086	0.218	0.426	0.126	0.330	0.679
	(0.040)	(0.088)	(0.176)	(0.060)	(0.193)	(0.401)	(0.082)	(0.274)	(0.619)
$\lvert DF_{01}^{f\,\mathrm{inf}} - DF_{10}^{r\,\mathrm{inf}}\rvert$	0.058	0.082	0.146	0.056	0.137	0.271	0.048	0.121	0.252
$DF_{10}^{f\,\mathrm{inf}}$	0.048	0.036	0.035	0.039	0.029	0.027	0.042	0.049	0.067
$DF_{01}^{r\,\mathrm{inf}}$	0.117	0.212	0.338	0.154	0.363	0.613	0.164	0.418	0.748
$\min(DF_{10}^{f\,\mathrm{inf}}, DF_{01}^{r\,\mathrm{inf}})$	0.088	0.136	0.228	0.104	0.245	0.462	0.113	0.303	0.623
	(0.062)	(0.117)	(0.213)	(0.083)	(0.233)	(0.452)	(0.093)	(0.279)	(0.600)
$\lvert DF_{10}^{f\,\mathrm{inf}} - DF_{01}^{r\,\mathrm{inf}}\rvert$	0.066	0.090	0.136	0.077	0.150	0.271	0.087	0.181	0.328

Note: The bracketed figures beneath the rejection frequency of $\min(DF_{01}^{f\,\mathrm{inf}}, DF_{10}^{r\,\mathrm{inf}})$ is the frequency with which $\min(DF_{01}^{f\,\mathrm{inf}}, DF_{10}^{r\,\mathrm{inf}}) = DF_{01}^{f\,\mathrm{inf}}$, conditional on the fact that $\min(DF_{01}^{f\,\mathrm{inf}}, DF_{10}^{r\,\mathrm{inf}})$ rejects the null hypothesis and expressed as a proportion of the total number of replications. That is, it measures the frequency with which the correct decision H_1^{01} is made. The figure under the rejection frequency of $\min(DF_{10}^{f\,\mathrm{inf}}, DF_{01}^{r\,\mathrm{inf}})$ is analogously defined. The same applies for all subsequent tables of this form.

Table 2.4(b)

Simulated rejection frequencies of tests at the nominal 0.05 level under H_1^{01}

	$\tau^* = 0.3$			$\tau^* = 0.5$			$\tau^* = 0.7$		
$T = 200$	$\rho = 0.9$	0.8	0.7	$\rho = 0.9$	0.8	0.7	$\rho = 0.9$	0.8	0.7
DF	0.049	0.065	0.080	0.113	0.177	0.205	0.260	0.401	0.487
$DF_{01}^{f\,\inf}$	0.128	0.366	0.663	0.238	0.727	0.968	0.361	0.941	0.999
$DF_{10}^{r\,\inf}$	0.043	0.040	0.047	0.058	0.080	0.077	0.128	0.260	0.324
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.090	0.264	0.529	0.162	0.591	0.915	0.276	0.872	0.998
	(0.070)	(0.245)	(0.515)	(0.141)	(0.570)	(0.909)	(0.224)	(0.824)	(0.984)
$\left\|DF_{01}^{f\,\inf} - DF_{10}^{r\,\inf}\right\|$	0.072	0.200	0.388	0.100	0.381	0.657	0.090	0.378	0.664
$DF_{10}^{f\,\inf}$	0.046	0.034	0.037	0.035	0.031	0.033	0.052	0.100	0.140
$DF_{01}^{r\,\inf}$	0.199	0.472	0.727	0.340	0.811	0.980	0.376	0.936	0.999
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.142	0.344	0.581	0.225	0.684	0.943	0.276	0.857	0.995
	(0.122)	(0.330)	(0.570)	(0.210)	(0.676)	(0.939)	(0.252)	(0.839)	(0.991)
$\left\|DF_{10}^{f\,\inf} - DF_{01}^{r\,\inf}\right\|$	0.092	0.176	0.297	0.131	0.366	0.624	0.161	0.467	0.700

Table 2.4(c)

Simulated rejection frequencies of tests at the nominal 0.05 level under H_1^{01}

	$\tau^* = 0.3$			$\tau^* = 0.5$			$\tau^* = 0.7$		
$T = 300$	$\rho = 0.9$	0.8	0.7	$\rho = 0.9$	0.8	0.7	$\rho = 0.9$	0.8	0.7
DF	0.060	0.070	0.077	0.139	0.191	0.216	0.345	0.439	0.500
$DF_{01}^{f\,\inf}$	0.213	0.627	0.917	0.469	0.965	1.000	0.692	0.999	1.000
$DF_{10}^{r\,\inf}$	0.053	0.040	0.043	0.070	0.083	0.098	0.182	0.298	0.366
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.146	0.485	0.852	0.347	0.915	0.999	0.562	0.998	1.000
	(0.123)	(0.472)	(0.847)	(0.321)	(0.906)	(0.999)	(0.505)	(0.982)	(0.998)
$\left\|DF_{01}^{f\,\inf} - DF_{10}^{r\,\inf}\right\|$	0.116	0.360	0.662	0.210	0.663	0.911	0.207	0.654	0.898
$DF_{10}^{f\,\inf}$	0.043	0.036	0.038	0.037	0.040	0.037	0.074	0.115	0.150
$DF_{01}^{r\,\inf}$	0.333	0.737	0.945	0.598	0.979	1.000	0.710	0.999	1.000
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.214	0.580	0.877	0.439	0.937	1.000	0.561	0.996	1.000
	(0.197)	(0.567)	(0.869)	(0.424)	(0.933)	(0.999)	(0.540)	(0.989)	(0.999)
$\left\|DF_{10}^{f\,\inf} - DF_{01}^{r\,\inf}\right\|$	0.128	0.270	0.513	0.228	0.604	0.869	0.291	0.711	0.910

Table 2.5(a)

Estimated means and standard deviations of the breakpoint under H_1^{01}

$T=100$	$\tau^*=0.3$					
	$\rho=0.9$		$\rho=0.8$		$\rho=0.7$	
$\hat{\tau}$	*mean*	*std*	*mean*	*std*	*mean*	*std*
$DF_{01}^{f\,\inf}$	0.461	0.199	0.439	0.187	0.407	0.167
$DF_{10}^{r\,\inf}$	0.478	0.216	0.481	0.220	0.481	0.220
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.468	0.192	0.461	0.191	0.423	0.173
$DF_{10}^{f\,\inf}$	0.529	0.196	0.524	0.194	0.529	0.195
$DF_{01}^{r\,\inf}$	0.494	0.177	0.480	0.169	0.457	0.160
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.498	0.180	0.488	0.174	0.465	0.164

Table 2.5(b)

Estimated means and standard deviations of the breakpoint under H_1^{01}

$T=100$	$\tau^*=0.5$					
	$\rho=0.9$		$\rho=0.8$		$\rho=0.7$	
$\hat{\tau}$	*mean*	*std*	*mean*	*std*	*mean*	*std*
$DF_{01}^{f\,\inf}$	0.517	0.190	0.534	0.162	0.544	0.138
$DF_{10}^{r\,\inf}$	0.451	0.225	0.426	0.233	0.402	0.233
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.492	0.192	0.516	0.173	0.528	0.149
$DF_{10}^{f\,\inf}$	0.537	0.193	0.532	0.199	0.528	0.208
$DF_{01}^{r\,\inf}$	0.554	0.175	0.577	0.155	0.581	0.132
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.547	0.178	0.572	0.159	0.577	0.136

Table 2.5(c)

Estimated means and standard deviations of the breakpoint under H_1^{01}

$T=100$	$\tau^*=0.7$					
	$\rho=0.9$		$\rho=0.8$		$\rho=0.7$	
$\hat{\tau}$	*mean*	*std*	*mean*	*std*	*mean*	*std*
$DF_{01}^{f\,\inf}$	0.589	0.201	0.660	0.152	0.689	0.118
$DF_{10}^{r\,\inf}$	0.380	0.216	0.312	0.188	0.286	0.176
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.511	0.215	0.592	0.200	0.645	0.170
$DF_{10}^{f\,\inf}$	0.513	0.215	0.493	0.227	0.464	0.239
$DF_{01}^{r\,\inf}$	0.590	0.192	0.658	0.155	0.691	0.124
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.565	0.198	0.633	0.174	0.673	0.147

Table 2.5(d)

Estimated means and standard deviations of the breakpoint under H_1^{01}

	$\tau^* = 0.3$					
$T = 200$	$\rho = 0.9$		$\rho = 0.8$		$\rho = 0.7$	
$\hat{\tau}$	*mean*	*std*	*mean*	*std*	*mean*	*std*
$DF_{01}^{f\,\inf}$	0.431	0.183	0.391	0.150	0.361	0.119
$DF_{10}^{r\,\inf}$	0.482	0.220	0.482	0.220	0.482	0.222
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.455	0.187	0.407	0.159	0.372	0.130
$DF_{10}^{f\,\inf}$	0.532	0.196	0.533	0.191	0.537	0.189
$DF_{01}^{r\,\inf}$	0.474	0.166	0.441	0.147	0.403	0.128
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.485	0.170	0.448	0.151	0.409	0.133

Table 2.5(e)

Estimated means and standard deviations of the breakpoint under H_1^{01}

	$\tau^* = 0.5$					
$T == 200$	$\rho = 0.9$		$\rho = 0.8$		$\rho = 0.7$	
$\hat{\tau}$	*mean*	*std*	*mean*	*std*	*mean*	*std*
$DF_{01}^{f\,\inf}$	0.540	0.162	0.541	0.121	0.534	0.094
$DF_{10}^{r\,\inf}$	0.437	0.236	0.412	0.239	0.402	0.240
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.521	0.174	0.533	0.129	0.532	0.096
$DF_{10}^{f\,\inf}$	0.548	0.200	0.538	0.208	0.529	0.215
$DF_{01}^{r\,\inf}$	0.586	0.150	0.586	0.115	0.569	0.097
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.579	0.155	0.584	0.118	0.570	0.097

Table 2.5(f)

Estimated means and standard deviations of the breakpoint under H_1^{01}

	$\tau^* = 0.7$					
$T = 200$	$\rho = 0.9$		$\rho = 0.8$		$\rho = 0.7$	
$\hat{\tau}$	*mean*	*std*	*mean*	*std*	*mean*	*std*
$DF_{01}^{f\,\inf}$	0.656	0.160	0.703	0.097	0.714	0.066
$DF_{10}^{r\,\inf}$	0.332	0.205	0.274	0.172	0.259	0.158
$\min(DF_{01}^{f\,\inf}, DF_{10}^{r\,\inf})$	0.581	0.206	0.675	0.141	0.707	0.086
$DF_{10}^{f\,\inf}$	0.495	0.231	0.452	0.250	0.422	0.253
$DF_{01}^{r\,\inf}$	0.657	0.163	0.715	0.100	0.728	0.068
$\min(DF_{10}^{f\,\inf}, DF_{01}^{r\,\inf})$	0.629	0.182	0.705	0.118	0.726	0.074

Table 2.6

Augmented Dickey-Fuller tests for inflation rate series

Country	ADF: full sample	pre-break	post-break
US	-2.205	n/a	n/a
CANADA	-1.860	-5.858***	-1.870
JAPAN	-1.333	-6.490***	-1.843
UK	-1.614	n/a	n/a
GERMANY	-2.672	n/a	n/a
FRANCE	-1.022	-5.724***	-1.410
ITALY	-1.407	-4.950***	-1.430

Note: The symbols *, ** and *** imply significance at the 10%, 5% and 1% level respectively; n/a denotes "not applicable". Estimated break dates are only reported when a test statistic rejects H_0^{11}.

Table 2.7

Persistence change tests for inflation rate series

Country	$DF_{10}^{f\,\text{inf}}$	break date	$DF_{01}^{r\,\text{inf}}$	break date	min(.,.)	break date
US	-3.406	n/a	-3.488	n/a	-3.488	n/a
CAN	-2.362	n/a	-4.604**	1973:Q2	-4.604***	1973:Q2
JAPAN	-3.224	n/a	-4.524***	1973:Q2	-4.524**	1973:Q2
UK	-2.966	n/a	-2.947	n/a	-2.966	n/a
GER	-1.915	n/a	-2.825	n/a	-2.825	n/a
FR	-2.904	n/a	-4.786***	1973:Q4	-4.786***	1973:Q4
ITALY	-2.083	n/a	-3.900**	1973:Q4	-3.900*	1973:Q4

Note: The symbols *, ** and *** imply significance at the 10%, 5% and 1% level respectively using the critical values from Table 1(b). The number of lags is that utilised in the ADF tests; n/a denotes "not applicable". Estimated break dates are only reported when a test statistic rejects H_0^{11}.

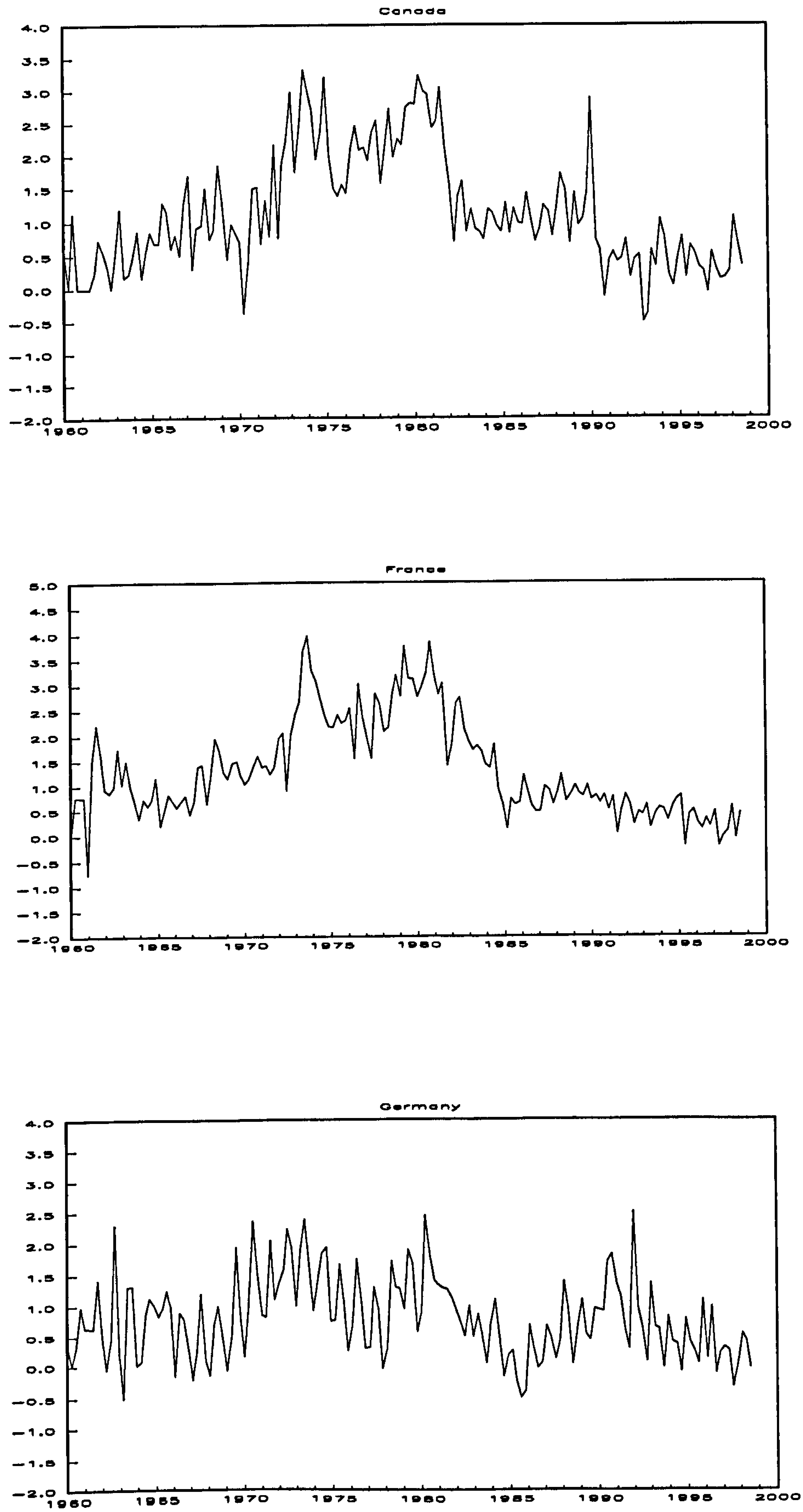
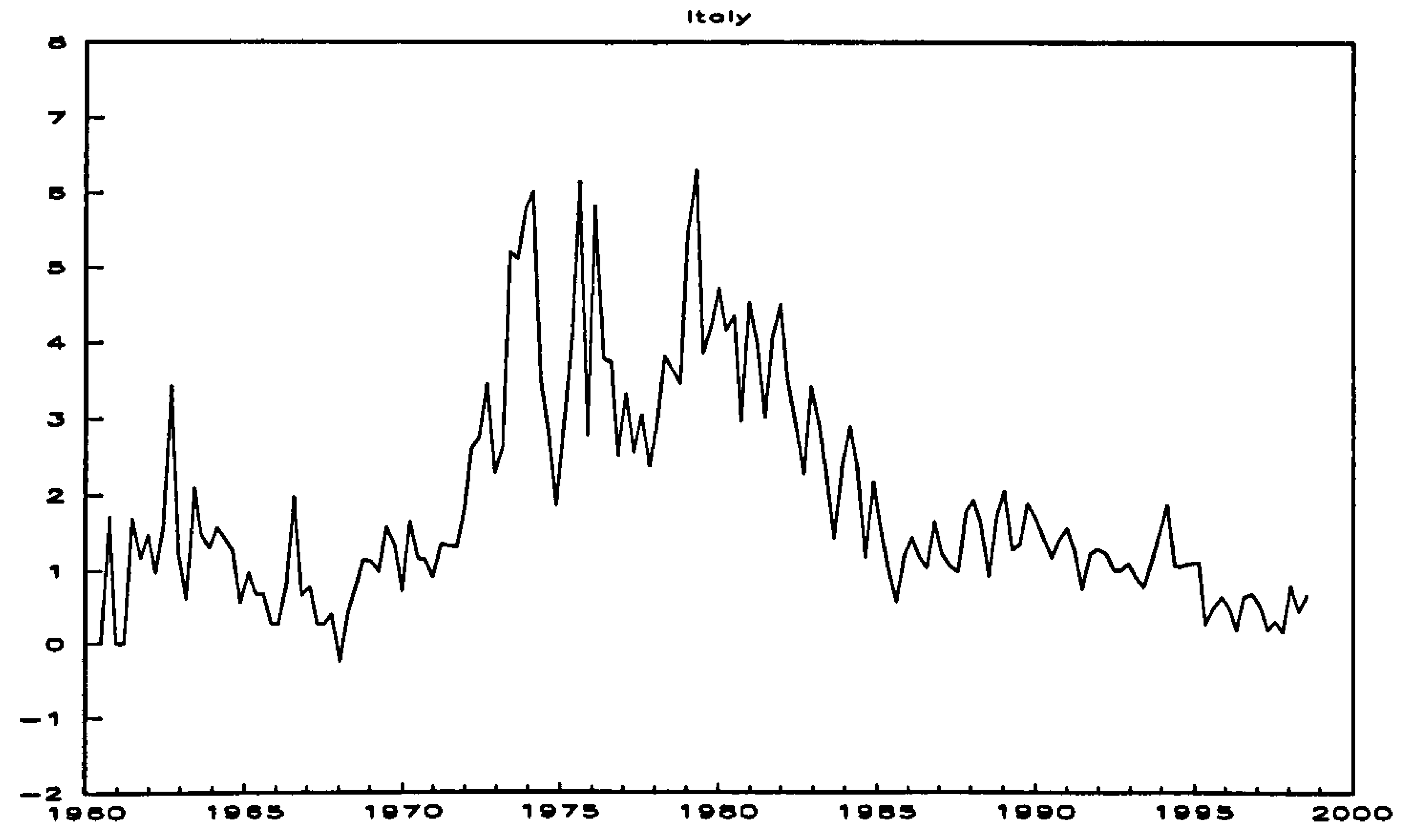


Figure 2.1. Plots of inflation rates of the G-7 countries.



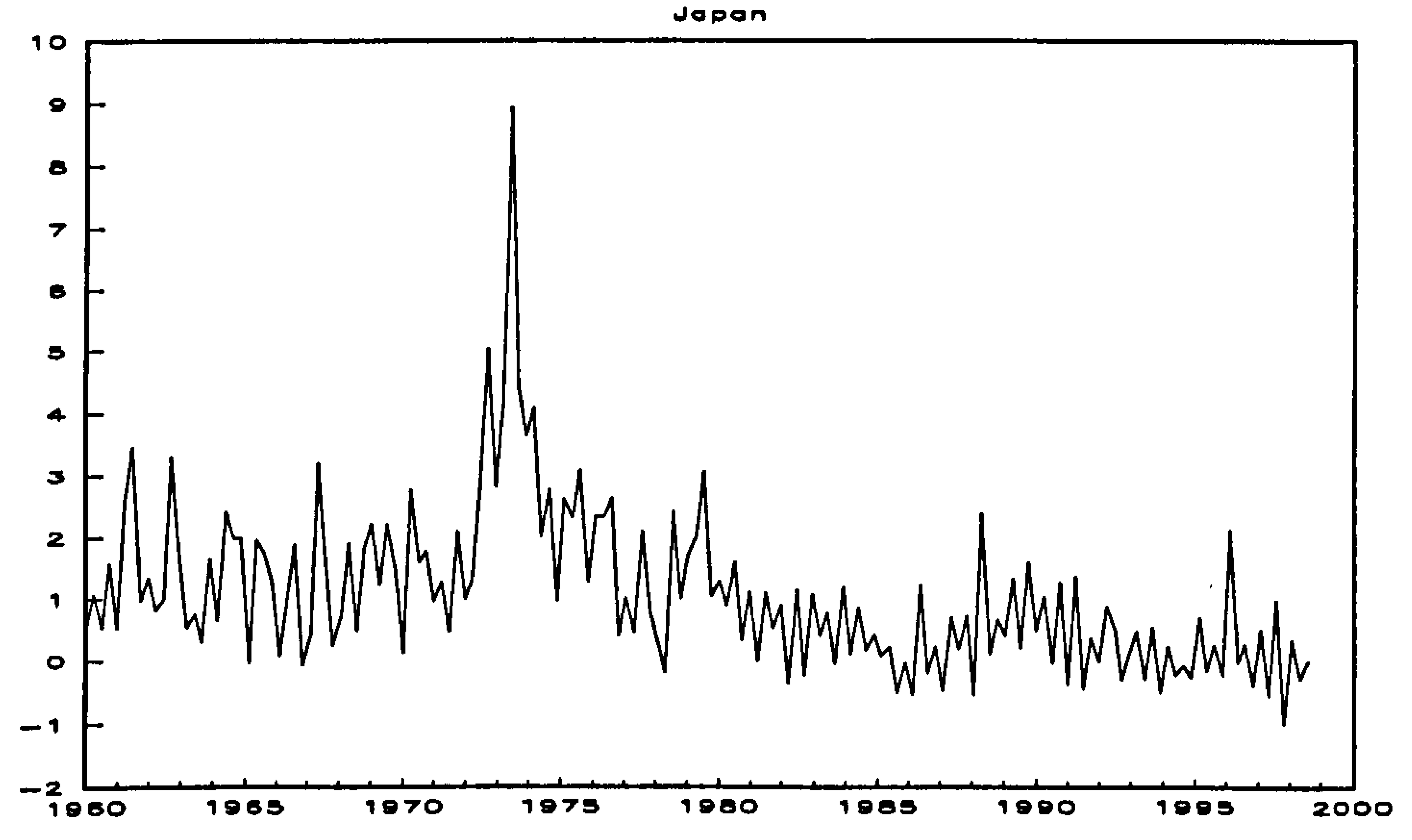
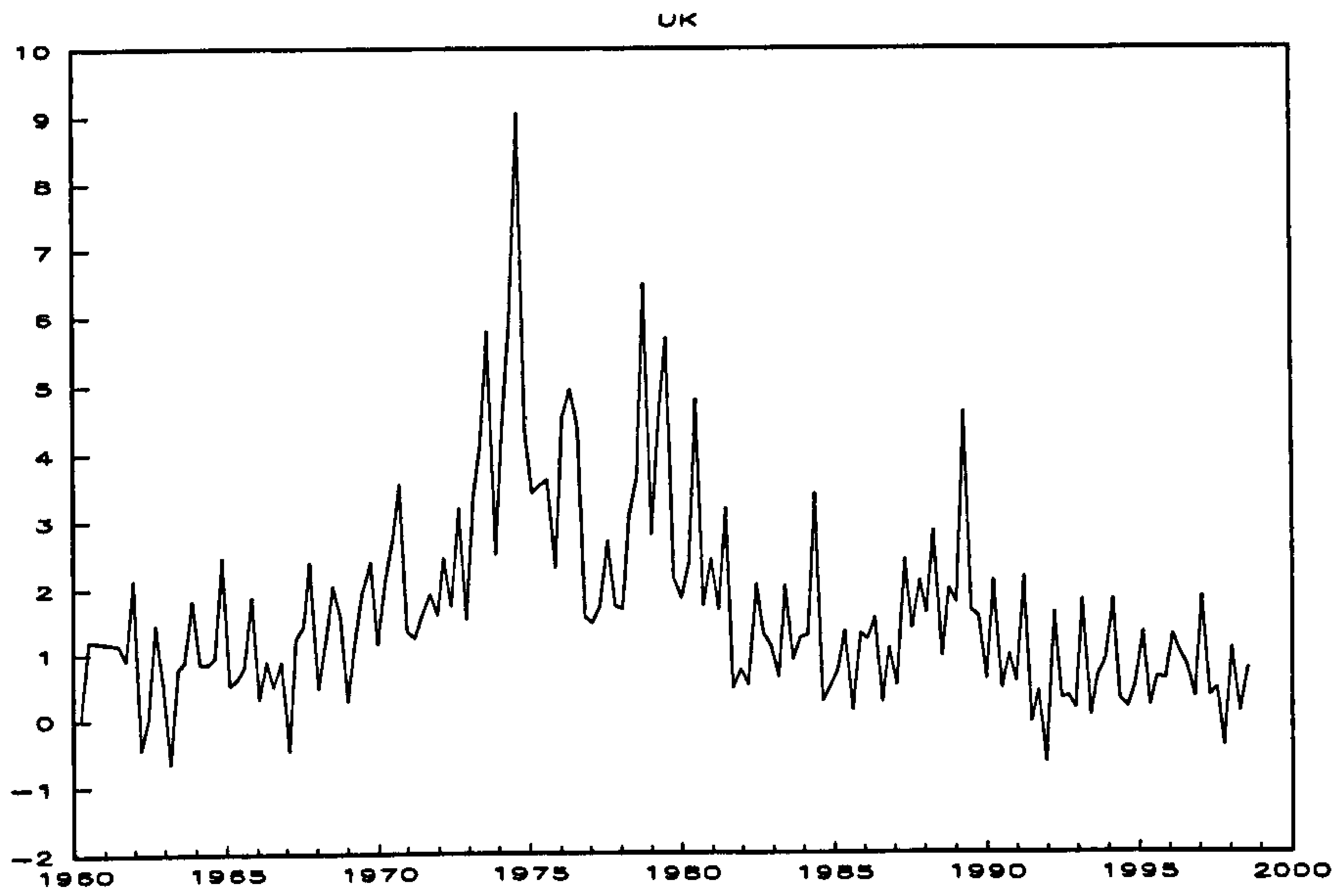


Figure 2.1. (continued)



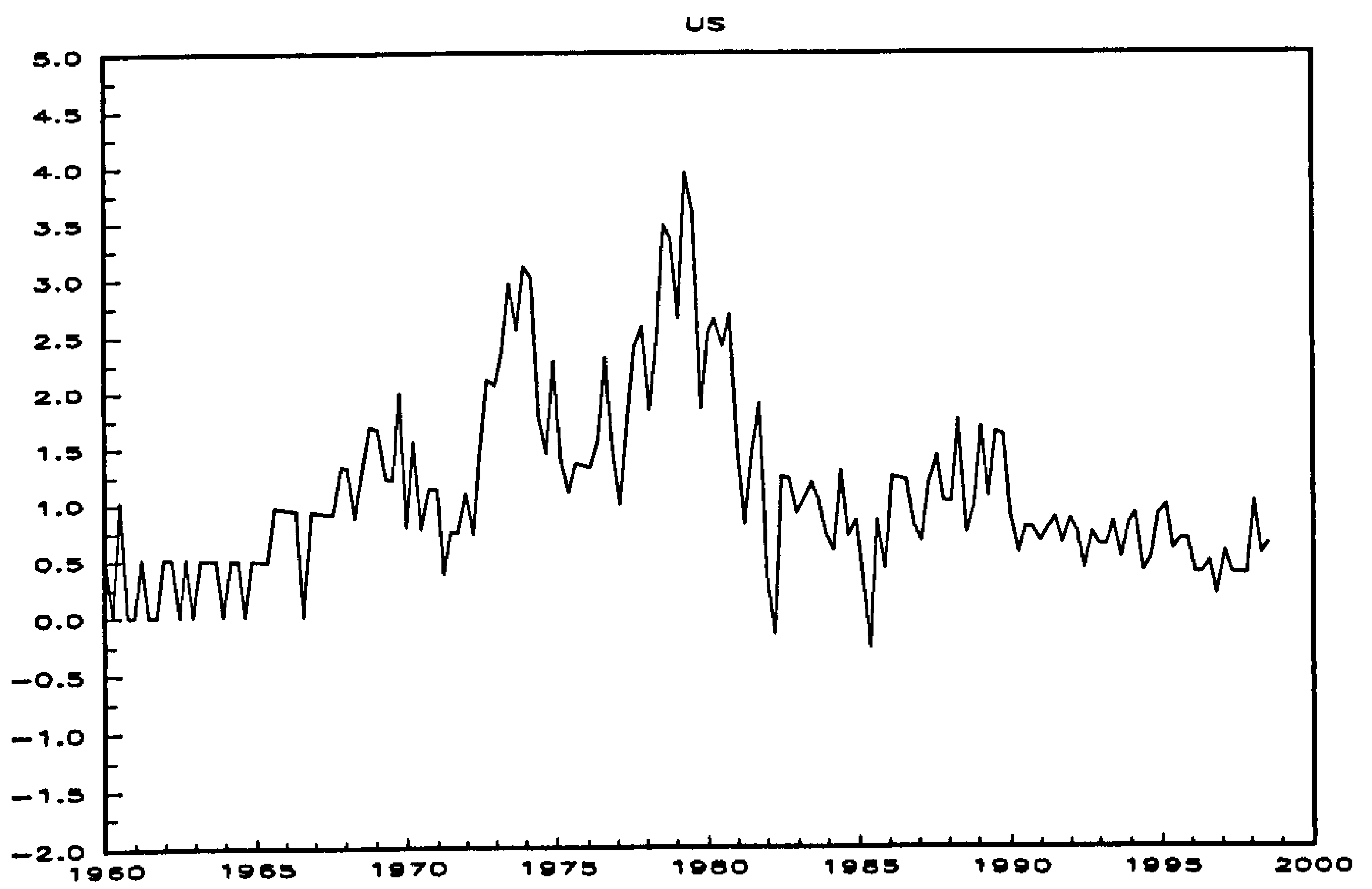


Figure 2.1. (continued)

Chapter 3

Unit Roots and Smooth Transitions Revisited: The Effect of Power- Enhancing Dickey-Fuller Type Tests

3.1 Introduction

Many macroeconomic variables usually appear to exhibit some tendency to increase over time, which could suggest that a linear deterministic trend be included in the time series model to account for this fact. Our attempt to better understand the nature of the trending mechanism(s) underlying such variables, and whether they are better modelled as stationary around a linear deterministic trend or difference stationary exhibiting some sort of stochastic trend behaviour, has led to the development of unit root tests and has provided the stimulus for research in a number of directions.

While the presence of the deterministic trend in macroeconomic series is important, its functional form plays an essential role in the unit root testing procedure and is also closely related to the power and size of the unit root tests. A deterministic linear trend is viewed by many as too restrictive especially if the time period under consideration is fairly long. On this ground, a number of studies challenge the view of the linear time trend hypothesis being the appropriate one for modelling the deterministic component of an economic time series and propose alternative flexible specifications of the trend function. These involve structural breaks, smooth transitions and Markov regime-switching type behaviour, as was mentioned in the previous chapter. Demonstrating the importance of a flexible specification for the deterministic component has justifiably been the issue of much investigation, particularly when one acknowledges the important consequences of trend misspecification, namely inconsistent estimates of the autoregressive parameter and unit root tests that can be misleading.

Recently, Leybourne *et al.* (1998) proposed unit root tests with the alternative specified as stationarity around a smooth transition in linear trend. They were recommended as complementary to those tests which consider as alternatives stationarity around a simple linear trend, to be adopted when one suspects the possibility of a structural change in the trend function. Such a suspicion is natural when time series data are considered over a long time span.

There are sometimes good reasons for feeling that the trend should be relatively smooth since it represents long run movements. One only has to consider that changes in economic aggregates are influenced by the behaviour of a large number of agents who may not act uniformly at the same time but respond to news that requires action with different time lags. Furthermore, smooth transition regression models are locally linear allowing often easy interpretation. It is based on these arguments that such models can be legitimately regarded as an attractive and intuitively plausible specification and will be adopted in this study.

Inherent in the two step testing procedure suggested by Leybourne *et al.* (1998) is the use of the conventional augmented Dickey-Fuller (ADF) t-test. Such a test has been in widespread use in practical applications, partly due to its computational simplicity and also because it displays the least sensitivity to model misspecification compared to existing competitors as shown by Schwert (1989).[1] However, one cannot overlook the problem of low power that such a test encounters against meaningful alternatives as addressed in the studies of Agiakoglou and Newbold (1992) and DeJong *et al.* (1992a) among others. To this end, parallel to research carried out involving unit root testing and flexible trend specification, a great amount of attention has been devoted in the literature

[1] In particular, when the underlying process contains an MA component they maintain size close to nominal level for all values of the MA parameter.

to devising unit root tests with higher power from which a number of modified Dickey-Fuller (DF) type t-tests have emerged.

The aim of this chapter is to explore and assess the performance of the power-enhancing DF type t-test statistics that have emerged in the literature, in the context of unit root testing when the alternative hypothesis is characterised by the more flexible specification of stationarity around a smooth transition in linear trend.

The rest of this chapter is organised as follows. Section 3.2 presents a brief overview of the recent advances in devising more powerful elaborations of DF-type unit root tests. Section 3.3 attempts to give some intuition behind the increased power performance of the two prevailing tests, namely the weighted symmetric, WS, and MAX tests, relative to the conventional DF test. Section 3.4 presents the smooth transition regression models and demonstrates how the WS and MAX test procedures can be incorporated into the testing of unit roots against smooth transition alternatives. Estimated percentiles of the limiting null distributions of the modified tests are obtained via simulation. In Section 3.5 extensive Monte Carlo simulations are performed to study the power of the proposed tests in finite samples. An empirical application based on the Nelson-Plosser data is provided in Section 3.6. Section 3.7 contains some concluding comments.

3.2 Recent Developments of Power-Enhancing Unit Root Tests: A Review

The most widely employed tests for a unit root, developed by Dickey-Fuller (1979,1981) are based on the t-statistic for $\rho = 1$ in the OLS regression

$$y_t = \gamma' z_t + \rho y_{t-1} + e_t, \qquad t = 1, 2, ..., T \tag{3.1}$$

where $e_t \tilde{} iid(0, \sigma^2)$, $z_t = 1$ or $z_t = (1, t)$ and γ' is a conformable vector of unknown parameters.

An asymptotically equivalent way of calculating the t-statistic is by removal of the trend by ordinary least squares (OLS), followed by OLS estimation of

(3.1) for $\gamma = 0$ applied to the detrended series. Such tests extend naturally to their augmented counterparts, namely the *ADF* tests, when lagged dependent variables are included in (3.1) to capture possible dynamics in the error term.

A solution to the well-known problem of low power these routinely applied unit root tests encounter, apart from increasing the sample size which is not always feasible, is the development of more powerful tests. In this direction of research a number of tests have been proposed.

Elliott *et al.* (1996) show that no uniformly most powerful test of the unit root hypothesis exists and using the theory of point optimal tests as second best, they derive the asymptotic Gaussian power envelope for unit root tests. In doing so, they propose a simple modification of the standard Dickey-Fuller t-test, that we denote the *GLS* test, such that this modified test can nearly achieve the power envelope using generalised least squares (GLS) estimation. Specifically, GLS detrending is performed under persistent local alternatives $\overline{\rho} = 1 + c/T$ for a particular value of $c = \bar{c}$, attempting in this way to improve the power of unit root tests by efficient estimation of the trend parameters.[2] The value of $\bar{c} \in (-\infty, 0]$ is chosen by default to be such that the test achieves the power envelope at 50% power, which corresponds to $\bar{c} = -7$ and $\bar{c} = -13.5$ in the demeaned and detrended cases, respectively. The test involves detrending by regressing $[y_1, y_2 - \overline{\rho}y_1, ..., y_T - \overline{\rho}y_{T-1}]'$ on $[z_1, z_2 - \overline{\rho}z_1, ..., z_T - \overline{\rho}z_{T-1}]'$ and then applying (3.1) to the estimated residuals $\tilde{y}_t$ with $\gamma = 0$, where the initial observation is assumed to be fixed. The *GLS* test statistic is then the conventional t-statistic for testing $\rho = 1$ against the alternative $\rho < 1$ in the same regression. This approach is often termed the conditional case.[3]

Elliott (1999) derives the asymptotic power envelope under the assumption that the initial observation is drawn from its unconditional distribution. This is the so-called unconditional case where under the alternative the deviation from trend of the first observation has the same variance, $\sigma^2(1-\rho^2)^{-1}$, as all remaining

[2]Lee and Phillips (1996) quantify the efficiency gains derived from generalised least squares detrending. These authors prefer the term quasi-difference detrending as full generalised least squares detrending is not used in the detrending procedure.

[3]Hwang and Schmidt (1996) proceed in the same fashion and propose essentially the same unit root tests. Although their aim is to maximise power, they do not base their approach on the theory of optimal tests but rather employ an empirically plausible value of the autoregressive root to perform GLS detrending and thus lack in asymptotic interpretation. Their results indicate that for approximately the same values of the autoregressive root that correspond to the value of $\bar{c}$ mentioned in the study of Elliot (1994), these tests are more powerful than the Dickey-Fuller tests.

observations. The impact of the initial observation on the power of unit roots tests particularly in finite samples has been addressed in a number of studies *inter alia* Evans and Savin (1981, 1984), Schmidt and Phillips (1992) and DeJong *et al.* (1992b). Under this approach the estimated residuals $\tilde{y}_t$ are generated from the regression of $[(1-\overline{\rho}^2)^{1/2}y_1, y_2 - \overline{\rho}y_{1,}, ..., y_T - \overline{\rho}y_{T-1}]'$ on $[(1-\overline{\rho}^2)^{1/2}z_1, z_2 - \overline{\rho}z_{1,}, ..., z_T - \overline{\rho}z_{T-1}]'$ and the resultant $GLSu$ test statistic is derived as the corresponding t-statistic in the same manner as described above. Elliott concentrates primarily on $\bar{c} = -10$.

An alternative deterministic trend removal procedure - recursive OLS detrending - was recently proposed by So and Shan (1999) who argued that mean adjustment using the overall sample mean in autoregressive time series causes biases of the estimator of the autoregressive coefficient and other statistics, especially when the sample size is small and the data are positively autocorrelated. The first step of this procedure involves OLS regression of y_j on z_j, $j \leq t$. The estimated residuals $\tilde{y}_t$ are then employed in regression (3.1) in the place of y_t with $\gamma = 0$. Leybourne *et al.* (2003) examine the performance of the unit root t- statistic based on this method which they refer to as REC. Such a procedure has also been applied by Taylor (2002) in the case of seasonal unit roots, where recursive sample means are used for adjusting the seasonal means.

Pantula *et al.* (1994) first employ OLS detrending to generate residuals $\tilde{y}_t$. They then recommend a test based on weighted symmetric least squares (WSLS) estimation of ρ, through the minimization of

$$Q(\rho) = \sum_{t=2}^{T} w_t(\tilde{y}_t - \rho\tilde{y}_{t-1})^2 + \sum_{t=1}^{T-1}(1 - w_{t+1})(\tilde{y}_t - \rho\tilde{y}_{t+1})^2 \qquad (3.2)$$

where $w_t = \frac{t-1}{T}$ are weights.[4]

This approach exploits the time reversibility of a stationary autoregressive process.[5] That is, the first order autoregressive process in (3.1) with $|\rho| < 1$ can

[4]Pantula *et al.* (1994) proposed weighted symmetric least squares estimation applied to the detrended series, as a computationally convenient approximation to maximization of the Gaussian likelihood. This method of estimation assigns alternative weights to intermediate observations and treats the initial and terminal observations in the same manner.

[5]See Corollary 2.6.1.3 in Fuller (1976).

be given either the forward representation

$$y_t = \gamma' z_t + \rho y_{t-1} + e_t, \qquad t = 1, 2, ..., T$$

or the backward representation

$$y_t = \gamma' z_t + \rho y_{t+1} + u_t, \qquad t = 1, 2, ..., T$$

where $\{e_t\}$ and $\{u_t\}$ are sequences of serially uncorrelated $(0, \sigma^2)$ random variables.

This symmetry led Park and Fuller (1995), who initially proposed WSLS, to consider estimators of ρ that minimize (3.2). The weighted symmetric estimator is given by

$$\widehat{\rho}_{ws} = \frac{\sum_{t=2}^{T} \tilde{y}_{t-1}\tilde{y}_t}{\sum_{t=2}^{T-1} \tilde{y}_t^2 + T^{-1} \sum_{t=1}^{T} \tilde{y}_t^2}.$$

The pivotal t-statistic based on such an estimator, which we denote WS, is then

$$WS = \widehat{\sigma}_{ws}^{-1}[\widehat{\rho}_{ws} - 1][\sum_{t=2}^{T-1} \tilde{y}_t^2 + T^{-1} \sum_{t=1}^{T} \tilde{y}_t^2]^{1/2}$$

where $\widehat{\sigma}_{ws} = (T - p)^{-1} Q(\widehat{\rho}_{ws})$ and $p = \{2, 3\}$ for the cases $z_t = 1$ and $z_t = (1, t)$ respectively. The ordinary least squares estimator is obtained by setting $w_t = 1$.

A further approach is due to Leybourne (1995) who exploits the time reversibility of stationary series in a more explicit manner. He proposes OLS estimation of (3.1) together with OLS estimation of the corresponding model for the reversed series, that is

$$v_t = \delta' z_t + \rho v_{t-1} + \eta_t, \qquad t = 1, 2, 3, ..., T \qquad (3.3)$$

where $v_t = y_{T+1-t}$, $t = 0, 1, 2, ..., T + 1$, *i.e* $v_0 = y_{T+1}$, $v_1 = y_T, ...,$ $v_T = y_1$, $v_{T+1} = y_0$. The Dickey-Fuller t-ratio from (3.1) is denoted DF_f, while that from the reverse regression is defined DF_r. The statistic proposed is then the MAX statistic with $MAX = \max(DF_f, DF_r)$. Under both the null hypothesis of a unit root and the trend-stationary alternative, Leybourne shows that DF_f and DF_r have the same asymptotic marginal distributions for *i.i.d.* errors. He derives the critical values of the MAX test via simulation.

Hansen (1995) argues that we rarely observe the time series y_t in isolation, observing at least one related time-series, x_t. He exploits information provided by these additional times series by including correlated stationary covariates in regression equation (3.1) and shows that large power gains can be achieved in this way over conventional unit root tests. He derives the asymptotic distribution of ordinary least squares (OLS) estimates of the largest autoregressive root and its t-statistic, which turns out to be a convex combination of the Dickey-Fuller distribution and the standard normal, the mixture depending on the correlation between the equation error and the regression covariates. While the power gains from inclusion of covariates is shown to be quite substantial, it is important to get the "correct" covariates, for major losses can be obtained by inclusion or exclusion of covariates. It is perhaps for this reason that such a test has not received much attention in the literature and we therefore only mention it at the outset.

The asymptotic and finite sample properties of the above modified DF-type tests have been the subject of a number of studies. Extensive simulation results on the size and power of the GLS tests and standard DF-type tests are reported in Stock (1994), for both the conditional and unconditional cases. The superior power demonstrated by the GLS test relative to the DF t-test is clearly evident in all cases. This finding is supported by the results in Phillips and Xiao (1998), who compare the effects of OLS and GLS detrending procedures on Perron and ADF tests for the conditional case.

Pantula *et al.* (1994) compare the power of a wider range of unit root test criteria, amongst which the GLS and the WS tests. They entertain two cases, that where the initial observation is drawn from a standard normal distribution and the unconditional case. Their results suggest that the modified tests are considerably more powerful than the traditional DF t-test. In particular, the WS test displays the highest power when the initial observation is drawn from the unconditional distribution while it also maintains good power in the alternative case.

Among the tests examined in Leybourne *et al.* (2003), the REC, WS and MAX tests are found to exhibit local asymptotic power close to the envelope, performing as well as and on a number of occasions better than the tests based on generalized least squares detrending. Based on their finite sample simulations, which include the performance of the WS and MAX tests in conjunction with GLS and recursive detrending, overall findings favour the WS and MAX tests with little to choose between the two. Alternative detrending procedures do not appear to provide these tests with any further power advantage.

The finding that emerges from all the above investigations is unanimous: substantial power gains are achieved by employing the modified tests over conventional DF tests when testing for a unit root in the presence of a linear trend. Evidence, however, points to preference for the WS and MAX tests which are the tests we adopt in our subsequent analysis. These two tests were described above in the case of a first order autoregressive process with i.i.d. errors. When the errors are in addition serially correlated the MAX test becomes $MAX = \max(ADF_f, ADF_r)$ where (3.1) and (3.3) are augmented to include lag changes as in Said and Dickey (1984). Details of this approach for the case of the WS test are provided in Appendix 3.A.

In view of the fact that the above two power-inducing test statistics are pervasive in the remaining of this thesis, we consider it appropriate to attempt to gain some insight into their increased power performance relative to the DF test in the standard case of testing for a unit root against stationarity around a simple linear trend, before exploring their performance within any alternative more complex setting.

3.3 Insight into the Increased Power Performance of the MAX and WS Test Statistics

Elliott *et al.* (1996) proposed the GLS test with the expectation that modified estimates of the deterministic function would improve the performance of the standard unit root t-test.[6] This implies that OLS detrending is responsible for

[6]Canjels and Watson (1997) investigate the efficient extraction of deterministic trends when the error term follows an $AR(1)$ process with largest root local to unity. They find that the

the suboptimal performance of the conventional DF tests. However, as pointed out by Leybourne *et al.* (2003) the WS and MAX statistics implicitly involve such detrending and are not improved by alternative detrending procedures. Furthermore, Burridge and Taylor (2000) who conduct finite sample Monte Carlo power investigations on the performance of the GLS and $GLSu$ tests find that efficiency of the mean estimator cannot take all the credit for the power advantage that these tests enjoy over the standard DF test. In seeking a complete explanation to the increased power performance of these tests, they demonstrate that such a favourable outcome is derived from the fact that the null distributions of the tests are shifted toward the origin, relative to that of the DF test, to a greater extent than is the distribution under the alternative.[7] We show that the same effect is present in the case of the WS and MAX test statistics, providing further insight into their superior power performance relative to the DF t-test.

Consider the times series process $\{y_t\}_{t=1}^{T}$ generated as

$$\begin{aligned} y_t &= d_t + v_t, \\ v_t &= \rho v_{t-1} + \varepsilon_t, \end{aligned} \tag{3.4}$$

where d_t is the deterministic component satisfying either (a) $d_t = 0$, (b) $d_t = \mu$ or (c) $d_t = \mu + \beta t$ and ε_t are independently and identically distributed disturbances with mean zero and standard deviation σ. Without loss of generality we focus on case (b).

A linear reparameterization of (3.4) yields the following equation

$$y_t = \mu(1-\rho) + \rho y_{t-1} + \varepsilon_t, \qquad t = 1, 2, ..., T. \tag{3.5}$$

For convenience we transform (3.5) to

$$y_t = a + \rho y_{t-1} + \varepsilon_t, \qquad t = 1, 2, ..., T \tag{3.6}$$

---

preferred estimator is the feasible Prais-Winsten GLS estimator.

[7]The limit theory for the GLS test depends on the value of c which as noted earlier is chosen by default to be the value for which local asymptotic power is 50%. While good power properties are observed against alternatives close to the null for this value of c, by using the same value in finite samples we may not always be achieving maximum power.

where our interest lies in the t-statistic for the null hypothesis that $\rho = 1$.

Figure 3.1 displays Monte Carlo generated plots of the probability density functions of the DF, MAX, and WS t-statistics under the null and alternative hypotheses, where ρ was set to 1 and 0.85 in (3.5) respectively, for a sample size of $T = 100$. The plots were generated with $\varepsilon_t \tilde{} NID(0,1)$ based on 50,000 replications. Under the alternative, we generated $T + 200$ observations and discarded the first 200 in order to obtain samples which closely resemble stationarity. It clearly becomes apparent from the plots that the null distributions of the MAX and WS test statistics are located closer to the origin relative to that of the DF statistic. A similar though smaller shift in magnitude is observed for the distributions of the tests under the alternative.

The limiting distribution of the test statistics displays the same pattern. Under the null and as $T \to \infty$, Park and Fuller (1995) derive the following result for the WS statistic

$$WS \Rightarrow \frac{0.5(W(1)^2 - 1) - HW(1) - G + 2H^2}{(G - H^2)^{1/2}} \tag{3.7}$$

where

$$H = \int_0^1 W(r)dr$$

$$G = \int_0^1 W(r)^2 dr$$

and $W(r)$ is standard Brownian motion.

For the MAX test, Leybourne *et al.* (2003) derive the limiting null distribution as

$$MAX \Rightarrow \max\{F_{n0}, R_{n0}\} \tag{3.8}$$

where

$$F_{n0} = \frac{0.5(W(1)^2 - 1) - HW(1)}{(G - H^2)^{1/2}} \tag{3.9}$$

and

$$R_{n0} = \frac{-0.5(W(1)^2 + 1) + HW(1)}{(G - H^2)^{1/2}}.$$

It follows from (3.8) that the null asymptotic distribution of the DF statistic under (3.6), for $\rho = 1$, is given by (3.9).

Assumption 3.1. (i) $y_1$is distributed with mean zero and variance $\sigma^2(1-\rho^2)^{-1}$, (ii) ε_t is i.i.d.$(0, \sigma^2)$ and (iii) y_1 is uncorrelated with ε_t, $t \geq 2$.

We next evaluate the behaviour of the tests under the local alternative. The limiting distribution of the MAX test in this case when y_t is generated through

$$\begin{aligned} y_t &= \rho y_{t-1} + \varepsilon_t \\ \rho &= 1 + \frac{c}{T} \end{aligned}$$

where $c \in (-\infty, 0)$ and Assumption 3.1 holds, while furthermore the fitted model is that in (3.6) is given in Leybourne *et al.* (2003) by

$$MAX \Rightarrow \max(F_0, R_0) \tag{3.10}$$

where

$$F_0 = \frac{0.5(J_c(1)^2 - 1) - H_c J_c(1)}{(G_c - H_c^2)^{1/2}}$$

$$R_0 = \frac{-0.5(J_c(1)^2 + 1) + H_c J_c(1)}{(G_c - H_c^2)^{1/2}}$$

and

$$J_c(r) = W_c(r) + (e^{rc} - 1)Z_c$$

$$H_c = \int_0^1 J_c(r)dr$$

$$G_c = \int_0^1 J_c(r)^2 dr.$$

Here, $W_c(r)$ is an Ornstein-Uhlenbeck process defined as

$$W_c(r) = c\int_0^r e^{c(r-s)}W(s)ds + W(r),$$

$W(r)$ is a standard Brownian motion process defined as the limit of $\sigma^{-1}T^{-1/2}\sum_{t=1}^{rT}\varepsilon_t$ and Z_c is a random variable with mean zero and variance $(-2c)^{-1}$.

The limiting distribution of the DF test under the local alternative is then

$$DF \Rightarrow F_0, \tag{3.11}$$

while the limiting distribution for the WS test can be directly inferred from the above and by using results in Elliott (1999) as

$$WS \Rightarrow \frac{0.5(J_c(1)^2-1) - H_cJ_c(1) - G_c + 2H_c^2}{(G_c - H_c^2)^{1/2}}. \tag{3.12}$$

Figure 3.2 illustrates the asymptotic densities of the test statistics under both the null and local alternative. These were generated by simulating 50,000 replications of the limiting functionals in (3.7), (3.8), (3.9) and (3.10), (3.11), (3.12) respectively, approximated by their sample moment analogues, using series of 5,000 Gaussian white noise innovations.[8] On inspecting Figure 3.2, the shift in the location of the distributions of the MAX and WS test statistics closer to the origin relative to that of the DF statistic is immediately discernable under the null, which is the case depicted in the upper plot. A similar though smaller shift is demonstrated under the alternative, as can be seen in the bottom plot. Thus the limiting null distribution of the test statistics exhibits a similar pattern to the simulated finite-sample distributions. These plots corroborate the findings of Leybourne *et al.* (2003), whereby the DF test is outperformed in terms of asymptotic local power by the MAX and WS tests.

In what follows we will see that by including additional trend terms in the deterministic function, as with the smooth transition models, the critical values shift further away from the origin requiring larger values of the test statistics to achieve a rejection of the null.

[8] For example, as $T \to \infty$, $T^{-1/2}\sum_1^n e_t \Rightarrow W(1)$, $T^{-3/2}\sum_1^n\sum_{j=1}^t e_j \Rightarrow \int_0^1 W(r)dr$, $T^{-1}\sum_1^n(\sum_{j=1}^{t-1} e_j)e_t \Rightarrow \frac{1}{2}(W(1)^2-1)$, and so forth.

3.4 Testing for a Smooth Transition in Linear Trend Using More Powerful Dickey-Fuller Type Tests

The idea of smooth transition dates back to Bacon and Watts (1971) who illustrated how a locally linear equation changes from the one extreme linear parameterisation to the other as a function of a continuous transition variable. It was then Granger and Teräsvirta (1993) and Teräsvirta (1994,1998) who promoted the family of univariate models called smooth transition autoregressive (STAR) models. During the last decade, such models have been employed widely mainly to investigate the non-linearity in the conditional mean of various macroeconomic aggregates. Adopting these models, Teräsvirta and Anderson (1992) describe industrial production in a number of countries, Granger and Teräsvirta (1993) analyse a non-linear relationship between US GNP growth and leading indicators, Skalin and Teräsvirta (1999) examine Swedish business cycles, while Öcal and Osborn (2000) investigate non-linearities in UK consumption and industrial production. An excellent overview of extensions of the STAR models as well as issues relating to their evaluation by means of out-of-sample forecasting and impulse response analysis can be found in van Dijk *et al.* (2002).

Following Leybourne *et al.* (1998) we consider three parameterizations of stationarity around a smooth transition in linear trend that give rise to the following regression models

$(A)\qquad y_t = a_1 + a_2 F_t(\gamma, \tau) + \varepsilon_t$

$(B)\qquad y_t = a_1 + b_1 t + a_2 F_t(\gamma, \tau) + \varepsilon_t$

$(C)\qquad y_t = a_1 + b_1 t + a_2 F_t(\gamma, \tau) + b_2 t F_t(\gamma, \tau) + \varepsilon_t$

where ε_t is a zero mean stationary process, not necessarily white noise.

The function operating on the parameters of the above models is a transition function bounded by convention by zero and one. The most common adopted transition function in the literature is the logistic function

$$F(t;\gamma,\tau) = \{1 + \exp(-\gamma(t - \tau T))\}^{-1}. \qquad (3.13)$$

Such a function is continuous, monotonically increasing in t, where $\gamma > 0$ constitutes an identifying restriction. The slope parameter γ indicates how rapid the transition from zero to unity is as a function of t and the location parameter τ determines where the transition occurs. If γ is small then $F_t(\gamma, \tau)$ traverses the interval (0,1) at a slow pace. As the value of γ increases $F_t(\gamma, \tau)$ traverses the interval at a more rapid pace, with the transition from zero to one becoming virtually instantaneous as γ approaches infinity. This is the case of an abrupt structural break with the shift point located at $t = \tau T$. Finally, when $\gamma = 0$, $F_t(\gamma, \tau) = 0.5$ for all t. The flexibility of a smooth transition model is thus evident, allowing for no transition, instantaneous transition and all smooth intermediate cases. Function (3.13) therefore assumes a continuum of states between two extremes, which correspond to $F_{-\infty}(\gamma, \tau) = 0$ and $F_{+\infty}(\gamma, \tau) = 1$. For model A this implies stationarity of y_t around a mean which changes from initial value a_1 to final value $a_1 + a_2$. Similarly for model B, the difference being that a fixed slope term is also included, while for model C it implies stationarity of y_t around an intercept and slope that are changing simultaneously from initial values a_1 and b_1 to $a_1 + a_2$ and $b_1 + b_2$ respectively.

Different assumptions regarding the transition function $F(.)$ imply different forms of smooth transition models and so different degrees of smoothness and different types of non-linearity, possibly involving nonmonotonic and non-symmetric transitions. Allowance can also be made for different type of transition functions to govern the constant and slope parameters. For practical purposes and for ease of exposition we confine ourselves as in Leybourne *et al.* (1998) to the above setting, since it provides a reasonably adequate understanding of the underlying mechanisms involved in such models.

The unit root null hypotheses considered are stated below, from which similar test statistics readily follow.[9]

$H_o : y_t = v_t, \qquad v_t = \rho v_{t-1} + e_t, \qquad t = 1, 2, ..., T.$

$H_1 :$ *Model A, B or C*

$H_o : y_t = v_t, \qquad v_t = \delta + \rho v_{t-1} + e_t, \qquad t = 1, 2, ..., T.$

$H_1 :$ *Model B or C*

for $\rho = 1$, where v_0 is either a fixed constant or a random variable and e_t is a zero mean stationary process.

[9] Similarity implies that the t-statistic associated with the parameter $\widehat{\rho}$ is not affected by the value, under the null, of any nuisance parameter and the critical values are the same as the ones that would apply if the nuisance parameter(s) were set to zero.

Testing for a unit root in the smooth transition framework consists of a two-step procedure. The first step involves estimating the deterministic component by use of a nonlinear least squares (NLS) algorithm and computing the corresponding residuals. For the three models this amounts to

(A) $\tilde{\varepsilon}_t = y_t - \tilde{a}_1 - \tilde{a}_2 F_t(\tilde{\gamma}, \tilde{\tau})$

(B) $\tilde{\varepsilon}_t = y_t - \tilde{a}_1 - \tilde{b}_1 t - \tilde{a}_2 F_t(\tilde{\gamma}, \tilde{\tau})$

(C) $\tilde{\varepsilon}_t = y_t - \tilde{a}_1 - \tilde{b}_1 t - \tilde{a}_2 F_t(\tilde{\gamma}, \tilde{\tau}) - \tilde{b}_2 t F_t(\tilde{\gamma}, \tilde{\tau})$

Once the parameters of the deterministic component of the desired model have been estimated, the unit root hypothesis is tested using the t-statistics, MAX and WS, associated with ρ in the following augmented regression of the estimated residuals

$$\tilde{\varepsilon}_t = \rho \tilde{\varepsilon}_{t-1} + \sum_{j=1}^{k} \zeta_j \Delta \tilde{\varepsilon}_{t-j} + v_t.$$

Augmentation of the regression by the use of lagged values of the first differenced dependent variable accounts for the possible autocorrelation exhibited in the error structure e_t.

The linearity in the parameters a_1, b_1, a_2 and b_2 reduces the nonlinear least squares (NLS) estimation problem to minimising the sum of squared residuals with respect to just the two parameters $\tilde{\gamma}$ and $\tilde{\tau}$. To see this in the case of the most general model, (C), consider the vector of regressors $z_t = z_t(\gamma, \tau) = \{1, t, F_t(\gamma, \tau), tF_t(\gamma, \tau)\}'$ and the vector of corresponding parameters $\xi = \{a_1, b_1, a_2, b_2\}'$. The residual sum of squares is given by

$$RSS = \sum_{t=1}^{T} (y_t - \tilde{\xi}' z_t)^2 \tag{3.14}$$

where $\tilde{\xi} = \{a_1, b_1, a_2, b_2\}' = \left(\sum_{t=1}^{T} z_t z_t'\right)^{-1} \sum_{t=1}^{T} z_t y_t$ as in the standard multiple regression model. Thus, the RSS is indeed a function of the two unknown parameters γ and τ. As Leybourne, Newbold and Vougas (1998) point out, the linearity property in the intercept and trend terms ensures that estimated residuals, $\tilde{\varepsilon}_t$, from all the models are invariant to the choice of starting value v_0 and models B and C are invariant to both the starting value and the drift term δ.

Given the complex nature of non-linearity involved in this context, Monte

Carlo simulation methods are employed in order to approximate the null distribution of the test statistics. The data generating process is specified as

$$y_t = v_t, \qquad v_t = v_{t-1} + e_t, \qquad v_{-201} = 0, \quad t = -200, ..., T$$

with e_t being independently distributed standard normal innovations. We set $\delta = 0$ and $\sigma^2 = 1$ without loss of generality due to the asymptotic invariance of the test statistics to these values under the null hypothesis. Sample series of $T + 200$ were generated for y_t discarding the first 200 observations to remove the effect of the initial conditions. By putting the initial condition that determines v_0 into the increasingly distant past as $T \to \infty$ is like making assumptions about v_0 which give it properties that are analogous to those of v_t itself.

In each experiment, for given values of $T \in \{25, 50, 100, 200, 500\}$ the 1%, 5% and 10% critical values of the null distributions of the MAX and WS statistics are computed as corresponding percentiles of the empirical finite sample distribution based on 10,000 replications. Critical values for the case of the conventional DF t-test are also reported for reasons of comparison and completeness.

In terms of increased efficiency in searching for the global minimum, as many iterative minimization procedures may converge to a local minimum while consistency results apply to the global minimum, we perform a grid search on the starting values of the iterative procedure and choose those values that minimise the objective function among the converged values. Throughout the grid search γ ranges from 0.05 to 5 in steps of 0.01 and τ from 0.1 to 0.9 in steps of 0.01. It should be noted that increasing the range of γ beyond 5 leads to no information gain since at this stage the transition is already instantaneous. For the NLS estimation we employed the OPTMUM subroutine library of GAUSS 3.1.

Tables 3.2(a)-(c) contain the approximated lower tail percent points of the limiting null distributions under the different forms of smooth transition, namely of a smooth transition in constant only excluding any trend, in constant only including a trend term and in both constant and trend respectively. The subscripts a, b and c are used respectively to characterise the statistics under such forms of transition. We observe that in all cases the critical values do not change significantly when increasing the sample size past the value of $T = 100$, after which they appear to be converging rather steadily to limiting values. Another feature perhaps worth noting throughout these tables is that the critical values of the MAX and WS tests under smooth transition are shifted further towards the origin relative to the DF test, as in the standard case when a linear trend term is included under the alternative.

3.5 Finite Sample Power Simulations

In this section finite sample power behaviour of the modified test statistics is analysed. The purpose is to determine how their power is affected by detrending under smooth transitions.

As a starting point we consider a stationary AR(1) data generating process

$$y_t = v_t, \qquad v_t = \rho v_{t-1} + e_t, \qquad v_{-201} = 0, \qquad t = -200, ..., T$$

for $\rho < 1$ and $e_t \sim iid\ N(0,1)$ random deviates. The same set of random deviates was used in all simulations with common sample size. In all cases, the number of replications was 2000 with power estimates reported and compared across tests using the relevant critical values in Tables 3.2(a)-3.2(c).

Empirical powers of the smooth transition DF, MAX and WS tests for sample sizes of $T = \{50, 100, 200\}$ and for associated values of $\rho = \{(0.60, 0.70, 0.80), (0.70, 0.80, 0.90), (0.85, 0.90, 0.95)\}$ are reported in Table 3.3. All tests appear to be consistent with power increasing rapidly when moving away from the unit root null. An interesting feature of the results is that power decreases monotonically the more complex the trend function employed in estimation and therefore the lowest power entries are observed in the case of a smooth transition in both the constant and trend term. Further results in this setting reveal similar power performance for the modified MAX and WS tests, however only fairly moderate power gains appear to be achieved with the use of these tests relative to the DF t-test. These are more pronounced in the smooth transition in constant only case.[10]

[10]In one sense, comparison of tests based on a first order autoregression with Gaussian innovations seems somewhat limited. However, it has the considerable advantage of guaranteeing that precisely correct critical values, those of Table 3.1 can be used. Thus, any differences found

We next consider simulations in which the DGP is a stationary AR(1) process around a smooth transition in mean

$$y_t = 1.0 + 2.0F_t(\gamma, \tau) + v_t, \qquad v_t = \rho v_{t-1} + e_t$$

where $v_{-200} = 0$ and $e_t \tilde{} NID(0,1)$. We generated samples of size $T = \{50, 100, 200\}$ with $\rho = \{(0.50, 0.60, 0.70), (0.60, 0.70, 0.80), (0.70, 0.80, 0.90)\}$ respectively. The values for $a_1 = 1.0$ and $a_2 = 2.0$ are chosen rather arbitrarily, while a variety of values are examined for both the speed of transition parameter, γ, and the location parameter, τ. Empirical powers of the 5% lower tail tests for this case are summarized in Table 3.4(a). To asses the impact of the magnitude of transition on the finite sample power properties of the tests we also present results for alternative values of $a_2 = \{5.0, 10.0\}$, which are given in Tables 3.4(b) and 3.4(c).

An equivalent picture to the above emerges. The MAX and WS tests under all type of smooth transitions exhibit rather similar finite sample power with neither one substantially more powerful than the other. For fixed ρ, power increases with T, reflecting the consistency of the tests. Similarly for a fixed sample size T, power increases the further we move away from the unit root null. The power advantage of the modified tests over that of the common DF test applied to the detrended series is quite evident, maintaining on average the value of approximately seven percent. Looking more closely at the results, no clear cut pattern emerges when the magnitude of the transition is small as observed in Table 3.4(a). In particular, as the speed of transition γ increases and the location of the transition shifts from the beginning of the series to the midpoint governed by the parameter τ, empirical powers appear to alternate from higher to lower values and vice versa though exhibiting only very small fluctuations.

As the magnitude of the transition increases as illustrated in Tables 3.4(b) and 3.4(c), power entries increase in value while average power gains of the MAX_a and WS_a tests over DF_a remain at the same level. In general a rather clearer pattern emerges. When the transition occurs at the midpoint of the sample the power of the tests is higher the slower the transition, while it tends to decrease

in rejection probabilities must correspond to power differences among the tests, rather than for example size distortions brought about by non-normality or the addition of further lags.

as the transition becomes virtually instantaneous for values of $\gamma = 1$ and $\gamma = 5$. When the transition is of a slow nature, that is $\gamma = 0.10$, power estimates of the tests reach their highest level and drop gradually thereafter reaching their lowest value when the transition becomes instantaneous. When the transition occurs fairly early in the sample, that is $\tau = 0.2$, power entries are slightly more erratic, while a similar overall pattern can be discerned. On the whole, results indicate that power estimates are higher when the transition occurs midway along the sample rather than early on.

Proceeding in the same fashion, we consider simulations based on the DGP of a stationary AR(1) process around a smooth transition in mean that involves a trend term

$$y_t = 1.0 + 2.0F_t(\gamma, \tau) + 1.0t + v_t, \qquad v_t = \rho v_{t-1} + e.$$

Table 3.5(a) evaluates the power performance of the 5% lower tail tests under the same specifications employed in the previous experiment, while Tables 3.5(b) and 3.5(c) evaluate the power performance for alternative transition magnitudes. Such tables permit comparisons across the tests and doing so indicates that in terms of power the MAX_b and WS_b tests behave once again very similar. What is worth noting however, is that in the presence of a trend term the power differences among these tests and DF_b are somewhat reduced compared to the smooth transition in mean only case. Apart from this, no other important differences are observed. The same results apply as above, which become more apparent the larger the magnitude of the transition.

We also conducted further simulation experiments when the true model is characterised by a smooth transition in intercept and trend. We do not repeat the results for this case as they are qualitatively similar although, quantitatively, the power differences among the DF_c and MAX_c and WS_c tests are further reduced.

Overall, comparing the results across the figures highlights the finding that the less elaborate is the deterministic component of the first stage regressions, the more substantial are the gains in power that can be achieved in comparison with the DF smooth transition tests. Power differences among the tests are therefore more pronounced in the case of a smooth transition in constant only compared to when a linear trend is incorporated in the model. This result is similar to the marked reduction in power associated with the estimation of additional trend terms in the standard DF tests. Thus, even though the critical values of the MAX and WS tests under smooth transition detrending are shifted closer to the origin relative to the corresponding DF test ones, the power differences between them are only modest.[11] This is possibly due to the more complex nature of the trend function. Furthermore, in the presence of structural change of a more abrupt nature the tests are slightly more powerful when breaks occur midway through the series rather than early on.

3.6 Empirical Application

In this section we apply the preceding test statistics to the data analysed by Nelson-Plosser (1982). These are US annual data ranging from 62 to 111 observations and ending in 1970. Evidence from a number of studies analysing such data are to an extent unanimous in there findings that most macroeconomic time series are characterised by the presence of a unit root. Exceptions hold for the cases of unemployment and industrial production whereby the unit root hypothesis is generally rejected. Following convention, we take natural logarithms of each time series except for the bond yield as in Nelson and Plosser (1982). Plots of the data can be found in Figure 3.3.

Since many macroeconomic variables exhibit some tendency to increase over time we examine the two versions of the tests that include a time trend namely, MAX_b, WS_b and MAX_c, WS_c, the underlying logic being that if the hypothesis of stationarity around a smooth transition in linear trend is correct, we should be able to reject the null hypothesis using such tests. For purposes of completeness we also report the reverse counterparts of the MAX test denoted ADF_b^r, ADF_c^r respectively, along with the ADF_b^f, ADF_c^f tests for comparison, as well as the standard ADF t-statistic for the case of a model that includes a constant and linear trend term.

[11]The same principles can be applied to tests against asymmetric smooth transition alternatives, and we conjecture that since these involve a further parameter in the trend function, power gains are likely to be even more modest.

The autoregressive truncation lag k for all tests was selected using the sequential (10%-level) t-tests for the significance of the coefficient on the longest lag, taking into account the available number of observations for each series in selecting $k$max. This resulted in fairly similar values of k chosen for the series across tests. Theoretical support for the adopted method as well as information based model selection rules such as the Akaike information criterion (AIC) and the Schwarz Bayesian information criterion (BIC) was provided by Hall (1994) for the pure AR case and by Ng and Perron (1995) for the $ARMA$ case. Overall, the general-to-specific lag selection procedure has been found to control well for size, though it does tend to choose higher values for k compared to AIC and BIC, resulting in a loss of power. However, such an effect is only moderate as illustrated in DeJong (1992a).[12] Nelson and Plosser (1982) argue in favour of the use of higher order models on the grounds that leaving out relevant terms might bias the results, while inclusion of irrelevant ones will only reduce efficiency.

Table 3.6 summarizes the results for each series. The tests were carried out using the 5% critical values in Tables 3.2(b)-(c). According to the standard ADF test as reported in column three, all series contain a unit root apart from the unemployment series for which the unit root hypothesis is rejected at the 5% level.[13] This result corroborates the findings of alternative empirical studies that have employed the Nelson-Plosser dataset.

Of the tests based on the smooth transition alternatives, we focus initially on the results pertaining to the MAX_b and WS_b statistics that correspond to a smooth transition in constant only with the inclusion of a fixed trend term, as shown in columns four to seven. In this case, the unit root null hypothesis cannot be rejected for 11 out of the 14 reported series. For industrial production there is evidence of a 5% rejection of the unit root null using the WS_b statistic. A rejection at the 10% level is reported for the ADF_b^f and MAX_b tests respectively, although the latter is very close to rejecting at the 5% level. For the real wages series, the WS_b statistic rejects the null at the 10% level. While no rejection is

[12]Stock (1994), in investigating the finite sample size and power properties of the GLS and $GLSu$ unit root tests, reports results based on the Schwartz Bayesian information criterion. However, he acknowledges the use of the sequential testing procedure and proposes either in his study. Pantula *et al.* (1994) choose the lag order based on the Akaike Information criterion (AIC), modified as AIC+2 to offset the size distortions that such a procedure is known to give rise to when choosing a small value for the lag parameter.

[13]Critical values for the standard Dickey-Fuller test were used from Table B.6 in Hamilton (1994).

displayed by the ADF_b^f test or the MAX_b test, the latter is once again very close to rejecting the unit root hypothesis at the 10% level. For the common stock price series all three tests namely, the ADF_b^f, the MAX_b and the WS_b tests reject the unit root hypothesis at the 5% level.

Turning to the second set of statistics, the case of a smooth transition in both the intercept and trend term, a rejection of the unit root null at the 10% level is only observed for the WS_c test for real per capita GNP. However, the MAX_c test is very near rejecting the null at the 10% level. The exact same outcome is observed for the industrial production series. For the common stock price series a similar picture to the earlier results emerges, that is, the MAX_b and WS_b reject at the 0.05 level, only now the ADF_b^f displays a rejection of the unit root hypothesis at the 10% level.

Typically it appears that the MAX test marginally misses out in attaining rejections at the same level of significance as the WS test. Such a result could be reflected in the slightly better power performance demonstrated by the latter at times as can be seen from the previous section. In any case, sharper results are obtained by using the more powerful $MAX_{b,c}$ and $WS_{b,c}$ test statistics relative to the $ADF_{b,c}$ tests that include either stronger rejections of the unit root null for certain series or additional rejections in favour of stationarity around smooth transition in linear trend.

For the industrial production series and that of common stock prices, rejections are observed for both the (b) and (c) type of smooth transition alternatives, the latter being the more elaborate of the two and encompassing the former. In view of this result, we conduct a likelihood ratio test for the significance of the smooth transition in the trend term. Consider model (C):

$$y_t = a_1 + b_1 t + a_2 F_t(\gamma, \tau) + b_2 t F_t(\gamma, \tau) + \varepsilon_t$$

and model (B):

$$y_t = a_1 + b_1 t + a_2 F_t(\gamma, \tau) + \varepsilon_t.$$

Testing for the significance of a smooth transition in the trend term translates into testing the hypothesis H_0: $b_2 = 0$ versus H_1: $b_2 \neq 0$. The likelihood ratio statistic is given by

$$LR = n \ln(RSS_b / RSS_c)$$

where RSS_b and RSS_c are the residual sum of squares as described in Section 3.4 from the first step of the unit root testing procedure that corresponds to models (B) and (C) respectively.[14] The LR test statistic asymptotically has a χ^2-distribution with one degree of freedom.[15] For the industrial series $LR = 0.002$, which implies that we cannot reject the null at the 5% level of significance. This result suggests that model (B) should be used in the initial detrending of the series, a finding which is not surprising given the sharper results displayed by the WS_b, MAX_b and ADF_b^f statistics for this series. For the common stock price series $LR = 2.886$. At the 5% level of significance we cannot reject the null, implying that the overall outcome of a unit root versus stationarity around a smooth transition in linear trend should be determined by the WS_b, MAX_b and ADF_b^f statistics for this series. However, at the 10% level there is evidence in favour of a smooth transition in the trend term which points to the results of the WS_c, MAX_c and ADF_c^f statistics in determining the overall outcome.

Finally it should be pointed out that while the unit root hypothesis is rejected for the unemployment series using the standard ADF t-statistic, in support of what is generally found in the literature, there is no evidence of a similar rejection for any of the alternative tests. This result as pointed out by Leybourne *et al.* (1998) serves to illustrate the usefulness of the smooth transition unit root tests as complementary to existing standard tests and not as direct substitutes.

In summary, while the majority of the macroeconomic series considered are found to be characterised by the presence of a unit root, employing the more powerful tests uncovers further evidence in favour of the alternative of stationarity around smooth transition in linear trend. Such an alternative appears to be an attractive characterisation for series such as industrial production, common stock prices, real per capita GNP and real wages, although the evidence is somewhat weaker for the latter two series. These results support the findings of Chu and White (1992) who reject the null of trend-stationarity against the alternative of

[14] For model (B), $z_t = \{1, t, F_t(\gamma, \tau)\}'$ and the vector of corresponding parameters is $\xi = \{a_1, b_1, a_2\}'$. The residual sum of squares is then $RSS = \sum_{t=1}^{T} (y_t - \tilde{\xi}' z_t)^2$.

[15] The upper 5% and 10% critical values for the $\chi^2(1)$ distribution are 3.841 and 2.706 respectively.

a trend-shift at the 10% level for real per capita GNP and real wages, though they do not find evidence in favour of the trend-shift alternative in the case of common stock prices. Similarly, results for the industrial production series lends support to studies such as those of Teräsvirta and Anderson (1992) and Öcal and Osborn (2000) who describe industrial production based on smooth transition type models.

3.7 Concluding Remarks

In the last decade, the literature on unit root testing has witnessed the emergence of more powerful test statistics as a solution to the power deficiency of conventional Dickey-Fuller tests. We showed in this chapter that the substantial power advantage associated with two such tests, namely the MAX and the WS tests, when testing for a unit root against the simple trend-stationary alternative, lies in the shift of their null distributions toward the origin, relative to that of the DF test, to a greater extent than is the distribution under the alternative. We subsequently investigated the performance of the elaboration principles associated with these modified Dickey-Fuller type t-tests in the context of testing for a unit root against the more flexible and intuitively plausible smooth transition alternatives.

Extensive finite sample Monte Carlo results illustrated that the substantial power gains offered by the power-enhancing tests over the conventional Dickey-Fuller type in the standard case of stationarity around a simple linear trend function, are modest when adopted in the smooth transition setting. Higher gains in power appeared to be offset by the complexity of the smooth transition type of structural change characterising the trend function under the alternative. It is therefore not surprising that the power gains were more prominent in the simpler case of a smooth transition in constant only. Notwithstanding, the use of the modified tests is worth the while as it signifies extra power at the minimum cost of a little more computational complexity. Moreover, the power performance of the MAX and WS tests under the smooth transition alternatives was very similar, with neither being substantially more powerful in all cases. Greater power was generally achieved when the transition was of a slower nature and occurred earlier rather than later in the sample.

An application of the more powerful smooth transition tests to US macroeconomic series showed stronger evidence against the unit root null for series such as industrial production and common stock prices. Furthermore, stationarity around smooth transition in linear trend was found an attractive characterisation for two additional series, real per capita GNP and real wages. However, because rejecting the null hypothesis does not necessarily imply smooth transition in the trend function of the series, care is warranted in drawing conclusion about the true data generating process.

References

Agiakoglou, C., Newbold, P., 1992. Empirical evidence on Dickey-Fuller type tests. *Journal of Time Series Analysis*, 13, 471-483.

Akaike, H., 1974. A new look at the statistical model identification. IEEE *Transactionson Automatic Control*, AC-19, 716-723.

Bacon, D.W., Watts, D.G., 1971. Estimating the transition between two intersecting straight lines. *Biometrika*, 58, 525-534.

Burridge, P., Taylor, A.M.R., 2000. On the power of GLS-type unit root tests. *Oxford Bulletin of Economics and Statistics*, 62, 633-645.

Canjels, E., Watson, M.W., 1997. Estimating deterministic trends in the presence of serially correlated errors. *The Review of Economics and Statistics*, 184-200.

Chu, C-S. J., White, H., 1992. A Direct Test for Changing Trend. *Journal of Business & Economic Statistics*, 10, 289-300.

DeJong, D.N., Nankervis, J.C., Savin, N.E., Whiteman, C.H., 1992a. The power problems of unit root tests in time series with autoregressive errors. *Journal of Econometrics*, 53, 323-343.

DeJong, D.N., Nankervis, J.C., Savin, N.E., Whiteman, C.H., 1992b. Integration versus trend stationarity in time series. *Econometrica*, 60, 423-433.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.

Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49, 1057-1072.

Elliott, G., 1999. Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review*, 40, 767-783.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. *Econometrica*, 64, 813-836.

Evans, G.B.A., Savin, N.E., 1981. Testing for unit roots: 1, *Econometrica*, 49, 753-779.

Evans, G.B.A., Savin, N.E., 1984. Testing for unit roots: 2, *Econometrica*, 52, 1241-1270.

Fuller, W.A., 1976. Introduction to statistical time series. New York: John Wiley.

Granger C.W.J., Teräsvirta, T., 1993. Modelling non-linear economic relationships, Oxford: Oxford University Press.

Hall, A., 1994. Testing for a unit root in time series with pretest data based model selection. *Journal of Business & Economic Statistics*, 12, 461-470.

Hamilton, J.D., 1994. Time series analysis. *Princeton University Press.*

Hansen, B.E., 1995. Rethinking the univariate approach to unit root testing: Using covariates to increase power. *Econometric Theory*, 11, 1148-1172.

Hwang, J., Schmidt, P., 1996. Alternative methods of detrending and the power of unit root tests. *Journal of Econometrics*, 71, 227-248.

Leybourne, S.J., 1995. Testing for unit roots using forward and reverse Dickey-Fuller Regressions. *Oxford Bulletin of Economics and Statistics,* 57, 559-571.

Leybourne, S.J., Kim, T-H., Newbold, P., 2003. Examination of some more powerful modifications of the Dickey-Fuller test. Forthcoming in the *Journal of Time Series Analysis.*

Leybourne, S., Newbold, P., Vougas, D., 1998. Unit roots and smooth transitions. *Journal of Time Series Analysis*, 19, 83-97.

Nelson, C.R., Plosser C.I., 1982. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10, 139–162.

Ng, S., Perron, P., 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.

Öcal, N., Osborn, D., 2000. Business cycle nonlinearities in UK consumption and production. *Journal of Applied Econometrics,* 15, 27-43.

Pantula, S.G., Gonzalez-Farias, G., Fuller, W.A., 1994. A comparison of unit-root test criteria. Journal of Business & Economic Statistics, 12, 449-459.

Park, H.J., Fuller, W.A., 1995. Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis*, 16, 415-429.

Phillips, P.C.B., and Lee, C.C., 1996. Efficiency gains from quasi-differencing under nonstationarity. In: Robinson, P.M., Rosenblatt, M., (Eds.). *Essays in Memory of E.J. Hannan.* New York: Springer-Verlag.

Phillips, C.B., Xiao, Z., 1998. A primer on unit root testing. *Journal of*

Economic Surveys, 12, 423-470.

Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika*, 71, 599-608.

Schmidt, P., Phillips, P.C.B., 1992. LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, 54, 257-287.

Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.

Schwert, G.W., 1989. Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economic Statistics*, 7, 147-159.

Skalin J., Teräsvirta, T., 1999. Another look at Swedish business cycles. *Journal of Applied Econometrics*, 14, 359-378.

So, B.S., Shin, D.W., 1999. Recursive mean adjustment in time-series inferences. *Statistics and Probability Letters*, 43, 65-73.

Stock, J.H., 1994. Unit roots structural breaks and trends. In: Engle, R.F., McFadden, D.L. (Eds.), *Handbook of Econometrics*, Vol.4. New York: North Holland, 2740-2841.

Taylor, A.M.R., 2002. Regression-based unit root tests with recursive mean adjustment for seasonal and non-seasonal time series. *Journal of Business & Economic Statistics*, 20, 269-281.

Teräsvirta, T., 1994. Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of American Statistical Association*, 89, 208-218.

Teräsvirta, T., 1998. Modeling economic relationships with smooth transition regressions. In: Ullah, A., Giles, D.E.A. (Eds.), *Handbook of applied economic statistics*, Statistics: Textbooks and Monographs, vol. 155. New York; Basel and Hong Kong: Dekker, 507-552.

Teräsvirta, T., Anderson, H.M., 1992. Characterizing Nonlinearities in Business Cycles using Smooth Transition Autoregressive Models. *Journal of Applied Econometrics*, 7, Supplement, S119-S136.

van Dijk, D., Teräsvirta, T., Franses, P.H., 2002. Smooth Transition Autoregressive Models - A Survey of Recent Developments. *Econometric Reviews*, 21, 1-47.

Appendix 3.A Derivation of the WS Test Statistic in the Case of Serially Correlated Errors

Assume that the errors in (3.1) are serially correlated and that the underlying process is generated by

$$y_t = a + \sum_{j=1}^{k} \psi_j y_{t-j} + \varepsilon_t, \qquad t = 1, 2, ..., T \tag{3.15}$$

which can be written equivalently as the $ADF(k-1)$ regression

$$y_t = a + \rho y_{t-1} + \sum_{j=1}^{k-1} \zeta_j \Delta y_{t-j} + \varepsilon_t$$

with $\rho = \psi_1 + \psi_2 + ... + \psi_k$, $\psi_0 = 1$ and $\zeta_j = -[\psi_{j+1} + \psi_{j+2} + ... + \psi_{j+k}]$ for $j = 1, 2, ..., k-1$.

If all roots of $\Psi(z) = 1 - \sum_{j=1}^{k} \psi_j z^j$ lie outside the unit circle equation then the stationary autoregressive series (3.15) can also be given the forward representation

$$y_t = a + \rho y_{t+1} - \sum_{j=1}^{k-1} \zeta_j \Delta y_{t+j+1} + \epsilon_t$$

with $\{\varepsilon_t\}$ and $\{\epsilon_t\}$ representing sequences of uncorrelated $(0, \sigma^2)$ random variables.

In the first stage the series is demeaned and the estimated residuals $\widetilde{y}_t$ are obtained. The weighted symmetric estimator of ρ is obtained by solving the following weighted least squares problem

$$\begin{aligned} Q(\rho, \zeta) &= \sum_{t=k+1}^{T} w_{t-} \left(\widetilde{y}_t - \rho \widetilde{y}_{t-1} - \sum_{j=1}^{k-1} \zeta_j \Delta \widetilde{y}_{t-j} \right)^2 \\ &+ \sum_{t=k+1}^{T} (1 - w_{t-k+1}) \left(\widetilde{y}_{t-k} - \rho \widetilde{y}_{t-k+1} + \sum_{j=1}^{k-1} \zeta_j \Delta \widetilde{y}_{t-k+j+1} \right)^2 \end{aligned}$$

where $\zeta = \{\zeta_1, \zeta_2, ..., \zeta_{k-1}\}$ and w_t is specified as

$$w_t = \begin{cases} 0, & t = 1, 2, ..., p \\ (t-k)/(T-2k+2), & t = p+1, p+2, ..., T-k+1 \\ 1, & t = T-k+2, T-k+3, ..., T \end{cases}$$

assuming that $T \geq 2k$.

Equivalently one could minimize

$$\begin{aligned} Q(\rho, \zeta) &= \sum_{t=k+1}^{T} w_t \left(\widetilde{y}_t - \rho \widetilde{y}_{t-1} - \sum_{j=1}^{k-1} \zeta_j \Delta \widetilde{y}_{t-j} \right)^2 \\ &\quad + \sum_{t=1}^{T-k} (1 - w_{t+1}) \left(\widetilde{y}_t - \rho \widetilde{y}_{t+1} + \sum_{j=1}^{k-1} \zeta_j \Delta \widetilde{y}_{t+j+1} \right)^2 . \end{aligned}$$

The associated t-statistic for testing the hypothesis $H_0 : \rho = 1$ against that of the alternative $H_1 : \rho < 1$ is $WS = (\hat{\rho} - 1)/(\hat{V}(\hat{\rho}))^{1/2}$, where $\hat{V}(\hat{\rho})$ is the estimated variance of $\hat{\rho}$ corresponding to the first element of the estimated parameter vector $\hat{\vartheta} = (\hat{\rho}, \widehat{\zeta}) = (X'WX)^{-1}X'WY$.[16] Referring to Table 3.1 below the estimated parameter vector $\hat{\vartheta}$ is obtained from the regression of the dependent variable Y, a $(2T - 2k) \times 1$ dimensional column vector, on the independent variables X, a $(2T - 2k) \times k$ matrix below the headings $\rho, \zeta_1, \zeta_{2,...,}\zeta_{k-1}$, multiplied by the appropriate weights W, a $(2T-2k)$ diagonal matrix. The estimated error variance is $\hat{\sigma} = Q(\hat{\vartheta})/(T - k - 2)$ for a model with an intercept or $\hat{\sigma} = Q(\hat{\vartheta})/(T - k - 3)$ for a model with a linear trend.

[16] In particular, $\hat{V}(\hat{\rho}) = \hat{\sigma}^2 a_{\rho\rho}$, where $a_{\rho\rho}$ is the element (1,1) in the inverse of $\partial^2 Q(\hat{\vartheta})/\partial\vartheta\partial\vartheta'$.

Table 3.1

Data arrangement for regression estimation of autoregressive parameters by the weighted symmetric procedure

Weight (W)	Dependent Variable (Y)	ρ	ζ_1	ζ_2	$\cdots$	ζ_{k-1}
w_{k+1}	$\breve{y}_{k+1}$	$\breve{y}_k$	$\Delta\breve{y}_k$	$\Delta\breve{y}_{k-1}$	$\cdots$	$\Delta\breve{y}_2$
w_{k+2}	$\breve{y}_{k+2}$	$\breve{y}_{k+1}$	$\Delta\breve{y}_{k+1}$	$\Delta\breve{y}_k$	$\cdots$	$\Delta\breve{y}_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
w_{k-p}	$\breve{y}_T$	$\breve{y}_{T-1}$	$\Delta\breve{y}_{T-1}$	$\Delta\breve{y}_{T-2}$	$\cdots$	$\Delta\breve{y}_{T-k+1}$
$1-w_2$	$\breve{y}_1$	$\breve{y}_2$	$-\Delta\breve{y}_3$	$-\Delta\breve{y}_4$	$\cdots$	$-\Delta\breve{y}_{k+1}$
$1-w_3$	$\breve{y}_2$	$\breve{y}_3$	$-\Delta\breve{y}_4$	$-\Delta\breve{y}_5$	$\cdots$	$-\Delta\breve{y}_{k+2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$1-w_{n-k+1}$	$\breve{y}_{1T-k}$	$\breve{y}_{T-k+1}$	$-\Delta\breve{y}_{T-k+2}$	$-\Delta\breve{y}_{T-k+1}$	$\cdots$	$-\Delta\breve{y}_T$

Note: The weighted symmetric estimator treats observations at the beginning of the sample period in the same way as observations at the end of the sample period.

Appendix 3.B Tables and Figures

Table 3.2(a)

Critical values for unit root tests against stationarity around a smooth transition in mean

	DF_a			MAX_a			WS_a		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
25	-4.322	-4.775	-5.811	-4.095	-4.551	-5.580	-4.207	-4.669	-5.704
50	-4.052	-4.405	-5.107	-3.807	-4.179	-4.870	-3.881	-4.245	-4.958
100	-3.939	-4.262	-4.882	-3.690	-4.022	-4.661	-3.766	-4.106	-4.718
200	-3.887	-4.187	-4.808	-3.644	-3.957	-4.600	-3.708	-4.010	-4.643
500	-3.832	-4.177	-4.728	-3.589	-3.906	-4.489	-3.658	-3.963	-4.529

Table 3.2(b)

Critical values for unit root tests against stationarity around a smooth transition in constant only,

fixed trend term included

	DF_b			MAX_b			WS_b		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
25	-5.174	-5.676	-6.611	-4.983	-5.492	-6.457	-5.115	-5.624	-6.595
50	-4.742	-5.095	-5.853	-4.543	-4.902	-5.647	-4.633	-4.975	-5.735
100	-4.560	-4.877	-5.483	-4.377	-4.680	-5.282	-4.441	-4.752	-5.351
200	-4.465	-4.762	-5.375	-4.271	-4.581	-5.182	-4.331	-4.639	-5.235
500	-4.414	-4.678	-5.257	-4.209	-4.504	-5.071	-4.256	-4.532	-5.115

Table 3.2(c)

Critical values for unit root tests against stationarity around a smooth transition in both the intercept and trend term simultaneously

	DF_c			MAX_c			WS_c		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
25	-5.630	-6.149	-7.243	-5.476	-5.999	-7.063	-5.596	-6.128	-7.216
50	-5.099	-5.444	-6.165	-4.929	-5.286	-6.009	-5.026	-5.378	-6.116
100	-4.888	-5.186	-5.860	-4.699	-5.028	-5.670	-4.774	-5.084	-5.737
200	-4.768	-5.059	-5.670	-4.603	-4.888	-5.478	-4.660	-4.945	-5.519
500	-4.673	-4.979	-5.453	-4.505	-4.786	-5.334	-4.561	-4.833	-5.362

Table 3.3

Empirical powers of smooth transition tests for a stationary AR(1) generating process at the 5% nominal level

	$T = 50$			$T = 100$			$T = 200$		
ρ	0.80	0.70	0.60	0.90	0.80	0.70	0.95	0.90	0.85
DF_a	0.110	0.226	0.423	0.111	0.393	0.821	0.108	0.393	0.789
MAX_a	0.116	0.254	0.478	0.130	0.472	0.886	0.124	0.457	0.858
WS_a	0.122	0.270	0.483	0.128	0.471	0.880	0.126	0.472	0.858
DF_b	0.090	0.179	0.315	0.102	0.293	0.636	0.100	0.284	0.635
MAX_b	0.090	0.200	0.334	0.109	0.329	0.695	0.106	0.309	0.690
WS_b	0.092	0.200	0.337	0.111	0.332	0.688	0.109	0.304	0.691
DF_c	0.081	0.155	0.272	0.088	0.275	0.588	0.093	0.238	0.573
MAX_c	0.081	0.164	0.292	0.100	0.290	0.626	0.107	0.267	0.621
WS_c	0.083	0.168	0.295	0.099	0.298	0.637	0.102	0.266	0.622

Table 3.4(a)

Empirical powers of smooth transition tests for a stationary AR(1) generating process around a smooth transition in mean changing from $\alpha_1 = 1$ to $\alpha_1 + \alpha_2 = 2$

γ	0.01		0.10		0.50		1.00		5.00	
τ	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2
	$T = 50,\ \rho = 0.7,\ 0.6,\ 0.5$									
DF_a	0.205	0.219	0.245	0.241	0.214	0.202	0.236	0.210	0.227	0.207
MAX_a	0.240	0.252	0.258	0.268	0.244	0.241	0.269	0.249	0.254	0.236
WS_a	0.254	0.262	0.271	0.275	0.256	0.246	0.279	0.258	0.272	0.244
DF_a	0.396	0.434	0.446	0.438	0.430	0.400	0.408	0.384	0.426	0.404
MAX_a	0.452	0.498	0.482	0.474	0.474	0.476	0.456	0.440	0.458	0.472
WS_a	0.488	0.522	0.500	0.502	0.478	0.494	0.460	0.452	0.466	0.498
DF_a	0.672	0.680	0.706	0.706	0.692	0.664	0.652	0.670	0.634	0.622
MAX_a	0.718	0.716	0.774	0.744	0.750	0.714	0.718	0.718	0.702	0.696
WS_a	0.736	0.742	0.782	0.760	0.744	0.750	0.726	0.738	0.718	0.708
	$T = 100,\ \rho = 0.8,\ 0.7,\ 0.6$									
DF_a	0.397	0.397	0.403	0.393	0.406	0.400	0.391	0.384	0.418	0.366
MAX_a	0.464	0.447	0.482	0.465	0.468	0.465	0.462	0.450	0.475	0.439
WS_a	0.463	0.444	0.482	0.463	0.468	0.459	0.460	0.444	0.482	0.432
DF_a	0.832	0.832	0.804	0.822	0.802	0.826	0.818	0.788	0.796	0.782
MAX_a	0.880	0.880	0.862	0.874	0.852	0.874	0.882	0.840	0.856	0.842
WS_a	0.888	0.880	0.864	0.876	0.846	0.876	0.870	0.836	0.854	0.836
DF_a	0.978	0.990	0.980	0.980	0.972	0.976	0.976	0.984	0.986	0.982
MAX_a	0.992	0.998	0.996	0.992	0.990	0.996	0.990	0.992	0.992	0.994
WS_a	0.994	0.996	0.996	0.992	0.992	0.994	0.986	0.990	0.986	0.994
	$T = 200,\ \rho = 0.9,\ 0.8,\ 0.7$									
DF_a	0.394	0.378	0.377	0.395	0.378	0.404	0.385	0.371	0.373	0.414
MAX_a	0.469	0.452	0.454	0.463	0.455	0.478	0.450	0.435	0.436	0.473
WS_a	0.482	0.465	0.456	0.472	0.475	0.478	0.459	0.445	0.452	0.486
DF_a	0.970	0.972	0.968	0.982	0.978	0.972	0.976	0.980	0.980	0.972
MAX_a	0.980	0.988	0.992	0.994	0.990	0.982	0.986	0.986	0.988	0.982
WS_a	0.986	0.990	0.996	0.994	0.988	0.984	0.986	0.986	0.982	0.976
DF_a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MAX_a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
WS_a	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: In this and all subsequent tables of this kind values of ρ in descending order correspond to the first, second and third set of test statistics DF, MAX and WS, respectively.

Table 3.4(b)

Empirical powers of smooth transition tests for a stationary AR(1) generating process around a smooth transition in mean changing from $\alpha_1 = 1$ to $\alpha_1 + \alpha_2 = 6$

γ	0.01		0.10		0.50		1.00		5.00	
τ	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2
				$T = 50$, $\rho = 0.7$, 0.6, 0.5						
DF_a	0.245	0.230	0.301	0.271	0.264	0.196	0.218	0.170	0.189	0.162
MAX_a	0.263	0.253	0.337	0.315	0.305	0.246	0.243	0.206	0.219	0.195
WS_a	0.274	0.270	0.350	0.330	0.312	0.261	0.252	0.222	0.228	0.203
DF_a	0.448	0.422	0.496	0.430	0.454	0.368	0.420	0.370	0.364	0.288
MAX_a	0.516	0.480	0.556	0.492	0.492	0.456	0.458	0.430	0.432	0.378
WS_a	0.514	0.502	0.568	0.512	0.508	0.460	0.486	0.454	0.444	0.390
DF_a	0.670	0.708	0.698	0.712	0.696	0.650	0.726	0.608	0.608	0.532
MAX_a	0.720	0.762	0.746	0.752	0.732	0.722	0.768	0.686	0.672	0.618
WS_a	0.738	0.776	0.752	0.774	0.754	0.728	0.762	0.698	0.674	0.632
				$T = 100$, $\rho = 0.8$, 0.7, 0.6						
DF_a	0.426	0.406	0.458	0.419	0.377	0.325	0.346	0.302	0.313	0.297
MAX_a	0.507	0.493	0.530	0.487	0.437	0.402	0.418	0.380	0.396	0.366
WS_a	0.505	0.489	0.526	0.488	0.441	0.409	0.418	0.388	0.392	0.375
DF_a	0.804	0.854	0.842	0.828	0.814	0.734	0.772	0.726	0.754	0.764
MAX_a	0.858	0.882	0.888	0.898	0.870	0.808	0.846	0.824	0.822	0.816
WS_a	0.856	0.872	0.894	0.884	0.858	0.816	0.842	0.830	0.816	0.810
DF_a	0.980	0.986	0.980	0.972	0.976	0.980	0.960	0.974	0.968	0.954
MAX_a	0.988	0.992	0.994	0.984	0.990	0.992	0.984	0.992	0.988	0.970
WS_a	0.986	0.992	0.994	0.984	0.986	0.992	0.982	0.994	0.988	0.968
				$T = 200$, $\rho = 0.9$, 0.8, 0.7						
DF_a	0.422	0.419	0.418	0.380	0.345	0.288	0.342	0.293	0.324	0.302
MAX_a	0.490	0.487	0.479	0.442	0.417	0.376	0.404	0.366	0.400	0.378
WS_a	0.502	0.503	0.492	0.445	0.427	0.390	0.418	0.384	0.413	0.390
DF_a	0.976	0.982	0.978	0.964	0.958	0.958	0.960	0.960	0.950	0.962
MAX_a	0.988	0.988	0.986	0.974	0.982	0.974	0.992	0.978	0.978	0.978
WS_a	0.990	0.986	0.984	0.980	0.984	0.978	0.990	0.986	0.980	0.980
DF_a	1.000	1.000	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000
MAX_a	1.000	1.000	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000
WS_a	1.000	1.000	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.4(c)

Empirical powers of smooth transition tests for a stationary AR(1) generating process around a smooth transition in mean changing from $\alpha_1 = 1$ to $\alpha_1 + \alpha_2 = 11$

γ	0.01		0.10		0.50		1.00		5.00	
τ	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2
					$T=50,\ \rho=0.7,\ 0.6,\ 0.5$					
DF_a	0.237	0.245	0.302	0.274	0.271	0.214	0.197	0.166	0.143	0.125
MAX_a	0.270	0.281	0.327	0.316	0.303	0.249	0.224	0.192	0.162	0.156
WS_a	0.282	0.293	0.344	0.333	0.308	0.263	0.233	0.203	0.171	0.166
DF_a	0.470	0.484	0.488	0.490	0.480	0.368	0.370	0.372	0.346	0.290
MAX_a	0.524	0.534	0.544	0.582	0.564	0.450	0.426	0.424	0.402	0.368
WS_a	0.552	0.552	0.584	0.588	0.572	0.476	0.436	0.440	0.414	0.368
DF_a	0.704	0.690	0.694	0.676	0.718	0.642	0.648	0.626	0.554	0.562
MAX_a	0.782	0.748	0.760	0.744	0.792	0.734	0.726	0.692	0.604	0.628
WS_a	0.786	0.762	0.780	0.754	0.798	0.746	0.740	0.712	0.612	0.644
					$T=100,\ \rho=0.8,\ 0.7,\ 0.6$					
DF_a	0.461	0.433	0.453	0.406	0.365	0.330	0.304	0.294	0.300	0.289
MAX_a	0.528	0.505	0.533	0.479	0.429	0.412	0.372	0.368	0.356	0.368
WS_a	0.526	0.505	0.532	0.469	0.426	0.415	0.373	0.374	0.358	0.371
DF_a	0.838	0.834	0.834	0.810	0.788	0.738	0.784	0.740	0.724	0.708
MAX_a	0.910	0.886	0.890	0.876	0.852	0.832	0.848	0.804	0.804	0.812
WS_a	0.908	0.888	0.896	0.870	0.846	0.832	0.858	0.806	0.814	0.814
DF_a	0.986	0.986	0.982	0.968	0.980	0.982	0.962	0.964	0.964	0.946
MAX_a	0.994	0.998	0.994	0.986	0.986	0.990	0.970	0.976	0.982	0.972
WS_a	0.996	0.998	0.992	0.984	0.986	0.994	0.968	0.976	0.982	0.978
					$T=200,\ \rho=0.9,\ 0.8,\ 0.7$					
DF_a	0.447	0.451	0.434	0.350	0.301	0.281	0.284	0.260	0.281	0.258
MAX_a	0.526	0.512	0.492	0.444	0.365	0.354	0.345	0.322	0.349	0.338
WS_a	0.532	0.519	0.503	0.453	0.384	0.371	0.360	0.340	0.369	0.352
DF_a	0.956	0.954	0.962	0.956	0.962	0.960	0.952	0.966	0.962	0.926
MAX_a	0.970	0.970	0.986	0.986	0.978	0.984	0.976	0.980	0.984	0.960
WS_a	0.964	0.972	0.990	0.990	0.984	0.984	0.982	0.978	0.984	0.960
DF_a	0.994	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MAX_a	0.994	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
WS_a	0.994	0.984	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.5(a)

Empirical powers of smooth transition tests for a stationary AR(1) generating process around a smooth transition in the intercept term changing from $\alpha_1 = 1$ to $\alpha_1 + \alpha_3 = 2$, fixed trend term included

γ	0.01		0.10		0.50		1.00		5.00	
τ	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2
					$T = 50,\ \rho = 0.7,\ 0.6,\ 0.5$					
DF_b	0.169	0.168	0.172	0.177	0.169	0.168	0.172	0.177	0.170	0.173
MAX_b	0.188	0.173	0.178	0.192	0.188	0.173	0.178	0.192	0.178	0.185
WS_b	0.192	0.179	0.181	0.195	0.192	0.179	0.181	0.195	0.176	0.188
DF_b	0.270	0.296	0.310	0.324	0.276	0.304	0.276	0.282	0.286	0.282
MAX_b	0.298	0.334	0.346	0.350	0.314	0.346	0.316	0.296	0.302	0.316
WS_b	0.308	0.340	0.352	0.354	0.320	0.362	0.320	0.298	0.308	0.314
DF_b	0.512	0.472	0.508	0.514	0.470	0.504	0.510	0.502	0.502	0.464
MAX_b	0.554	0.506	0.572	0.548	0.534	0.564	0.518	0.554	0.548	0.510
WS_b	0.572	0.526	0.572	0.546	0.530	0.552	0.514	0.562	0.552	0.508
					$T = 100,\ \rho = 0.8,\ 0.7,\ 0.6$					
DF_b	0.293	0.293	0.300	0.300	0.301	0.276	0.266	0.268	0.257	0.295
MAX_b	0.329	0.325	0.333	0.338	0.337	0.323	0.310	0.305	0.294	0.327
WS_b	0.332	0.324	0.329	0.337	0.328	0.320	0.303	0.302	0.290	0.327
DF_b	0.640	0.692	0.628	0.654	0.658	0.652	0.648	0.658	0.616	0.582
MAX_b	0.718	0.710	0.684	0.696	0.692	0.710	0.686	0.710	0.668	0.648
WS_b	0.728	0.704	0.668	0.694	0.690	0.700	0.674	0.732	0.670	0.656
DF_b	0.948	0.912	0.930	0.910	0.924	0.914	0.918	0.916	0.932	0.902
MAX_b	0.972	0.930	0.948	0.956	0.948	0.952	0.936	0.932	0.968	0.948
WS_b	0.974	0.938	0.940	0.952	0.942	0.946	0.938	0.932	0.962	0.942
					$T = 200,\ \rho = 0.9,\ 0.8,\ 0.7$					
DF_b	0.284	0.294	0.276	0.284	0.273	0.281	0.264	0.257	0.293	0.274
MAX_b	0.308	0.325	0.308	0.322	0.300	0.314	0.297	0.302	0.327	0.316
WS_b	0.305	0.324	0.314	0.323	0.302	0.316	0.294	0.299	0.328	0.313
DF_b	0.906	0.906	0.902	0.912	0.906	0.906	0.902	0.908	0.914	0.908
MAX_b	0.944	0.922	0.954	0.936	0.938	0.922	0.932	0.940	0.944	0.948
WS_b	0.932	0.932	0.946	0.934	0.938	0.936	0.934	0.940	0.944	0.942
DF_b	1.000	1.000	0.998	1.000	0.998	1.000	0.998	1.000	0.998	1.000
MAX_b	1.000	1.000	0.998	1.000	0.998	1.000	0.998	1.000	0.998	1.000
WS_b	1.000	1.000	0.998	1.000	0.998	1.000	1.000	1.000	0.998	1.000

Table 3.5(b)

Empirical powers of smooth transition tests for a stationary AR(1) generating process around a smooth transition in the intercept term changing from $\alpha_1 = 1$ to $\alpha_1 + \alpha_3 = 6$, fixed slope term included

γ	0.01		0.10		0.50		1.00		5.00	
τ	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2
	$T = 50,\ \rho = 0.7,\ 0.6,\ 0.5$									
DF_b	0.169	0.168	0.171	0.186	0.142	0.143	0.113	0.108	0.080	0.085
MAX_b	0.188	0.173	0.177	0.198	0.156	0.164	0.119	0.122	0.088	0.100
WS_b	0.192	0.180	0.178	0.197	0.157	0.165	0.124	0.127	0.089	0.101
DF_b	0.318	0.322	0.304	0.280	0.292	0.288	0.198	0.244	0.164	0.150
MAX_b	0.338	0.362	0.322	0.298	0.308	0.306	0.226	0.290	0.204	0.196
WS_b	0.348	0.374	0.324	0.302	0.312	0.314	0.240	0.306	0.206	0.202
DF_b	0.496	0.502	0.534	0.528	0.458	0.430	0.402	0.386	0.346	0.326
MAX_b	0.530	0.550	0.582	0.566	0.500	0.478	0.436	0.450	0.398	0.376
WS_b	0.560	0.562	0.590	0.584	0.500	0.486	0.456	0.464	0.400	0.402
	$T = 100,\ \rho = 0.8,\ 0.7,\ 0.6$									
DF_b	0.293	0.292	0.284	0.306	0.228	0.219	0.170	0.173	0.179	0.183
MAX_b	0.329	0.325	0.318	0.351	0.253	0.257	0.201	0.212	0.203	0.217
WS_b	0.332	0.324	0.315	0.351	0.248	0.265	0.203	0.215	0.199	0.216
DF_b	0.640	0.692	0.610	0.676	0.572	0.552	0.530	0.514	0.490	0.446
MAX_b	0.718	0.708	0.680	0.710	0.616	0.616	0.578	0.578	0.546	0.512
WS_b	0.728	0.704	0.672	0.706	0.614	0.614	0.572	0.580	0.546	0.522
DF_b	0.948	0.910	0.922	0.930	0.884	0.886	0.870	0.866	0.824	0.838
MAX_b	0.972	0.930	0.948	0.956	0.920	0.924	0.894	0.888	0.880	0.888
WS_b	0.974	0.938	0.934	0.960	0.920	0.930	0.902	0.888	0.896	0.896
	$T = 200,\ \rho = 0.9,\ 0.8,\ 0.7$									
DF_b	0.283	0.295	0.256	0.275	0.190	0.208	0.187	0.199	0.211	0.211
MAX_b	0.306	0.323	0.293	0.314	0.221	0.237	0.209	0.225	0.239	0.239
WS_b	0.304	0.326	0.289	0.319	0.224	0.242	0.210	0.227	0.239	0.233
DF_b	0.906	0.906	0.906	0.884	0.822	0.848	0.794	0.842	0.850	0.856
MAX_b	0.946	0.922	0.930	0.944	0.876	0.892	0.872	0.876	0.890	0.896
WS_b	0.932	0.930	0.936	0.938	0.886	0.896	0.874	0.886	0.892	0.892
DF_b	1.000	1.000	0.998	1.000	0.994	0.998	0.996	0.998	0.998	1.000
MAX_b	1.000	1.000	0.998	1.000	0.996	0.998	0.998	1.000	0.998	1.000
WS_b	1.000	1.000	0.998	1.000	0.998	0.998	0.998	0.998	0.998	1.000

Table 3.5(c)

Empirical powers of smooth transition tests for a stationary AR(1) generating process around a smooth transition in the intercept term changing from $\alpha_1 = 1$ to $\alpha_1 + \alpha_3 = 11$, fixed slope term included

γ	0.01		0.10		0.50		1.00		5.00	
τ	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2	0.5	0.2
					$T = 50$, $\rho = 0.7$, 0.6, 0.5					
DF_b	0.169	0.168	0.165	0.183	0.132	0.124	0.092	0.081	0.059	0.065
MAX_b	0.188	0.173	0.172	0.199	0.138	0.143	0.101	0.090	0.061	0.071
WS_b	0.192	0.180	0.173	0.195	0.140	0.148	0.105	0.094	0.061	0.074
DF_b	0.294	0.348	0.278	0.332	0.288	0.254	0.204	0.226	0.176	0.148
MAX_b	0.332	0.380	0.316	0.366	0.326	0.298	0.234	0.248	0.190	0.168
WS_b	0.340	0.380	0.314	0.374	0.318	0.300	0.232	0.250	0.196	0.168
DF_b	0.456	0.482	0.452	0.540	0.442	0.394	0.418	0.390	0.348	0.340
MAX_b	0.512	0.536	0.502	0.594	0.486	0.452	0.456	0.444	0.368	0.386
WS_b	0.518	0.536	0.508	0.594	0.478	0.460	0.462	0.448	0.368	0.388
					$T = 100$, $\rho = 0.8$, 0.7, 0.6					
DF_b	0.294	0.292	0.271	0.297	0.196	0.185	0.144	0.132	0.130	0.134
MAX_b	0.329	0.325	0.298	0.344	0.218	0.224	0.167	0.164	0.155	0.156
WS_b	0.333	0.321	0.296	0.341	0.216	0.225	0.162	0.164	0.151	0.152
DF_b	0.640	0.696	0.594	0.652	0.556	0.506	0.496	0.492	0.442	0.398
MAX_b	0.718	0.712	0.652	0.690	0.578	0.574	0.550	0.540	0.498	0.446
WS_b	0.728	0.698	0.646	0.690	0.574	0.576	0.538	0.550	0.504	0.454
DF_b	0.948	0.914	0.910	0.922	0.880	0.872	0.854	0.864	0.790	0.802
MAX_b	0.972	0.932	0.938	0.948	0.910	0.910	0.882	0.888	0.848	0.850
WS_b	0.972	0.938	0.936	0.944	0.918	0.918	0.894	0.888	0.862	0.852
					$T = 200$, $\rho = 0.9$, 0.8, 0.7					
DF_b	0.282	0.298	0.233	0.246	0.135	0.145	0.125	0.125	0.148	0.139
MAX_b	0.306	0.325	0.266	0.276	0.160	0.179	0.141	0.150	0.175	0.160
WS_b	0.305	0.327	0.260	0.281	0.160	0.180	0.143	0.156	0.179	0.165
DF_b	0.906	0.896	0.898	0.860	0.796	0.820	0.766	0.798	0.814	0.802
MAX_b	0.948	0.922	0.918	0.914	0.844	0.866	0.834	0.838	0.846	0.850
WS_b	0.932	0.928	0.922	0.908	0.856	0.864	0.834	0.848	0.860	0.850
DF_b	1.000	1.000	0.998	1.000	0.994	0.998	0.996	0.996	0.996	1.000
MAX_b	1.000	1.000	0.998	1.000	0.996	1.000	0.998	1.000	0.998	1.000
WS_b	1.000	1.000	0.998	1.000	0.996	1.000	0.998	0.998	0.998	1.000

Table 3.6

Empirical application to the Nelson-Plosser data

Series	T	ADF	ADF_b^f	ADF_b^r	MAX_b	WS_b	ADF_c^f	ADF_c^r	MAX_c	WS_c
Real GNP	62	-2.994	-3.616	-3.230	-3.230	-3.409	-4.329	-4.339	-4.329	-4.467
Nominal GNP	62	-2.195	-3.700	-3.730	-3.700	-3.830	-3.743	-3.832	-3.743	-3.914
Real per capita GNP	62	-3.045	-3.368	-3.747	-3.368	-3.470	-4.623	-4.692	-4.623	-4.776*
Industrial production	111	-2.203	-4.623*	-4.700	-4.623*	-4.838**	-4.625	-4.700	-4.625	-4.840*
Employment	81	-3.356	-3.139	-3.156	-3.139	-3.377	-2.962	-2.888	-2.888	-3.069
Unemployment rate	81	-3.553**	-3.338	-3.405	-3.338	-3.567	-3.285	-3.436	-3.285	-3.521
GNP deflator	82	-2.516	-3.643	-3.238	-3.238	-3.357	-3.505	-3.254	-3.254	-3.449
Consumer prices	111	-2.369	-3.472	-3.521	-3.472	-3.656	-3.370	-3.395	-3.370	-3.548
Wages	71	-2.616	-3.782	-3.340	-3.340	-3.445	-2.914	-2.991	-2.914	-3.019
Real wages	71	-3.049	-4.523	-4.339	-4.339	-4.669*	-4.277	-4.266	-4.266	-4.344
Money stock	82	-3.078	-3.011	-3.266	-3.011	-3.160	-3.512	-3.386	-3.386	-3.496
Velocity	102	-1.663	-4.272	-4.294	-4.272	-4.419	-4.326	-4.329	-4.326	-4.463
Bond yield	71	0.686	-3.104	-3.351	-3.104	-2.818	-4.949	-5.046	-4.949	-4.943
Common Stock	100	-2.653	-5.114**	-5.048	-5.048**	-5.212**	-5.099*	-5.195	-5.099**	-5.120**

Note: *,** and*** denote rejection of the unit root null at the 10%, 5% and 1% significance level respectively. The 10%, 5% and 1% critical values for the standard augmented Dickey-Fuller (ADF) t-statistic employed in column three are those related to the case of a constant and trend term included in the model, taken from Table B.6 in Hamilton (1994).

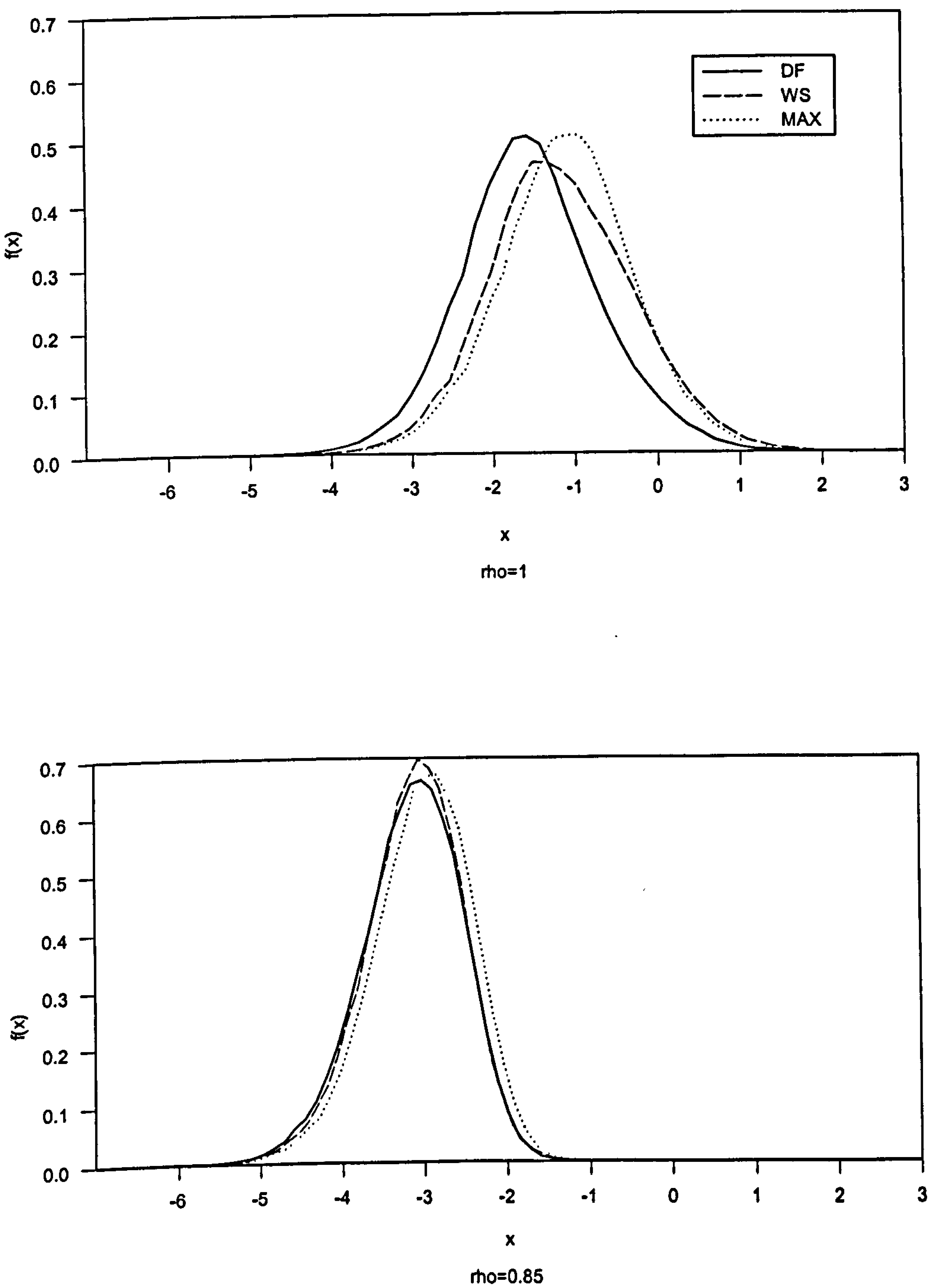


Figure 3.2. Finite sample density of unit root t-tests, $T = 100$, $AR(1)$ data. Top plot, $\rho = 1$, bottom plot $\rho = 0.85$.

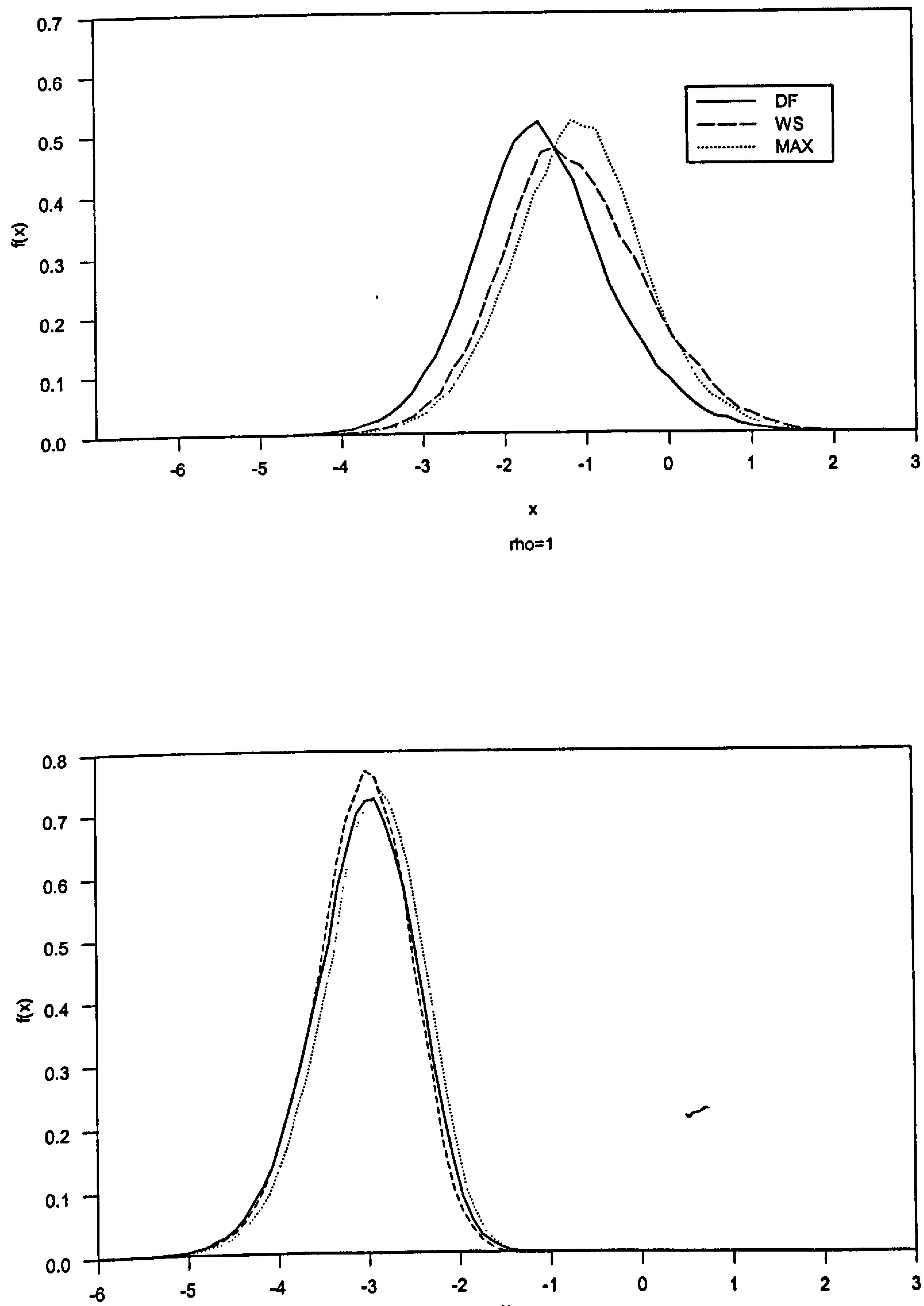


Figure 3.2. Asymptotic density of unit root t-tests. Top plot $\rho = 1$, bottom plot $\rho = 1 + c/T$, where $c = -15$.

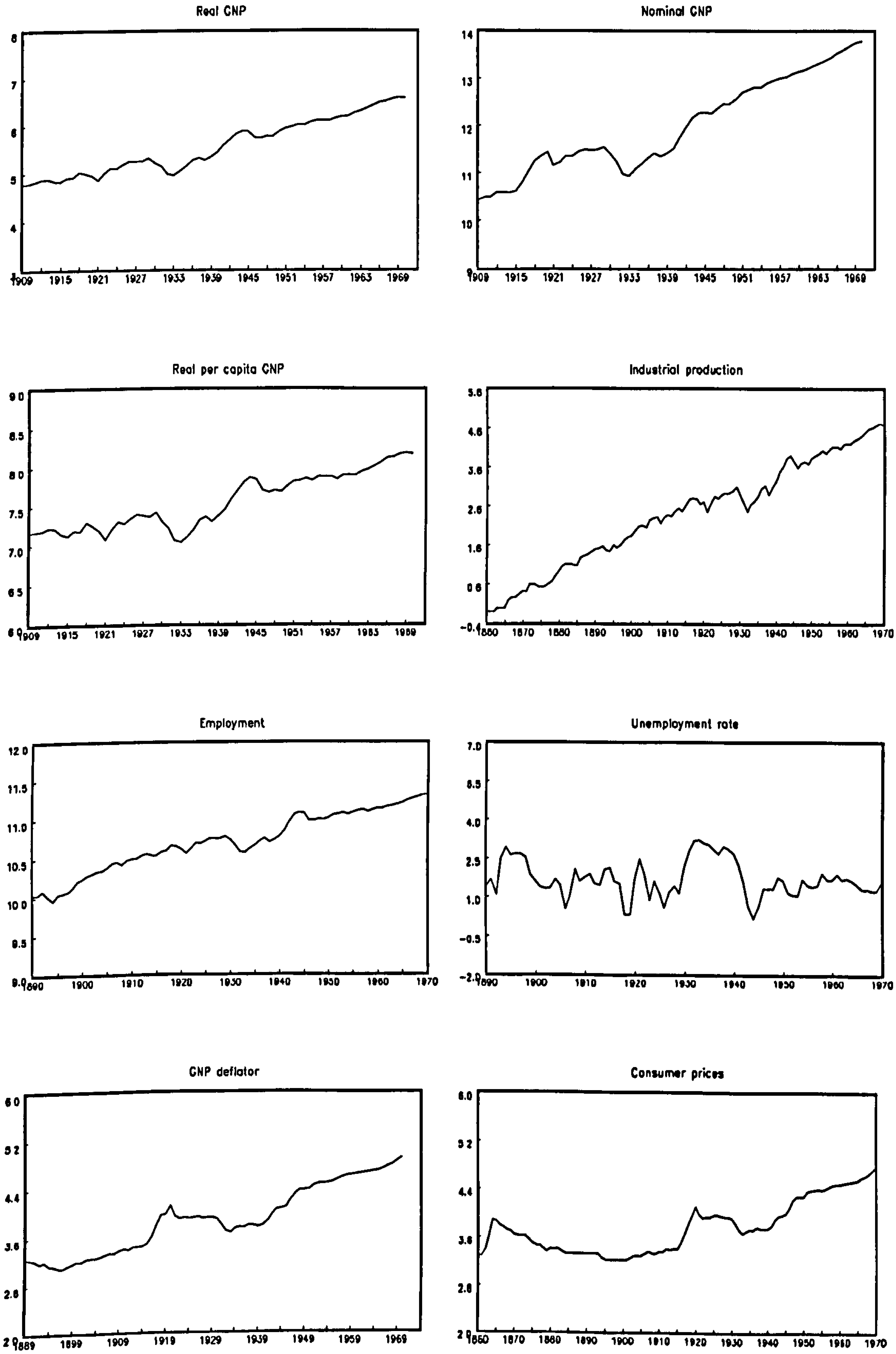


Figure 3.3. Plots of the Nelson-Plosser data.

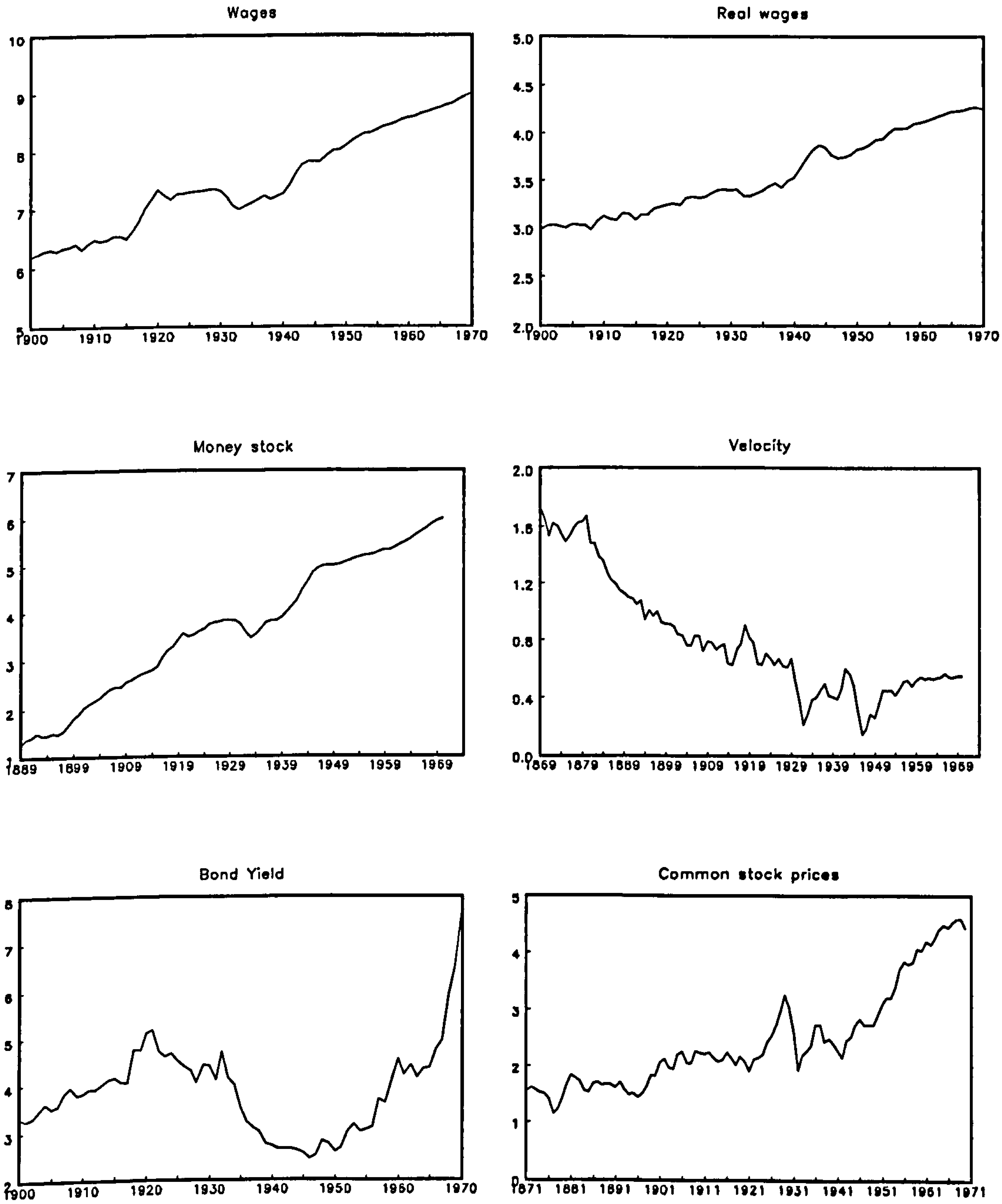


Figure 3.3. (continued)

Chapter 4

More Powerful Panel Data Unit Root Tests with an Application to Mean Reversion in Real Exchange Rates

4.1 Introduction

There is no doubt that panel data offer researchers more possibilities than purely cross section data or time series data, in that they are able to distinguish effects that time series or cross section data alone cannot identify. This fact, along with the availability of a number of important panel data sets covering different individuals, industries, countries over a relatively long period has seen the development of various panel models with both theoretical and empirical research in this area receiving increased attention. Such models involve monitoring variations in characteristics of individuals over time, and therefore allow for the better control of the effects of individual heterogeneity.

A wide range of economic issues have been investigated using panel data. Considerable interest lies in inter-country comparisons that make use of multi-country macroeconomic data, in which case an appreciable time series dimension (T) is typically encountered as well as the existence of a large cross-section dimension (N). While most previous panel research has focused on stationary panels, the feature of nonstationarity known to characterise macroeconomic data, along with the very different medium to long run implications for the way in which the data series are expected to evolve in the presence or absence of unit roots, has prompted research on nonstationary panel data and unit root tests. One possible test might be of the null hypothesis that every individual series is integrated of order one, against the alternative that one panel member series is stationary around a fixed mean or linear trend. As we subsequently note, such a concern has been expressed in the literature on purchasing power parity, where the issue of interest is the possible mean reversion of real exchange rates.

A number of contributions have been made in the econometric literature for testing the unit root hypothesis within a panel data framework. Panel data unit root tests have been developed by, among others, Levin and Lin (1992) and Im, Pesaran and Shin (1997) - hereafter IPS - as extensions of standard time series unit root tests, such as the Dickey-Fuller test or a Lagrange Multiplier test. In particular, the tests put forward by the former involve a dynamic pooled model with fixed effects relying on the assumption of homogeneity of the autoregressive coefficient that indicates the presence or absence of a unit root. However, individuals may react differently to changes in explanatory variables and the individual's reaction may change over time giving rise to spatial heterogeneity that cannot be captured completely by the variable intercept. Thus, the natural generalisation of allowing for the slope parameters of the regressors to vary across panel members seems a more plausible specification. Pesaran and Smith (1995) argue that the dynamic pooled model could be biased because of heterogeneity in the parameters across individual units and suggest that an average of the individual regressions can lead to consistent estimates of the parameters as long as N and T tend to infinity. Building on such an argument IPS put forward more general panel data unit root tests that rely on combining evidence from individual regressions.

Of the empirical applications found in the literature, a large number of which employ the Levin and Lin (1992) and IPS tests, the issue most commonly investigated is the validity of the long run purchasing power parity (PPP) by examining the mean reversion in real exchange rates, see *inter alia* Abuaf and Jorion (1990), Jorion and Sweeney (1996), Frankel and Rose (1996), MacDonald (1996), Oh (1996) and Wu (1996), Coakley and Fuertes (1997), Papell (1997), Higgins and Zakrajsek (1999), Maddala and Wu (1999) and Fleissig and Strauss (2000). Additional applications that reveal the diversity of their use include investigation of mean reversion in inflation rates (Culver and Papell, 1997), in real wage data (Fleissig and Strauss, 1997 ; Lee and Wu, 2001), the convergence of per capita output (Fleissig and Strauss, 1999), the real interest rate parity hypothesis (Wu and Chen, 1998) and stock market efficiency (Boumahdi and Thomas, 1991).

The previous chapter reviewed a number of power-enhancing unit root tests that have emerged in the literature as a result of the power deficiency of traditional such tests. The modified tests were subsequently applied to the problem of testing for a unit root under the alternative of structural change of a smooth transition type. We proceed here along the same lines. Given that the most widely applied panel unit root tests are extensions of traditional unit root tests, we explore whether the two modifications, due to Pantula *et al.* (1994) and Leybourne (1995) in the basic time series context, maintain their power gains when applied to the panel data unit root testing context. To this end, we extend the t-bar and LM-bar panel unit root tests proposed by IPS and provide some illustrative Monte Carlo evidence to assess their finite sample performance. In doing so, we seek to contribute to the substantial amount of simulation work necessary to establish systematically the role of asymptotic theory in estimates and tests derived from finite samples in this area, as pointed out by Banerjee (1999).

This chapter unfolds as follows. Section 4.2 briefly reviews the most widely applied panel unit root tests documented in the literature. In Section 4.3, modified panel unit root tests are introduced in the base case, where independence over panel data members is assumed. However, as stressed, for example, by O'Connell (1998) difficulties, particularly spurious rejections of the null hypothesis, can arise when individual panel series are generated by cross-sectionally correlated innovations. An important special case, which is readily dealt with through the subtraction of time-specific means, was considered by IPS, and is briefly discussed in Section 4.4. Section 4.5 reports results of a simulation exercise both for the case of independence across the units of the panel and allowing for a common time-specific component, demonstrating that, while the modified panel data unit root tests retain size reliability, they can produce appreciable gains in power. In Section 4.6, we analyse a panel of series of real exchange rates against the US dollar. Preliminary analysis suggests that the data generating process implicitly assumed in Section 4.5 provides an inadequate description of the data. Accordingly, we follow Maddala and Wu (1999) in applying a bootstrap approach to testing. However, we consider also our modifications applied to the bootstrap approach, finding that the application of these tests yields appreciably stronger evidence against the unit root null hypothesis for our data than do the unmodified tests. Simulation evidence in Section 4.7 reveals that the modified bootstrap tests retain the power gains noted in simpler cases in Section 4.5. Conclusions are presented in Section 4.8.

4.2 Review of Panel Data Unit Root Tests

The econometric literature has produced a large variety of tests for unit roots in panel data. In this section we review the most commonly applied to date that consider asymptotics in both dimensions of the panel ($T \to \infty$ and $N \to \infty$).[1,2] They are based on the well known Dickey-Fuller type tests and differ in the assumptions made about the heterogeneity of the regression parameters for the observed units and the derived test statistics.

While it was Quah (1992,1994)[3] who initiated research in this area by considering tests for panel data based on models without fixed effects, the first widely used panel unit root test was introduced by Levin and Lin (1992), hereafter LL, who considered more general models allowing for fixed effects and heterogeneous serially correlated errors. LL develop unit root tests for

$$\Delta y_{it} = a_i + \varphi_i y_{i,t-1} + \varepsilon_{it} \qquad i = 1, ..., N; t = 1, ..., T \qquad \varepsilon_{it} \tilde{\ } iid(0, \sigma^2)$$

[1]Phillips and Moon (2000), on the theoretical side of the ongoing research on non-stationary large N and T panels, develop a limit theory that allows for sequential limits where $T \to \infty$ followed by $N \to \infty$ and joint limits where N, T tend to infinity simultaneously. They also give a condition under which the sequential limit becomes the joint limit. Their limit distribution theory relies on $N/T \to 0$, a condition that indicates that it is most likely to be useful in practice when N is moderate and T is large.

[2]Panel unit root tests have also been developed for fixed T and $N \to \infty$, as in Harris and Tzavalis (1999), to account for the case of small T frequently observed in for example microeconomic panel data. As it is true that sometimes series are rather short, we are well advised to be concerned about the small sample properties of estimators and tests whose asymptotic properties depend on $T \to \infty$.

[3]In particular, Quah (1992,1994) considered testing for a unit root, that is $\rho = 1$, in the following models

$y_{it} = \rho y_{it-1} + \varepsilon_{it}, i = 1, ..., N; t = 1, ..., T \qquad \varepsilon_{it} \tilde{\ } iid(0, \sigma^2)$

$y_{it} = a + \rho y_{it-1} + \varepsilon_{it}, i = 1, ..., N; t = 1, ..., T \qquad \varepsilon_{it} \tilde{\ } iid(0, \sigma^2)$

However, the limited nature of such models renders them of trivial empirical interest and they are therefore only mentioned at the outset.

where N and T are the cross-section and time series dimensions of the panel, respectively. The parameters a_i allow for the possibility of member-specific fixed effects that account for heterogeneity in individual behaviour. The error term ε_{it} is uncorrelated with y_{it-1}, and is initially assumed to follow an *iid* process. Such an assumption is relaxed in the sequel to allow for the possibility of serial correlation in the disturbances[4] and heteroskedasticity, under which the limiting distribution of the test statistics remains unaffected. Their procedure is based on the estimation of the ADF test equation,

$$\Delta y_{it} = a_i + \varphi_i y_{i,t-1} + \sum_{j=1}^{k_i} \vartheta_{ij} \Delta y_{i,t-j} + \varepsilon_{it} \qquad i = 1, ..., N; t = 1, ..., T, \quad \varepsilon_{it} \tilde{} iid(0, \sigma_i^2) \tag{4.1}$$

applied to each individual series, first subtracting cross sectional averages $\overline{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$ from the observed data to ensure that they are asymptotically independent across units. The above ADF regression is equivalent to performing two auxiliary regressions of Δy_{it} and y_{it-1} on the remaining part of equation (4.1), that is,

$$\Delta y_{it} = \widehat{a}_{1i} + \sum_{j=1}^{k_i} \widehat{\vartheta}_{1ij} \Delta y_{i,t-j} + \widehat{e}_{it}$$

and

$$y_{i,t-1} = \widehat{a}_{2i} + \sum_{j=1}^{k_i} \widehat{\vartheta}_{2ij} \Delta y_{i,t-j} + \widehat{V}_{i,t-1}.$$

To control for heterogeneity across individuals the estimated innovations $\widehat{e}_{it}$ and $\widehat{V}_{it-1}$ that arise from the above two auxiliary regressions are normilized by the regression standard error for equation (4.1). If the normalized residuals are denoted $\widetilde{e}_{it}$ and $\widetilde{V}_{i,t-1}$, respectively, then the test they propose is conducted on the regression of the following form, which involves pooling the now homoskedastic disturbances ε_{it} for all the cross-sections and time periods, where the parameter

[4]Papell (1997) highlights the importance of accounting for serially correlated disturbances. He finds that the LL test, the critical values of which do not incorporate serial correlation in the disturbances, suffers from size distortions biasing the results toward rejection of the unit root null.

φ is assumed to be identical for all units

$$\widetilde{e}_{it} = \varphi \widetilde{V}_{i,t-1} + \varepsilon_{it}.$$

LL test the following hypothesis

$$H_0: \ \varphi_1 = \varphi_2 = ... = \varphi_i = \varphi = 0, \qquad i = 1, ..., N$$

against the homogeneous alternative

$$H_1: \ \varphi_i = \varphi < 0, \qquad i = 1, ..., N.$$

Under the null they show that the estimator of φ in the pooled regression converges to a non-central normal distribution, where the degree of non-centrality depends on T. Hence, the t-statistic $t_{\widehat{\varphi}}$ for a test of the hypothesis $\varphi = 0$ adjusted for the mean and standard deviation of the asymptotic distribution of $\widehat{\varphi}$, has a standard normal distribution. This result is established under the assumption that both T and N tend to infinity, but that T increases faster than N, so that $N/T \rightarrow 0$.[5]

Various extensions of the above models are considered by LL that include homogeneous and heterogeneous deterministic trends. In such cases, the resulting test statistics are shown to also follow standard normal distributions asymptotically. A drawback, however, of their study is the restrictive nature of the alternative hypothesis, which for example in the case of testing for the PPP hypothesis would mean confining the speed of convergence across countries to be the same.

IPS relax the assumption of a common φ under the alternative, allowing for more heterogeneity of behaviour across individuals, and base their approach on the use of separate unit root tests for the N cross-section units. They consider

[5]Some of their results involve the assumption that $\frac{\sqrt{N}}{T} \rightarrow 0$ as N and T tend to infinity.

the following data generating process

$$y_{it} = (1 - \rho_i)\mu_i + \rho_i y_{i,t-1} + \varepsilon_{it}, \qquad i = 1, ..., N; t = 1, ..., T \tag{4.2}$$

where ε_{it}, is assumed independently distributed across both groups and time periods with mean zero and finite heterogeneous variance σ_i^2, thus accounting for heteroskedasticity in the error terms.[6] The initial values, y_{i0}, are assumed known. The corresponding model is

$$\Delta y_{it} = a_i + \varphi_i y_{i,t-1} + \varepsilon_{it}, \qquad i = 1, ..., N; t = 1, ..., T \tag{4.3}$$

where they test the null hypothesis that φ_i is zero for all cross-sectional units against the alternative that some of the φ_i's are less then zero. Under the null, there is no fixed effect, while under the alternative, each fixed effect is equal to $(1 - \rho_i)\mu_i$. The tests they propose are based on the average over the individual units of an ADF t-statistic, as well as a Lagrange-multiplier test statistic of the hypothesis that $\varphi_i = 0$. The ADF version of (4.2) is used to account for residual serial correlation across time periods. They further allow for the possibility of heteroskedasticity and a common time-specific component in the errors. To account for the latter, IPS express all variables in the equation as deviations from their time-specific means.

The t-bar test, $\bar{t}$, they propose is computed as the average of the individual t-statistics t_i, that is $\bar{t} = 1/N \sum_{i=1}^{N} t_i$. Standardized by the appropriate mean and variance values,

$$\frac{\sqrt{N}\{\bar{t} - E(t_i)\}}{\sqrt{Var(t_i)}} = \bar{t}_s \tag{4.4}$$

$\bar{t}_s$ is shown to be distributed as a standard normal distribution under the null as $N \to \infty$, where $E(t_i)$ and $Var(t_i)$ are the mean and variance respectively of the DF distribution. In the presence of residual serial correlation the same result is obtained under the assumption that both N and T tend to infinity such that $N/T \to q$, where q is a finite positive constant.[7] Monte Carlo techniques

[6] Specification of the DGP as in (4.2) follows from considering initially the model as $y_{it} = \mu_i + x_{it}, \quad x_{it} = \rho_i x_{it-1} + \varepsilon_{it}$.

[7] Weaker assumptions are therefore required in the derivation of the asymptotics by IPS as

are required to estimate the mean and variance adjustment factors, which are tailored to the number of ADF lags and are tabulated by IPS for the various values of the lag order k and the sample size T.[8] In their derivations, IPS assume that the second moment of t_i exists. They justify this assumption by extensive Monte Carlo studies and conjecture that the second moment of t_i is finite if $T \geq 5$ when the regression includes only a constant and $T \geq 6$ when a time trend is included.

Using the likelihood framework associated with the DF regressions in (4.2), IPS also propose the LM-bar test based on the average of the Lagrange multiplier (LM) statistics applied to each cross-section unit in the panel. The LM-bar statistic for testing $\varphi_i = 0$ is defined as $\overline{LM} = \frac{1}{N}\sum_{i=1}^{N} LM_i$, where $LM_i = \frac{T\Delta \mathbf{y}_i' P_i \Delta \mathbf{y}_i}{\Delta \mathbf{y}_i' M_\tau \Delta \mathbf{y}_i}$, $M_\tau = I_T - \tau_T(\tau_T'\tau_T)^{-1}\tau_T'$, $P_i = M_\tau \mathbf{y}_{i,-1}(\mathbf{y}_{i,-1}' M_\tau \mathbf{y}_{i,-1})^{-1}\mathbf{y}_{i,-1}' M_\tau$, $\tau_T = (1, 1, ..., 1)'$, $\mathbf{y}_{i,-1} = (y_{i0}, y_{i1}, ..., y_{iT-1})'$ and $\Delta \mathbf{y}_i = (\Delta y_{i1}, \Delta y_{i2}, ..., \Delta y_{iT})'$.[9] Invoking the Lindeberg-Levy central limit theorem, IPS show that

$$\frac{\sqrt{N}\{\overline{LM} - E(LM_i)\}}{\sqrt{Var(LM_i)}} = \overline{LM}_s \tag{4.5}$$

that is, $\overline{LM}_s$ follows a standard normal distribution under the null as $N \to \infty$, where $E(LM_i)$ and $Var(LM_i)$ are the mean and variance of the distribution of the LM_i statistic respectively, tabulated also in IPS for various values of k and T. Existence of the second moment is assured from the equation $LM_i = TR_i^2$. As R_i^2 is the square of a correlation coefficient bounded between 0 and 1, its second moment exists for any finite T and therefore the second moment of LM_i exists for finite T. As in the case of the t-bar statistic, in the presence of residual serial correlation, where the individual LM_i statistics for finite T depend on nuisance parameters under the null hypothesis, T and N are both required to tend to infinity at similar rates such that $N/T \to q$, where q is a finite positive constant. This is needed to ensure that $\overline{LM}$ is asymptotically normally distributed in this

opposed to the stronger, $N/T \to 0$, required for the asymptotic validity of the LL test.

[8]While the lag order can in principle differ across individuals, most studies hold it fixed for expository purposes.

[9]When the disturbances ε_{it} are serially correlated, $LM_i = \frac{T(\Delta \mathbf{y}_i' P_i^* \Delta \mathbf{y}_i)}{\Delta \mathbf{y}_i' M_{Q_i} \Delta \mathbf{y}_i}$ where $P_i^* = M_{Q_i}\mathbf{y}_{i,-1}(\mathbf{y}_{i,-1}' M_{Q_i}\mathbf{y}_{i,-1})^{-1}\mathbf{y}_{i,-1}' M_{Q_i}$, $M_{Q_i} = I_T - Q_i(Q_i'Q_i)^{-1}Q_i'$ and $Q_i = (\tau_T, \Delta \mathbf{y}_{i,-1}, \Delta \mathbf{y}_{i,-2}, ..., \Delta \mathbf{y}_{i,-k})$.

situation. Similar results hold for models including deterministic trends where IPS allow for the trend coefficients to differ across individuals.

Lastly, it would be an oversight not to mention another method that relies on combining the evidence from several independent tests, proposed by Maddala and Wu (1999) who follow Fisher (1932). The so-called Fisher test is non-parametric as opposed to the parametric tests of LL and IPS, and is based on the same model (4.2) and null hypothesis as in IPS, namely that all individuals exhibit a unit root with the alternative stating that at least one is stationary. Assuming that there are N individual units with p_i the observed p-value of the ADF test for the ith individual, the Fisher test involves calculating the p_i-values from the t-statistics of the N units using a simulated distribution of the t-statistic under the null hypothesis. These p-values are uniformly distributed random variables $p_i \tilde{} U(0,1)$ assumed to be independent so the usual chi-square statistic is computed as

$$-2\sum \log p_i \tilde{} \chi^2_{DF=2N}.$$

The idea of combining p-values from a unit root test applied to each individual unit in the panel is also shared in the tests proposed by Choi (2001). He considers a broader range of test statistics than just the Fisher test considered by Maddala and Wu (1999), which appear to have better finite sample size and power properties. Such tests however, have not received much attention in the literature. We will concentrate on the tests of IPS, which are widely applicable and permit greater heterogeneity across panel members under the alternative hypothesis than the LL test.

4.3 Motivation and Proposed Panel Data Unit Root Tests

Investigations to solve the well-known power deficiency problem of traditional unit root tests were seen in the previous chapter to have brought to light a number of more powerful test statistics, most of which are elaborations of standard Dickey-Fuller tests. Our suspicion is that since the IPS t-bar statistic is obtained as a linear combination of the ADF statistics of the individual time series, we could expect a similar pattern of increased power performance when such test statistics are adopted in the panel unit root framework. To this end, we extend the most powerful of the modified unit root tests, namely the MAX-test of Leybourne (1995) and the weighted symmetric (WS) test of Park and Fuller (1995), to the panel setting.

Let us consider the data generating process expressed as (4.2). We define a vector of error terms for later use; $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$. The t-statistic is obtained by running

$$\Delta y_{it} = \widehat{a}_i + \widehat{\varphi}_i y_{i,t-1} + \widehat{\varepsilon}_{it}, \tag{4.6}$$

over each individual cross-section unit in the panel. Let DF_{f_i} be the t-statistic associated with $\widehat{\varphi}_i$ for testing $\varphi_i = 0$.[10] Then the standard IPS t-bar statistic is given by $\bar{t} \equiv \frac{1}{N}\sum_{i=1}^{N} DF_{f_i}$.

Following Leybourne (1995), we now consider the reverse series of y_{it}, that is $z_{it} \equiv y_{i,T+1-t}$. The corresponding reverse regression is given by

$$\Delta z_{it} = \widetilde{a}_i + \widetilde{\varphi}_i z_{i,t-1} + \widetilde{\varepsilon}_{it} \tag{4.7}$$

for each i. Let DF_{r_i} be the t-statistic associated with $\widetilde{\varphi}_i$ for testing $\varphi_i = 0$. The subscript r indicates that the statistic is based on the reverse regressions. The MAX statistic for series i is then $Max_i = Max(DF_{f_i}, DF_{r_i})$. With $\overline{Max} = N^{-1}\sum_{i=1}^{N} Max_i$ a panel data unit root test statistic is given by

$$\frac{\sqrt{N}\{\overline{Max_i} - E(Max_i)\}}{\sqrt{Var(Max_i)}} = \overline{Max_s} \tag{4.8}$$

which is compared with critical values from a standard normal distribution.[11] The validity of this approach is established in the following Proposition.[12] The proof is outlined in Appendix 4.A.

Proposition 4.3.1 *Suppose y_{it} are generated by (4.2) with $\rho_i = 1$. Further*

[10] We interchange the notation of the individual t-statistic t_i with that of DF_{f_i} to comply with the notation of the Max test as initially introduced in chapter two. The subscript f_i indicates that the statistic is based on the forward regressions of (4.6).

[11] The subscript s is used throughtout to denote "standardized".

[12] It is referred to as a proposition, as the underlying assumption of the existence of finite moments for DF_{f_i} and DF_{r_i} is based on IPS's conjecture that has not been formally proved.

assume that the errors ε_{it} are independent and identically distributed across $i = 1, 2, ..., N$. Then the statistic (4.8) tends to the standard normal variate as $N \to \infty$.

We now turn to the weighted symmetric estimator and the associated t-statistic, WS. As mentioned in the previous chapter both the WS and MAX statistics exploit the time reversibility property of stationary autoregressive processes, namely that a stationary autoregressive series can be given either a forward or a backward representation.

Consider the first-order stochastic difference equation

$$y_{it} = a_i + \rho_i y_{i,t-1} + \varepsilon_{it}, \qquad i = 1, ..., N; t = 1, ..., T \qquad (4.9)$$

where $a_i \equiv (1 - \rho_i)\mu_i$. The forward representation of such a series is given by

$$y_{it} = a_i + \rho_i y_{i,t+1} + u_{it}.$$

The weighted symmetric estimator of the autoregressive parameter ρ_i for each individual unit i follows from minimizing the following weighted sum of squares of the estimated ε_{it} and u_{it}

$$Q(\rho_i) = \sum_{t=2}^{T} w_t \left(\breve{y}_{it} - \rho_i \breve{y}_{i,t-1}\right)^2 + \sum_{t=1}^{T-1} (1 - w_{t+1}) \left(\breve{y}_{it} - \rho_i \breve{y}_{i,t+1}\right)^2$$

where $\breve{y}_{it} = y_{it} - T^{-1} \sum_{t=1}^{T} y_{it}$ and $w_t = \frac{t-1}{T}$.

The weighted symmetric estimator is given by

$$\widehat{\rho}_{iws} = \frac{\sum_{t=2}^{T} \breve{y}_{it}\breve{y}_{i,t-1}}{\sum_{t=2it}^{T-1} \breve{y}_{it}^2 + T^{-1} \sum_{t=1}^{T} \breve{y}_{it}^2}$$

and the corresponding t-statistic is obtained by running (4.9) over each individual cross-section unit in the panel

$$WS_i = \widehat{\sigma}_{iws}^{-1}(\widehat{\rho}_{iws} - 1)(\sum_{t=2}^{T-1} \breve{y}_{it}^2 + T^{-1}\sum_{t=1}^{T} \breve{y}_{it}^2)^{1/2}$$

where $\sigma_{iws}^2 = Q(\widehat{\rho}_{iws})/(T-2)$.

For further details on the individual WS_i statistics, the reader is referred to chapter three where WS_i is equivalent to WS without the subscript i. The ensuing panel data unit root statistic is obviously

$$\frac{\sqrt{N}\{\overline{WS} - E(WS_i)\}}{\sqrt{Var(WS_i)}} = \overline{WS}_s. \tag{4.10}$$

Limiting null standard normality of this statistic can be established along lines just followed, as simulation evidence establishes the existence of second moments of WS_i for precisely those sample sizes found for t_i by IPS.

Finally, a more powerful variant of the Lagrange Multiplier statistic, the latter originally introduced by Solo (1984), can be found. This is based on forward and reverse regressions, as above, which yield the statistics LM_{f_i} and LM_{r_i}. However, noting that these are necessarily positive (see Schmidt and Phillips, 1992), tests are based on their minima $Min_i \equiv Min(LM_{f_i}, LM_{r_i})$. In particular, consider once again model (4.2) as the data generating process. For each i, the following two regressions are run.

(1) Restricted regression with $\varphi_i = 0$

$$\Delta y_{it} = \widetilde{a}_i + \widetilde{\varepsilon}_{it} \tag{4.11}$$

(2) Auxiliary regression

$$\widetilde{\varepsilon}_{it} = \widehat{a}_i + \widehat{\varphi}_i y_{i,t-1} + \widehat{\varepsilon}_{it}. \tag{4.12}$$

Then, the Langrange multiplier statistic denoted by LM_{f_i} for testing $\varphi_i = 0$ is $TR_{f_i}^2$, where $R_{f_i}^2$ is the square of the correlation coefficient from the auxiliary regression (4.12). The LM-bar statistic is given by

$$\overline{LM} \equiv \frac{1}{N}\sum_{i=1}^{N} LM_{f_i}.$$

Considering the reverse series z_{it}, the corresponding reverse regressions are given as follows.

(1) Restricted regression with $\varphi_i = 0$

$$\Delta z_{it} = \widetilde{a}_i + \widetilde{\varepsilon}_{it} \tag{4.13}$$

(2) Auxiliary regression

$$\widetilde{\varepsilon}_{it} = \widehat{a}_i + \widehat{\varphi}_i z_{i,t-1} + \widehat{\varepsilon}_{it}, \tag{4.14}$$

where the Lagrange multiplier statistic in this case is $LM_{r_i} = TR^2_{ri}$ and R^2_{ri} is the square of the correlation coefficient from the auxiliary regression in (4.14).

The test statistic is then, with $\overline{Min} = N^{-1}\sum_{i=1}^{N} Min_i$,

$$\frac{\sqrt{N}\{\overline{Min} - E(Min_i)\}}{\sqrt{Var(Min_i)}} = \overline{Min_s}. \tag{4.15}$$

The following theorem establishes the limiting standard normal null distribution of this statistic. The proof can be found in Appendix 4.A.

Theorem 4.3.1 *Suppose y_{it} are generated by (4.2) with $\rho_i = 1$. Under the assumption that the standardised errors ε_i/σ_i are independent and identically distributed across $i = 1, 2, ..., N$, the statistic (4.15) tends to the standard normal variate as $N \to \infty$.*

In the case where the the null generating process incorporates a non-zero drift, and the alternative contains a linear trend, following Bhargava (1986) we consider the model

$$\begin{aligned} y_{it} &= \xi_i + \mu_i t + x_{it} \\ x_{it} &= \rho_i x_{i,t-1} + \varepsilon_{it}, \qquad i = 1, ..., N; t = 1, ..., T \end{aligned}$$

which can be written as

$$y_{it} = \xi_i(1 - \rho_i) + \mu_i \rho_i + \mu_i(1 - \rho_i)t + \rho_i y_{i,t-1} + \varepsilon_{it}$$

where ε_{it} are independently distributed across units and over time. By setting $\xi_i = \mu_i$ we obtain

$$y_{it} = \mu_i + \mu_i(1 - \rho_i)t + \rho_i y_{i,t-1} + \varepsilon_{it}. \tag{4.16}$$

Specifying the data generating process as in (4.16) allows to evaluate the ability of the tests to reject the null hypothesis of a unit root against the trend stationary alternative where the rate of trend is the same under the null and alternative hypotheses.

The estimated model then becomes

$$\Delta y_{it} = a_i + \beta_i t + \varphi_i y_{i,t-1} + \varepsilon_{it}.$$

The preceding test statistics are calculated in exactly the same way as in the intercept only case, with the time trend variable included in both the forward (4.6) and reverse regressions (4.7) for the $\overline{Max_s}$ statistic and in the restricted (4.11), (4.13) and auxiliary regressions (4.12), (4.14) for the $\overline{Min_s}$ statistic. For the $\overline{WS_s}$ statistic the series y_{it} is detrended and $\breve{y}_{it}$ becomes $\breve{y}_{it} = y_{it} - \widehat{a}_i - \widehat{b}_i t$. The limiting distribution for all statistics can be shown as above to follow the standard normal distribution. Similarly so for the case of serially correlated errors. In such an event the corresponding augmented Dickey-Fuller regressions are adopted with $Max_i \equiv Max(ADF_{f_i}, ADF_{r_i})$ and $Min_i \equiv Min(LM_{f_i}, LM_{r_i})$, where LM_{f_i}, LM_{r_i} are adjusted as in footnote 9. Details on the derivation of the individual WS_i statistic in the case of lag-supplementation can be found in Appendix 3.A of chapter three. The values for the mean and variance of the Max_i , WS_i, Min_i, t_i and LM_i statistics depend on T and the lag order k and are computed via simulations from independent normal samples.[13] Tables 4.1(a) and 4.1(b) provide such values for $T = \{15, 25, 50, 75, 100\}$ and $k = \{1, 2, 3, 4\}$. The requirement that $N/T \to q$ as $T \to \infty$ and $N \to \infty$ is needed in order for the $\overline{Max_s}$, $\overline{WS_s}$ and $\overline{Min_s}$ tests to be distributed asymptotically as a standard normal variate.

4.4 Allowing a Common Time-Specific Component

The above results were derived under the assumption that the disturbances ε_{it} in (4.2) are independent across units, which implies that the error covariance matrix Σ_ε is diagonal. One way of relaxing this assumption common in the panel literature is to allow for the presence of a single aggregate common factor which has an identical impact on all the individuals in the panel. We can think of the error term in this case consisting of a stationary time-specific common effect v_t and an idiosyncratic random effect ε_{it}.

In such an event, (4.3) becomes

$$\begin{aligned} \Delta y_{it} &= a_i + \varphi_i y_{i,t-1} + u_{it}, \\ u_{it} &= v_t + \varepsilon_{it} \end{aligned} \tag{4.17}$$

[13]The mean and variance of the individual test statistics are computed using 50,000 replications under the data generating process given by $y_t = y_{t-1} + e_t$, $y_0 = 0$ $e_t \tilde{} N(0,1)$, $t = 1, 2, ..., T$. For every replication b we estimate the model $\Delta y_t = a + \varphi y_{t-1} + \sum_{j=1}^{k} \zeta_j \Delta y_{t-j} + v_t$, $t = k+2, k+3, ..., T$ and obtain the $Max_{t_i}^{(b)}$, $WS_i^{(b)}$, $Min_{LM_i}^{(b)}$, $t_i^{(b)}$ and $LM_i^{(b)}$ statistics for testing $\varphi = 0$. The quantities of interest are then calculated as $E(TS_i) = \frac{1}{b} \sum_{b=1}^{\#rep} TS_i^{(b)}$ and $Var(TS_i) = \frac{1}{b} \sum_{b=1}^{\#rep} (TS_i^{(b)} - E(TS_i))^2$, where TS_i is the corresponding test statistic. Similarly so for the case of a time trend, where a in the above regression equation is replaced by $a + \beta t$. In calculating the mean and variance of the test statistics Max_i, WS_i, Min_i, t_i and LM_i we have made use of the assumption that the error terms are independent across groups, and therefore these values will be the same across individuals. It should also be noted that our results are not directly comparable with IPS as their "T" would in effect be $T - k - 1$ in our notation.

where v_t is generated by some stationary process whose disturbance term is independent of ε_{it}, and ε_{it} itself is independently distributed across time and units.[14] The adverse effects of cross-section correlation of this type have been reported by O'Connell (1998), who finds that failure to control for such can lead to spurious rejection of the unit root null.

The typical procedure in removing the common component v_t in order to apply the panel unit root tests of the previous section that are valid when the disturbances are independent across units, involves subtracting the cross-section averages from the data. Then, equation (4.17) becomes

$$\Delta\widetilde{y}_{it} = \widetilde{a}_i + \varphi_i\widetilde{y}_{i,t-1} + \eta_{it} \tag{4.18}$$

where $\widetilde{y}_{it} = y_{it} - N^{-1}\sum_{j=1}^{N} y_{jt}$, $\widetilde{a}_i = a_i - N^{-1}\sum_{j=1}^{N} a_j$ and

$$\eta_{it} = N^{-1}\sum_{j=1}^{N}(\varphi_i - \varphi_j)y_{j,t-1} + (\varepsilon_{it} - N^{-1}\sum_{j=1}^{N}\varepsilon_{jt}). \tag{4.19}$$

Under the null of a unit root where $\varphi_i = 0\ \forall i$ it follows that

$$\eta_{it} = \varepsilon_{it} - N^{-1}\sum_{j=1}^{N}\varepsilon_{jt}.$$

It is apparent from the above that the act of cross-section demeaning introduces dependence across the errors η_{it} in the demeaning regressions (4.18) and thus Proposition 1 and Theorem 1 will not be applicable in this context. In the case however where the idiosyncratic effects, ε_{it}, are serially uncorrelated IPS prove that the LM-bar test is asymptotically distributed as standard normal under the null as $N \rightarrow \infty$.[15] In the same way, the $\overline{Max_s}$, $\overline{Min_s}$ and $\overline{WS_s}$ test statistics can be shown to follow asymptotically the standard normal distribution.

[14] If the variances across the units of the panel are about the same, i.e. $\sigma_i^2 \cong \sigma_j^2$, then it can easily be seen that the $(N \times N)$ covariance matrix $E(u_t u_t')$, where $u_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$, produces after standardization a correlation matrix whose off-diagonal terms are identical. Thus the cross-section units are correlated to the same degree.

[15] It is easy to show that for large N the off diagonal elements of the $(N \times N)$ covariance matrix $E(\eta_t\eta_t')$, where $\eta_t = (\eta_{1t}, \eta_{2t}, ..., \eta_{Nt})'$, are zero.

The panel unit root tests based on the demeaned regresssions, (4.18), will be robust to general but common patterns of residual serial correlation across the groups.[16] Moreover, it should be pointed out that if countries are groupwise highly correlated then the removal of the global mean from each series will do little to reduce the amount of cross-sectional dependence that is present.

Though restrictive in nature, cross-sectional dependence of the former type provides an interesting starting point for such an important issue regarding panel data before we follow up the more realistic case of cross-section correlation of a heterogeneous form.

In the simulation study that proceeds we investigate both the case of independent cross-section units and that of a common time-specific effect included in the model of interest. In the event of the latter, any effects induced by cross-section demeaning are expected to be uncovered. To make earlier results operational, we briefly describe the procedure involved in the computation of the test statistics. In the first instance the $ADF(k)$ regression is selected. In the presence of cross-section correlation of the type illustrated above we subtract the cross-section averages from the data in order to remove the common component. The Max_i, WS_i, Min_i, t_i and LM_i statistics for each group are then obtained and the group averages $\overline{Max}$, $\overline{WS}$, $\overline{Min}$, $\bar{t}$ and $\overline{LM}$ are computed as $\overline{Max} = N^{-1}\sum_{i=1}^{N} Max_i$, $\overline{WS} = N^{-1}\sum_{i=1}^{N} WS_i$, $\overline{Min} = N^{-1}\sum_{i=1}^{N} Min_{LM_i}$, $\bar{t} = N^{-1}\sum_{i=1}^{N} t_i$ and $\overline{LM} = N^{-1}\sum_{i=1}^{N} LM_i$, respectively. Subsequently, the group averages are standardized using the values of the mean and variance of the Max_i , WS_i, Min_i, t_i and LM_i statistics provided in Tables 4.1(a) and 4.1(b). The resulting statistics are then compared to the critical values of the standard normal distribution for one-sided tests.

4.5 Some Simulation Results

In this section we present the results of Monte Carlo simulations to judge the finite sample performance of the modified test statistics. Initially we investigate the case where the errors are independent across units, both in the presence and absence of serial correlation. Results then follow for the case of cross-section

[16] In simulations that follow it will be seen that any effect of cross-sectional dependence across the error terms induced by the act of demeaning is minimal for small N. The same applies when allowing for somewhat different patterns of residual serial correlation across the groups.

correlation of the type described in the previous section. To complement the tabulated simulation results, empirical power graphs of the tests are presented in the base case where the errors are independent both across units and across time. The performance of the LM-bar ($\overline{LM}_s$) and t-bar ($\overline{t}_s$) tests are reported for means of comparison and completeness. For all experiments we report both the size and the power of the test statistics for different combinations of N and T. A total of 5000 and 2000 iterations are used respectively in computing the empirical size and power of the tests, at $N = \{5, 10, 25, 50, 100\}$ and $T = \{15, 25, 50, 75, 100\}$. Nominal size for the simulation results was set at 0.05.

For the first set of results, where the errors are independent across units, we consider the following data generating process

$$y_{it} = (1-\rho_i)\mu_i + \rho_i y_{i,t-1} + \varepsilon_{it}; \qquad \varepsilon_{it} = \lambda_i \varepsilon_{i,t-1} + e_{it}; \qquad i = 1, ..., N; t = 1, ..., T \tag{4.20}$$

where ε_{it} is generated as a stationary process independently distributed across $i = 1, ..., N$ and $e_{it} \sim N(0, \sigma_i^2)$.

The remaining parameters that control for the effects of heterogeneity across groups and the serial correlation of the disturbance term are generated in the following manner once and for all at the initial stage of the experiments and are held fixed throughout

$$\mu_i \sim N(0,1)$$
$$\lambda_i \sim U(0.2, 0.4)$$
$$\sigma_i^2 \sim U(0.5, 1.5)$$

The null hypothesis considered involves testing

$$H_0: \ \rho_i = 1, \quad \text{for all } i$$

against

$$H_1: \ \rho_i < 1, \quad \text{for at least one } i$$

While the tests permit different values for ρ_i under the alternative, our power comparisons are based on $\rho_i = 0.9 \ \forall \ i$. To obtain samples which closely resemble stationarity under the alternative we generate $T + 200$ observations and discard the first 200. In this case the initialization becomes unimportant and we choose zeros for the initial values.

To compute the unit root test statistics we estimate the following ADF regression for each individual series

$$\Delta y_{it} = a_i + \varphi_i y_{i,t-1} + \sum_{j=1}^{k} \zeta_{ij} \Delta y_{i,t-j} + residual \tag{4.21}$$

We further allow for the presence of heterogeneous trends under the alternative, in which case the generating process is taken to be

$$y_{it} = \mu_i + (1-\rho_i)\mu_i t + \rho_i y_{i,t-1} + \varepsilon_{it}, \qquad i = 1, ..., N; t = 1, ..., T. \tag{4.22}$$

The corresponding estimated model is

$$\Delta y_{it} = a_i + \beta_i t + \varphi_i y_{i,t-1} + \sum_{j=1}^{k} \zeta_{ij} \Delta y_{i,t-j} + residual \tag{4.23}$$

For the second set of results that allow for a common time-specific component we consider the data generating process

$$y_{it} = (1-\rho_i)\mu_i + \rho_i y_{i,t-1} + u_{it}; \qquad u_{it} = v_t + \varepsilon_{it}; \qquad i = 1, ..., N; t = 1, ..., T \tag{4.24}$$

where v_t is the common time-specific effect and ε_{it} is the group-specific random component. The former is generated as the stationary process

$$v_t = 0.9 v_{t-1} + \omega_t, \qquad \omega_t \tilde{} N(0,1)$$

with ω_t independent of ε_{it}, while the rest of the parameters are generated in exactly the same manner as above. The main difference is that in computing the unit root test statistics the following *ADF* regression is run for each individual series with cross-section averages subtracted from the data, in order to remove the common time-specific effects v_t.

$$\Delta\widetilde{y}_{it} = \widetilde{a}_i + \varphi_i \widetilde{y}_{i,t-1} + \sum_{j=1}^{k} \widetilde{\zeta}_{ij} \Delta\widetilde{y}_{i,t-1} + residual \tag{4.25}$$

where $\widetilde{y}_{it} = y_{it} - N^{-1}\sum_{j=1}^{N} y_{jt}$, $\widetilde{a}_i = a_i - N^{-1}\sum_{j=1}^{N} a_j$, $\widetilde{\zeta}_{ij} = \zeta_{ij} - N^{-1}\sum_{q=1}^{N} \zeta_{qj}$.

When allowing for heterogeneous trends we consider the data generating process

$$y_{it} = \mu_i + (1-\rho_i)\mu_i t + \rho_i y_{i,t-1} + u_{it}, \qquad i = 1, ..., N; t = 1, ..., T \tag{4.26}$$

and the estimated model is

$$\Delta\widetilde{y}_{it} = \widetilde{a}_i + \widetilde{\beta}_i t + \varphi_i \widetilde{y}_{i,t-1} + \sum_{j=1}^{k} \widetilde{\zeta}_{ij} \Delta\widetilde{y}_{i,t-j} + residual \tag{4.27}$$

Tables 4.2(a) and 4.2(b) summarize the results for the case of independent cross-section units when fixed effects only and when deterministic trend components are included in the model, respectively. Similarly, the results in the presence of a common time-specific component are summarized in Tables 4.3(a)-4.3(b). The layout of the tables is the same for both cases. The size and power of the tests are reported as measured by the empirical frequency with which the null hypothesis is rejected using the nominal one sided 5% critical value of the standard normal distribution.[17]

To evaluate the size of the tests in the case where fixed effects only are included in the data generating process, we set $\rho_i = 1$ in (4.20) and (4.24). Similarly so for (4.22) and (4.26) when heterogeneous trends are included. In evaluating power, we set $\rho_i = 0.9 \; \forall \; i$ in the same models. Columns four and five in the tables give the size and power of the tests in the benchmark case where the errors, ε_{it}, are uncorrelated, which corresponds to setting $\lambda_i = 0$ in (4.20), (4.24) and (4.22), (4.26) and $k = 0$ in the estimated models (4.21), (4.25) and (4.23), (4.27) for the

[17]The upper 5% tail of the standard normal distribution is used for the LM based statistics, while the lower is used for the remaining tests.

case of a constant only and trend term, respectively. The ensuing columns report results in the presence of residual serial correlation for autoregressive lag lengths of $k = 1, 2, 3, 4$, in order to assess the sensitivity of the tests to overspecification.[18] We do not report the results for $k = 0$ as in IPS, since the findings corroborate what holds in the single series case namely, underspecification of the lag length relative to the DGP leads to very undersized tests.

We begin by commenting on the results that correspond to the simplest case where the errors are independent across units and a constant only is included in the model as displayed in Table 4.2(a). When the errors are serially uncorrelated results in column four under the heading $\lambda_i = 0$ and $k = 0$ show that for a given T and N all tests have empirical sizes reasonably close to nominal size. A noteworthy feature regarding the power results given in the adjacent column is the considerable power advantage enjoyed by the $\overline{WS}_s$, $\overline{Max}_s$ and $\overline{Min}_s$ tests relative to the $\bar{t}_s$ and $\overline{LM}_s$ tests. In general, for fixed T and as N increases all tests demonstrate higher power which justifies the use of panel data. The power increase due to an increase in the time-series dimension is larger than the corresponding increase in N as found in most simulation studies of this sort that deal with panel data. The overall performance of $\bar{t}_s$ is superior to the $\overline{LM}_s$ test in accordance with the results of IPS. Of the modified panel unit root tests, it is the $\overline{WS}_s$ test that outperforms all remaining tests in terms of power with the $\overline{Max}_s$ test very close in performance. In particular, for $N = 25$ and $T = 25$, the power of $\overline{WS}_s$ is 0.906 while, the powers of $\overline{Max}_s$, $\bar{t}_s$ and $\overline{Min}_s$ are 0.846, 0.519 and 0.735 respectively. The corresponding sizes are 0.054, 0.056, 0.055 and 0.060. Power gains of $\overline{WS}_s$ over $\bar{t}_s$ are on average in the range of 25%, and are particularly striking for moderate values of N.

A slightly different picture emerges when the errors are considered to be serially correlated as follows from the results in subsequent columns of Table 4.2(a). When the correct model $k = 1$ is selected, results show that in the case of small T or large T and small N the $\overline{LM}_s$ and $\overline{Min}_s$ tests tend to be slightly oversized. On the other hand, for small values of T and as N increases the $\overline{Max}_s$ and $\overline{WS}_s$ tests appear undersized, a phenomenon which is somewhat more discernible in the case of the former test. However, as T increases past the value of 25 correct

[18] We impose without loss of generality the same order of augmentation in the ADF regression for each group for simplicity, although this can easily be relaxed.

size fairly quickly restores for both tests. On the whole, results are in agreement with asymptotic theory's prediction that the tests should keep nominal size well when N is large. Increasing the number of lags in the ADF regressions produces mild size distortions for small T in the case of the LM based tests. However, for larger T adding extra lags has trivial impact on the size of the tests.

Regarding power, once again results highlight two statistics that consistently outperform the remaining, namely the $\overline{WS}_s$ and $\overline{Max}_s$ tests. The power of the $\overline{WS}_s$, $\overline{Max}_s$, and $\overline{Min}_s$ tests rises considerably relative to that of $\bar{t}_s$ and $\overline{LM}_s$ as the cross-sectional dimension of the panel rises. In particular, for $T = 25$ and $N = 25$ and when the estimated model is correct the powers of $\overline{WS}_s$, $\overline{Max}_s$ and $\overline{Min}_s$ are 0.760, 0.700 and 0.604 respectively, relative to that of 0.437 and 0.386 for $\bar{t}_s$ and $\overline{LM}_s$. Moreover, the test statistics suffer only a moderate reduction in empirical power at the five-percent level when the true lag order is overstated.

In general, the tests have better small sample properties when both N and T are of comparable size allowing us to confirm the relevance in finite samples of the requirement $N/T \rightarrow q$ as $N, T \rightarrow \infty$ that underpins their asymptotic validity.

When heterogeneous trends are included in the autoregressive models, results as illustrated in Table 4.2(b) appear to be qualitatively similar to the preceding fixed effects only case. However, once the deterministic component is elaborated to include a trend term there is an apparant decrease in the power of all tests, which is particularly noticeable for moderate values of T and N. A similar phenomenon is encountered with conventional unit root tests. Thus, power gains are now less substantial. An alternative feature worth noting is that all tests have almost correct size even for small T in the case of serially correlated residuals.

Turning to the case of a common time-specific component, Tables 4.3(a) and 4.3(b) show that a similar picture holds, qualitatively and quantitatively, as under the assumption of independent errors. To avoid repetition we merely point out the implication of such a result. The act of cross-sectional demeaning works well for small N. Any potential consequences induced by the non-diagonality of the covariance matrix of the residuals on the size and power of the tests for N as small as 5, appear to be trivial.

Empirical powers of the tests for various combinations of N and T are illustrated in Figure 1. They correspond to the benchmark case of no serial correlation present in the errors when a constant only is included in the model, complementing the results in Table 4.2(a).[19] Consistency of the tests shows up well in the graphs with power increasing monotonically with ρ for fixed N and T. Similarly, for fixed ρ power tends to unity as N and T tend to infinity. The superior power performance of the $\overline{WS}_s$ test is easily observed for all values of N and T, with the $\overline{Max}_s$ test appearing nearly as powerful. From the simulation results for alternative values of the autoregressive coefficient ρ_i not presented here, it follows that in general when the power of the $\overline{WS}_s$ test is in the range of 50% the power of the $\bar{t}_s$ and $\overline{LM}_s$ tests is 20% lower. This result indicates a substantial gain in mid-range power being achieved over the IPS tests and better overall power in distinguishing alternatives closer to the null.

To summarize our results, it turns out that whether in the absence or presence of cross-section correlation in the residuals the modified panel test statistics offer dramatic power gains over the standard t-bar and LM-bar tests. In particular, the $\overline{WS}_s$ test outperform the others in the sense that it does not present important size distortions and it demonstrates considerable power even when the true lag order is overstated. Thus, the weighted symmetric estimator as in the case of a single time series, when adopted in the panel framework it provides higher power than the t-bar test for models in which the alternative is a stationary process. The $\overline{Max}_s$ test displays power in the same range as $\overline{WS}_s$ and therefore little would be lost in applying one or the other, the former having the advantage that it is easier to compute.

Finally, it should be noted that although the alternative hypothesis is specified as at least one of the autoregressive parameters ρ_i less than one, our simulation study has been limited to the case where all the series are stationary under the alternative. Maddala and Wu (1999) and Karlsson and Löthgren (1999) investigate the behaviour of the LL and IPS tests when a subset of the series is stationary and the remainder have unit roots. The latter point out the potential risk for small T panels, of the whole panel being erroneously modelled as non-stationary due to the relatively low power of the tests even for large proportions of stationary series in the panel. We would expect the adoption of the above modified panel unit root statistics to reduce the risk involved in such a situation, benefiting from their increased power properties.

[19] We omit the case $N = 5$ as no interesting power differences are discernable.

4.6 Bootstrap Tests with Application to Purchasing Power Parity

As recently noted in the literature, cross-sectional correlation of a more complicated nature than that accounted for by common time-specific components is often encountered in many panel data sets in macroeconomics, international economics and finance.[20] Maddala and Wu (1999) report substantial size distortions for the LL, IPS and Fisher tests in such an event and resort to bootstrap methods to derive the empirical distribution of the unit root test statistics. Following these authors, we adopt the bootstrap in an application of the proposed panel unit root tests to real exchange rate data.[21] We anticipate the use of the more powerful tests to point to sharper conclusions regarding the stationarity of the particular time series.

We use quarterly data extracted from the OECD Main Economic Indicators over the period 1973:I-1998:IV following the collapse of the fixed exchange rate system.[22] Nominal exchange rates and consumer prices have been selected for 18 OECD countries namely Australia, New Zealand, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, UK and the United States. All variables are measured in

[20] An alternative multivariate test for unit roots has been adopted in the literature by Abuaf and Jorion (1990) and Taylor and Sarno (1998), which accommodates an arbitrary pattern of contemporaneous cross-sectional dependence. The so-called SUR test is based on systems estimation of autoregressive processes with each cross-section unit constituting an equation, forming in this way a system of seemingly unrelated regression equations

$$y_{it} = \mu_i + \rho_i y_{it-1} + \varepsilon_{it}, \qquad i = 1,...,N;\ t = 1,...,T$$

where $E[\varepsilon_i \varepsilon_i']=\Omega$ and $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})$. Estimates of the parameters are obtained by iterating the feasible GLS procedure, where in every iteration the residuals of the seperate regressions are used to update the elements of the covariance matrix Ω. However, the estimators can be difficult to compute as they require matrix inversions whose orders can be quite large depending on the number of individual units. Therefore, it is only of value when T is large and N is small. Empirical applications typically limit the number of cross section units N, in order to achieve computational feasibility.

[21] The bootstrap method to a large extent is employed to improve inferences for estimated models for which available asymptotic theory may be expected to be unreliable. Indeed considerable improvement over the asymptotic results is possible when the underlying statistics are asymptotically pivotal (see Beran, 1988). In the present context the method is appropriate as the distribution of the test statistics in the limit is not invariant to the covariance matrix of the error terms. The method rests on the idea of using a single data set in a similar to the Monte Carlo set up, in which the data themselves are used to approximate random quantities of interest in the model. For a detailed outline of the bootstrap see *inter alia* Efron and Tibshirani (1993) and Veall (1998).

[22] We investigate the period up until the introduction of the single currency (January 1, 1999) as from then on the countries participating in the Monetary Union, a number of which are included in the panel we investigate, lock their currencies together at a fixed rate.

natural logarithms. We calculate the real exchange rate of each country relative to the US dollar according to $s_i + p_{us} - p_i$, where s_i is the logarithm of the nominal exchange rate against the US dollar, p_i denotes the logarithm of the consumer price index in country i and p_{us} denotes the logarithm of the consumer price index for the United States, obtaining in this way 17 bilateral real exchange rates as the variables of interest.

A graphical inspection of the real exchange rates for several countries as shown in Figure 4.2 reveals a common structural break occurring around 1985.[23] While this is the case for the majority of the countries, for expositional clarity we restrict ourselves to the illustration of only a subset. As our main purpose is in demonstrating the superiority of the modified panel unit root tests that are not specifically designed to account for structural breaks, we would like to avoid plaguing our results with the potential adverse effects of what looks like a severe break. For this reason our investigation covers the period 1987:IV-1998:IV, with a total of 45 observations, allowing for the foreign exchange market to absorb the repercussions of that year's event. Thus, in the notation previously employed, we have $N = 17$ and $T = 45$. The null hypothesis of interest is that the generating process for all y_{it}, $i = 1, 2, ..., N$ contain a unit autoregressive root. Since stationarity around any linear trend with non-zero slope would not be consistent with the economic content of the relative purchasing power parity hypothesis, our test equations include a constant but no slope term, corresponding to part (a) of Tables 4.1-4.3.

We begin the analysis by fitting to each individual series ADF-style regressions

$$\Delta y_{it} = a_i + \varphi_i y_{i,t-1} + \sum_{j=1}^{k_i} \zeta_{ij} \Delta y_{i,t-1} + e_{it} \tag{4.28}$$

where different orders k_i are permitted for each series, and these were chosen according to the general-to-specific testing procedure at the 10% level of significance. Table 4.4 illustrates the correlation matrix of the ADF residuals across countries. While the variances of the generated residuals $\widehat{e}_{it}$ appeared to differ

[23] Within the history of flexible exchange rates such a date corresponds to the intervention of the large industrial countries (US, Japan, Germany, France and UK) in the foreign exchange market in an attempt to bring the dollar down, which was agreed to have risen too much. This endeavour was subsequently achieved.

only negligibly, their cross correlations are far from equal. This latter finding is unsurprising as some subsets of the countries in the sample retained over the period relatively strong currency linkages, as for example some of the members of the European Union.[24] Thus, while cross-section correlation exists, following from inclusion of the common US benchmark, it appears not to be of the special type that can be accounted for using the procedures of Section 4.4. We employ instead a bootstrap approach, the steps of which are described below.

1. Determine the order (k_i) of the autoregressive process for each cross-section unit. For this purpose we adopt the general-to-specific approach which is essentially a recursive t-statistic procedure that tests the significance of the last coefficient based on the 10% value of the asymptotic normal distribution. As discussed by Campbell and Perron (1991) and Ng and Perron (1995), this procedure has better size and power properties than alternative information-based model selection methods.

2. Estimate the residuals from the following regression

$$\Delta y_{it} = \sum_{j=1}^{k_i} \zeta_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \qquad t = k_i + 2, ..., T \; ; \; i = 1, ..., N \tag{4.29}$$

where k_i is the lag order of the individual cross-section unit determined above.[25]

3. Recenter the estimated residuals $\widehat{\varepsilon}_t = (\widehat{\varepsilon}_{1t}, \widehat{\varepsilon}_{2t}, ..., \widehat{\varepsilon}_{Nt})'$ from (4.29) in the following manner

$$\widetilde{\varepsilon}_t = \widehat{\varepsilon}_t - (T - k - 1)^{-1} \sum_{t=k+2}^{T} \widehat{\varepsilon}_t, \quad t = k + 2, ..., T$$

where $\widetilde{\varepsilon}_t = (\widetilde{\varepsilon}_{1t}, \widetilde{\varepsilon}_{2t}, ..., \widetilde{\varepsilon}_{Nt})'$ and $k = \max_i(k_i)$. This ensures that the mean of the resulting residuals remains zero as a constant term is not included in (4.29), avoiding possible distortions of the bootstrap estimates (see Stine, 1987 and Berkowitz

[24]One might for example expect more correlation for France and Germany than, for example, Canada and Japan.

[25]We estimate the regression by least-squares as is commonly the case. Fleissig and Strauss (1999) adopt iterative SUR as a more efficient procedure in the presence of contemporaneous correlation. However, we believe that the computational burden involved in SUR estimation particularly when $N \geq 5$ will outweigh the any efficiency gains achieved, thus not making the use of such an approach worthwhile.

and Kilian, 2000).

4. Generate bootstrap innovations $\varepsilon_t^* = (\varepsilon_{1t}^*, \varepsilon_{2t}^*, ..., \varepsilon_{Nt}^*)'$ by resampling with replacement from the empirical residuals $\widetilde{\varepsilon}_t$ keeping the cross-section index fixed in order to maintain the correlations within each cross-section unit. This ensures that $E^*[\varepsilon_t^* \varepsilon_t^{*\prime}] = \widetilde{\Sigma}$, where E^* denotes the expectation under bootstrap sampling, that is, the expectation relative to the empirical distribution of the estimation data and $\widetilde{\Sigma}$ is the $(N \times N)$ estimated covariance matrix of the recentered residuals, while $E^*[\varepsilon_t^*] = 0$.

5. Generate the bootstrap sample y_{it}^* as

$$y_{it}^* = y_{i,t-1}^* + u_{it}^*, \quad y_{i0}^* = 0, \quad i = 1, ..., N; t = 1, ..., T$$

with

$$u_{it}^* = \sum_{j=1}^{k_i} \widehat{\zeta}_{ij} u_{i,t-j}^* + \varepsilon_{it}^*, \quad t = 1, ..., T$$

initialising by setting

$$u_{it}^{0*} = \sum_{s=0}^{m} \widehat{\psi}_{is} \varepsilon_{i,t-s}^*, \quad t = -(k_i - 1), ..., -1, 0 \tag{4.30}$$

where $\widehat{\zeta}_{ij}$ are from the estimation results of (4.29).[26]

6. Apply the panel unit root tests of Section 4.3 to the bootstrap sample by estimating the model

$$\Delta y_{it}^* = \delta_i + \varphi_i y_{i,t-1}^* + \sum_{j=1}^{k_i} \vartheta_{ij} \Delta y_{i,t-j}^* + residual \tag{4.31}$$

and repeat the above steps 3-4 many times to obtain the p-values of the tests.[27]

[26]The bootstrap samples y_{it}^* are generated with the unit root imposed. Basawa *et al.* (1991) show that if this is not the case the samples will not necessarily have the unit root property rendering the bootstrap procedure inconsistent.

[27]The p-values are calculated as $\#(TS^* \geqslant \text{observed value of test})/B$ and $\#(TS^* \leq \text{observed value of test})/B$ for the LM based tests and the remaining tests respectively, where TS^* is the corresponding boostrap test statistic for testing $\varphi_i = 0$ in (4.31) and B is the number of bootstrap iterations. The 'observed value of the test' is the value of the test statistic from regression (4.28) computed for the given data series.

The starting values for u^*_{it} in step 5 are built up from the estimated $MA(\infty)$ representation, $u^*_{it} = \widehat{\Psi}(L)\varepsilon^*_{it}$, of the stationary $AR(k_i)$ process in equation (4.30).[28] To see how the estimated weights $\widehat{\Psi}(L)$ are obtained we write equation (4.30) in terms of lag operators as $(1 - \widehat{\zeta}_{i1}L - \widehat{\zeta}_{i2}L^2 - ... - \widehat{\zeta}_{ik_i}L^{k_i})u^*_{it} = \varepsilon^*_{it}$. Omitting the index i for simplicity, it then follows that

$$\widehat{\Psi}(L) = \widehat{\psi}_0 + \widehat{\psi}_1 L + \widehat{\psi}_2 L^2 + ... = 1/(1 - \widehat{\zeta}_1 L - \widehat{\zeta}_2 L^2 - ... - \widehat{\zeta}_k L^k)$$

or

$$(1 - \widehat{\zeta}_1 L - \widehat{\zeta}_2 L^2 - ... - \widehat{\zeta}_k L^k)(\psi_0 + \psi_1 L + \psi_2 L^2 + ...) = 1$$

Writing out explicitly the above equation leads to a recursive algorithm for generating the $MA(\infty)$ weights $\widehat{\psi}_0$, $\widehat{\psi}_1$, $\widehat{\psi}_{2,}$... for each individual cross section, where $\widehat{\psi}_0 = 1$. To make such an exercise operational the infinite sum is truncated with m in equation (4.30) representing the truncation value. While there is no guide in the literature as to the appropriate value of m, Rayner (1990) in bootstrapping p-values in the stationary first-order autoregressive model, chooses m to ensure that the starting values are sufficiently close to their true distributions. Experimenting with various values of m, we find no significant effect of the choice of this parameter on the reported results.

Furthermore, the initialisation of u^*_{it} in step 5 involves generating in every bootstrap replication a new sample of $(m+1)$ bootstrap innovations ε^*_{it} from the empirical residuals $\widetilde{\varepsilon}_{it}$ for each individual i. For every additional lag in (4.30) a further initial value is required decreasing the $(m+1)$ sample by one. As a result, an extra bootstrap innovation is generated each time and appended to the sample keeping it of a fixed size $(m+1)$ for any given lag and corresponding initial value.

We applied the bootstrap approach to the real exchange rate data, employing 5000 bootstrap replications, and computed p-values for the tests of the null hypothesis that the generating process for every series contains a unit root. The results are shown in Table 4.5. Row three reports the results corresponding to a truncation value of 30. We experimented with a number of alternative truncation

[28] Alternative methods of obtaining the initial values are mentioned in Berkowitz and Kilian (2000).

values, the results remaining virtually unchanged even for a value of 100 as seen in row four of the same table. Considering the number of observations available in our dataset, the maximum number of lags was set to $k_{\max} = 4$, although the results did not seem to be sensitive to a moderate increase in such value. Results indicate that by using the modified bootstrap panel unit root tests $\overline{Max}^*$, $\overline{WS}^*$ and $\overline{Min}^*$, the null hypothesis of a unit root in the real exchange rates for the set of countries investigated is strongly rejected at the 5% significance level. The conventional bootstrap t-bar and LM-bar tests, that is $\bar{t}^*$ and $\overline{LM}^*$, on the other hand fail to reject the null at the 5% level.

It is tempting particularly given the results from more straightforward settings of Section 4.5, to conclude that the results of Table 4.5 simply reflect continued power advantages of the modified statistics in a bootstrap setting. In the ensuing section, we investigate through simulation whether the modified bootstrap tests do indeed possess additional power while retaining size reliability.

4.7 Monte Carlo Results with Bootstrap Critical Values

A set of Monte Carlo simulations follow in which the objects of interest are the bootstrap panel unit root tests. We seek to evaluate their finite sample size and power performance.

The simulation experiments are based on data generated as in (4.20), that is

$$y_{it} = \mu_i(\rho_i - 1) + \rho_i y_{i,t-1} + \varepsilon_{it} \tag{4.32}$$

where ε_{it} follows an $AR(1)$ process

$$\varepsilon_{it} = \lambda_i \varepsilon_{i,t-1} + e_{it}$$

with $\mu_i \tilde{} N(0,1)$ and the AR coefficients λ_i's $\tilde{}$ $U(0.2, 0.4)$. The only difference is that now the innovations e_{it} are drawn from an N-dimensional multivariate normal distribution with mean zero and covariance matrix Σ, $e_{it} \tilde{} N(0, \Sigma)$. The parameter values for the $(N \times N)$ covariance matrix $\Sigma = (\sigma_{ij})$ are drawn randomly, though in such a way as to ensure that Σ is a symmetric positive definite matrix so as to avoid the problem of near singularity. We briefly describe the procedure in generating Σ below.

1. Generate an $(N \times N)$ matrix $P = (p_{ij})$, $i = 1, ..., N; j = 1, ..., N$ where p_{ij} are drawn from the uniform distribution $U(0,1)$.

2. Construct from P an orthogonal matrix $X = P(P'P)^{-1/2}$.[29]

3. Generate a set of N eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$ where $\lambda_1 = r$ and $\lambda_N = 1$ and the intermediate eigenvalues are generated from the uniform distribution $U(r, 1)$.[30]

4. Create a diagonal matrix Λ with the eigenvalues on the diagonal.

5. Obtain the covariance matrix Σ as the spectral representation $\Sigma = X\Lambda X'$.[31]

As before, the parameters of the data generating process were generated once and for all at the outset. The setup of the above experiment is designed to replicate the 'heterogenous' residual cross-section correlation structure displayed by the exchange rate series as reported in the previous section. We derive the empirical size and power of the tests for different combinations of N and T, specifically $N = \{5, 10, 25\}$ and $T = \{25, 50\}$ using 1000 Monte Carlo replications and 1000 bootstrap samples. Each bootstrap sample is generated following the steps as described earlier using the now simulated data as input. As in the previous simulations, we report the results for autoregressive lag lengths of $k = 1, 2, 3, 4$ so as to assess the sensitivity of the tests to overspecification. The lag order of the autoregressive process is specified a priori.

In analyzing size we set $\rho_i = 1$ in (4.32), and in calculating power we set

---

[29]The orthogonal matrix X arises as follows. We begin with the symmetric matrix $P'P$ which we denote Ω. Since Ω is symmetric its characteristic roots are real. Providing they are also positive (i.e Ω is a positive definite matrix) there exists a nonsingular matrix L such that $\Omega = LL'$, where L which can be considered as the "square root" of Ω is a lower triangular matrix with positive diagonal elements. It follows that $L^{-1}\Omega L^{-1'} = I$ and substituting L with $\Omega^{1/2}$ we obtain $(\Omega^{1/2})^{-1}\Omega(\Omega^{1/2})^{-1'} = I$. Replacing Ω with $P'P$ we then get $(P'P)^{-1/2}P'P(P'P)^{-1/2'} = (P'P)^{-1/2}P'P(P'P)^{-1/2} = I$. From this last expression it readily follows that $P(P'P)^{-1/2}$ constitutes an orthogonal matrix.

[30]In generating a positive definite matrix it is imperative that the corresponding eigenvalues be positive, while their scale is irrelevant. We therefore set $\lambda_N = 1$ without loss of generality. For the parameter r we experimented with a variety of values, those close to one rendering the covariance matrix almost spherical, while those close to zero almost singular. In any case, the results were quanitatively similar and r was set to 0.1 throughout.

[31]If the orthogonal matrix X is the matrix of eigenvectors, one for each eigenvalue $\lambda_1, \lambda_2, ..., \lambda_N$, corresponding to the unknown matrix Σ then X diagonalises Σ, i.e. $X'\Sigma X = \Lambda$, where Λ is a diagonal matrix with the eigenvalues of Σ on the diagonal.

$\rho_i = 0.8 \; \forall i$. Table 4.6 reports simulation results for the bootstrap test statistics $\overline{WS}^*$, $\overline{Max}^*$, $\overline{Min}^*$, $\bar{t}^*$ and $\overline{LM}^*$. The setup is as in Tables 4.2, 4.3, except that the underspecified case $k = 0$ is not included here. The correct lag specification, $ADF(1)$, is taken as the most favourable situation for the tests with respect to dynamic specification. Results indicate that in general the bootstrap test statistics maintain correct size when N is small for all T. However, as N increases there is some tendency for the tests to be a little under-sized more so for smaller T, but generally confirm our findings in more straightforward cases. In most cases including additional lags in the model improves the size of the tests.

When examining the power results, what stands out once again at first glance are the dramatic power gains offered this time by the modified bootstrap tests over the bootstrap t-bar and LM-bar tests. For a correctly specified model these are in the range of 25%. The bootstrap test based on the weighted symmetric estimator emerges as the most powerful of them all, albeit little there is to choose between this test and the $\overline{Max}^*$ test. The bootstrap tests are well-behaved in that their power tends to unity for fixed ρ, as N and T tend to infinity. The increase in power owing to the time dimension is greater than the corresponding increase in N. Furthermore, including additional lags in the model does not result in a sharp reduction in test power. In short, the modified bootstrap tests are a great deal more powerful than the unmodified tests.

These results strongly support our conjecture that the stronger rejections obtained in Table 4.5 for the modified tests are a reflection of additional power in those tests. They supplement and strengthen findings in Oh (1996), Wu (1996) and Papell (1997).

4.8 Conclusions

Increasingly, panel data, often consisting of relatively short series, are available for analysis, and an issue of considerable interest concerns the reversion or otherwise of series to a fixed mean or trend. This issue can be addressed through an extension of commonly applied unit root tests, such as the Dickey-Fuller test. However, it has become well recognised in the basic time series context that more powerful modifications of commonly applied unit root tests are available. We have seen in this chapter that such increased power persists when two simple modification principles are applied to panel data unit root tests. To this end, modified panel unit root tests were introduced and the resultant test statistics were shown to follow asymptotically a standard normal distribution as is customary in the panel context.

An extensive Monte Carlo simulation study was conducted to investigate the small sample properties of the proposed panel unit root tests, when the residuals were assumed independent across units as well as in the presence of a common time-specific component. Results were fairly similar in both cases, qualitatively and quantitatively. In particular, the size of all tests appeared in general to be well behaved apart from the LM based tests which were found to be slightly oversized for small T, mainly when the true lag was overstated. In the same situation, the $\overline{Max}_s$ and $\overline{WS}_s$ tests appeared to be somewhat undersized for small T and increasing N. However, this phenomenon fairly quickly restored with increasing T. The most striking finding were the substantial power gains offered by the modified tests over the unmodified tests. In estimating autoregressive models with additional lags the tests were found to suffer only a moderate reduction in test power.

On the whole, considering the trade-off between size and power, the $\overline{WS}_s$ test was found to outperform the other tests in that it did not show important size distortions while it demonstrated considerable power even in the case when the lag order was overstated. Little should be lost in applying the $\overline{Max}_s$ over the $\overline{WS}_s$ test in empirical applications as their power differences are very small and the former has the advantage of being easier to compute. When a linear time trend was included in the model the power gains of the modified tests were less pronounced, as in the case of conventional unit root tests for a single time series. Nevertheless, these gains remain worthwhile as and both N and T are required to be sufficiently large for reliable inference to be conducted. Similar results followed, qualitatively and quantitatively, when a common time-specific component was allowed for in the model.

As an illustration, the modified panel unit root tests were applied to the real exchange rates of 17 OECD countries using the bootstrap method to accommodate the heterogeneous nature of cross-section correlation that appeared in the data. Results showed evidence against the unit root null when the modified panel unit root tests were employed, while the standard t-bar and LM-bar statistics failed to reject the null at the 5% level. The finite sample performance of the bootstrap tests was investigated through simulation. The tests appeared moderately undersized as the number of cross-section units N increased, a phenomenon which subsided with increasing T. An interesting finding was that the modified tests retained the significant power gains over the standard t-bar and LM-bar tests demonstrated in the simpler cases. The Monte Carlo evidence on the modified test statistics provided further support in favour of the results derived from the empirical application, where the unit root hypothesis was found to be rejected for the real exchange rate series for the period under investigation.

References

Abuaf, N., Jorion, P., 1990. Purchasing power parity in the long run. *Journal of Finance*, 45, 157-145.

Banerjee, A., 1999. Panel data unit roots and cointegration: An overview. *Oxford Bulletin of Economics and Statistics*, 61, 607-630.

Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H., Taylor, R.L., 1991. Bootstrapping unstable first-order autoregressive processes. *Annals of Statistics,* 19, 1098-1101.

Beran, R., 1988. Prepivoting test statistics: A bootstrap view of asymptotic refinements. *Journal of American Statistical Association*, 83, 687-697.

Berkowitz, J., Kilian, L., 2000. Recent developments in bootstrapping time series. *Econometric Reviews,* 19, 1-48.

Bhargava, A., 1986. On the theory of testing for unit roots in observed time series. *Review of Economic Studies*, 53, 369-84.

Boumahdi, R., Thomas, A., 1991. Testing for unit roots using panel data: Application to the French stock market efficiency. *Economics Letters*, 37, 77-79.

Campbell, J., Perron, P., 1991. Pitfalls and opportunities: What macroeconomists should know about unit roots. NBER Macroeconomics Annual, 141-201.

Choi, I., 2001. Unit root tests for panel data. *Journal of International Money and Finance*, 20, 249-272.

Coakley, J., Fuertes, A.M., 1997. New panel unit root tests of PPP. *Economics Letters*, 57, 17-22.

Culver, S.E., Papell, D.H., 1997. Is there a unit root in the inflation rate? Evidence from sequential break and panel data models. *Journal of Applied Econometrics*, 12, 435-444.

Efron, B., Tibshirani, R., 1990. *An introduction to the bootstrap.* Chapman and Hall, New York.

Fisher, R.A., 1932. Statistical methods for research workers, 4th edition, Oliver and Boyd, Edinburgh.

Fleissig, A.R., Strauss, J., 1997. Unit root tests on real wage panel data for the G-7. *Economics Letters,* 56, 149-155.

Fleissig, A.R., Strauss, J., 1999. Is OECD real per capita GDP trend or difference stationary? Evidence from panel unit root tests. *Journal of Macroecono-*

mics, 21, 489-506.

Fleissig, A.R., Strauss, J., 2000. Panel unit root tests of purchasing power parity for price indices. *Journal of International Money and Finance*, 19, 673-690.

Frankel, J.A., Rose, A.K., 1996. A panel project on purchasing power parity: Mean reversion within and between countries. *Journal of International Economics*, 40, 209-224.

Harris, R.D.F., Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics*, 91, 201-226.

Higgins, M., Zakrajšek, E., 1999. Purchasing power parity: Three stakes through the heart of the unit root null. Working Paper, Federal Reserve Bank of New York.

Im, K.S., Pesaran, M.H., Shin, Y., 1997. Testing for unit roots in heterogeneous panels. Working Paper, Department of Applied Economics, University of Cambridge.

Jorion, P., Sweeney, R.J., 1996. Mean reversion in real exchange rates: evidence and implications for forecasting. *Journal of International Money and Finance*, 15, 535-550.

Karlsson, S., Löthgren, M., 1999. On the power and interpretation of panel unit root tests. Working Paper, Stockholm School of Economics.

Lee, H.Y., Wu, J.L., 2001. Mean reversion of inflation rates: Evidence from 13 OECD countries. *Journal of Macroeconomics*, 23, 477-487.

Levin, A., Lin, C.F., 1992. Unit root tests in panel data: Asymptotic and finite sample properties, Discussion Paper, Department of Economics, University of California San Diego.

Leybourne, S.J., 1995. Testing for unit roots using forward and reverse Dickey-Fuller Regressions. *Oxford Bulletin of Economics and Statistics*, 57, 559-571.

MacDonald, R., 1996. Panel unit root tests and real exchange rates. *Economics Letters*, 50, 7-11.

Maddala, G.S., Wu, S., 1999. A comparative study of unit root tests with panel data and a new simple approach. *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

Ng, S., Perron, P., 1995. Unit root tests in ARMA models with data dependent methods of for the selection of the truncation lag. *Journal of the American*

Statistical Association, 90, 268-281.

O'Connell, P.G.J., 1998. The overvaluation of purchasing power parity. *Journal of International Economics*, 44, 1-19.

Oh, K.Y., 1996. Purchasing power parity and unit root tests using panel data. *Journal of International, Money and Finance*, 15, 405-418.

Pantula, S.G., Gonzalez-Farias, G., Fuller, W.A., 1994. A comparison of unit-root test criteria. *Journal of Business & Economic Statistics*, 12, 449-459.

Papell, D.H., 1997. Searching for stationarity: Purchasing power parity under the current float. *Journal of International Economics*, 43, 313-332.

Park, H.J., Fuller, W.A., 1995. Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis*, 16, 415-429.

Pesaran, M.H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68, 79-113.

Phillips, P.C.B., Moon, H.R., 2000. Linear regression limit theory for nonstationary panel data. *Econometrica*, 67, 1057-1111.

Quah, D., 1992. International patterns of growth: I. persistency in cross-country disparities. Unpublished manuscript, London School of Economics, London.

Quah, D., 1994. Exploiting cross-section variation for unit root inference in dynamic data. *Economics Letters*, 44, 9-19.

Rayner, R.K., 1990. Bootstrapping p-values and power in the first order autoregression: A Monte Carlo investigation. *Journal of Business & Economic Statistics*, 8, 251-263.

Schmidt, P., Phillips, P.C.B., 1992. LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, 54, 257-287.

Solo, V., 1984. The order of differencing in ARIMA models. *Journal of the American Statistical Association*, 79, 916-921.

Stine, R.A., 1987. Estimating properties of autoregressive forecasts. *Journal of the American Statistical Association*, 82, 1072-1078.

Taylor, M.P., Sarno, L., 1998. The behavior of real exchange rates during the post-Bretton Woods period. *Journal of International Economics*, 46, 281-312.

Veall, M.R., 1998. Applications of the bootstrap in econometrics and economic statistics, in *Handbook of applied economic statistics* (eds.) Ullah, A., Giles, D.E.A., New York, Dekker, pp. 421-441.

Wu, Y., 1996. Are real exchange rates nonstationary? Evidence from a panel-data test. *Journal of Money, Credit and Banking*, 28, 54-63.

Wu, J.L., Chen, S.L., 1998. A re-examination of real interest rate parity. *Canadian Journal of Economics*, 31, 837-851.

Appendix 4.A Mathematical Proofs

Proof of Proposition 4.3.1. Since ε_{it} is assumed independent and identically distributed with mean zero and finite heterogeneous variance σ_i^2, DF_{f_i} is independent and identically distributed across $i = 1, ..., N$. Similarly so, DF_{r_i} is *iid* across $i = 1, ..., N$ under the same assumption for ε_{it}, which implies that $Max_i \equiv Max(DF_{f_i}, DF_{r_i})$ is also *iid*. Therefore, we only need to show that $E(Max_i^2) < \infty$ in order to apply the Lindeberg-Levy CLT. If we assume that the second moment of DF_{f_i} exists following IPS, then it in turn implies that the second moment of DF_{r_i} also exists under the null hypothesis that $\rho_i = 1$. This is because the reverse series z_{it} is the same kind of pure random walk as y_{it} under the null and DF_{r_i} is based on the same type of regression as DF_{f_i} using z_{it} instead of y_{it}. Then, we have that

$$|Max_i| \leq |DF_{f_i}| + |DF_{r_i}|$$

which implies

$$Max_i^2 \leq (|DF_{f_i}| + |DF_{r_i}|)^2$$

and so

$$E[Max_i^2] \leq E[(|DF_{f_i}| + |DF_{r_i}|)^2].$$

By Minkowski's inequality

$$(E[(|DF_{f_i}| + |DF_{r_i}|)^2])^{1/2} \leq (E[DF_{f_i}^2])^{1/2} + (E[DF_{r_i}^2])^{1/2}$$

Taking squares on both sides of the above inequality

$$E[(|DF_{f_i}| + |DF_{r_i}|)^2] \leq ((E[DF_{f_i}^2])^{1/2} + E([DF_{r_i}^2])^{1/2})^2$$

and thus

$$E(Max_i^2) \leq ((E[DF_{f_i}^2])^{1/2} + (E[DF_{r_i}^2])^{1/2})^2 < \infty.$$

Then the result follows from the Lindeberg-Levy CLT. ∎

Proof of Theorem 4.3.1. It is shown in IPS (1997) that LM_{f_i} can be expressed as a function of ε_i/σ_i, which under the assumption that ε_i/σ_i is independent and identically distributed across $i = 1, 2, ..., N$ implies that LM_{f_i} is also *iid.* In a similar manner, LM_{r_i} is *iid* under the same assumption, which implies that $Min_i \equiv Min(LM_{f_i}, LM_{r_i})$ is also *iid.* Hence, as in the proof of Proposition 1 we only need to show that $E(|LM_{f_i}|^2) < \infty$. Since $R^2_{f_i}$ is the square of a correlation coefficient bounded between 0 and 1, the second moment of $LM_{f_i} = TR^2_{f_i}$ exists for any finite T. By the same reasoning, it can be shown that the second moment of LM_{r_i} also exists for any finite T. Then, since

$$Min_i \leq LM_{f_i} + LM_{r_i}$$

we have as above by using Minkowski's inequality that

$$E(Min_i^2) \leq ((E[LM^2_{f_i}])^{1/2} + (E[LM^2_{r_i}])^{1/2})^2 < \infty$$

which completes the proof. ■

Appendix 4.B Tables and Figures

Table 4.1(a)

Means and variances of unit root statistics under the null hypothesis, constant only case

	k	0		1		2		3		4	
T		mean	var	mean	var	mean	var	mean	var	mean	var
15	Min_i	1.709	2.926	1.966	3.851	2.019	4.323	2.476	6.726	2.999	9.476
	LM_i	2.731	4.351	3.121	5.541	3.367	6.733	3.977	9.478	4.717	12.857
	Max_i	-1.019	0.802	-0.969	0.935	-0.753	1.031	-0.687	1.309	-0.446	1.638
	t_i	-1.514	0.933	-1.497	1.065	-1.358	1.219	-1.317	1.539	-1.166	2.080
	WS_i	-1.253	0.821	-1.297	0.895	-1.193	0.860	-1.242	0.976	-1.134	0.922
25	Min_i	1.765	3.338	1.911	3.886	1.898	4.042	2.107	4.953	2.159	5.505
	LM_i	2.834	5.188	3.070	6.009	3.188	6.698	3.493	7.961	3.689	9.247
	Max_i	-1.054	0.734	-1.023	0.797	-0.901	0.845	-0.874	0.918	-0.750	0.976
	t_i	-1.516	0.822	-1.503	0.880	-1.428	0.940	-1.408	1.010	-1.325	1.089
	WS_i	-1.211	0.794	-1.231	0.819	-1.163	0.817	-1.197	0.843	-1.140	0.831
50	Min_i	1.803	3.654	1.873	3.951	1.862	3.977	1.954	4.393	1.952	4.459
	LM_i	2.941	5.952	3.064	6.459	3.109	6.764	3.258	7.330	3.309	7.763
	Max_i	-1.073	0.690	-1.059	0.717	-1.002	0.742	-0.992	0.774	-0.936	0.795
	t_i	-1.526	0.754	-1.520	0.779	-1.481	0.810	-1.480	0.827	-1.435	0.866
	WS_i	-1.181	0.770	-1.188	0.779	-1.155	0.783	-1.166	0.797	-1.139	0.794
75	Min_i	1.827	3.806	1.872	3.980	1.856	3.986	1.920	4.258	1.915	4.273
	LM_i	2.979	6.224	3.062	6.535	3.085	6.777	3.187	7.285	3.196	7.345
	Max_i	-1.083	0.681	-1.073	0.697	-1.036	0.706	-1.031	0.729	-0.993	0.746
	t_i	-1.529	0.735	-1.526	0.747	-1.499	0.768	-1.500	0.783	-1.465	0.802
	WS_i	-1.172	0.769	-1.175	0.773	-1.154	0.774	-1.164	0.784	-1.139	0.790
100	Min_i	1.824	3.832	1.860	3.994	1.866	4.066	1.909	4.228	1.904	4.277
	LM_i	2.988	6.424	3.052	6.733	3.092	6.927	3.162	7.196	3.179	7.321
	Max_i	-1.082	0.673	-1.074	0.689	-1.052	0.703	-1.049	0.714	-1.021	0.726
	t_i	-1.526	0.730	-1.523	0.744	-1.510	0.760	-1.511	0.764	-1.490	0.775
	WS_i	-1.163	0.771	-1.166	0.777	-1.154	0.781	-1.164	0.780	-1.144	0.781

Table 4.1(b)

Means and variances of unit root statistics under the null hypothesis, linear trend case

	k	0		1		2		3		4	
T		mean	var	mean	var	mean	var	mean	var	mean	var
15	Min_i	3.468	4.882	3.834	6.548	3.699	7.809	4.262	11.711	4.810	15.771
	LM_i	4.427	5.363	4.982	7.005	5.156	8.866	5.940	12.660	6.726	17.012
	Max_i	-1.785	0.808	-1.722	1.003	-1.422	1.121	-1.329	1.640	-1.008	2.675
	t_i	-2.163	0.883	-2.154	1.095	-1.948	1.278	-1.919	1.933	-1.735	3.500
	WS_i	-2.249	0.766	-2.374	0.798	-2.250	0.639	-2.308	0.668	-2.102	0.540
25	Min_i	3.689	6.090	3.927	7.219	3.837	7.740	4.118	9.576	4.058	10.723
	LM_i	4.745	7.122	5.121	8.358	5.219	9.272	5.663	11.334	5.825	13.297
	Max_i	-1.824	0.693	-1.791	0.766	-1.639	0.810	-1.589	0.916	-1.415	0.995
	t_i	-2.170	0.732	-2.166	0.796	-2.059	0.843	-2.036	0.950	-1.911	1.046
	WS_i	-2.118	0.671	-2.191	0.670	-2.148	0.602	-2.206	0.586	-2.141	0.514
50	Min_i	3.867	7.164	3.991	7.751	3.938	8.016	4.127	8.942	4.061	9.187
	LM_i	5.005	8.858	5.210	9.564	5.270	10.064	5.527	11.059	5.546	11.650
	Max_i	-1.849	0.617	-1.834	0.644	-1.764	0.666	-1.759	0.705	-1.682	0.726
	t_i	-2.173	0.637	-2.170	0.659	-2.128	0.677	-2.130	0.703	-2.070	0.729
	WS_i	-2.024	0.622	-2.058	0.612	-2.040	0.591	-2.086	0.586	-2.063	0.557
75	Min_i	3.947	7.731	4.030	8.132	3.995	8.220	4.095	8.681	4.059	8.897
	LM_i	5.117	9.552	5.251	10.019	5.295	10.343	5.451	11.004	5.460	11.275
	Max_i	-1.860	0.605	-1.849	0.621	-1.806	0.629	-1.798	0.645	-1.753	0.655
	t_i	-2.178	0.611	-2.176	0.621	-2.150	0.629	-2.149	0.645	-2.113	0.649
	WS_i	-1.995	0.620	-2.017	0.612	-2.007	0.594	-2.033	0.580	-2.021	0.568
100	Min_i	3.973	8.010	4.038	8.317	4.014	8.384	4.089	8.706	4.071	8.818
	LM_i	5.163	9.945	5.272	10.376	5.313	10.627	5.413	11.087	5.442	11.299
	Max_i	-1.863	0.592	-1.855	0.606	-1.824	0.612	-1.818	0.623	-1.787	0.632
	t_i	-2.178	0.598	-2.177	0.610	-2.160	0.615	-2.155	0.628	-2.134	0.632
	WS_i	-1.977	0.617	-1.994	0.611	-1.990	0.595	-2.008	0.587	-2.002	0.577

Table 4.2(a)

Empirical sizes and powers of panel data unit root tests at the nominal 5% level with no cross-section correlation, constant only case

			$\lambda_i = 0$		$\lambda_i \sim U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
15	5	$\overline{Min}_s$	0.074	0.156	0.064	0.128	0.071	0.124	0.078	0.111	0.084	0.121
		$\overline{LM}_s$	0.062	0.103	0.062	0.093	0.063	0.096	0.064	0.091	0.077	0.099
		$\overline{Max}_s$	0.060	0.141	0.046	0.117	0.048	0.100	0.055	0.102	0.058	0.105
		$\bar{t}_s$	0.054	0.095	0.051	0.081	0.054	0.081	0.054	0.085	0.063	0.089
		$\overline{WS}_s$	0.060	0.150	0.053	0.129	0.052	0.117	0.053	0.108	0.055	0.110
	10	$\overline{Min}_s$	0.059	0.200	0.066	0.179	0.070	0.139	0.075	0.129	0.088	0.109
		$\overline{LM}_s$	0.056	0.135	0.065	0.118	0.069	0.100	0.070	0.111	0.084	0.096
		$\overline{Max}_s$	0.050	0.219	0.045	0.168	0.040	0.132	0.050	0.132	0.049	0.106
		$\bar{t}_s$	0.048	0.133	0.057	0.113	0.054	0.090	0.057	0.108	0.062	0.083
		$\overline{WS}_s$	0.050	0.249	0.053	0.187	0.046	0.150	0.051	0.139	0.052	0.119
	25	$\overline{Min}_s$	0.055	0.355	0.063	0.280	0.072	0.222	0.079	0.194	0.096	0.152
		$\overline{LM}_s$	0.056	0.201	0.065	0.195	0.077	0.155	0.091	0.151	0.090	0.138
		$\overline{Max}_s$	0.049	0.432	0.032	0.286	0.036	0.240	0.040	0.207	0.049	0.183
		$\bar{t}_s$	0.055	0.232	0.051	0.183	0.054	0.146	0.065	0.156	0.074	0.132
		$\overline{WS}_s$	0.053	0.490	0.039	0.338	0.042	0.294	0.046	0.231	0.048	0.211
	50	$\overline{Min}_s$	0.058	0.560	0.060	0.402	0.071	0.300	0.083	0.269	0.111	0.219
		$\overline{LM}_s$	0.054	0.320	0.073	0.297	0.079	0.201	0.095	0.217	0.109	0.184
		$\overline{Max}_s$	0.051	0.692	0.030	0.443	0.030	0.364	0.037	0.322	0.043	0.251
		$\bar{t}_s$	0.052	0.368	0.050	0.298	0.051	0.224	0.069	0.215	0.078	0.182
		$\overline{WS}_s$	0.055	0.778	0.039	0.538	0.039	0.456	0.040	0.366	0.049	0.318
	100	$\overline{Min}_s$	0.055	0.812	0.055	0.637	0.074	0.465	0.100	0.404	0.142	0.297
		$\overline{LM}_s$	0.051	0.479	0.079	0.440	0.088	0.290	0.111	0.313	0.147	0.248
		$\overline{Max}_s$	0.050	0.927	0.017	0.677	0.017	0.565	0.031	0.482	0.037	0.390
		$\bar{t}_s$	0.053	0.583	0.048	0.446	0.054	0.317	0.069	0.321	0.079	0.266
		$\overline{WS}_s$	0.053	0.961	0.031	0.779	0.029	0.681	0.036	0.569	0.040	0.510

Table 4.2(a) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
25	5	$\overline{Min}_s$	0.068	0.274	0.063	0.230	0.066	0.204	0.067	0.184	0.078	0.164
		$\overline{LM}_s$	0.066	0.168	0.059	0.153	0.065	0.141	0.069	0.141	0.073	0.119
		$\overline{Max}_s$	0.051	0.279	0.041	0.231	0.040	0.202	0.045	0.172	0.042	0.157
		$\bar{t}_s$	0.052	0.153	0.045	0.139	0.049	0.120	0.052	0.113	0.049	0.098
		$\overline{WS}_s$	0.051	0.307	0.044	0.246	0.042	0.225	0.044	0.185	0.044	0.175
	10	$\overline{Min}_s$	0.058	0.399	0.066	0.344	0.070	0.285	0.064	0.245	0.071	0.198
		$\overline{LM}_s$	0.057	0.241	0.066	0.216	0.059	0.175	0.062	0.154	0.065	0.132
		$\overline{Max}_s$	0.043	0.495	0.049	0.378	0.045	0.327	0.043	0.271	0.043	0.231
		$\bar{t}_s$	0.049	0.250	0.053	0.211	0.048	0.180	0.048	0.146	0.046	0.134
		$\overline{WS}_s$	0.046	0.538	0.051	0.410	0.052	0.364	0.045	0.297	0.044	0.257
	25	$\overline{Min}_s$	0.060	0.735	0.064	0.604	0.069	0.508	0.065	0.426	0.066	0.343
		$\overline{LM}_s$	0.062	0.441	0.076	0.386	0.075	0.294	0.081	0.261	0.079	0.198
		$\overline{Max}_s$	0.056	0.846	0.041	0.700	0.041	0.621	0.043	0.535	0.038	0.451
		$\bar{t}_s$	0.055	0.519	0.057	0.437	0.055	0.338	0.058	0.288	0.058	0.229
		$\overline{WS}_s$	0.054	0.906	0.045	0.760	0.047	0.696	0.049	0.602	0.042	0.512
	50	$\overline{Min}_s$	0.064	0.930	0.060	0.849	0.068	0.755	0.068	0.664	0.071	0.519
		$\overline{LM}_s$	0.055	0.691	0.074	0.599	0.077	0.469	0.081	0.420	0.080	0.309
		$\overline{Max}_s$	0.052	0.987	0.038	0.922	0.037	0.877	0.036	0.800	0.033	0.715
		$\bar{t}_s$	0.054	0.814	0.056	0.682	0.054	0.564	0.054	0.501	0.051	0.406
		$\overline{WS}_s$	0.052	0.996	0.043	0.964	0.048	0.937	0.043	0.869	0.039	0.788
	100	$\overline{Min}_s$	0.057	0.997	0.060	0.986	0.066	0.955	0.071	0.905	0.078	0.786
		$\overline{LM}_s$	0.058	0.924	0.086	0.870	0.085	0.744	0.093	0.659	0.090	0.520
		$\overline{Max}_s$	0.050	1.000	0.027	0.998	0.029	0.992	0.026	0.971	0.021	0.938
		$\bar{t}_s$	0.058	0.978	0.057	0.938	0.053	0.855	0.056	0.789	0.053	0.662
		$\overline{WS}_s$	0.051	1.000	0.033	1.000	0.043	0.998	0.036	0.990	0.033	0.968

Table 4.2(a) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
50	5	$\overline{Min}_s$	0.067	0.698	0.072	0.609	0.073	0.541	0.073	0.479	0.070	0.443
		$\overline{LM}_s$	0.066	0.428	0.064	0.363	0.066	0.297	0.065	0.271	0.063	0.243
		$\overline{Max}_s$	0.048	0.747	0.048	0.647	0.050	0.585	0.048	0.513	0.045	0.459
		$\bar{t}_s$	0.052	0.432	0.049	0.357	0.046	0.289	0.044	0.250	0.043	0.217
		$\overline{WS}_s$	0.050	0.776	0.045	0.679	0.048	0.605	0.048	0.545	0.045	0.497
	10	$\overline{Min}_s$	0.062	0.927	0.066	0.861	0.069	0.815	0.067	0.749	0.064	0.692
		$\overline{LM}_s$	0.058	0.701	0.059	0.593	0.062	0.521	0.061	0.448	0.060	0.388
		$\overline{Max}_s$	0.047	0.970	0.050	0.903	0.053	0.867	0.046	0.823	0.048	0.778
		$\bar{t}_s$	0.046	0.758	0.046	0.640	0.049	0.562	0.048	0.477	0.046	0.416
		$\overline{WS}_s$	0.050	0.983	0.050	0.934	0.053	0.899	0.051	0.861	0.051	0.823
	25	$\overline{Min}_s$	0.064	0.999	0.068	0.998	0.067	0.990	0.065	0.973	0.064	0.943
		$\overline{LM}_s$	0.060	0.965	0.067	0.918	0.062	0.845	0.059	0.792	0.061	0.693
		$\overline{Max}_s$	0.052	1.000	0.053	1.000	0.047	0.998	0.043	0.994	0.043	0.983
		$\bar{t}_s$	0.047	0.989	0.054	0.967	0.051	0.923	0.046	0.876	0.048	0.806
		$\overline{WS}_s$	0.055	1.000	0.049	1.000	0.049	1.000	0.046	0.998	0.046	0.993
	50	$\overline{Min}_s$	0.053	1.000	0.055	1.000	0.058	1.000	0.058	1.000	0.059	0.999
		$\overline{LM}_s$	0.058	1.000	0.061	0.997	0.063	0.988	0.057	0.976	0.062	0.932
		$\overline{Max}_s$	0.050	1.000	0.039	1.000	0.041	1.000	0.036	1.000	0.037	1.000
		$\bar{t}_s$	0.051	1.000	0.050	1.000	0.049	0.999	0.045	0.994	0.052	0.983
		$\overline{WS}_s$	0.047	1.000	0.042	1.000	0.042	1.000	0.046	1.000	0.042	1.000
	100	$\overline{Min}_s$	0.060	1.000	0.066	1.000	0.063	1.000	0.061	1.000	0.061	1.000
		$\overline{LM}_s$	0.055	1.000	0.071	1.000	0.069	1.000	0.064	1.000	0.072	0.999
		$\overline{Max}_s$	0.055	1.000	0.043	1.000	0.044	1.000	0.041	1.000	0.037	1.000
		$\bar{t}_s$	0.049	1.000	0.051	1.000	0.054	1.000	0.049	1.000	0.053	1.000
		$\overline{WS}_s$	0.052	1.000	0.045	1.000	0.047	1.000	0.048	1.000	0.041	1.000

Table 4.2(a) (continued)

			$\lambda_i = 0$		λ_i˜$U(0.2, 0.4)$							
			$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
75	5	$\overline{Min}_s$	0.071	0.947	0.070	0.903	0.071	0.872	0.069	0.827	0.069	0.783
		$\overline{LM}_s$	0.062	0.782	0.069	0.686	0.071	0.624	0.065	0.571	0.068	0.525
		$\overline{Max}_s$	0.049	0.972	0.052	0.941	0.054	0.903	0.050	0.858	0.049	0.817
		$\bar{t}_s$	0.047	0.801	0.052	0.705	0.052	0.637	0.048	0.576	0.052	0.512
		$\overline{WS}_s$	0.048	0.977	0.049	0.947	0.050	0.924	0.050	0.878	0.053	0.839
	10	$\overline{Min}_s$	0.066	0.999	0.066	0.994	0.067	0.987	0.070	0.973	0.071	0.955
		$\overline{LM}_s$	0.060	0.965	0.062	0.927	0.063	0.879	0.063	0.833	0.071	0.783
		$\overline{Max}_s$	0.054	1.000	0.048	0.999	0.049	0.995	0.050	0.989	0.050	0.978
		$\bar{t}_s$	0.050	0.985	0.048	0.955	0.052	0.917	0.049	0.880	0.055	0.825
		$\overline{WS}_s$	0.054	1.000	0.050	0.999	0.052	0.998	0.049	0.994	0.054	0.988
	25	$\overline{Min}_s$	0.059	1.000	0.057	1.000	0.052	1.000	0.053	1.000	0.051	1.000
		$\overline{LM}_s$	0.053	1.000	0.061	1.000	0.063	1.000	0.057	0.999	0.064	0.987
		$\overline{Max}_s$	0.050	1.000	0.038	1.000	0.042	1.000	0.039	1.000	0.039	1.000
		$\bar{t}_s$	0.041	1.000	0.049	1.000	0.047	1.000	0.041	1.000	0.047	0.999
		$\overline{WS}_s$	0.047	1.000	0.039	1.000	0.043	1.000	0.040	1.000	0.040	1.000
	50	$\overline{Min}_s$	0.064	1.000	0.064	1.000	0.063	1.000	0.062	1.000	0.057	1.000
		$\overline{LM}_s$	0.062	1.000	0.058	1.000	0.058	1.000	0.058	1.000	0.069	1.000
		$\overline{Max}_s$	0.052	1.000	0.045	1.000	0.043	1.000	0.041	1.000	0.042	1.000
		$\bar{t}_s$	0.054	1.000	0.052	1.000	0.050	1.000	0.049	1.000	0.048	1.000
		$\overline{WS}_s$	0.052	1.000	0.044	1.000	0.044	1.000	0.041	1.000	0.049	1.000
	100	$\overline{Min}_s$	0.055	1.000	0.055	1.000	0.060	1.000	0.058	1.000	0.057	1.000
		$\overline{LM}_s$	0.056	1.000	0.064	1.000	0.065	1.000	0.062	1.000	0.073	1.000
		$\overline{Max}_s$	0.048	1.000	0.041	1.000	0.044	1.000	0.041	1.000	0.042	1.000
		$\bar{t}_s$	0.056	1.000	0.052	1.000	0.054	1.000	0.049	1.000	0.057	1.000
		$\overline{WS}_s$	0.049	1.000	0.043	1.000	0.046	1.000	0.044	1.000	0.046	1.000

Table 4.2(a) (continued)

			$\lambda_i = 0$		$\lambda_i\text{˜}U(0.2, 0.4)$							
			$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
100	5	$\overline{Min}_s$	0.067	0.998	0.066	0.991	0.067	0.984	0.068	0.971	0.069	0.950
		$\overline{LM}_s$	0.061	0.959	0.065	0.913	0.061	0.860	0.067	0.812	0.064	0.765
		$\overline{Max}_s$	0.049	1.000	0.048	0.997	0.046	0.991	0.045	0.986	0.046	0.971
		$\bar{t}_s$	0.047	0.967	0.047	0.925	0.044	0.869	0.049	0.826	0.045	0.774
		$\overline{WS}_s$	0.047	1.000	0.046	0.995	0.046	0.993	0.044	0.984	0.047	0.972
	10	$\overline{Min}_s$	0.065	1.000	0.065	1.000	0.062	1.000	0.063	0.999	0.060	0.998
		$\overline{LM}_s$	0.061	1.000	0.062	0.995	0.059	0.987	0.058	0.977	0.062	0.957
		$\overline{Max}_s$	0.051	1.000	0.049	1.000	0.050	1.000	0.047	0.999	0.047	1.000
		$\bar{t}_s$	0.047	1.000	0.045	0.999	0.047	0.996	0.046	0.989	0.048	0.972
		$\overline{WS}_s$	0.047	1.000	0.049	1.000	0.047	1.000	0.043	1.000	0.045	1.000
	25	$\overline{Min}_s$	0.063	1.000	0.062	1.000	0.059	1.000	0.060	1.000	0.061	1.000
		$\overline{LM}_s$	0.061	1.000	0.061	1.000	0.056	1.000	0.059	1.000	0.060	1.000
		$\overline{Max}_s$	0.055	1.000	0.050	1.000	0.049	1.000	0.047	1.000	0.047	1.000
		$\bar{t}_s$	0.051	1.000	0.053	1.000	0.049	1.000	0.046	1.000	0.051	1.000
		$\overline{WS}_s$	0.048	1.000	0.047	1.000	0.042	1.000	0.043	1.000	0.050	1.000
	50	$\overline{Min}_s$	0.062	1.000	0.060	1.000	0.054	1.000	0.053	1.000	0.055	1.000
		$\overline{LM}_s$	0.057	1.000	0.055	1.000	0.051	1.000	0.050	1.000	0.055	1.000
		$\overline{Max}_s$	0.056	1.000	0.047	1.000	0.040	1.000	0.038	1.000	0.040	1.000
		$\bar{t}_s$	0.045	1.000	0.047	1.000	0.046	1.000	0.040	1.000	0.046	1.000
		$\overline{WS}_s$	0.055	1.000	0.048	1.000	0.046	1.000	0.041	1.000	0.046	1.000
	100	$\overline{Min}_s$	0.062	1.000	0.059	1.000	0.047	1.000	0.047	1.000	0.046	1.000
		$\overline{LM}_s$	0.056	1.000	0.057	1.000	0.046	1.000	0.048	1.000	0.052	1.000
		$\overline{Max}_s$	0.052	1.000	0.043	1.000	0.039	1.000	0.035	1.000	0.037	1.000
		$\bar{t}_s$	0.050	1.000	0.054	1.000	0.040	1.000	0.039	1.000	0.043	1.000
		$\overline{WS}_s$	0.049	1.000	0.049	1.000	0.043	1.000	0.038	1.000	0.042	1.000

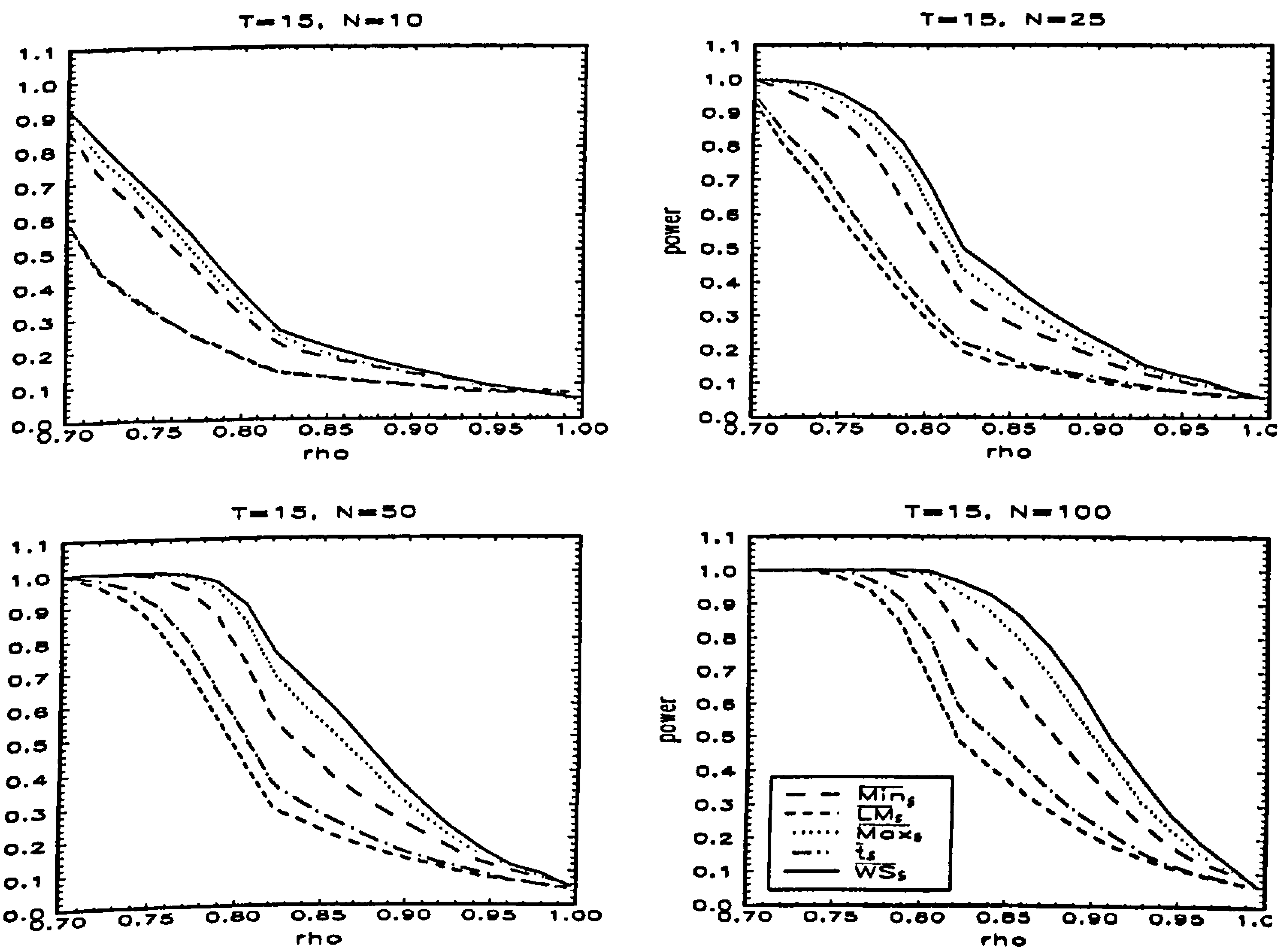


Figure 4.1. Empirical powers of panel data unit root tests at the nominal 5% level with no cross-section correlation, constant only case.

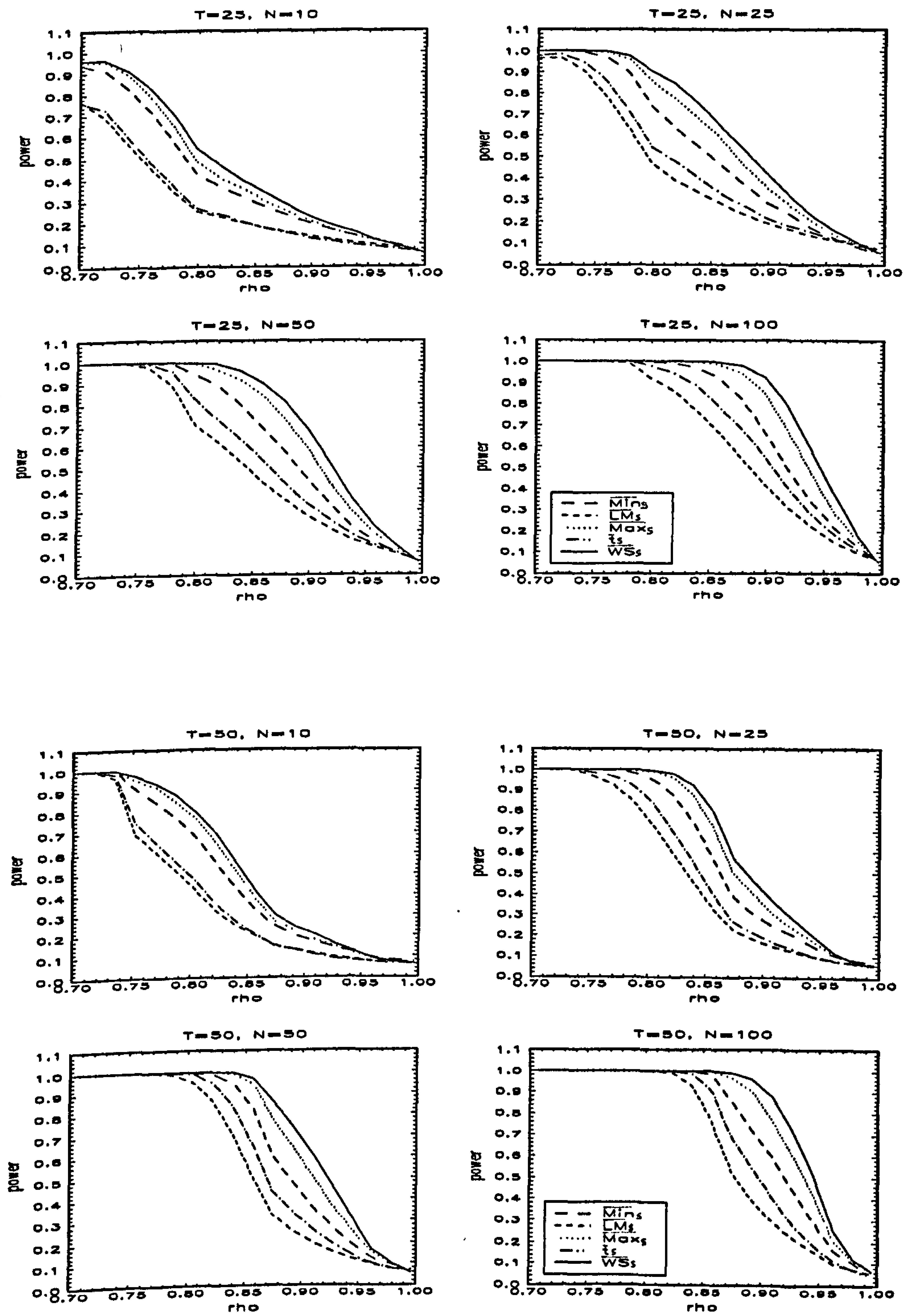


Figure 4.1. (continued)

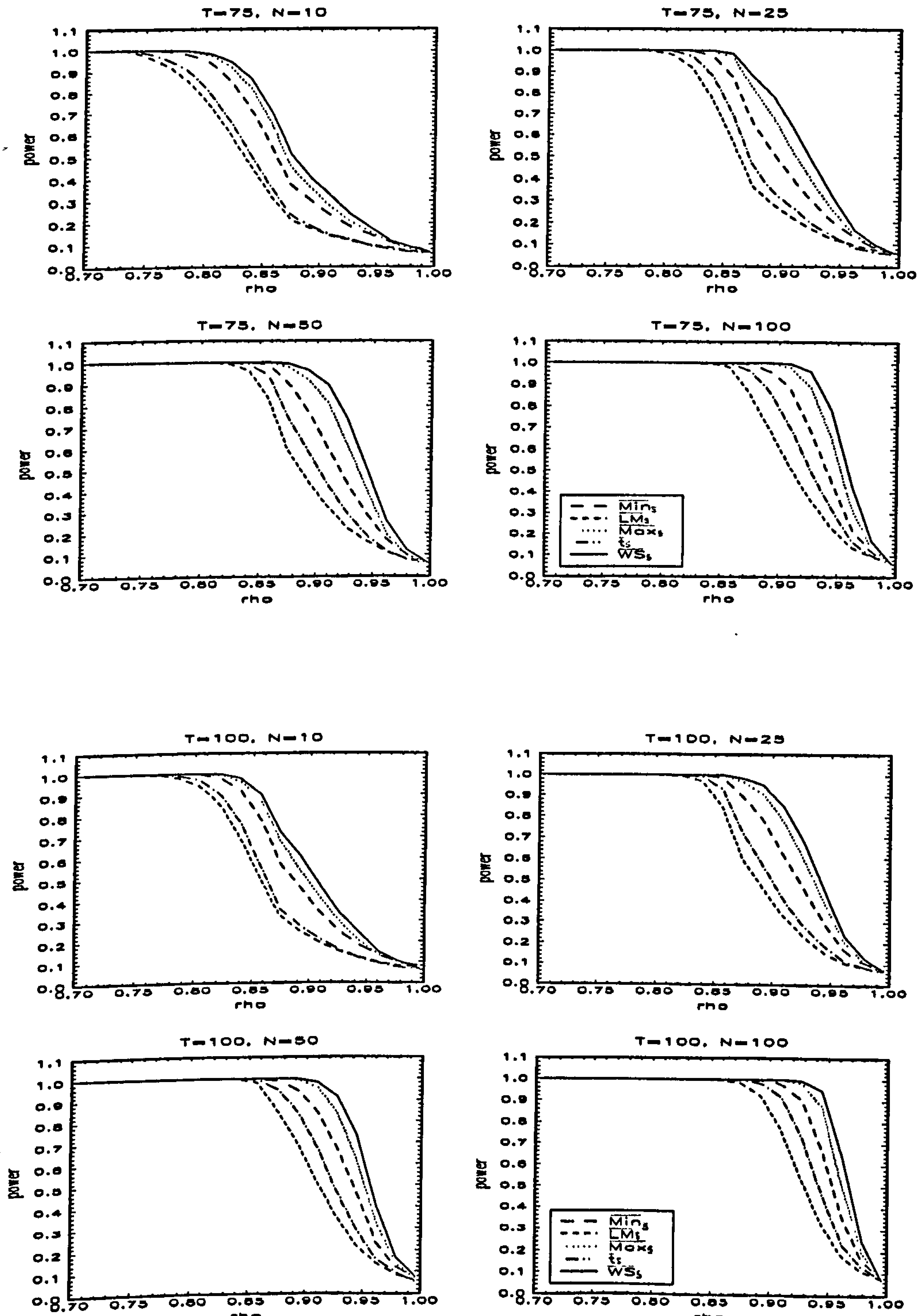


Figure 4.1. (continued)

Table 4.2(b)

Empirical sizes and powers of panel data unit root tests at the nominal 5% level with no cross-section correlation, linear ternd case

			$\lambda_i = 0$		$\lambda_i\tilde{}U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
15	5	$\overline{Min}_s$	0.061	0.088	0.057	0.064	0.061	0.068	0.063	0.067	0.061	0.064
		$\overline{LM}_s$	0.052	0.067	0.057	0.067	0.056	0.062	0.062	0.059	0.060	0.057
		$\overline{Max}_s$	0.058	0.084	0.054	0.066	0.052	0.063	0.059	0.069	0.049	0.050
		$\bar{t}_s$	0.056	0.072	0.060	0.067	0.060	0.066	0.065	0.068	0.059	0.064
		$\overline{WS}_s$	0.059	0.086	0.067	0.077	0.060	0.073	0.064	0.075	0.060	0.073
	10	$\overline{Min}_s$	0.057	0.084	0.051	0.073	0.054	0.068	0.061	0.066	0.067	0.079
		$\overline{LM}_s$	0.052	0.070	0.056	0.070	0.056	0.064	0.059	0.068	0.068	0.064
		$\overline{Max}_s$	0.056	0.082	0.050	0.069	0.047	0.059	0.051	0.061	0.044	0.056
		$\bar{t}_s$	0.052	0.074	0.064	0.074	0.058	0.064	0.063	0.071	0.060	0.065
		$\overline{WS}_s$	0.056	0.082	0.065	0.085	0.057	0.073	0.058	0.067	0.050	0.061
	25	$\overline{Min}_s$	0.055	0.098	0.045	0.075	0.049	0.059	0.061	0.074	0.070	0.076
		$\overline{LM}_s$	0.054	0.084	0.062	0.081	0.061	0.075	0.065	0.080	0.072	0.088
		$\overline{Max}_s$	0.054	0.096	0.041	0.066	0.042	0.051	0.047	0.064	0.047	0.059
		$\bar{t}_s$	0.057	0.083	0.063	0.084	0.058	0.070	0.065	0.083	0.072	0.079
		$\overline{WS}_s$	0.055	0.102	0.063	0.097	0.059	0.078	0.056	0.062	0.049	0.058
	50	$\overline{Min}_s$	0.059	0.128	0.036	0.073	0.040	0.069	0.061	0.086	0.079	0.098
		$\overline{LM}_s$	0.057	0.093	0.066	0.095	0.063	0.080	0.071	0.086	0.087	0.103
		$\overline{Max}_s$	0.055	0.131	0.030	0.061	0.035	0.066	0.046	0.066	0.047	0.062
		$\bar{t}_s$	0.057	0.091	0.061	0.089	0.064	0.074	0.070	0.088	0.071	0.077
		$\overline{WS}_s$	0.054	0.133	0.060	0.105	0.063	0.087	0.054	0.074	0.046	0.059
	100	$\overline{Min}_s$	0.058	0.165	0.027	0.073	0.035	0.060	0.059	0.094	0.087	0.115
		$\overline{LM}_s$	0.054	0.124	0.066	0.121	0.061	0.086	0.077	0.100	0.110	0.135
		$\overline{Max}_s$	0.058	0.167	0.021	0.061	0.021	0.050	0.037	0.062	0.034	0.056
		$\bar{t}_s$	0.057	0.133	0.061	0.105	0.056	0.081	0.070	0.100	0.078	0.092
		$\overline{WS}_s$	0.058	0.169	0.058	0.124	0.051	0.087	0.057	0.085	0.038	0.054

Table 4.2(b) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
25	5	$\overline{Min}_s$	0.066	0.104	0.056	0.088	0.063	0.092	0.059	0.084	0.065	0.085
		$\overline{LM}_s$	0.062	0.088	0.061	0.075	0.063	0.075	0.059	0.075	0.067	0.064
		$\overline{Max}_s$	0.057	0.091	0.049	0.080	0.050	0.080	0.049	0.071	0.052	0.067
		$\bar{t}_s$	0.056	0.085	0.056	0.070	0.054	0.068	0.052	0.068	0.056	0.056
		$\overline{WS}_s$	0.059	0.099	0.052	0.087	0.056	0.087	0.058	0.077	0.054	0.074
	10	$\overline{Min}_s$	0.055	0.137	0.055	0.113	0.053	0.104	0.052	0.080	0.058	0.085
		$\overline{LM}_s$	0.057	0.099	0.062	0.102	0.065	0.092	0.065	0.085	0.062	0.081
		$\overline{Max}_s$	0.051	0.133	0.048	0.108	0.042	0.087	0.042	0.069	0.046	0.070
		$\bar{t}_s$	0.054	0.091	0.053	0.098	0.058	0.084	0.058	0.075	0.058	0.072
		$\overline{WS}_s$	0.052	0.143	0.057	0.121	0.051	0.103	0.049	0.083	0.051	0.068
	25	$\overline{Min}_s$	0.056	0.205	0.052	0.152	0.046	0.108	0.050	0.114	0.053	0.095
		$\overline{LM}_s$	0.052	0.142	0.066	0.135	0.062	0.112	0.069	0.109	0.062	0.088
		$\overline{Max}_s$	0.053	0.206	0.046	0.144	0.039	0.098	0.040	0.106	0.039	0.078
		$\bar{t}_s$	0.049	0.145	0.060	0.129	0.054	0.101	0.060	0.099	0.056	0.089
		$\overline{WS}_s$	0.055	0.219	0.055	0.157	0.048	0.120	0.046	0.112	0.042	0.085
	50	$\overline{Min}_s$	0.057	0.317	0.046	0.194	0.047	0.152	0.047	0.150	0.055	0.120
		$\overline{LM}_s$	0.057	0.194	0.066	0.174	0.064	0.132	0.066	0.132	0.067	0.104
		$\overline{Max}_s$	0.056	0.337	0.036	0.179	0.032	0.141	0.035	0.135	0.039	0.116
		$\bar{t}_s$	0.053	0.202	0.054	0.166	0.053	0.120	0.058	0.118	0.056	0.100
		$\overline{WS}_s$	0.054	0.348	0.053	0.229	0.050	0.168	0.044	0.144	0.045	0.110
	100	$\overline{Min}_s$	0.058	0.490	0.038	0.317	0.038	0.235	0.045	0.201	0.045	0.164
		$\overline{LM}_s$	0.052	0.331	0.069	0.268	0.066	0.210	0.078	0.173	0.067	0.141
		$\overline{Max}_s$	0.053	0.543	0.030	0.310	0.027	0.222	0.030	0.183	0.031	0.146
		$\bar{t}_s$	0.050	0.345	0.059	0.261	0.053	0.203	0.064	0.166	0.053	0.138
		$\overline{WS}_s$	0.055	0.565	0.043	0.380	0.041	0.284	0.040	0.210	0.035	0.135

Table 4.2(b) (continued)

			$\lambda_i = 0$		$\lambda_i \sim U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
50	5	$\overline{Min}_s$	0.064	0.275	0.063	0.242	0.060	0.205	0.054	0.174	0.059	0.159
		$\overline{LM}_s$	0.059	0.214	0.064	0.182	0.063	0.144	0.064	0.133	0.067	0.128
		$\overline{Max}_s$	0.055	0.267	0.051	0.232	0.050	0.195	0.042	0.156	0.041	0.138
		$\bar{t}_s$	0.051	0.196	0.054	0.169	0.053	0.130	0.053	0.117	0.055	0.111
		$\overline{WS}_s$	0.055	0.278	0.051	0.239	0.051	0.207	0.043	0.169	0.046	0.150
	10	$\overline{Min}_s$	0.061	0.445	0.058	0.359	0.056	0.310	0.054	0.254	0.057	0.233
		$\overline{LM}_s$	0.056	0.302	0.059	0.248	0.057	0.192	0.057	0.170	0.058	0.161
		$\overline{Max}_s$	0.051	0.464	0.050	0.360	0.047	0.299	0.044	0.245	0.046	0.226
		$\bar{t}_s$	0.051	0.304	0.052	0.245	0.049	0.180	0.049	0.160	0.047	0.149
		$\overline{WS}_s$	0.050	0.486	0.050	0.379	0.049	0.323	0.046	0.258	0.050	0.232
	25	$\overline{Min}_s$	0.057	0.779	0.057	0.641	0.058	0.562	0.049	0.458	0.050	0.394
		$\overline{LM}_s$	0.055	0.577	0.067	0.466	0.065	0.373	0.055	0.284	0.063	0.254
		$\overline{Max}_s$	0.053	0.824	0.051	0.670	0.047	0.585	0.041	0.472	0.043	0.413
		$\bar{t}_s$	0.052	0.613	0.059	0.487	0.056	0.378	0.050	0.300	0.059	0.257
		$\overline{WS}_s$	0.050	0.849	0.056	0.709	0.052	0.624	0.042	0.501	0.047	0.433
	50	$\overline{Min}_s$	0.052	0.964	0.047	0.885	0.050	0.819	0.041	0.703	0.042	0.641
		$\overline{LM}_s$	0.059	0.854	0.067	0.700	0.059	0.579	0.051	0.479	0.063	0.425
		$\overline{Max}_s$	0.048	0.983	0.037	0.912	0.039	0.850	0.031	0.741	0.032	0.675
		$\bar{t}_s$	0.056	0.891	0.059	0.740	0.052	0.618	0.044	0.508	0.052	0.447
		$\overline{WS}_s$	0.047	0.989	0.044	0.935	0.047	0.880	0.035	0.770	0.036	0.695
	100	$\overline{Min}_s$	0.054	1.000	0.054	0.992	0.054	0.975	0.038	0.925	0.044	0.869
		$\overline{LM}_s$	0.059	0.989	0.069	0.938	0.060	0.854	0.050	0.752	0.066	0.663
		$\overline{Max}_s$	0.051	1.000	0.041	0.997	0.045	0.988	0.032	0.949	0.035	0.903
		$\bar{t}_s$	0.054	0.995	0.060	0.959	0.051	0.890	0.046	0.800	0.055	0.713
		$\overline{WS}_s$	0.050	1.000	0.046	0.999	0.050	0.992	0.037	0.964	0.040	0.915

Table 4.2(b) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
75	5	$\overline{Min_s}$	0.058	0.572	0.061	0.471	0.063	0.420	0.059	0.375	0.061	0.338
		$\overline{LM_s}$	0.059	0.414	0.066	0.350	0.063	0.313	0.059	0.265	0.059	0.229
		$\overline{Max_s}$	0.049	0.578	0.050	0.464	0.050	0.408	0.047	0.355	0.049	0.315
		$\overline{t}_s$	0.052	0.406	0.054	0.335	0.050	0.299	0.048	0.237	0.049	0.206
		$\overline{WS_s}$	0.046	0.589	0.053	0.491	0.052	0.427	0.050	0.390	0.052	0.345
	10	$\overline{Min_s}$	0.051	0.828	0.051	0.733	0.062	0.668	0.056	0.615	0.058	0.545
		$\overline{LM_s}$	0.051	0.658	0.055	0.563	0.054	0.484	0.053	0.412	0.064	0.372
		$\overline{Max_s}$	0.045	0.859	0.043	0.766	0.051	0.691	0.046	0.628	0.047	0.560
		$\overline{t}_s$	0.045	0.680	0.048	0.572	0.045	0.491	0.044	0.404	0.049	0.371
		$\overline{WS_s}$	0.047	0.878	0.048	0.784	0.050	0.712	0.048	0.661	0.048	0.575
	25	$\overline{Min_s}$	0.053	0.997	0.050	0.968	0.049	0.945	0.051	0.906	0.050	0.858
		$\overline{LM_s}$	0.058	0.957	0.060	0.897	0.054	0.812	0.060	0.742	0.060	0.679
		$\overline{Max_s}$	0.048	1.000	0.043	0.982	0.039	0.966	0.042	0.929	0.040	0.897
		$\overline{t}_s$	0.050	0.971	0.053	0.917	0.046	0.852	0.048	0.778	0.050	0.710
		$\overline{WS_s}$	0.048	1.000	0.046	0.990	0.047	0.975	0.044	0.946	0.044	0.911
	50	$\overline{Min_s}$	0.055	1.000	0.056	1.000	0.054	1.000	0.055	0.998	0.056	0.989
		$\overline{LM_s}$	0.052	1.000	0.065	0.991	0.060	0.973	0.059	0.945	0.069	0.907
		$\overline{Max_s}$	0.052	1.000	0.047	1.000	0.044	1.000	0.045	0.999	0.048	0.994
		$\overline{t}_s$	0.049	1.000	0.057	0.996	0.052	0.985	0.049	0.961	0.057	0.931
		$\overline{WS_s}$	0.054	1.000	0.049	1.000	0.048	1.000	0.051	1.000	0.050	0.998
	100	$\overline{Min_s}$	0.051	1.000	0.045	1.000	0.048	1.000	0.048	1.000	0.047	1.000
		$\overline{LM_s}$	0.059	1.000	0.068	1.000	0.059	1.000	0.055	0.998	0.066	0.996
		$\overline{Max_s}$	0.047	1.000	0.041	1.000	0.042	1.000	0.038	1.000	0.036	1.000
		$\overline{t}_s$	0.058	1.000	0.060	1.000	0.052	1.000	0.048	0.999	0.053	0.999
		$\overline{WS_s}$	0.045	1.000	0.040	1.000	0.042	1.000	0.039	1.000	0.039	1.000

Table 4.2(b) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
100	5	$\overline{Min}_s$	0.063	0.855	0.063	0.759	0.060	0.710	0.064	0.641	0.061	0.591
		$\overline{LM}_s$	0.058	0.708	0.065	0.594	0.061	0.526	0.066	0.467	0.063	0.410
		$\overline{Max}_s$	0.053	0.872	0.049	0.777	0.051	0.715	0.050	0.644	0.048	0.583
		$\bar{t}_s$	0.047	0.711	0.054	0.587	0.047	0.517	0.051	0.446	0.048	0.387
		$\overline{WS}_s$	0.052	0.882	0.052	0.786	0.053	0.726	0.049	0.663	0.050	0.604
	10	$\overline{Min}_s$	0.056	0.987	0.058	0.954	0.062	0.929	0.062	0.892	0.062	0.836
		$\overline{LM}_s$	0.056	0.924	0.058	0.839	0.060	0.772	0.057	0.716	0.059	0.647
		$\overline{Max}_s$	0.049	0.992	0.052	0.965	0.051	0.943	0.052	0.907	0.050	0.856
		$\bar{t}_s$	0.049	0.937	0.050	0.858	0.050	0.788	0.051	0.737	0.052	0.658
		$\overline{WS}_s$	0.050	0.996	0.052	0.972	0.054	0.958	0.054	0.920	0.053	0.880
	25	$\overline{Min}_s$	0.053	1.000	0.061	1.000	0.059	1.000	0.057	1.000	0.056	0.998
		$\overline{LM}_s$	0.055	1.000	0.061	0.996	0.057	0.990	0.058	0.984	0.058	0.953
		$\overline{Max}_s$	0.050	1.000	0.052	1.000	0.049	1.000	0.050	1.000	0.047	0.999
		$\bar{t}_s$	0.050	1.000	0.054	0.999	0.051	0.994	0.055	0.993	0.052	0.967
		$\overline{WS}_s$	0.050	1.000	0.057	1.000	0.053	1.000	0.051	1.000	0.049	1.000
	50	$\overline{Min}_s$	0.046	1.000	0.044	1.000	0.049	1.000	0.050	1.000	0.050	1.000
		$\overline{LM}_s$	0.051	1.000	0.054	1.000	0.054	1.000	0.058	1.000	0.057	1.000
		$\overline{Max}_s$	0.046	1.000	0.042	1.000	0.043	1.000	0.042	1.000	0.042	1.000
		$\bar{t}_s$	0.048	1.000	0.049	1.000	0.047	1.000	0.047	1.000	0.051	1.000
		$\overline{WS}_s$	0.049	1.000	0.043	1.000	0.043	1.000	0.044	1.000	0.044	1.000
	100	$\overline{Min}_s$	0.048	1.000	0.048	1.000	0.047	1.000	0.046	1.000	0.040	1.000
		$\overline{LM}_s$	0.049	1.000	0.054	1.000	0.046	1.000	0.050	1.000	0.048	1.000
		$\overline{Max}_s$	0.048	1.000	0.038	1.000	0.037	1.000	0.038	1.000	0.034	1.000
		$\bar{t}_s$	0.048	1.000	0.047	1.000	0.039	1.000	0.045	1.000	0.046	1.000
		$\overline{WS}_s$	0.049	1.000	0.043	1.000	0.039	1.000	0.039	1.000	0.036	1.000

Table 4.3(a)

Empirical sizes and powers of panel data unit root tests at the nominal 5% level with fixed effects included, constant only case

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
15	5	$\overline{Min}_s$	0.076	0.166	0.076	0.138	0.082	0.129	0.087	0.129	0.089	0.102
		$\overline{LM}_s$	0.069	0.118	0.088	0.126	0.077	0.108	0.081	0.107	0.085	0.092
		$\overline{Max}_s$	0.069	0.162	0.058	0.125	0.059	0.115	0.065	0.115	0.061	0.097
		$\bar{t}_s$	0.064	0.109	0.073	0.114	0.066	0.096	0.068	0.098	0.067	0.079
		$\overline{WS}_s$	0.068	0.177	0.063	0.136	0.066	0.128	0.068	0.127	0.062	0.108
	10	$\overline{Min}_s$	0.064	0.214	0.069	0.176	0.080	0.158	0.083	0.142	0.093	0.126
		$\overline{LM}_s$	0.066	0.136	0.074	0.130	0.073	0.111	0.078	0.117	0.087	0.105
		$\overline{Max}_s$	0.057	0.241	0.047	0.162	0.050	0.150	0.056	0.142	0.057	0.122
		$\bar{t}_s$	0.060	0.135	0.059	0.121	0.060	0.101	0.062	0.109	0.071	0.095
		$\overline{WS}_s$	0.059	0.260	0.055	0.189	0.059	0.178	0.057	0.151	0.056	0.136
	25	$\overline{Min}_s$	0.057	0.361	0.059	0.273	0.064	0.206	0.077	0.197	0.097	0.149
		$\overline{LM}_s$	0.055	0.195	0.069	0.207	0.071	0.151	0.074	0.155	0.094	0.141
		$\overline{Max}_s$	0.050	0.437	0.039	0.275	0.036	0.229	0.041	0.211	0.045	0.172
		$\bar{t}_s$	0.051	0.217	0.054	0.197	0.054	0.156	0.054	0.153	0.068	0.141
		$\overline{WS}_s$	0.055	0.497	0.041	0.329	0.043	0.279	0.042	0.235	0.045	0.204
	50	$\overline{Min}_s$	0.062	0.541	0.057	0.406	0.068	0.312	0.087	0.273	0.113	0.202
		$\overline{LM}_s$	0.050	0.292	0.071	0.263	0.071	0.195	0.096	0.199	0.119	0.168
		$\overline{Max}_s$	0.057	0.680	0.026	0.443	0.032	0.370	0.034	0.305	0.045	0.267
		$\bar{t}_s$	0.048	0.358	0.052	0.276	0.052	0.225	0.066	0.205	0.078	0.172
		$\overline{WS}_s$	0.053	0.755	0.039	0.524	0.040	0.444	0.043	0.357	0.049	0.324
	100	$\overline{Min}_s$	0.060	0.815	0.063	0.646	0.076	0.479	0.102	0.416	0.143	0.299
		$\overline{LM}_s$	0.054	0.486	0.080	0.463	0.085	0.315	0.114	0.317	0.152	0.253
		$\overline{Max}_s$	0.052	0.933	0.019	0.684	0.024	0.583	0.031	0.504	0.036	0.404
		$\bar{t}_s$	0.048	0.584	0.043	0.464	0.045	0.337	0.066	0.344	0.077	0.271
		$\overline{WS}_s$	0.052	0.967	0.031	0.806	0.036	0.695	0.037	0.591	0.042	0.506

Table 4.3(a) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
25	5	$\overline{Min}_s$	0.068	0.278	0.076	0.237	0.075	0.217	0.078	0.209	0.078	0.178
		$\overline{LM}_s$	0.069	0.184	0.088	0.160	0.084	0.150	0.079	0.142	0.078	0.116
		$\overline{Max}_s$	0.062	0.293	0.063	0.226	0.060	0.221	0.062	0.191	0.060	0.171
		$\bar{t}_s$	0.061	0.171	0.075	0.146	0.068	0.132	0.064	0.125	0.061	0.100
		$\overline{WS}_s$	0.064	0.311	0.067	0.250	0.064	0.244	0.064	0.211	0.061	0.189
	10	$\overline{Min}_s$	0.070	0.424	0.067	0.344	0.077	0.309	0.077	0.262	0.073	0.229
		$\overline{LM}_s$	0.068	0.256	0.076	0.226	0.076	0.187	0.074	0.176	0.078	0.147
		$\overline{Max}_s$	0.061	0.483	0.052	0.371	0.056	0.334	0.052	0.285	0.051	0.256
		$\bar{t}_s$	0.062	0.265	0.062	0.223	0.062	0.187	0.057	0.172	0.056	0.144
		$\overline{WS}_s$	0.061	0.547	0.054	0.425	0.060	0.384	0.053	0.315	0.054	0.287
	25	$\overline{Min}_s$	0.055	0.736	0.058	0.612	0.065	0.532	0.064	0.444	0.075	0.366
		$\overline{LM}_s$	0.062	0.466	0.070	0.391	0.079	0.305	0.075	0.277	0.078	0.224
		$\overline{Max}_s$	0.049	0.855	0.040	0.700	0.041	0.630	0.038	0.540	0.036	0.470
		$\bar{t}_s$	0.058	0.541	0.060	0.434	0.057	0.353	0.055	0.320	0.053	0.251
		$\overline{WS}_s$	0.047	0.902	0.040	0.764	0.050	0.708	0.042	0.620	0.042	0.528
	50	$\overline{Min}_s$	0.058	0.943	0.058	0.853	0.064	0.766	0.071	0.680	0.075	0.559
		$\overline{LM}_s$	0.057	0.706	0.077	0.630	0.072	0.488	0.082	0.445	0.085	0.344
		$\overline{Max}_s$	0.053	0.993	0.035	0.923	0.040	0.880	0.032	0.811	0.035	0.749
		$\bar{t}_s$	0.054	0.822	0.061	0.703	0.052	0.582	0.054	0.530	0.056	0.411
		$\overline{WS}_s$	0.050	0.997	0.042	0.962	0.049	0.931	0.037	0.872	0.038	0.809
	100	$\overline{Min}_s$	0.053	0.996	0.063	0.985	0.073	0.958	0.067	0.900	0.077	0.780
		$\overline{LM}_s$	0.057	0.918	0.094	0.873	0.088	0.723	0.092	0.670	0.099	0.503
		$\overline{Max}_s$	0.050	1.000	0.029	0.999	0.030	0.993	0.024	0.976	0.025	0.942
		$\bar{t}_s$	0.059	0.977	0.064	0.934	0.055	0.836	0.055	0.783	0.056	0.659
		$\overline{WS}_s$	0.048	1.000	0.033	1.000	0.042	0.998	0.035	0.990	0.035	0.973

Table 4.3(a) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
50	5	$\overline{Min}_s$	0.073	0.683	0.072	0.584	0.076	0.535	0.074	0.490	0.070	0.443
		$\overline{LM}_s$	0.068	0.435	0.068	0.373	0.074	0.337	0.079	0.304	0.080	0.281
		$\overline{Max}_s$	0.058	0.728	0.059	0.618	0.061	0.566	0.057	0.509	0.058	0.474
		$\bar{t}_s$	0.057	0.444	0.059	0.364	0.062	0.327	0.063	0.296	0.064	0.268
		$\overline{WS}_s$	0.059	0.747	0.061	0.645	0.065	0.594	0.061	0.546	0.064	0.495
	10	$\overline{Min}_s$	0.068	0.908	0.071	0.838	0.074	0.783	0.073	0.727	0.073	0.671
		$\overline{LM}_s$	0.063	0.691	0.069	0.593	0.076	0.505	0.077	0.461	0.073	0.414
		$\overline{Max}_s$	0.060	0.953	0.055	0.899	0.059	0.860	0.058	0.812	0.055	0.750
		$\bar{t}_s$	0.056	0.748	0.057	0.650	0.057	0.560	0.058	0.495	0.059	0.453
		$\overline{WS}_s$	0.060	0.973	0.056	0.916	0.058	0.886	0.056	0.850	0.057	0.790
	25	$\overline{Min}_s$	0.061	0.999	0.062	0.996	0.063	0.988	0.061	0.972	0.062	0.947
		$\overline{LM}_s$	0.057	0.968	0.068	0.914	0.068	0.861	0.062	0.805	0.066	0.716
		$\overline{Max}_s$	0.054	1.000	0.047	1.000	0.046	0.997	0.043	0.993	0.046	0.982
		$\bar{t}_s$	0.054	0.986	0.055	0.960	0.055	0.930	0.048	0.880	0.055	0.824
		$\overline{WS}_s$	0.052	1.000	0.046	1.000	0.044	0.999	0.044	0.997	0.045	0.992
	50	$\overline{Min}_s$	0.060	1.000	0.060	1.000	0.059	1.000	0.052	1.000	0.054	0.998
		$\overline{LM}_s$	0.062	1.000	0.067	0.996	0.070	0.989	0.066	0.971	0.065	0.936
		$\overline{Max}_s$	0.055	1.000	0.041	1.000	0.043	1.000	0.039	1.000	0.036	1.000
		$\bar{t}_s$	0.058	1.000	0.056	1.000	0.055	0.998	0.050	0.995	0.056	0.986
		$\overline{WS}_s$	0.048	1.000	0.041	1.000	0.044	1.000	0.045	1.000	0.038	1.000
	100	$\overline{Min}_s$	0.061	1.000	0.057	1.000	0.062	1.000	0.062	1.000	0.062	1.000
		$\overline{LM}_s$	0.054	1.000	0.064	1.000	0.073	1.000	0.072	1.000	0.068	0.999
		$\overline{Max}_s$	0.059	1.000	0.039	1.000	0.042	1.000	0.035	1.000	0.038	1.000
		$\bar{t}_s$	0.046	1.000	0.053	1.000	0.055	1.000	0.053	1.000	0.055	1.000
		$\overline{WS}_s$	0.055	1.000	0.041	1.000	0.047	1.000	0.044	1.000	0.041	1.000

Table 4.3(a) (continued)

			$\lambda_i = 0$		λ_i˜$U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
75	5	$\overline{Min}_s$	0.072	0.939	0.073	0.873	0.080	0.828	0.074	0.805	0.081	0.752
		$\overline{LM}_s$	0.076	0.763	0.078	0.659	0.078	0.603	0.075	0.540	0.080	0.506
		$\overline{Max}_s$	0.063	0.962	0.061	0.911	0.068	0.875	0.061	0.838	0.064	0.788
		$\bar{t}_s$	0.063	0.780	0.063	0.665	0.063	0.610	0.062	0.541	0.063	0.504
		$\overline{WS}_s$	0.066	0.971	0.064	0.918	0.065	0.884	0.063	0.856	0.065	0.814
	10	$\overline{Min}_s$	0.069	0.999	0.074	0.993	0.074	0.984	0.069	0.974	0.074	0.952
		$\overline{LM}_s$	0.068	0.964	0.072	0.912	0.070	0.878	0.063	0.818	0.070	0.773
		$\overline{Max}_s$	0.055	1.000	0.057	0.998	0.060	0.994	0.056	0.989	0.054	0.978
		$\bar{t}_s$	0.058	0.984	0.060	0.941	0.058	0.912	0.054	0.861	0.059	0.821
		$\overline{WS}_s$	0.058	1.000	0.057	0.999	0.060	0.996	0.057	0.990	0.058	0.984
	25	$\overline{Min}_s$	0.062	1.000	0.065	1.000	0.067	1.000	0.061	1.000	0.062	1.000
		$\overline{LM}_s$	0.061	1.000	0.065	1.000	0.062	0.999	0.060	0.994	0.067	0.987
		$\overline{Max}_s$	0.053	1.000	0.048	1.000	0.050	1.000	0.046	1.000	0.046	1.000
		$\bar{t}_s$	0.051	1.000	0.053	1.000	0.050	1.000	0.048	0.999	0.056	0.996
		$\overline{WS}_s$	0.052	1.000	0.047	1.000	0.043	1.000	0.046	1.000	0.044	1.000
	50	$\overline{Min}_s$	0.061	1.000	0.062	1.000	0.063	1.000	0.059	1.000	0.056	1.000
		$\overline{LM}_s$	0.064	1.000	0.066	1.000	0.069	1.000	0.061	1.000	0.071	1.000
		$\overline{Max}_s$	0.055	1.000	0.047	1.000	0.049	1.000	0.046	1.000	0.046	1.000
		$\bar{t}_s$	0.056	1.000	0.052	1.000	0.054	1.000	0.052	1.000	0.060	1.000
		$\overline{WS}_s$	0.055	1.000	0.051	1.000	0.050	1.000	0.046	1.000	0.050	1.000
	100	$\overline{Min}_s$	0.058	1.000	0.060	1.000	0.069	1.000	0.060	1.000	0.064	1.000
		$\overline{LM}_s$	0.065	1.000	0.069	1.000	0.069	1.000	0.061	1.000	0.079	1.000
		$\overline{Max}_s$	0.050	1.000	0.042	1.000	0.048	1.000	0.042	1.000	0.044	1.000
		$\bar{t}_s$	0.058	1.000	0.054	1.000	0.059	1.000	0.053	1.000	0.065	1.000
		$\overline{WS}_s$	0.049	1.000	0.044	1.000	0.048	1.000	0.043	1.000	0.052	1.000

Table 4.3(a) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
100	5	$\overline{Min}_s$	0.077	0.995	0.074	0.981	0.074	0.966	0.075	0.947	0.076	0.922
		$\overline{LM}_s$	0.076	0.939	0.071	0.866	0.072	0.823	0.074	0.782	0.077	0.728
		$\overline{Max}_s$	0.066	0.998	0.060	0.989	0.062	0.980	0.059	0.964	0.062	0.944
		$\bar{t}_s$	0.065	0.949	0.063	0.885	0.059	0.833	0.065	0.801	0.062	0.739
		$\overline{WS}_s$	0.066	0.999	0.063	0.993	0.062	0.984	0.060	0.976	0.064	0.955
	10	$\overline{Min}_s$	0.072	1.000	0.067	1.000	0.064	1.000	0.061	0.999	0.060	0.998
		$\overline{LM}_s$	0.068	0.999	0.065	0.993	0.066	0.987	0.064	0.969	0.062	0.950
		$\overline{Max}_s$	0.057	1.000	0.055	1.000	0.051	1.000	0.045	1.000	0.048	1.000
		$\bar{t}_s$	0.054	1.000	0.054	0.998	0.053	0.995	0.051	0.987	0.050	0.973
		$\overline{WS}_s$	0.058	1.000	0.054	1.000	0.052	1.000	0.051	1.000	0.050	1.000
	25	$\overline{Min}_s$	0.063	1.000	0.060	1.000	0.056	1.000	0.055	1.000	0.059	1.000
		$\overline{LM}_s$	0.063	1.000	0.067	1.000	0.058	1.000	0.060	1.000	0.066	1.000
		$\overline{Max}_s$	0.052	1.000	0.048	1.000	0.044	1.000	0.046	1.000	0.045	1.000
		$\bar{t}_s$	0.058	1.000	0.058	1.000	0.053	1.000	0.052	1.000	0.057	1.000
		$\overline{WS}_s$	0.054	1.000	0.051	1.000	0.044	1.000	0.045	1.000	0.047	1.000
	50	$\overline{Min}_s$	0.060	1.000	0.055	1.000	0.051	1.000	0.051	1.000	0.056	1.000
		$\overline{LM}_s$	0.058	1.000	0.062	1.000	0.055	1.000	0.056	1.000	0.058	1.000
		$\overline{Max}_s$	0.053	1.000	0.049	1.000	0.046	1.000	0.046	1.000	0.044	1.000
		$\bar{t}_s$	0.053	1.000	0.053	1.000	0.051	1.000	0.048	1.000	0.050	1.000
		$\overline{WS}_s$	0.057	1.000	0.051	1.000	0.048	1.000	0.046	1.000	0.049	1.000
	100	$\overline{Min}_s$	0.065	1.000	0.065	1.000	0.059	1.000	0.060	1.000	0.059	1.000
		$\overline{LM}_s$	0.059	1.000	0.065	1.000	0.054	1.000	0.052	1.000	0.056	1.000
		$\overline{Max}_s$	0.062	1.000	0.053	1.000	0.051	1.000	0.047	1.000	0.048	1.000
		$\bar{t}_s$	0.056	1.000	0.057	1.000	0.050	1.000	0.045	1.000	0.046	1.000
		$\overline{WS}_s$	0.059	1.000	0.055	1.000	0.052	1.000	0.046	1.000	0.051	1.000

Table 4.3(b)

Empirical sizes and powers of panel data unit root tests at the nominal 5% level with fixed effects included, linear trend case

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k = 0$		$k = 1$		$k = 2$		$k = 3$		$k = 4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
15	5	$\overline{Min}_s$	0.079	0.098	0.065	0.070	0.061	0.068	0.068	0.075	0.074	0.077
		$\overline{LM}_s$	0.070	0.077	0.076	0.083	0.070	0.085	0.073	0.079	0.079	0.086
		$\overline{Max}_s$	0.075	0.095	0.058	0.064	0.056	0.060	0.062	0.068	0.053	0.052
		$\bar{t}_s$	0.069	0.076	0.073	0.087	0.069	0.082	0.073	0.079	0.068	0.066
		$\overline{WS}_s$	0.078	0.099	0.074	0.086	0.068	0.079	0.070	0.079	0.069	0.064
	10	$\overline{Min}_s$	0.060	0.091	0.058	0.057	0.059	0.059	0.070	0.067	0.069	0.082
		$\overline{LM}_s$	0.063	0.080	0.067	0.073	0.071	0.080	0.068	0.084	0.070	0.086
		$\overline{Max}_s$	0.059	0.089	0.051	0.056	0.051	0.056	0.060	0.065	0.053	0.060
		$\bar{t}_s$	0.064	0.081	0.068	0.072	0.074	0.086	0.072	0.086	0.066	0.076
		$\overline{WS}_s$	0.066	0.095	0.069	0.083	0.064	0.071	0.069	0.071	0.059	0.064
	25	$\overline{Min}_s$	0.054	0.101	0.043	0.049	0.052	0.059	0.072	0.076	0.073	0.085
		$\overline{LM}_s$	0.056	0.092	0.065	0.070	0.066	0.077	0.067	0.075	0.080	0.091
		$\overline{Max}_s$	0.051	0.096	0.042	0.049	0.042	0.051	0.058	0.064	0.050	0.052
		$\bar{t}_s$	0.058	0.094	0.068	0.079	0.067	0.082	0.067	0.070	0.068	0.076
		$\overline{WS}$	0.053	0.099	0.067	0.083	0.064	0.085	0.065	0.083	0.050	0.056
	50	$\overline{Min}$	0.058	0.129	0.036	0.052	0.040	0.055	0.063	0.072	0.074	0.108
		$\overline{LM}$	0.054	0.097	0.069	0.097	0.066	0.083	0.073	0.098	0.091	0.124
		$\overline{Max}$	0.055	0.117	0.029	0.037	0.031	0.036	0.047	0.061	0.047	0.056
		$\bar{t}$	0.054	0.099	0.064	0.094	0.058	0.085	0.076	0.095	0.084	0.096
		$\overline{WS}_s$	0.059	0.125	0.063	0.114	0.059	0.095	0.057	0.086	0.044	0.047
	100	$\overline{Min}_s$	0.057	0.171	0.028	0.048	0.035	0.058	0.053	0.081	0.086	0.123
		$\overline{LM}_s$	0.053	0.126	0.069	0.102	0.061	0.094	0.076	0.111	0.108	0.153
		$\overline{Max}_s$	0.057	0.177	0.021	0.037	0.024	0.036	0.036	0.048	0.039	0.045
		$\bar{t}_s$	0.057	0.125	0.063	0.098	0.055	0.081	0.067	0.095	0.079	0.094
		$\overline{WS}_s$	0.057	0.180	0.063	0.137	0.058	0.102	0.056	0.088	0.037	0.050

Table 4.3(b) (continued)

			$\lambda_i = 0$		$\lambda_i \tilde{} U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
25	5	$\overline{Min}_s$	0.069	0.114	0.074	0.097	0.071	0.092	0.069	0.094	0.073	0.097
		$\overline{LM}_s$	0.067	0.105	0.085	0.098	0.079	0.104	0.079	0.098	0.076	0.089
		$\overline{Max}_s$	0.065	0.110	0.067	0.086	0.062	0.083	0.059	0.077	0.062	0.078
		$\bar{t}_s$	0.061	0.100	0.076	0.089	0.071	0.097	0.071	0.087	0.066	0.077
		$\overline{WS}_s$	0.066	0.111	0.072	0.111	0.072	0.101	0.066	0.085	0.066	0.082
	10	$\overline{Min}_s$	0.073	0.148	0.059	0.088	0.058	0.091	0.060	0.088	0.069	0.093
		$\overline{LM}_s$	0.068	0.113	0.067	0.095	0.074	0.099	0.067	0.087	0.080	0.094
		$\overline{Max}_s$	0.066	0.142	0.050	0.083	0.051	0.074	0.050	0.074	0.057	0.077
		$\bar{t}_s$	0.065	0.111	0.062	0.089	0.068	0.089	0.062	0.083	0.071	0.085
		$\overline{WS}_s$	0.068	0.145	0.058	0.113	0.061	0.106	0.058	0.091	0.060	0.084
	25	$\overline{Min}_s$	0.057	0.206	0.051	0.120	0.051	0.101	0.052	0.096	0.055	0.101
		$\overline{LM}_s$	0.054	0.147	0.068	0.130	0.065	0.114	0.063	0.106	0.068	0.099
		$\overline{Max}_s$	0.053	0.212	0.045	0.113	0.039	0.084	0.042	0.082	0.040	0.083
		$\bar{t}_s$	0.052	0.149	0.060	0.124	0.055	0.100	0.057	0.100	0.058	0.085
		$\overline{WS}_s$	0.053	0.226	0.055	0.159	0.051	0.132	0.051	0.105	0.048	0.087
	50	$\overline{Min}_s$	0.055	0.305	0.045	0.160	0.046	0.132	0.048	0.126	0.053	0.120
		$\overline{LM}_s$	0.054	0.199	0.065	0.181	0.065	0.149	0.068	0.141	0.067	0.128
		$\overline{Max}_s$	0.054	0.333	0.039	0.141	0.036	0.115	0.035	0.102	0.039	0.097
		$\bar{t}_s$	0.053	0.207	0.060	0.171	0.057	0.135	0.058	0.129	0.056	0.114
		$\overline{WS}_s$	0.053	0.345	0.054	0.226	0.048	0.183	0.046	0.144	0.044	0.117
	100	$\overline{Min}_s$	0.057	0.501	0.048	0.255	0.044	0.194	0.044	0.162	0.044	0.137
		$\overline{LM}_s$	0.050	0.331	0.078	0.280	0.071	0.209	0.071	0.176	0.069	0.154
		$\overline{Max}_s$	0.055	0.554	0.034	0.225	0.031	0.153	0.033	0.125	0.028	0.104
		$\bar{t}_s$	0.051	0.352	0.067	0.250	0.056	0.174	0.059	0.153	0.059	0.126
		$\overline{WS}_s$	0.052	0.575	0.053	0.375	0.048	0.271	0.044	0.190	0.036	0.136

Table 4.3(b) (continued)

			$\lambda_i = 0$		λ_i ~ $U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
50	5	$\overline{Min}_s$	0.073	0.297	0.068	0.199	0.073	0.192	0.065	0.167	0.066	0.173
		$\overline{LM}_s$	0.077	0.223	0.073	0.162	0.074	0.147	0.066	0.151	0.073	0.140
		$\overline{Max}_s$	0.066	0.290	0.059	0.190	0.062	0.174	0.054	0.155	0.055	0.148
		$\bar{t}_s$	0.071	0.213	0.065	0.153	0.066	0.136	0.058	0.135	0.059	0.126
		$\overline{WS}_s$	0.066	0.308	0.062	0.237	0.066	0.213	0.057	0.177	0.058	0.172
	10	$\overline{Min}_s$	0.058	0.440	0.066	0.307	0.069	0.276	0.059	0.244	0.058	0.234
		$\overline{LM}_s$	0.062	0.307	0.068	0.231	0.063	0.201	0.061	0.185	0.064	0.172
		$\overline{Max}_s$	0.055	0.452	0.058	0.302	0.056	0.272	0.052	0.236	0.049	0.223
		$\bar{t}_s$	0.057	0.305	0.061	0.226	0.058	0.189	0.054	0.175	0.054	0.162
		$\overline{WS}_s$	0.054	0.480	0.056	0.377	0.059	0.318	0.053	0.269	0.052	0.239
	25	$\overline{Min}_s$	0.062	0.772	0.059	0.579	0.061	0.520	0.052	0.416	0.050	0.383
		$\overline{LM}_s$	0.057	0.581	0.066	0.453	0.064	0.369	0.060	0.288	0.063	0.269
		$\overline{Max}_s$	0.062	0.824	0.051	0.607	0.050	0.536	0.042	0.423	0.040	0.382
		$\bar{t}_s$	0.054	0.614	0.062	0.469	0.057	0.365	0.054	0.297	0.054	0.262
		$\overline{WS}_s$	0.060	0.846	0.053	0.707	0.055	0.621	0.045	0.508	0.046	0.441
	50	$\overline{Min}_s$	0.053	0.965	0.052	0.848	0.053	0.773	0.041	0.666	0.046	0.597
		$\overline{LM}_s$	0.058	0.856	0.066	0.689	0.063	0.572	0.053	0.461	0.063	0.423
		$\overline{Max}_s$	0.051	0.981	0.041	0.870	0.043	0.799	0.032	0.691	0.033	0.628
		$\bar{t}_s$	0.055	0.887	0.060	0.716	0.054	0.596	0.046	0.489	0.052	0.435
		$\overline{WS}_s$	0.050	0.987	0.047	0.927	0.048	0.881	0.037	0.766	0.040	0.696
	100	$\overline{Min}_s$	0.054	1.000	0.051	0.983	0.055	0.962	0.036	0.896	0.043	0.852
		$\overline{LM}_s$	0.055	0.986	0.070	0.915	0.065	0.815	0.053	0.710	0.067	0.641
		$\overline{Max}_s$	0.050	1.000	0.040	0.988	0.041	0.974	0.029	0.925	0.033	0.873
		$\bar{t}_s$	0.053	0.994	0.062	0.940	0.057	0.845	0.049	0.743	0.056	0.676
		$\overline{WS}_s$	0.051	1.000	0.046	0.997	0.049	0.989	0.033	0.956	0.040	0.924

Table 4.3(b) (continued)

			$\lambda_i = 0$		λ_i ~ $U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
75	5	$\overline{Min}_s$	0.075	0.567	0.058	0.391	0.067	0.365	0.071	0.347	0.070	0.315
		$\overline{LM}_s$	0.071	0.426	0.071	0.292	0.074	0.274	0.069	0.250	0.079	0.223
		$\overline{Max}_s$	0.066	0.573	0.050	0.391	0.056	0.364	0.059	0.335	0.059	0.307
		$\bar{t}_s$	0.064	0.419	0.062	0.284	0.064	0.257	0.058	0.234	0.067	0.204
		$\overline{WS}_s$	0.068	0.592	0.059	0.469	0.061	0.418	0.062	0.383	0.064	0.355
	10	$\overline{Min}_s$	0.067	0.822	0.064	0.660	0.064	0.623	0.066	0.551	0.065	0.514
		$\overline{LM}_s$	0.064	0.664	0.068	0.493	0.066	0.437	0.065	0.392	0.073	0.360
		$\overline{Max}_s$	0.060	0.849	0.055	0.680	0.057	0.634	0.055	0.564	0.057	0.518
		$\bar{t}_s$	0.056	0.675	0.061	0.501	0.056	0.439	0.055	0.393	0.065	0.356
		$\overline{WS}_s$	0.058	0.867	0.060	0.767	0.064	0.713	0.062	0.639	0.061	0.581
	25	$\overline{Min}_s$	0.055	0.992	0.053	0.954	0.054	0.926	0.051	0.892	0.053	0.830
		$\overline{LM}_s$	0.054	0.954	0.065	0.848	0.065	0.776	0.064	0.701	0.071	0.637
		$\overline{Max}_s$	0.050	0.998	0.042	0.970	0.045	0.942	0.042	0.913	0.042	0.866
		$\bar{t}_s$	0.056	0.974	0.057	0.882	0.054	0.810	0.055	0.737	0.059	0.662
		$\overline{WS}_s$	0.050	1.000	0.055	0.989	0.052	0.971	0.053	0.950	0.045	0.905
	50	$\overline{Min}_s$	0.054	1.000	0.049	0.999	0.049	0.996	0.046	0.993	0.047	0.979
		$\overline{LM}_s$	0.056	0.999	0.070	0.984	0.064	0.958	0.063	0.937	0.066	0.896
		$\overline{Max}_s$	0.051	1.000	0.041	1.000	0.041	0.999	0.037	0.998	0.037	0.990
		$\bar{t}_s$	0.052	1.000	0.059	0.990	0.056	0.972	0.055	0.953	0.054	0.918
		$\overline{WS}_s$	0.048	1.000	0.051	1.000	0.046	1.000	0.044	1.000	0.045	0.997
	100	$\overline{Min}_s$	0.052	1.000	0.043	1.000	0.041	1.000	0.041	1.000	0.041	1.000
		$\overline{LM}_s$	0.055	1.000	0.071	1.000	0.062	1.000	0.061	0.999	0.071	0.993
		$\overline{Max}_s$	0.051	1.000	0.033	1.000	0.032	1.000	0.028	1.000	0.031	1.000
		$\bar{t}_s$	0.053	1.000	0.060	1.000	0.049	1.000	0.051	1.000	0.057	0.997
		$\overline{WS}_s$	0.048	1.000	0.046	1.000	0.043	1.000	0.040	1.000	0.040	1.000

Table 4.3(b) (continued)

			$\lambda_i = 0$		λ_i˜$U(0.2, 0.4)$							
			$k=0$		$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
100	5	$\overline{Min}_s$	0.075	0.803	0.060	0.617	0.069	0.597	0.078	0.568	0.075	0.519
		$\overline{LM}_s$	0.066	0.672	0.061	0.496	0.073	0.452	0.077	0.418	0.076	0.393
		$\overline{Max}_s$	0.068	0.818	0.054	0.631	0.061	0.594	0.066	0.564	0.065	0.516
		$\bar{t}_s$	0.059	0.676	0.055	0.487	0.063	0.447	0.065	0.412	0.064	0.379
		$\overline{WS}_s$	0.067	0.827	0.061	0.721	0.068	0.674	0.072	0.627	0.069	0.574
	10	$\overline{Min}_s$	0.059	0.981	0.051	0.899	0.059	0.877	0.061	0.839	0.061	0.786
		$\overline{LM}_s$	0.063	0.928	0.062	0.774	0.064	0.725	0.068	0.687	0.071	0.614
		$\overline{Max}_s$	0.056	0.987	0.043	0.917	0.051	0.892	0.051	0.855	0.054	0.805
		$\bar{t}_s$	0.054	0.939	0.053	0.796	0.053	0.741	0.060	0.692	0.062	0.625
		$\overline{WS}_s$	0.056	0.990	0.055	0.962	0.054	0.934	0.056	0.901	0.058	0.858
	25	$\overline{Min}_s$	0.056	1.000	0.053	1.000	0.051	1.000	0.053	0.998	0.049	0.990
		$\overline{LM}_s$	0.056	1.000	0.056	0.995	0.054	0.976	0.058	0.956	0.057	0.923
		$\overline{Max}_s$	0.050	1.000	0.044	1.000	0.044	1.000	0.045	1.000	0.043	0.997
		$\bar{t}_s$	0.051	1.000	0.051	0.998	0.050	0.988	0.053	0.971	0.050	0.943
		$\overline{WS}_s$	0.053	1.000	0.053	1.000	0.050	1.000	0.051	1.000	0.050	1.000
	50	$\overline{Min}_s$	0.052	1.000	0.045	1.000	0.047	1.000	0.049	1.000	0.045	1.000
		$\overline{LM}_s$	0.048	1.000	0.058	1.000	0.053	1.000	0.058	1.000	0.058	0.995
		$\overline{Max}_s$	0.051	1.000	0.040	1.000	0.040	1.000	0.039	1.000	0.039	1.000
		$\bar{t}_s$	0.048	1.000	0.052	1.000	0.046	1.000	0.052	1.000	0.050	0.999
		$\overline{WS}_s$	0.051	1.000	0.046	1.000	0.042	1.000	0.044	1.000	0.045	1.000
	100	$\overline{Min}_s$	0.053	1.000	0.047	1.000	0.046	1.000	0.044	1.000	0.041	1.000
		$\overline{LM}_s$	0.052	1.000	0.059	1.000	0.055	1.000	0.061	1.000	0.053	1.000
		$\overline{Max}_s$	0.051	1.000	0.041	1.000	0.039	1.000	0.037	1.000	0.033	1.000
		$\bar{t}_s$	0.050	1.000	0.051	1.000	0.047	1.000	0.054	1.000	0.047	1.000
		$\overline{WS}_s$	0.055	1.000	0.047	1.000	0.045	1.000	0.044	1.000	0.039	1.000

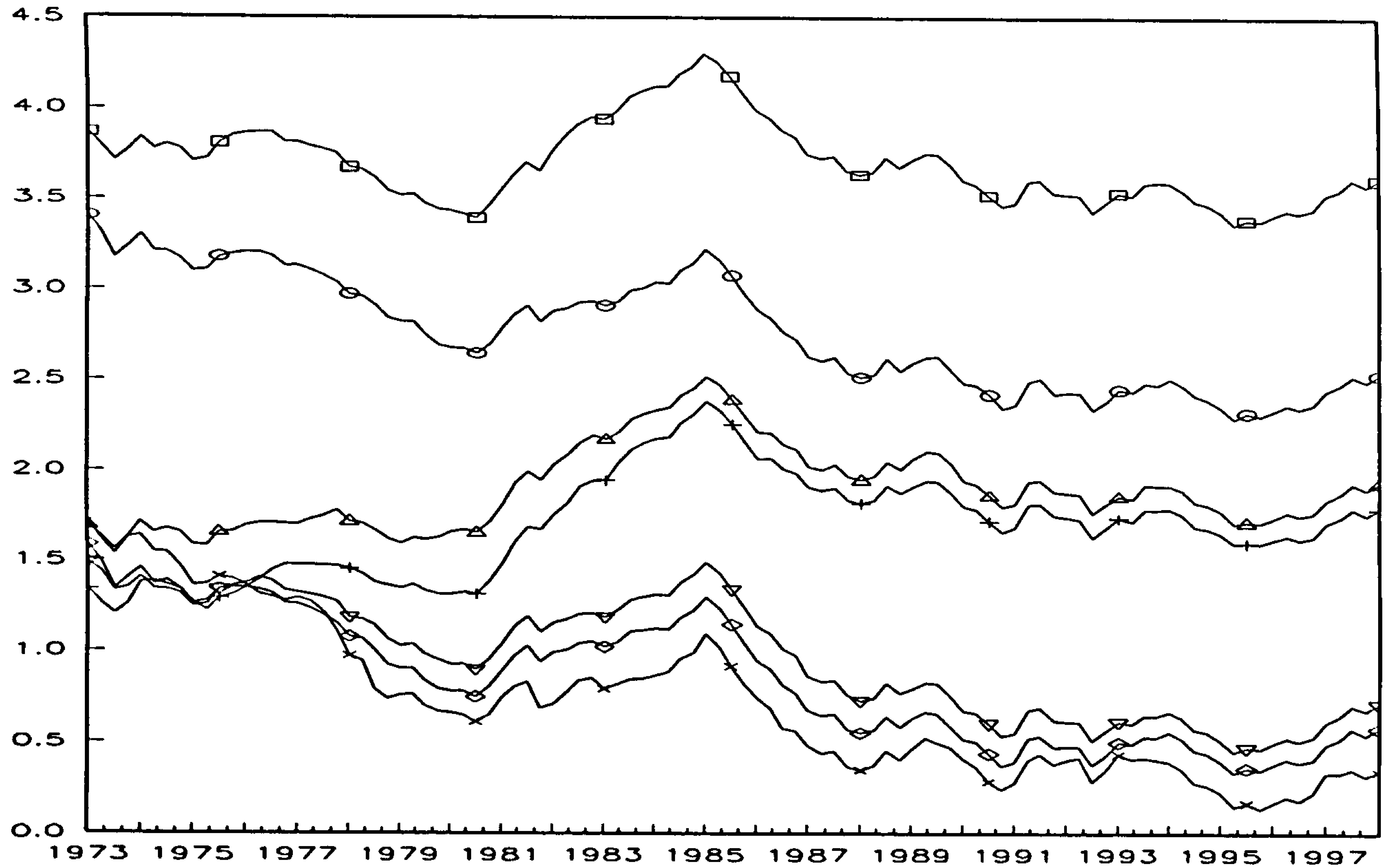


Figure 4.2. Real exchange rates against the US dollar.

□ BEL ∘ AUT △ NOR + FR ▽ NTH ◇ GER × DEN

Table 4.4

Correlation matrix of ADF residuals of the real exchange rates against the US dollar

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
C1	1.00																
C2	0.42	1.00															
C3	0.10	0.44	1.00														
C4	0.18	0.45	0.90	1.00													
C5	0.30	0.20	0.23	0.24	1.00												
C6	0.18	0.45	0.89	0.99	0.24	1.00											
C7	0.23	0.43	0.70	0.70	0.33	0.67	1.00										
C8	0.18	0.47	0.90	0.99	0.24	0.99	0.71	1.00									
C9	0.14	0.49	0.98	0.91	0.23	0.91	0.72	0.92	1.00								
C10	0.37	0.37	0.64	0.73	0.25	0.71	0.81	0.76	0.68	1.00							
C11	0.01	0.49	0.50	0.53	0.16	0.49	0.32	0.49	0.50	0.16	1.00						
C12	0.10	0.48	0.97	0.92	0.21	0.92	0.71	0.93	0.98	0.67	0.51	1.00					
C13	0.23	0.38	0.84	0.77	0.32	0.77	0.74	0.78	0.87	0.73	0.29	0.86	1.00				
C14	0.31	0.36	0.71	0.83	0.33	0.82	0.81	0.84	0.73	0.87	0.28	0.75	0.72	1.00			
C15	0.30	0.23	0.64	0.71	0.25	0.68	0.79	0.72	0.67	0.88	0.12	0.63	0.73	0.76	1.00		
C16	0.13	0.49	0.88	0.86	0.23	0.85	0.72	0.88	0.91	0.68	0.52	0.91	0.74	0.72	0.67	1.00	
C17	0.33	0.38	0.58	0.66	0.27	0.65	0.77	0.68	0.63	0.81	0.33	0.63	0.74	0.76	0.72	0.63	1.00

Note: C1-C17 correspond to the countries arranged in the order specified in Section 4.6.

Table 4.5

p-values for bootstrap panel data unit root tests applied to 17 quarterly real exchange rates ($T = 45$)

	Test				
Block Size	$\overline{t}^*$	$\overline{Max}^*$	$\overline{WS}^*$	$\overline{LM}^*$	$\overline{Min}^*$
30	0.124	0.036	0.023	0.103	0.034
100	0.123	0.033	0.020	0.104	0.029

Note: Estimates are based on 5000 replications.

Table 4.6

Empirical sizes and powers of bootstrap panel data unit root tests at the nominal 5% level

$\rho_i = 0.8\ \forall i$			$\lambda_i\tilde{}U(0.2, 0.4)$							
			$k=1$		$k=2$		$k=3$		$k=4$	
T	N		Size	Power	Size	Power	Size	Power	Size	Power
25	5	$\overline{Min}^*$	0.044	0.400	0.060	0.352	0.052	0.264	0.050	0.208
		$\overline{LM}^*$	0.053	0.236	0.049	0.185	0.056	0.163	0.040	0.113
		$\overline{Max}^*$	0.045	0.452	0.055	0.400	0.049	0.302	0.053	0.271
		$\bar{t}^*$	0.050	0.250	0.050	0.204	0.050	0.174	0.045	0.118
		$\overline{WS}^*$	0.045	0.460	0.056	0.430	0.054	0.318	0.052	0.284
	10	$\overline{Min}^*$	0.055	0.691	0.046	0.558	0.052	0.467	0.043	0.351
		$\overline{LM}^*$	0.046	0.400	0.056	0.305	0.040	0.252	0.035	0.167
		$\overline{Max}^*$	0.051	0.777	0.041	0.665	0.035	0.541	0.047	0.470
		$\bar{t}^*$	0.044	0.443	0.056	0.336	0.043	0.285	0.036	0.197
		$\overline{WS}^*$	0.057	0.813	0.051	0.703	0.037	0.574	0.044	0.507
	25	$\overline{Min}^*$	0.027	0.935	0.034	0.860	0.036	0.715	0.042	0.598
		$\overline{LM}^*$	0.021	0.670	0.030	0.532	0.031	0.377	0.030	0.309
		$\overline{Max}^*$	0.017	0.965	0.031	0.921	0.036	0.826	0.042	0.731
		$\bar{t}^*$	0.016	0.720	0.028	0.595	0.027	0.421	0.028	0.352
		$\overline{WS}^*$	0.018	0.978	0.027	0.935	0.038	0.859	0.042	0.792
50	5	$\overline{Min}^*$	0.053	0.922	0.049	0.861	0.051	0.789	0.057	0.722
		$\overline{LM}^*$	0.047	0.745	0.044	0.625	0.059	0.526	0.055	0.457
		$\overline{Max}^*$	0.055	0.952	0.046	0.913	0.047	0.858	0.053	0.792
		$\bar{t}^*$	0.051	0.786	0.045	0.672	0.056	0.563	0.056	0.505
		$\overline{WS}^*$	0.048	0.962	0.055	0.918	0.048	0.872	0.055	0.825
	10	$\overline{Min}^*$	0.044	0.998	0.049	0.993	0.037	0.969	0.040	0.941
		$\overline{LM}^*$	0.049	0.963	0.058	0.912	0.043	0.841	0.044	0.709
		$\overline{Max}^*$	0.044	1.000	0.054	0.999	0.041	0.988	0.038	0.978
		$\bar{t}^*$	0.046	0.974	0.055	0.944	0.045	0.878	0.040	0.781
		$\overline{WS}^*$	0.045	1.000	0.052	0.997	0.041	0.990	0.038	0.980
	25	$\overline{Min}^*$	0.036	1.000	0.037	1.000	0.042	1.000	0.046	1.000
		$\overline{LM}^*$	0.039	1.000	0.035	1.000	0.040	0.996	0.043	0.976
		$\overline{Max}^*$	0.030	1.000	0.028	1.000	0.036	1.000	0.033	1.000
		$\bar{t}^*$	0.033	1.000	0.027	1.000	0.043	0.999	0.037	0.990
		$\overline{WS}^*$	0.031	1.000	0.029	1.000	0.036	1.000	0.031	1.000

Chapter 5

On the Power and Efficiency of Unit Root Tests Based on Generalised Least Squares Detrending

5.1 Introduction

Due to the well-known power deficiency of commonly applied unit root tests, the last decade has seen efforts being directed toward the development of tests with greater power. This direction of research led Elliott, Rothenberg and Stock (1996) - henceforth ERS - to study the asymptotic power envelope for various unit root tests and propose a simple modification of the *ADF* test such that the modified test, which we will refer to as the *GLS* test, can nearly achieve the power envelope. In particular, they consider the time series process $\{y_t\}_{t=1}^{\infty}$ generated by

$$\begin{aligned} y_t &= z_t'\beta + u_t \qquad t = 1, 2, ..., T \\ u_t &= \rho u_{t-1} + v_t \end{aligned}$$

where either $z_t = 1$ or $z_t = (1, t)$, β is a conformable vector of unknown parameters and v_t is a potentially serially correlated stationary process. Based on the theory of point optimal testing, they analyse the sequence of Neyman-Pearson tests of the unit root null hypothesis, $\rho = 1$, against the local alternative of $\overline{\rho} = 1 + \overline{c}/T$ for $u_0 = 0$, where T is the sample size and $\overline{c} < 0$ is some constant associated with the power envelope. In particular, $\overline{c}$ is chosen so that the tests achieve the Gaussian power envelope at 50 percent power, which corresponds to setting $\overline{c} = -7$ in the demeaned case and $\overline{c} = -13.5$ in the detrended case. The *GLS* test is computed in two steps. Initially β is estimated by regressing $[y_1, (1-\overline{\rho}L)y_2, ..., (1-\overline{\rho}L)y_T]$ onto $[z_1, (1-\overline{\rho}L)z_2, ..., (1-\overline{\rho}L)z_T]$, with the resulting estimator denoted by $\widetilde{\beta}_{GLS}$. Then the local *GLS*-detrended series $\widetilde{y}_t = y_t - z_t'\widetilde{\beta}_{GLS}$ is used to compute the t-statistic in the following regression

$$\Delta\widetilde{y}_t = \varphi\widetilde{y}_{t-1} + \sum_{j=1}^{p} \delta_j \, \Delta\widetilde{y}_{t-j} + \varepsilon_t \tag{5.1}$$

where incorporation of the lags $\Delta\widetilde{y}_{t-j}$ in (5.1) is to account for any potential dynamics present in the error term v_t.

Elliott (1999) derives the asymptotic power envelope under the more natural stationarity alternative where the initial observation is drawn from the unconditional distribution with variance $\frac{\sigma^2}{1-\rho^2}$, where σ^2 is the variance of v_t, often referred to as the unconditional case. Generalised least square detrending under the local alternative in this case, referred to as *GLSu* detrending, would generate $\widetilde{y}_t$ as the residuals from the regression of $[(1-\overline{\rho}^2)^{1/2}y_1, (1-\overline{\rho}L)y_2, ..., (1-\overline{\rho}L)y_T]'$ on $[(1-\overline{\rho}^2)^{1/2}z_1, (1-\overline{\rho}L)z_2, ..., (1-\overline{\rho}L)z_T]'$. The resultant unit root test is then the *GLSu* test. While Elliott experimented with a number of values for $\overline{c}$, he finds that the preferred value both when a constant only and when a constant and a linear trend term are included in the model is $\overline{c} = -10$.

The above studies provide Monte Carlo evidence on the size and power performance of the modified *DF* tests. They report asymptotic and finite sample power results based on $T = 100$, the latter with main focus on the $MA(1)$ specification of the errors v_t, both when the initial observation is fixed and when it is drawn from the unconditional model. Use of the Bayesian information criterion (BIC) is made in selecting the lag length of the autoregression for estimation. Limited results are tabulated for the $AR(1)$ specification of the errors v_t by ERS who do not, however, consider the strictly stationary alternative under such a specification.

ERS expected the modified estimates of the trend parameters to improve the power performance of the standard *DF* tests. Evidence supporting this expectation comes from Phillips and Lee (1996) and Canjels and Watson (1997), who investigate the efficient extraction of deterministic trends in the case of near unit root non-stationary time series. They show that efficiency gains in the estimation of deterministic trends can be obtained by *GLS* detrending under the alternative using the unknown localizing parameter, and they quantify these gains.[1] In related work, Burridge and Taylor (2000) explore the source of the increased power performance of the *GLS* type tests relative to the *DF* test. They present Monte Carlo results on the power of the tests and the relative efficiencies of the mean estimators under *GLS* and *GLSu* detrending for a sample size of $T = 100$, when $z_t = 1$ and v_t follows an $MA(1)$ process in both the fixed initial observation case and the unconditional case.[2] Their simulation results, reported for values of the *MA* parameter not too close to the boundary[3], show that the power advantage of the tests based on *GLS* detrending cannot be solely attributable to the efficiency of the estimators used in constructing the test statistics as suggested by ERS. They find that the complete explanation lies in the shift of the null distribution of the *GLS* type tests closer to the origin, relative to that of the *DF* test, as opposed to a smaller shift of the distribution under the alternative. For later reference we will refer to this effect as the location effect.

A number of empirical studies have investigated the integration properties of data using the *GLS* univariate unit root tests of ERS, see among others Siklos and Granger (1997), Cheung and Lai (1998), Banerjee and Russell (2001), Cushman (2001), Barkoulas *et al.* (2002), McNown and Puttitanun (2002) and Rapach and Wohar (2002). Perhaps their wide use as opposed to other power-enhancing modifications of the Dickey-Fuller test that are available due to Park and Fuller (1995) and Leybourne (1995), namely the *WS* and *MAX* tests, resides in the unknown optimality properties of the latter as pointed out by Müller and Elliott (2001). This issue, however, has recently been addressed by Leybourne *et al.*

[1] The treatment of Phillips and Lee (1996) allows for general polynomial trends, while Canjels and Watson (1997) study the case of a linear trend extraction.

[2] They calculate the variances of the *OLS*, *GLS* and *GLSu* estimators as $(X'X)^{-1}X'\Omega X(X'X)^{-1}$, where Ω is the variance-covariance matrix of $[y_1, y_2, ..., y_T]'$, $[y_1, (1-\bar{\rho}L)y_2, ..., (1-\bar{\rho}L)y_T]'$ and $[(1-\bar{\rho}^2)^{1/2}y_1, (1-\bar{\rho}L)y_2, ..., (1-\bar{\rho}L)y_T]'$ for $\widehat{a}_{OLS}$, $\widehat{a}_{GLS}$ and $\widehat{a}_{GLSu}$, respectively and X is $\mathbf{1}'$, $[1, 1-\bar{\rho}, ..., 1-\bar{\rho}]'$ and $[(1-\bar{\rho}^2)^{1/2}, 1-\bar{\rho}, ..., 1-\bar{\rho}]'$, respectively.

[3] Their results are illustrated for values of the MA paramater θ between ± 0.7.

(2003), who find these alternative tests that explore the time reversibility of stationary autoregressive processes to have asymptotic local power very close to the envelope. Reported simulations based on finite samples clearly indicate their superior performance over the DF and GLS type tests, see also Pantula *et al.* (1994).

This work differs from that of Burridge and Taylor (2000), in that they investigate the direct connection between the power performance of the GLS based tests and the relative efficiencies of the mean estimators by utilising exact formulas for the latter based on the variance-covariance matrix of an $ARMA(1,1)$ model. In doing so, their aim is to disclose the source of increased power performance of these tests relative to the DF test. Thus, confronted with simulation evidence that postulates the GLS and $GLSu$ tests as having power greater than the DF test with a corresponding mean estimator that is less efficient than the OLS estimator, they devote a large part of their study in exploring the location effect they find to be associated with such tests. On the other hand, we attempt to identify cases whereby the GLS test performs poorly relative to the DF test by appealing to approximate theoretical results for the relative efficiencies of the mean estimator based on local to unity asymptotics. Focus is on the quality of the mean estimator per se, in order not to confound the efficiency issue with the location effect. We uncover significant finite sample effects. Burridge and Taylor hinted towards these, but we do a more comprehensive analysis for the $AR(2)$ model and also report large sample results. We concentrate purely on the autoregressive structure given the importance of the latter in characterizing key economic series and conjecture that our results hold for higher order autoregressive models.

For the purposes of our study we employ the values for $\bar{c}$ proposed by ERS (1996) and Elliott (1999) for the GLS and $GLSu$ tests, respectively, which are those most often used in the literature. Accommodating changes in the value of $\bar{c}$ will affect the critical values of the GLS and $GLSu$ tests, which in turn will affect their power performance. We briefly consider such cases towards the end of the chapter.

The remainder of this chapter is organised as follows. In Section 5.2 we obtain simple, useful approximations to the relative efficiencies of the trend estimators, $\widetilde{\beta}$, under both GLS and GLSu detrending when u_t is characterized by a stationary general linear process. The expressions that arise are then applied to individual autoregressive models and their limiting values are calculated. While our results are confined to the case of a deterministic function that includes only a constant, they can easily be extended to allow for a trend term as well. Section 5.3 contains large and finite sample size and power results based on an extensive simulation study tailored to the case of a second order autoregressive model, where the relative efficiencies of the mean estimators are calculated according to the derived formulas in Section 5.2. Exact results of the relative efficiencies of the mean estimators are also tabulated in order to assess the quality of the approximation. Standard $AR(1)$ model results are included for purposes of completeness. Section 5.4 presents some concluding remarks.

5.2 Relative efficiencies of the mean estimators under general linear processes

Consider the time series $\{y_t\}$ generated by

$$y_t = a + u_t, \qquad t = 1, 2, ..., T, \tag{5.2}$$

where u_t is a linear stochastic process that has the possibly infinite moving average representation given by

$$u_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(L)\varepsilon_t \tag{5.3}$$

with $\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$, $\psi(1) = \sum_{j\geqslant 0} \psi_j \neq 0$, $\psi_0 = 1$ and ε_t being zero mean white noise with variance σ_ε^2. It follows that $E(u_t) = 0$. Under the additional assumption $\sum_{j=0}^{\infty} j\left|\psi_j\right| < \infty$, the infinite sequence in (5.3) generates any stationary finite-order $ARMA$ model.

Define

$$\gamma_j = cov(u_t, u_{t-j}) = E(u_t u_{t-j})$$

The autocovariance generating function of u_t is

$$g_u(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j = \sigma_\varepsilon^2 \psi(z)\psi(z^{-1}) \tag{5.4}$$

and it is well known (Hamilton p.188) that $\lim_{T\to\infty}[TE(\overline{u}_T - E(\overline{u}_T))^2] = \sum_{j=-\infty}^{\infty} \gamma_j$, where $\overline{u}_T = T^{-1}(u_1 + u_2 + ...u_T)$, so that, as an approximation $var(\sum_{t=1}^{T} u_t) = T \sum_{j=-\infty}^{\infty} \gamma_j$.

Then, setting $z = 1$ in (5.4) we have

$$\sum_{j=-\infty}^{\infty} \gamma_j = \sigma_\varepsilon^2[\psi(1)]^2$$

and

$$var(\sum_{t=1}^{T} u_t) = T\sigma_\varepsilon^2[\psi(1)]^2. \tag{5.5}$$

Also, from (5.5) and $\gamma_{-j} = \gamma_j$ it follows that

$$\sum_{j=1}^{\infty} \gamma_j = \frac{1}{2}[\sigma_\varepsilon^2[\psi(1)]^2 - \gamma_0] \tag{5.6}$$

while

$$\gamma_0 = var(u_t) = \sigma_\varepsilon^2(1 + \sum_{j=1}^{\infty} \psi_j^2). \tag{5.7}$$

After setting out the preliminaries, we consider generalized least squares estimation of the mean a in (5.2). That is, a is estimated by regressing $[y_1, y_2 - \overline{\rho}y_1, y_3 - \overline{\rho}y_2, ..., y_T - \overline{\rho}y_{T-1}]$ on $[1, 1-\overline{\rho}, 1-\overline{\rho}, ..., 1-\overline{\rho}]$, where $\overline{\rho} = 1 + \overline{c}/T$ is the local alternative. We denote the resultant estimator by $\widehat{a}_{GLS}$, which can be written as

$$\widehat{a}_{GLS} = \frac{(1-\overline{\rho}+\overline{\rho}^2)y_1 + (1-\overline{\rho})^2 \sum_{t=2}^{T-1} y_t + (1-\overline{\rho})y_T}{1 + (T-1)(1-\overline{\rho})^2}$$

In what follows we retain T but still appeal to some "large T results". Calculating the variance of $\widehat{a}_{GLS}$ we obtain

$$var(\widehat{a}_{GLS}) = \frac{f(\overline{\rho})\gamma_0 + (1-\overline{\rho})^4 var(\sum_{t=2}^{T-1} u_t) + 2h(\overline{\rho}) \sum_{j=1}^{T-2} \gamma_j + 2g(\overline{\rho})\gamma_{T-1}}{[k(\overline{\rho}, T)]^2} \tag{5.8}$$

where

$$\begin{aligned} f(\overline{\rho}) &= (1-\overline{\rho}+\overline{\rho}^2)^2 + (1-\overline{\rho})^2 \\ h(\overline{\rho}) &= [(1-\overline{\rho}+\overline{\rho}^2) + (1-\overline{\rho})](1-\overline{\rho})^2 \\ g(\overline{\rho}) &= (1-\overline{\rho}+\overline{\rho}^2)(1-\overline{\rho}) \\ k(\overline{\rho}, T) &= 1 + (T-1)(1-\overline{\rho})^2. \end{aligned}$$

Assuming T sufficiently large so that $var(\sum_{t=2}^{T-1} u_t) = var(\sum_{t=1}^{T} u_t)$ and $\sum_{j=1}^{T-2} \gamma_j = \sum_{j=1}^{\infty} \gamma_j$, the variance of $\widehat{a}_{GLS}$ using (5.5) and (5.6) can be approximated by

$$var(\widehat{a}_{GLS}) \cong \frac{f(\overline{\rho})\gamma_0 + (1-\overline{\rho})^4 T\sigma_\varepsilon^2[\psi(1)] + h(\overline{\rho})[\sigma_\varepsilon^2[\psi(1)]^2 - \gamma_0]}{[k(\overline{\rho}, T)]^2} \tag{5.9}$$

where the last term in the numerator of (5.8) was omitted as being negligible. As the least squares estimator of a is simply the sample mean of y_t, it follows from (5.5) that

$$var(\widehat{a}_{OLS}) = T^{-1}\sigma_\varepsilon^2[\psi(1)]^2. \tag{5.10}$$

Then, by taking the variance ratio of the mean estimators in (5.9) and (5.10) we obtain

$$\frac{var(\widehat{a}_{GLS})}{var(\widehat{a}_{OLS})} \cong \frac{T^2(1-\overline{\rho})^4 + Th(\overline{\rho}) + T[f(\overline{\rho}) - h(\overline{\rho})]\gamma_0[\sigma_\varepsilon^2[\psi(1)]^2]^{-1}}{[k(\overline{\rho}, T)]^2}. \tag{5.11}$$

A similar expression can be obtained for $\widehat{a}_{GLS_u}$, in which case a is estimated by regressing $[y_1\sqrt{1-\overline{\rho}^2}, y_2 - \overline{\rho}y_1, y_3 - \overline{\rho}y_2, ..., y_T - \overline{\rho}y_{T-1}]$ on $[\sqrt{1-\overline{\rho}^2}, 1-\overline{\rho}, 1-$

$\overline{\rho}, ..., 1-\overline{\rho}]$,

$$\widehat{a}_{GLSu} = \frac{(1-\overline{\rho})y_1 + (1-\overline{\rho})^2 \sum_{t=2}^{T-1} y_t + (1-\overline{\rho})y_T}{(1-\overline{\rho}^2) + (T-1)(1-\overline{\rho})^2}.$$

The variance of $\widehat{a}_{GLSu}$ is given by

$$var(\widehat{a}_{GLSu}) = \frac{f_u(\overline{\rho})\gamma_0 + (1-\overline{\rho})^4 var(\sum_{t=2}^{T-1} u_t) + 2h_u(\overline{\rho}) \sum_{j=1}^{T-2} \gamma_j + f_u(\overline{\rho})\gamma_{T-1}}{[k_u(\overline{\rho}, T)]^2} \qquad (5.12)$$

where

$$\begin{aligned} f_u(\overline{\rho}) &= 2(1-\overline{\rho})^2 \\ h_u(\overline{\rho}) &= 2(1-\overline{\rho})^3 \\ k_u(\overline{\rho}, T) &= (1-\overline{\rho}^2) + (T-1)(1-\overline{\rho})^2. \end{aligned}$$

Then, as above by assuming T sufficiently large we can approximate the variance of $\widehat{a}_{GLSu}$ by

$$var(\widehat{a}_{GLSu}) \cong \frac{f_u(\overline{\rho})\gamma_0 + (1-\overline{\rho})^4 T\sigma_\varepsilon^2[\psi(1)] + h_u(\overline{\rho})[\sigma_\varepsilon^2[\psi(1)]^2 - \gamma_0]}{[k_u(\overline{\rho}, T)]^2}.$$

Taking the variance ratio of the $\widehat{a}_{GLS_u}$ and $\widehat{a}_{OLS}$ estimators we obtain

$$\frac{var(\widehat{a}_{GLS_u})}{var(\widehat{a}_{OLS})} \cong \frac{T^2(1-\overline{\rho})^4 + Th_u(\overline{\rho}) + T[f_u(\overline{\rho}) - h_u(\overline{\rho})]\gamma_0[\sigma_\varepsilon^2[\psi(1)]^2]^{-1}}{[k_u(\overline{\rho}, T)]^2}. \qquad (5.13)$$

Expressions (5.11) and (5.13) give a useful approximation to the relative efficiencies of the GLS and $GLSu$ estimators of the mean under the general linear process that governs the error term of the time series y_t in (5.2). For T and $\overline{\rho}$ fixed, (5.11) and (5.13) depend on the model parameters only through its final term. If $[f(\overline{\rho}) - h(\overline{\rho})] > 0$ and $[f_u(\overline{\rho}) - h_u(\overline{\rho})] > 0$, these variance ratios are an increasing function of the ratio $\frac{\gamma_0}{\sigma_\varepsilon^2[\psi(1)]^2}$, which from (5.7) and assuming $\sigma_\varepsilon^2 = 1$ throughout without loss of generality, can be written as $\frac{1+\sum_{j=1}^{\infty}\psi_j^2}{[1+\sum_{j=1}^{\infty}\psi_j]^2}$. This latter ratio is that of the short-run to long-run variance of u_t denoted by $Q = V_s/V_l$.

Substituting for $\overline{\rho} = 1 + \frac{\overline{c}}{T}$ in (5.11) and (5.13), we then have

$$\frac{var(\widehat{a}_{GLS})}{var(\widehat{a}_{OLS})} \cong \frac{\overline{c}^2T(\overline{c}^2T + T^2 + \overline{c}^2) + TQ(T + \overline{c})T(T^2 + \overline{c}T + 2\overline{c}^2)}{(T^2 + \overline{c}^2T - \overline{c}^2)^2} \tag{5.14}$$

and

$$\frac{var(\widehat{a}_{GLSu})}{var(\widehat{a}_{OLS})} \cong \frac{T^2(\overline{c}^2 - 2\overline{c} + 2TQ + 2\overline{c}Q)}{(-2T - 2\overline{c} + \overline{c}T)^2}. \tag{5.15}$$

We denote the ratios given by (5.14) and (5.15) R_{GLS} and R_{GLSu}, respectively. It follows that

$$R_{GLS} \cong \frac{TQ + 2\overline{c}Q + O(T^{-1})}{1 + O(T^{-1})} \tag{5.16}$$

and

$$R_{GLSu} \cong \frac{2TQ + 2\overline{c}Q + \overline{c}(\overline{c} - 2)}{(\overline{c} - 2)^2 + O(T^{-1})}. \tag{5.17}$$

To render R_{GLS} and R_{GLSu} non-divergent, we need to make Q inversely dependent on T. When Q is related to T, such that $Q = O(T^{-1})$ it is easily observed from (5.16) and (5.17) that the limits of R_{GLS} and R_{GLSu} will be determined (primarily) by the term TQ. It is also readily apparent that R_{GLSu} will further depend on $\overline{c}$ asymptotically through the terms $\overline{c}(\overline{c} - 2)$ and $(\overline{c} - 2)^2$, while for R_{GLS} all terms involving $\overline{c}$ are asymptotically negligible.[4] Moreover, since $Q > 0$ we find OLS detrending is more efficient than GLS detrending - asymptotically at least - for any model. The same result holds for the $GLSu$ estimator when $TQ > 2 - \overline{c}$.

We consider an $AR(p+1)$ model for u_t of the form

$$(1 - \sum_{j=1}^{p} \delta_j L^j)(1 - \rho_T L)u_t = \varepsilon_t \tag{5.18}$$

where $\rho_T = 1 + \frac{c}{T}$ for a fixed constant $c \leq 0$, so that the parameter space is a shrinking neighborhood of unity as the sample size grows. This parameterization is common in the local to unity setting see *inter alia* Chen and Wei (1987) and

[4] In small samples R_{GLS} will also depend on $\overline{c}$. This effect will become visible in graphs presented at a later stage.

Phillips (1987). Model (5.18) admits the explicit $AR(p+1)$ form

$$u_t = \sum_{j=1}^{p+1} \phi_j u_{t-j} + \varepsilon_t$$

where

$$\phi_1 = 1 + c/T \qquad p = 0$$

$$\left.\begin{aligned} \phi_1 &= 1 + c/T + \delta_1 \\ \phi_j &= \delta_j - (1 + c/T)\delta_{j-1}, \quad j = 2, ..., p \\ \phi_{p+1} &= -(1 + c/T)\delta_p \end{aligned}\right\} p > 0$$

Under such a specification, V_l obtains the simple expression

$$V_l = \frac{T^2}{c^2(1 - \sum_{j=1}^{p} \delta_j)^2}.$$

It is much more difficult, however, to obtain a neat expression for V_s, compelling us to look at a case-by-case analysis.

For $p = 0$, which corresponds to an $AR(1)$ process,

$$\begin{aligned} V_l &= \frac{1}{(1-\phi_1)^2} \\ &= \frac{T^2}{c^2} \end{aligned}$$

while,

$$\begin{aligned} V_s &= \frac{1}{1-\phi_1^2} \\ &= -\frac{T^2}{c(2T+c)}. \end{aligned}$$

Thus,

$$TQ = -T\frac{c}{2T+c}$$

and

$$\lim_{T\to\infty} R_{GLS} = -\frac{1}{2}c. \tag{5.19}$$

For $p = 1$, in which case the corresponding process is an $AR(2)$,

$$\begin{aligned} V_l &= \frac{1}{(1-\phi_1-\phi_2)^2} \\ &= \frac{T^2}{c^2(\delta_1-1)^2} \end{aligned}$$

and

$$\begin{aligned} V_s &= \frac{1-\phi_2}{(1+\phi_2)[(1-\phi_2)^2-\phi_1^2]} \\ &= -\frac{T^2(T+\delta_1 T+\delta_1 c)}{c(\delta_1^2-1)(2T+c)(-T+\delta_1 T+\delta_1 c)}. \end{aligned}$$

Thus,

$$\begin{aligned} TQ &= -T\frac{c(\delta_1-1)(T+\delta_1 T+\delta_1 c)}{(\delta_1+1)(2T+c)(-T+\delta_1 T+\delta_1 c)} \\ &= \frac{T^2 c(1-\delta_1^2)+O(T)}{2T^2(\delta_1^2-1)+O(T)} \end{aligned} \tag{5.20}$$

and

$$\lim_{T\to\infty} R_{GLS} = -\frac{1}{2}c.$$

For $p = 2$, the case of an $AR(3)$ process,

$$\begin{aligned} V_l &= \frac{1}{(1-\phi_1-\phi_2-\phi_3)^2} \\ &= \frac{T^2}{c^2(\delta_1+\delta_2-1)^2} \end{aligned}$$

while

$$V_s = \frac{\phi_2+\phi_3\phi_1+\phi_3^2-1}{(\phi_1-\phi_2+\phi_3+1)(\phi_1+\phi_2+\phi_3-1)(\phi_1\phi_3+\phi_2-\phi_3^2+1)}$$

$$= -T^2 \frac{T(\delta_2+1)[c\delta_1+T(\delta_1-\delta_2+1)]-c(2T+c)\delta_2(\delta_2-1)}{c(2T+c)(\delta_1-\delta_2+1)(\delta_1+\delta_2-1)(\delta_2+1)[c(2T+c)\delta_2+Tc\delta_1+T^2(\delta_1+\delta_2-1)]}$$

and

$$\begin{aligned} TQ &= -T\frac{c(\delta_1+\delta_2-1)\{T(1+\delta_2)[c\delta_1+T(\delta_1-\delta_2+1)]-c(2T+c)\delta_2(\delta_2-1)\}}{(2T+c)(\delta_1-\delta_2+1)(\delta_2+1)[c(2T+c)\delta_2+Tc\delta_1+T^2(\delta_1+\delta_2-1)]} \\ &= -\frac{T^3c(1+\delta_2)(\delta_1+\delta_2-1)(\delta_1-\delta_2+1)+O(T^2)}{2T^3(1+\delta_2)(\delta_1+\delta_2-1)(\delta_1-\delta_2+1)+O(T^2)} \end{aligned} \tag{5.21}$$

from which

$$\lim_{T\to\infty} R_{GLS} = -\frac{1}{2}c.$$

For $p = 3$, the expressions for V_s become very messy, but still terms involving the $\delta_j's$ cancel and we find the same result as in (5.19). Similarly so for $p = 4$. It therefore turns out that these limits in the case of GLS detrending, depend only on c, but not on the underlying model or $\overline{c}$.

Conjecture 5.2.1 *Suppose y_t is generated according to the $AR(p+1)$ model $y_t = a+u_t$, $t = 1,2,...,T$ with $(1-\sum_{j=1}^{p}\delta_jL^j)(1-\rho_TL)u_t = \varepsilon_t$, where all the roots of $\delta(z) = 1-\sum_{j=1}^{k}\delta_jz^j = 0$ lie outside the unit root circle and $|\rho_T| < 1$. Furthermore, if u_0 has properties analogous to u_t itself and GLS detrending is performed under the local alternative $\overline{\rho} = 1+\overline{c}/T$, ($\overline{c} < 0$), then $\lim_{T\to\infty} R_{GLS} = -\frac{1}{2}c$ so that $R_{GLS} > 1$ for $c < -2$.*

Without plotting the limiting function for R_{GLS_u} it can easily be seen that for $c = -30$ the relative efficiency of the mean estimator is 15, but even for values of c closer to the null it is still well above unity. The prediction is that the GLS estimator of the mean will always be a less efficient estimator relative to that of OLS for $c < -2$, more so for the more negative values of c, in which case the GLS test will perform poorly relative to the DF test.

We proceed in the same manner for the unconditional case.

For $p = 0$,

$$\begin{aligned} k_1(\overline{c})TQ + k_2(\overline{c}) &= -\frac{1}{(\overline{c}-2)}[\frac{2Tc}{(\overline{c}-2)(2T+c)} - \overline{c}] \\ &= -\frac{c+2\overline{c}-\overline{c}^2+O(T^{-1})}{(\overline{c}-2)^2+O(T^{-1})} \end{aligned}$$

where

$$
\begin{aligned}
k_1(\overline{c}) &= \frac{2}{(\overline{c}-2)^2} \\
k_2(\overline{c}) &= \frac{\overline{c}}{\overline{c}-2}
\end{aligned}
$$

and

$$
\lim_{T\to\infty} R_{GLSu} = -\frac{c+2\overline{c}-\overline{c}^2}{(\overline{c}-2)^2} = -\frac{1}{(\overline{c}-2)}[\frac{c}{(\overline{c}-2)} - \overline{c}]. \tag{5.22}
$$

For $p = 1$,

$$
\begin{aligned}
k_1(\overline{c})TQ + k_2(\overline{c}) &= -\frac{1}{(\overline{c}-2)}[\frac{2cT(\delta_1-1)(T+\delta_1 T+\delta_1 c)}{(\overline{c}-2)(\delta_1+1)(2T+c)(-T+\delta_1 T+\delta_1 c)} - \overline{c}] \\
&= \frac{-2T^2(c+2\overline{c}-\overline{c}^2)(\delta_1^2-1)+O(T)}{2T^2(\overline{c}-2)^2(\delta_1^2-1)+O(T)}
\end{aligned}
$$

$$
\lim_{T\to\infty} R_{GLSu} = -\frac{1}{(\overline{c}-2)}[\frac{c}{(\overline{c}-2)} - \overline{c}].
$$

For $p = 2$,

$$
\begin{aligned}
& k_1(\overline{c})TQ + k_2(\overline{c}) \\
&= -\tfrac{1}{(\overline{c}-2)}[T\tfrac{2c(\delta_1+\delta_2-1)\{T(1+\delta_2)[c\delta_1+T(\delta_1-\delta_2+1)]-c(2T+c)\delta_2(\delta_2-1)\}}{(\overline{c}-2)(2T+c)(\delta_1-\delta_2+1)(\delta_2+1)[c(2T+c)\delta_2+Tc\delta_1+T^2(\delta_1+\delta_2-1)]} - \overline{c}] \\
&= -\tfrac{2T^3(c+2\overline{c}-\overline{c}^2)(\delta_2+1)(\delta_1+\delta_2-1)(\delta_1-\delta_2+1)+O(T^2)}{2T^3(\overline{c}-2)^2(\delta_2+1)(\delta_1+\delta_2-1)(\delta_1-\delta_2+1)+O(T^2)}
\end{aligned}
$$

$$
\lim_{T\to\infty} R_{GLSu} = -\frac{1}{(\overline{c}-2)}[\frac{c}{(\overline{c}-2)} - \overline{c}].
$$

For $p = 3$ and likewise for $p = 4$, the expressions for V_s become quite unpleasant, however we end up with the same limit as in (5.22). Thus, under $GLSu$ detrending the limiting value of the relative efficiency of the mean estimator depends further on $\overline{c}$, but as in the GLS case it is independent of the underlying model.

Conjecture 5.2.2 *Suppose y_t is generated according to the $AR(p+1)$ model $y_t = a + u_t$, $t = 1, 2, ..., T$ with $(1 - \sum_{j=1}^{p} \delta_j L^j)(1 - \rho_T L)u_t = \varepsilon_t$, where all the roots of $\delta(z) = 1 - \sum_{j=1}^{k} \delta_j z^j = 0$ lie outside the unit root circle and $|\rho_T| < 1$. Furthermore, if u_0 has properties analogous to u_t itself and GLSu detrending is performed under the local alternative $\overline{\rho} = 1 + \overline{c}/T$, $(\overline{c} < 0)$, then $\lim_{T\to\infty} R_{GLSu} = -\frac{1}{(\overline{c}-2)}[\frac{c}{(\overline{c}-2)} - \overline{c}]$.*

The limiting function for R_{GLS_u} is graphed in Figure 5.1 for $-30 \leq c \leq -5$ and $-15 \leq \bar{c} \leq -5$. It follows that for $c \leq -15$ and for all values of $\bar{c}$ considered, the limiting function barely increases above unity, while the maximum it reaches for $c = -30$ and $\bar{c} = -5$ is about 1.33. The GLSu estimator of the mean is therefore on the whole more efficient than the OLS estimator and to the degree that the relative efficiencies predict power well, we expect the power of the $GLSu$ test to be greater than the DF test in all cases when $-30 \leq c \leq -5$ and $\bar{c} = -10$ as will be examined in the sequel. Such a test should not suffer from the shortcomings related to the GLS test.

As was mentioned earlier, for $[f(\bar{\rho}) - h(\bar{\rho})] > 0$ and $[f_u(\bar{\rho}) - h_u(\bar{\rho})] > 0$ the relative efficiencies of the mean estimators R_{GLS} and R_{GLS_u} respectively are an increasing function of the short run to long run variance, Q, of the model under consideration. For those values of $\bar{c}$ as determined by ERS and Elliott (1999), it can be easily verified that these inequalities are indeed satisfied. It follows from (5.20) and (5.21) that Q is a function of the sample size, the localizing parameter c and the higher order autoregressive parameters of the model, $Q = Q(T, c, \delta_1, \delta_2, ...)$. By keeping T and c fixed, the effect of the variation in the higher order autoregressive parameters δ_j, $j = 1, 2, ...$ on Q and the subsequent impact on the relative efficiencies of the mean estimators R_{GLS} and R_{GLS_u} would in principle become readily apparent. However, with Q being a complicated (nonlinear) function of the $\delta_j' s$ these effects need to be quantified through a Monte Carlo simulation study. As the sample size increases, we would expect any effect of the $\delta_j' s$ on the relative efficiencies of the mean estimators to dissipate as their limiting values where found above to be independent of these parameters.

For future reference we derive next the exact expressions for the variances in (5.10), (5.8) and (5.12) respectively, when u_t in (5.2) is generated by the stationary second-order autoregressive model

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t \tag{5.23}$$

where ε_t is $iid(0, \sigma_\varepsilon^2)$. The variances of the mean estimators in this case are given by

$$var(\widehat{a}_{OLS}) = Var(\frac{u_1 + u_2 + ... + u_T}{T})$$
$$= T^{-2}\{TVar(u_t) + 2[(T-1)Cov(u_T, u_{T-1}) + (T-2)Cov(u_T, u_{T-2})$$
$$+ ... + Cov(u_T, u_1)]\}$$
$$= T^{-1}\gamma_0 + 2T^{-2}\sum_{j=1}^{T-1}(T-j)\gamma_j \tag{5.24}$$

$$var(\widehat{a}_{GLS}) = [k(\overline{\rho}, T)]^{-2}[f(\overline{\rho}) + (1-\overline{\rho})^4(T-2)]\gamma_0 + 2(1-\overline{\rho})^4\sum_{j=1}^{T-3}(T-j-2)\gamma_j$$
$$+ 2h(\overline{\rho})\sum_{j=1}^{T-2}\gamma_j + 2g(\overline{\rho})\gamma_{T-1}] \tag{5.25}$$

$$var(\widehat{a}_{GLSu}) = [k_u(\overline{\rho}, T)]^{-2}[f_u(\overline{\rho}) + (1-\overline{\rho})^4(T-2)]\gamma_0 + 2(1-\overline{\rho})^4\sum_{j=1}^{T-3}(T-j-2)\gamma_j$$
$$+ 2h_u(\overline{\rho})\sum_{j=1}^{T-2}\gamma_j + f_u(\overline{\rho})\gamma_{T-1}] \tag{5.26}$$

where the autocovariances γ_j, $j \geq 2$ are obtained as the recursive solution to the second-order difference equation

$$\gamma_j = \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2}$$

with

$$\gamma_0 = \frac{(1-\phi_2)\sigma_\varepsilon^2}{(1+\phi_2)(1-\phi_2)^2 - \phi_1^2}$$
$$\gamma_1 = \frac{\phi_1}{1-\phi_2}\gamma_0.$$

For $\phi_2 = 0$ the model in (5.23) reduces to an $AR(1)$ with

$$\gamma_j = \phi^j \gamma_0, \quad j \geq 1$$

and

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \phi_1}$$

By taking the relevant ratios of the variances in (5.24), (5.25) and (5.26) and substituting for the corresponding autocovariances we derive the exact relative efficiencies of the mean estimator under GLS and $GLSu$ detrending in the case of either an $AR(2)$ or $AR(1)$ model.[5]

In the following section, we undertake an extensive Monte Carlo study to examine the performance of the GLS and $GLSu$ tests under the alternative of a "strictly stationary process", where we make use of the above derived formulas of the relative efficiencies of the mean estimators. Their performance is also assessed relative to existing rivals, namely the WS and MAX tests.

5.3 Monte Carlo results

The first order autoregressive nature of the GLS detrending implicit in the GLS and $GLSu$ tests, led us above to consider the model of the form

$$(1 - \sum_{j=1}^{p} \delta_j L^j)(1 - \rho_T L)u_t = \varepsilon_t \tag{5.27}$$

when entertaining higher order stationary autoregressions for u_t in (5.2). In the simulation results that follow we examine the case where $p = 1$, which corresponds to an $AR(1)$ model with $AR(1)$ errors

$$\begin{aligned} u_t &= \rho_T u_{t-1} + v_t \\ v_t &= \delta_1 v_{t-1} + \varepsilon_t. \end{aligned} \tag{5.28}$$

[5]The expressions for the variances in (5.24), (5.25) and (5.26) are an alternative formulation to those obtained in Burridge and Taylor (2000), where Ω is now the variance-covariance matrix of an $AR(2)$ model, see Footnote 2.

Alternatively (5.27) can be written in the form $u_t = (\rho_T + \delta_1)u_{t-1} - \rho_T\delta_1 u_{t-2} + \varepsilon_t$, which is easily seen to belong to the general second order stationary autoregression given by

$$u_t = \beta_1 u_{t-1} + \beta_2 u_{t-2} + \varepsilon_t \tag{5.29}$$

provided the appropriate stationarity conditions on the coefficients ρ_T and δ_1 are satisfied. We restrict our attention to the $AR(2)$ model, for expositional purposes.

Consider the data generating process given by

$$y_t = u_t; \quad (1-\delta_1 L)(1-\rho_T L)u_t = \epsilon_t, \ \rho_T = 1 + \frac{c}{T} \tag{5.30}$$

since the tests are invariant to any mean included in the y_t equation, with ϵ_t ~ $NID(0,1)$. The corresponding regression equation is

$$\Delta\widetilde{y}_t = \varphi\widetilde{y}_{t-1} + \vartheta\Delta\widetilde{y}_{t-1} + e_t \tag{5.31}$$

where $\widetilde{y}_t$ is the local GLS or $GLSu$-demeaned series. The hypothesis of interest is $H_0 : \varphi = 0$ versus $H_1 : \varphi < 0$.

We begin by inspecting the plots in Figures 5.2(a)-(e) that illustrate the simulated rejection frequencies of the DF, GLS, $GLSu$ and WS tests at the nominal 0.05 level under the alternative hypothesis.[6] The rejection frequencies are plotted against the second order autoregressive parameter δ_1 that takes values within the range $(-1, 1)$.[7] Sample sizes of $T = \{100, 200, 500, 1000, 4000\}$ are considered for $c \in \{-7, -10, -15, -20, -25, -30\}$ based on 5000 replications. Here and throughout the values used for GLS and $GLSu$ demeaning are those proposed by ERS and Elliott (1999), $\bar{c} = -7$ and $\bar{c} = -10$ respectively, unless otherwise stated. The lower power of the GLS test relative to the DF test for certain parameter configurations becomes immediately apparent from these figures, particularly more so for smaller sample sizes. The greatest discrepencies between

[6]Initially the results for the MAX test were also included in the graphs. However, for expositional purposes they were omitted as they were almost identical to those pertaining to the WS test. For the same purpose we use two dimensional graphs keeping c fixed in each individual graph.

[7]Note that to conform to the requirements of a covariance stationary process the parameters of the $AR(2)$ model (5.29) must satisfy $\beta_1 + \beta_2 < 1$, $\beta_2 - \beta_1 < 1$ and $|\beta_2| < 1$.

the power of the two tests in finite samples are mainly observed for values in the negative domain of δ_1. In general, the power for the GLS test maintains a concave shape. Power decreases as $\delta_1 \to -1$ in which case y_t approaches a nearly integrated seasonal model of order 2, while it decreases even further (approximating size) as $\delta_1 \to 1$, where y_t approaches an $I(2)$ process.

Figures 5.3(a)-(e) demonstrate the performance of the approximate relative efficiencies of the mean estimator under local GLS and $GLSu$ detrending, that is the ratio of the variance of the GLS and $GLSu$ mean estimator to the variance of the OLS estimator given by (5.14) and (5.15) respectively, where $Q = -\frac{c(-1+\delta_1)(T+\delta_1 T+\delta_1 c)}{(1+\delta_1)(2T+c)(-T+\delta_1 T+\delta_1 c)}$. In the case of the $AR(1)$ model for which $\delta_1 = 0$ in (5.30) and $\vartheta = 0$ in (5.31), Q is replaced by $-\frac{c}{2T+c}$. Inspecting these graphs in parallel to the power graphs in Figures 5.2(a)-(e) for the same values of T, c and δ_1, it turns out that the cases for which the GLS test performs poorly relative to the DF test, are indeed associated with highly inefficient mean estimates. The inefficiency of the GLS estimator relative to that of the OLS estimator manifests itself clearly in both finite and large sample sizes although it is more pronounced in the former case and in general for more negative values of δ_1. Moreover, the relative efficiency of the mean estimator under GLS detrending is close to its asymptotic value $-c/2$ for $T = 4000$ which is given in Table 5.1, apart from when δ_1 is extreme. The behaviour of the mean estimator based on $GLSu$ detrending is rather more subtle with a relative efficiency around the area of unity, in agreement with the asymptotic values, rising only moderately for the more extreme negative values of c and δ_1. In reporting the Monte Carlo results that follow we restrict our attention to a subset of the negative domain of δ_1, where the interesting cases appear to lie.

Table 5.2 reports the rejection frequencies of the test statistics at the 5% significance level under the null and alternative hypotheses, based on asymptotic critical values calculated by setting $\delta_1 = 0$ in (5.30) for 20000 replications. The range of values considered for the localizing parameter c are $\{0, -7, -10, -15, -20, -25, -30\}$ and $\delta_1 \in \{-0.98, -0.95, -0.90, -0.85, -0.80, -0.75, -0.70, -0.65, -0.60, 0.0\}$. Entries in the second column under the heading $c = 0$ pertain to the size of the tests, while the remaining columns illustrate power. Values in the upper brackets are the approximate relative efficiencies of the mean estimator under local GLS and $GLSu$ detrending. The corresponding lower bracketed values are the exact relative efficiencies of the mean estimator given by the ratios of (5.25) over (5.24) and (5.26) over (5.24) respectively, once substituted for the appropriate autocovariances. These latter values are provided as a way of assessing the quality of the approximate relative efficiencies.

The size of all tests is well behaved in that empirical size is found to be very close to the nominal 5% level. As regards power, adjacent columns show that for fixed T and δ_1 power increases for all tests the further we divert from the null, which reflects the consistency of the tests. The corresponding relative efficiencies of the mean estimators under GLS and $GLSu$ detrending increase, which is not surprising since it is for alternatives closer to the unit root null that GLS based detrending is more efficient. The most stricking feature of the results is the high degree of inefficiency of the GLS detrended mean estimator relative to the OLS estimator which is evident for all sample sizes, although it is particularly discernible in finite samples and for the more negative values of δ_1. Specifically, for $T = 100$, $\delta_1 = -0.85$ and $c = -30$, the GLS estimator demonstrates a variance of 22.23 times that of the OLS estimator which more than doubles for $\delta_1 = -0.95$. As the sample size increases, the effect of δ_1 on the inefficiency of the GLS estimator is mitigated. The mean estimator under $GLSu$ detrending is slightly more efficient relative to the OLS estimator for alternatives closer to the null, while its variance is on average 1.2 times that of the OLS estimator for more distant alternatives. Such a phenomenon is undoubtedly associated with the different treatment of the initial observation implicit in the two detrending procedures under the 'strict stationarity' of the data generating process considered.

For fixed T and increasing δ_1, the GLS and $GLSu$ tests exhibit an increase in terms of power in conjunction with improved efficiency of the corresponding mean estimators. This increase in power is observed for higher values of δ_1 in absolute terms, for which a sharp drop in the relative efficiencies of the mean estimators is noted, and for small sample sizes. The power of the alternative tests based on OLS detrending remains practically invariant under the same circumstances.[8]

[8] In fact, there is a slight decrease in the power of the DF statistic. This result corroborates the findings of DeJong *et al.* (1992), who examine the size and power performance of several conventional unit root tests in the presence of autoregressively correlated errors. The DGP they consider also includes a trend term in the deterministic function

$$y_t = a + bt + u_t, \quad u_t = \rho u_{t-1} + v_t, \quad v_t = \delta v_{t-1} + \varepsilon_t$$

Figures in Table 5.2 show that in total the *GLS* test demonstrates inferior power performance relative to the *DF* test for large negative values of c, that is for values of the localizing parameter $c < -15$, which implies that the relative efficiencies of the mean estimators predict powers well. In particular, the *GLS* test displays power 10 to 30 percent lower on average than the conventional *DF* test, the larger discrepancies observed for the more negative values of δ_1 and smaller sample sizes. Related to such power results are the highest relative efficiencies of the mean estimator. To assess the quality of the approximate relative efficiencies under *GLS* and *GLSu* detrending, we compare them to their exact counterparts as displayed in the lower bracketed figures of Tables 5.2(a)-(e). In turns out that the approximate relative efficiencies are very close to the exact values with the highest divergence observed below 1.00 on average.

When considering the theoretical limiting values of the approximate relative efficiencies of the mean estimator under local *GLS* and *GLSu* detrending provided in Table 5.1 relative to the upper bracketed figures in Tables 5.2(a)-(e), it emerges that for $T = 4000$ the *GLS* estimator is equal to its limiting value $-\frac{c}{2}$. In general, the rate of convergence to its limiting value is slow particularly for the more negative values of δ_1. The *GLSu* test on the other hand tends to follow rather closely its limiting value even for smaller sample sizes.

In comparing the power performance of the *DF* test with the *GLS* based tests and their existing competitors, namely the *WS* and *MAX* tests, the power superiority of the latter is overwhelmingly visible throughout Table 5.2, with little difference to choose between the two. The *WS* and *MAX* tests not only outperform the remaining tests, but they also maintain their power advantage for all combinations of c and δ_1 even for the very extreme values of δ_1. For the $AR(1)$ case the findings are qualitatively similar although not as pronounced. Thus it appears that exploiting the time reversibility of a Gaussian stationary autoregressive finite order series as do these tests with implicit *OLS* detrending, is far more successful than *GLS* or *GLSu* detrending in terms of power advantage over the *DF* test.

In summary, the above results suggest that unit root tests based on *GLS* detrending can have very low power for certain parameter values, and for some sample sizes, a fact that appears to be directly linked to highly inefficient mean

which explains the more pronounced decrease in the power of the *DF* statistic that they find.

estimates. Asymptotic effects are uncovered for the localising parameter c, as predicted by our theoretical attempt in exploring such an issue, which appear considerably pronouced in finite samples. What is perhaps more interesting though are the finite sample effects related to the second order autoregressive parameter, δ_1. An alternative result worth pointing out is that the GLS test performs better than the $GLSu$ test in terms of power for alternatives closer to the null despite the higher relative efficiency of the mean estimator associated with the latter.

Finally, diverting from our original focus on the quality of the mean estimator per se, we briefly present some selected graphical results on the effect of variation in $\bar{c}$. In doing so, we aim to highlight the dependence of the GLS and $GLSu$ tests on the value of $\bar{c}$ used for detrending, which can affect the outcome of finite sample results. In Figures 5.4(a)-(d), we present power functions and the corresponding relative efficiencies of the GLS and $GLSu$ mean estimators, for $T = 100, 200, 500$ and 1000 and $\bar{c} \in \{-7, -10, -15, -20\}$. The results are reported for $c = -20$, a value that is representative of the overall effects underway. Results for alternative values of c are qualitatively similar and in conjunction with Figures 5.2 and 5.3, they are not difficult to predict. Varying $\bar{c}$ will change the critical values of the GLS and $GLSu$ tests which in turn will affect their power, and based on the above results the relative efficiency of the corresponding mean estimator. This becomes readily apparent when studying the plots in Figures 5.4(a)-(d). In general, increasing the value of $\bar{c}$ in absolute terms under which local detrending is performed has the effect of augmenting the power of the GLS and $GLSu$ tests and improving the efficiency of the mean estimator. Such an effect is particularly discernible for the GLS test and less so for the $GLSu$ test. Specifically, in the case of the former by the time $\bar{c} = -20$ any poor power performance demonstrated by the test is almost fully restored.

5.4 Concluding remarks

This study provided a further investigation into the behaviour of the GLS and $GLSu$ tests in light of their widespread use in empirical applications, particularly of the former. We entertained solely the case where the initial observation is drawn from its unconditional distribution as a more realistic alternative. The principle issue was that the unit root test based on GLS detrending can have very low power for some models and parameter values, and for some sample sizes. A theoretical attempt was made, relying on local to unity asymptotics, to identify such "unsatisfactory behaviour" through the investigation of variance ratios, in the case of an unknown mean. To this end, approximate results were derived for the relative efficiencies of the mean estimator under a general linear process for the error term which appeared to take the form of simple expressions. To obtain their limiting values an autoregressive structure was considered for the error term. The limiting function of the relative efficiency of the mean estimator based on GLS detrending showed that for large negative values of the localising parameter c, the GLS estimator will be an inferior estimator relative to the OLS estimator. To the extent that such a result is associated with lower power it would be expected that the GLS test demonstrates on occasion lower power compared to the DF test, which involves OLS detrending. The $GLSu$ test on the other hand did not suffer from such shortcomings. Overall, asymptotics revealed that for both tests the underlying model will not affect the relative performance of the test. We can always select a value for c that will make GLS perform unfavourably. We conjectured that such results carries over to higher order autoregressive models.

Monte Carlo results on the size and power properties of the tests were presented tailored to a second order autoregressive model. The approximate relative efficiencies of the mean estimator appeared to predict powers well, demonstating at the same time good quality when compared to their exact counterparts. Findings highlighted the low rejection probabilities of the GLS test for certain parameter configurations, which were associated with higher inefficiency of the GLS mean estimator relative to the OLS estimator. While such a finding was moderately evident in large sample sizes, it was particularly pronounced in finite samples where higher power differences of the conventional DF test over the GLS test were noted mainly for the more negative values of second order autoregressive parameter δ_1. Contrary to the asymptotic and finite effects displayed by the localising parameter c, the second order autoregressive parameter δ_1 was found to exhibit only finite effects which were more prominent for the GLS test. Related to this phenomenon was the finding that the relative efficiency of the mean estimator under GLS detrending converged to its limiting value at a slower rate for the more negative values of δ_1. On the other hand, the relative efficiency of the $GLSu$ detrended mean estimator was more in line with its limiting value even in small samples.

The above results raise some doubts regarding the reliability of the GLS test as a more powerful unit root test in empirical applications. On the contrary they provide a compelling argument in favour of alternative power-enhancing unit root tests, namely the WS and MAX tests. These latter tests appeared to be more robust maintaining power superiority over all tests for all parameter configurations and all sample sizes.

References

Banerjee, A., Russell, B., 2001. The Relationship between the markup and inflation in the G-7 economies and Australia. *Review of Economics and Statistics,* 83, 377-84.

Barkoulas, J., Baum, C.F., Chakraborty, A., 2002. Forward premiums and market efficiency: Panel unit-root evidence from the term structure of forward premiums. *Journal of Macroeconomics,* forthcoming.

Burridge, P., Taylor, A.M.R., 2000. On the power of GLS-type unit root tests. *Oxford Bulletin of Economics and Statistics, 6*2, 633-45.

Canjels, E., Watson, M.W., 1997. Estimating deterministic trends in the presence of serially correlated errors. *Review of Economics and Statistics,* 79, 184-200.

Chan, N.H., Wei, C.Z., 1987. Asymptotic inference for nearly nonstationary AR(1) processes. *The Annals of Statistics*, 15, 1050-1063.

Cheung, Y.W., Lai, K.S., 1998. Parity reversion in real exchange rates during the post-Bretton Woods period. *Journal of International Money and Finance*, 17, 597-614.

Cushman, D.O., 2001. Bayesian and DF-GLS unit root tests of real exchange rates over the current floating period. *Economics Bulletin*, 6,1-7.

DeJong, D.N., Nankervis, J.C., Savin, N.E., Whiteman, C.H., 1992. The power problems of unit root tests in time series with autoregressive errors. *Journal of Econometrics*, 53, 323-343.

Elliott, G., 1999. Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review*, 40, 767-783.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. *Econometrica*, 64, 813-836.

Hamilton, J.D., 1994. Time series analysis. *Princeton University Press.*

Leybourne, S.J., 1995. Testing for unit roots using forward and reverse Dickey-Fuller Regressions. *Oxford Bulletin of Economics and Statistics,* 57, 559-571.

Leybourne, S.J., Kim, T-H., Newbold, P., 2003. Examination of some more powerful modifications of the Dickey-Fuller test. Forthcoming in the *Journal of*

Time Series Analysis.

McNown, R., Puttitanun, T., 2002. New unit root tests of the Nelson-Plosser Data. *Applied Economics Letters,* 9, 9-11.

Müller, U.K., Elliott, G., 2001. Tests for unit roots and the initial observation. Discussion paper 2001-19, University of California, San Diego.

Pantula, S.G., Gonzalez-Farias, G., Fuller, W.A., 1994. A comparison of unit-root test criteria. *Journal of Business & Economic Statistics*, 12, 449-459.

Park, H.J., Fuller, W.A., 1995. Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis*, 16, 415-429.

Phillips, P.C.B., 1987. Towards a unified asymptotic theory for autoregression. *Biometrika, 74, 535-47.*

Phillips, P.C.B., Lee, C.C., 1996. Efficiency gains from quasi-differencing under nonstationarity. In: Robinson, P.M., Rosenblatt, M., (Eds.). *Essays in Memory of E.J. Hannan.* New York: Springer-Verlag.

Rapach D.E., Wohar, M.E., 2002. Testing the monetary model of exchange rate determination: New evidence from a century of data. *Journal of International Economics,* 58, 359-385.

Siklos, P., and C.W.J. Granger, 1997. Regime-sensitive cointegration with an illustration from interest rate parity. *Macroeconomic Dynamics*, 1, 640-57.

Appendix 5.A Tables and Figures

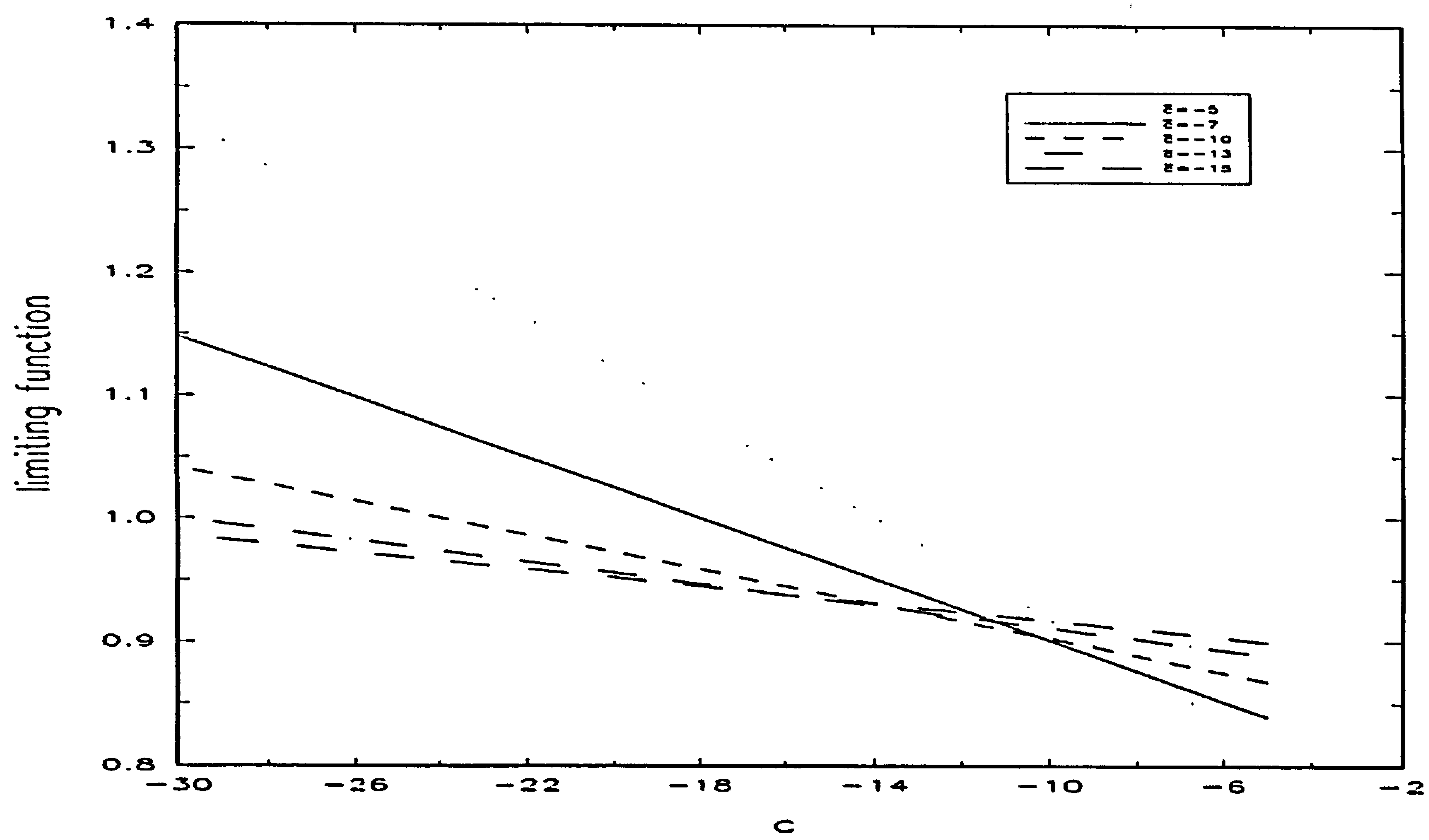


Figure 5.1. Limiting function of the relative efficiency of the mean estimator under *GLSu* detrending.

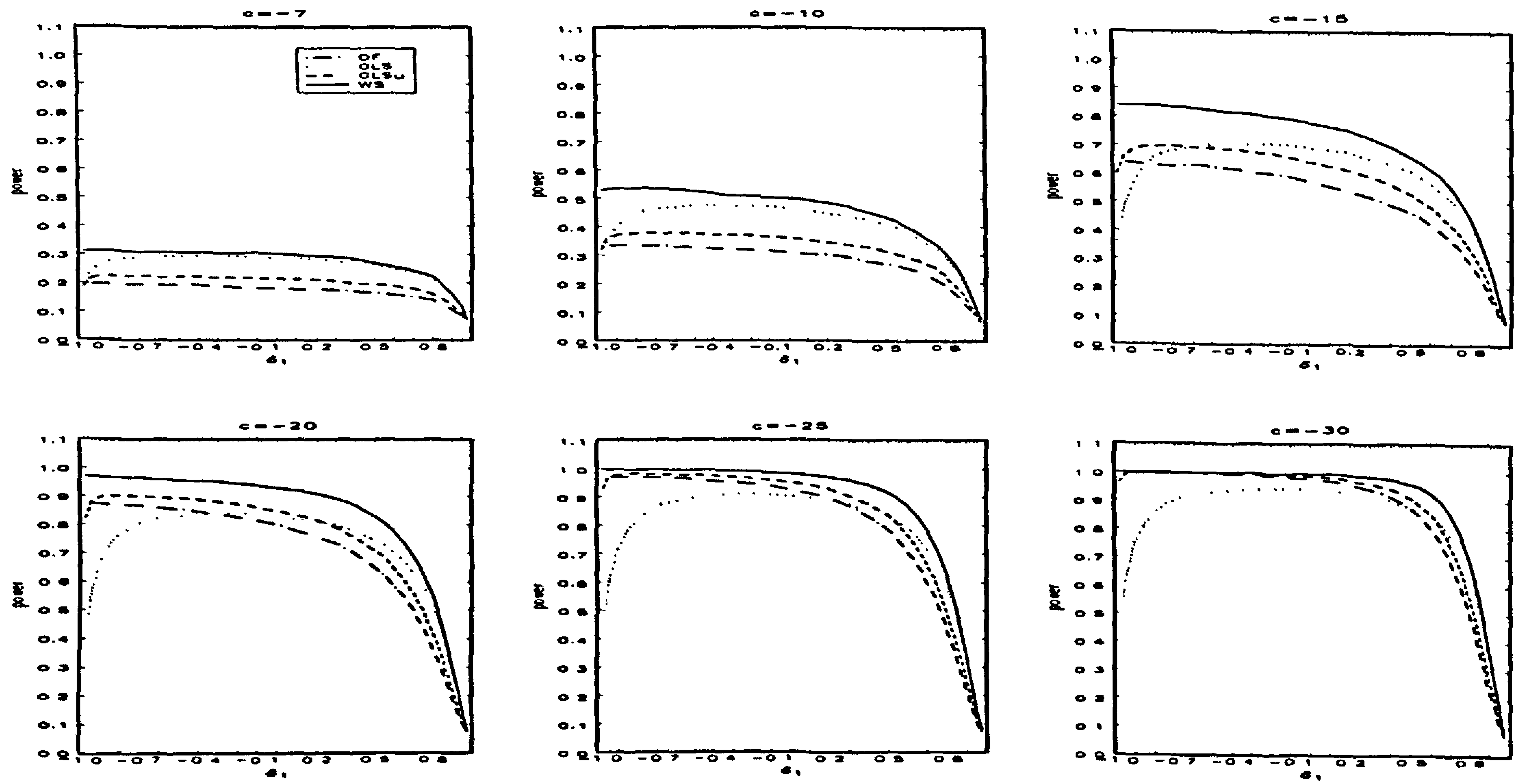


Figure 5.2(a). Power functions of unit root tests, $T = 100$. The graphs in this and all subsequent figures related to power estimates are based on the data-generating process $y_t = u_t$, $(1 - \delta_1 L)(1 - \rho_T L)u_t = \epsilon_t$, $\rho_T = 1 + \frac{c}{T}$. The corresponding regression equation is $\Delta\widetilde{y}_t = \varphi\widetilde{y}_{t-1} + \vartheta\Delta\widetilde{y}_{t-1} + e_t$, where $\widetilde{y}_t$ is the demeaned series. In the case of $\delta_1 = 0.0$, ϑ is set to zero in the regression equation.

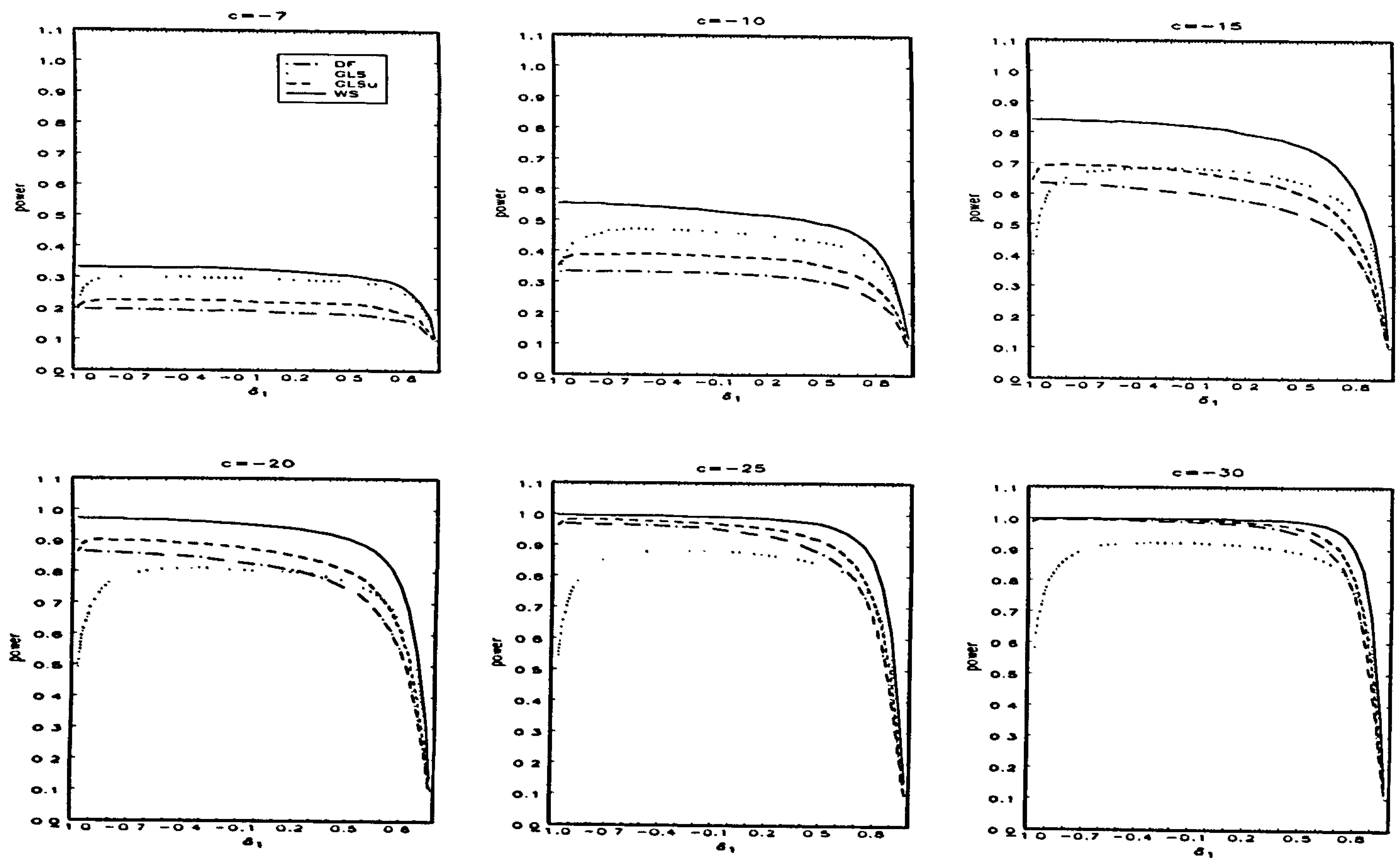


Figure 5.2(b). Power functions of unit root tests, $T = 200$.

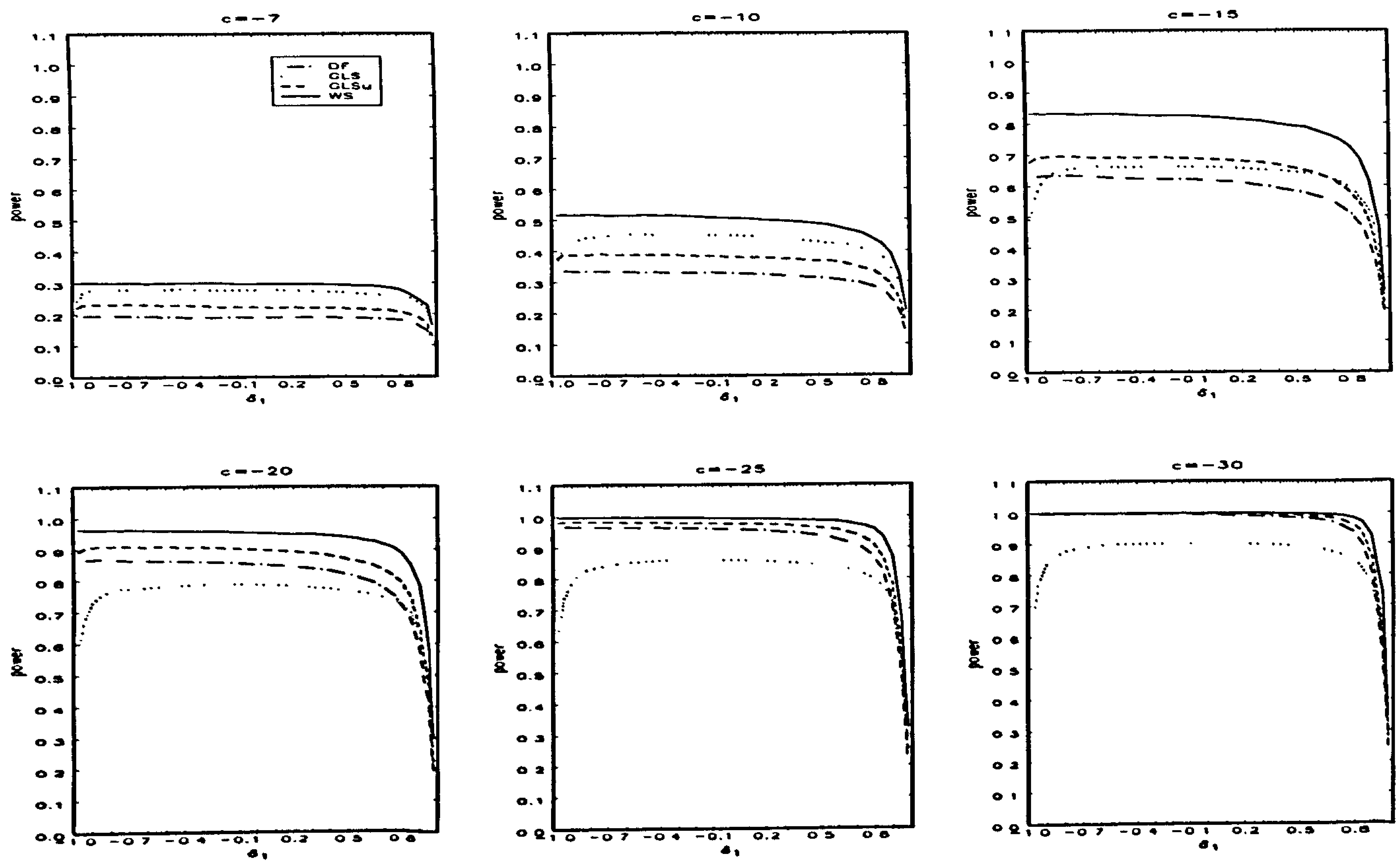


Figure 5.2(c). Power functions of unit root tests, $T = 500$.

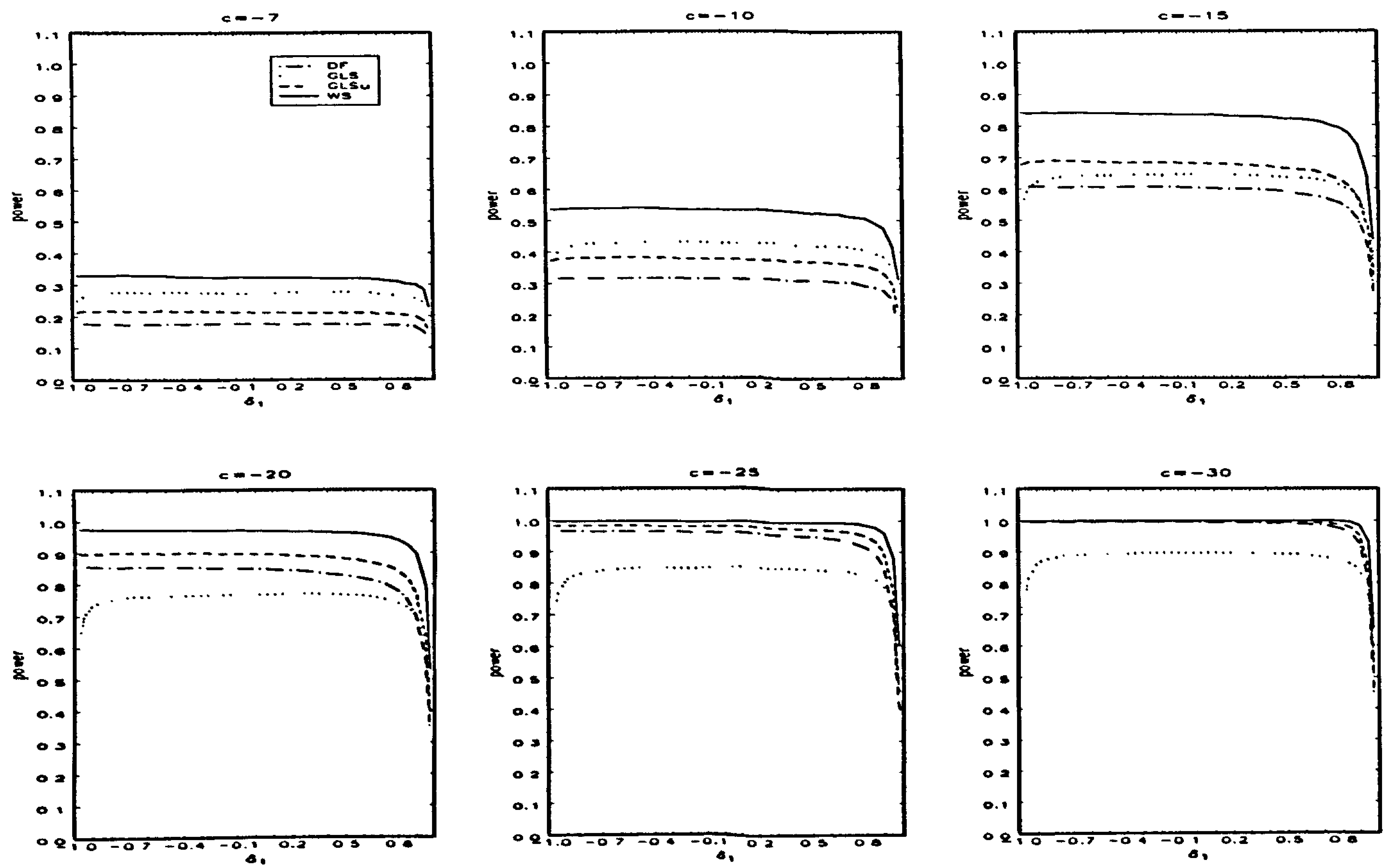


Figure 5.2(d). Power functions of unit root tests, $T = 1000$.

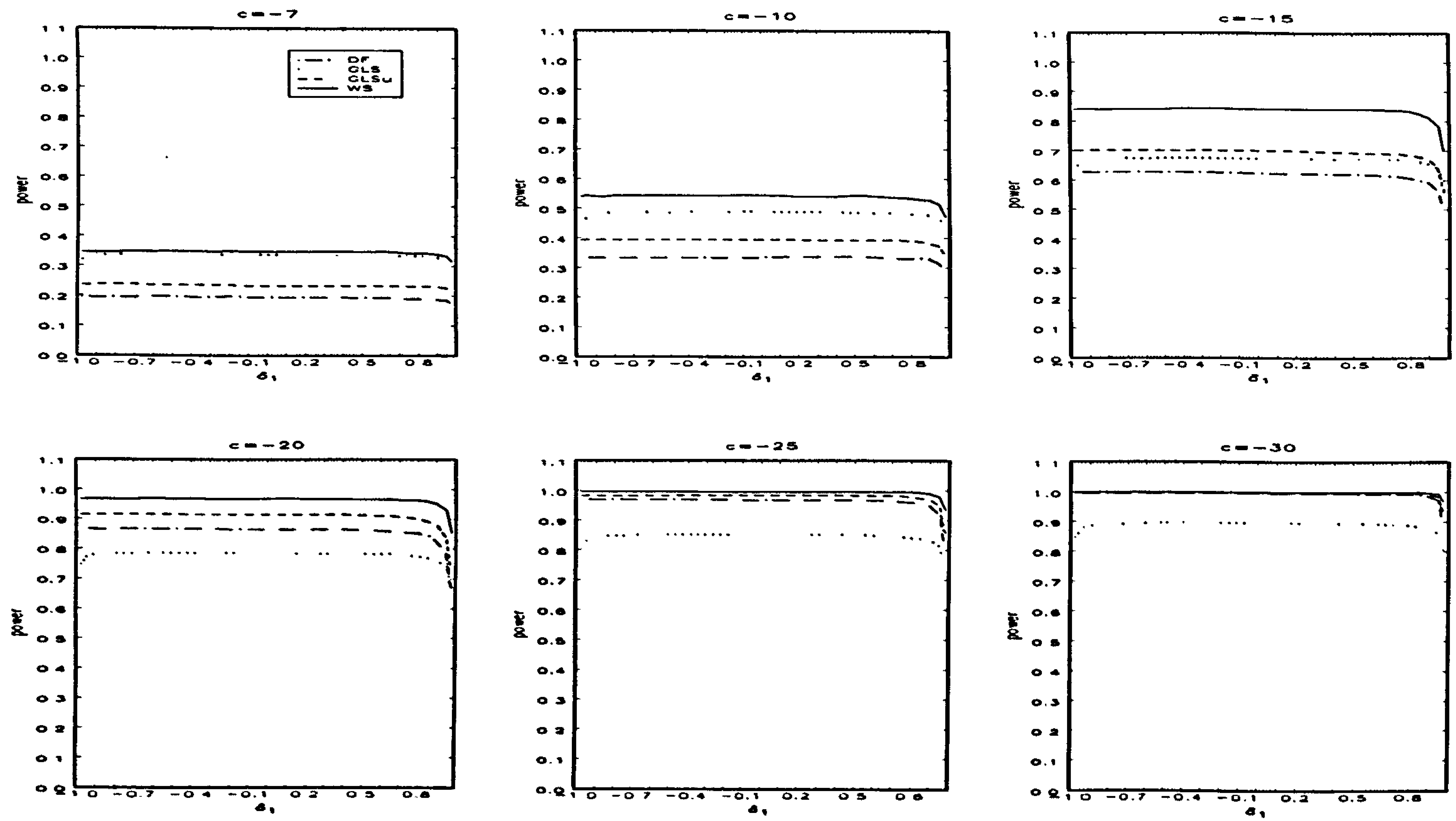


Figure 5.2(e). Power functions of unit root tests, $T = 4000$.

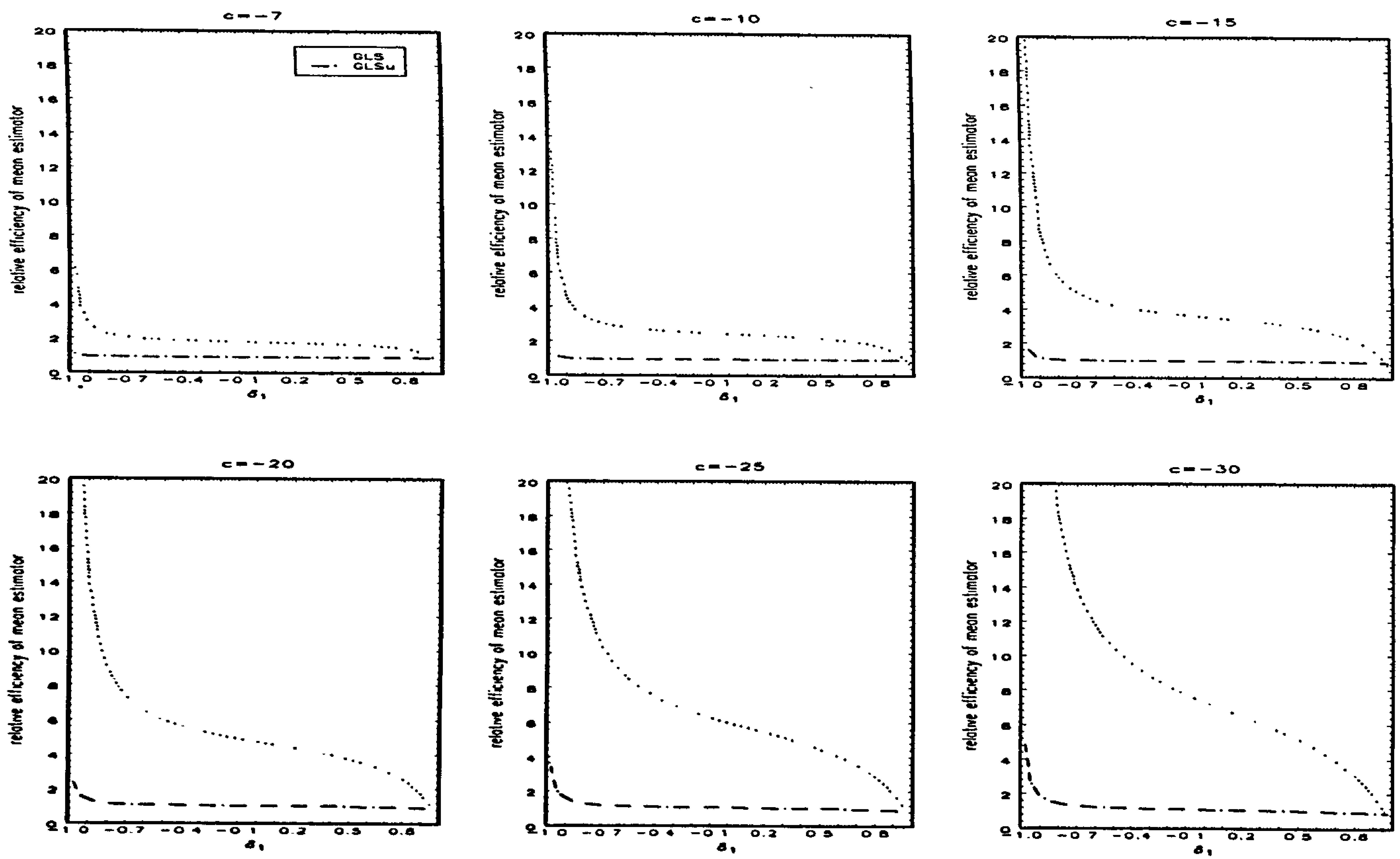


Figure 5.3(a). Relative efficiencies of mean estimators for the GLS and $GLSu$ tests, $T = 100$. Relative efficiences are calculated as the ratio of the variance of the GLS and $GLSu$ mean estimators to the variance of the OLS estimator.

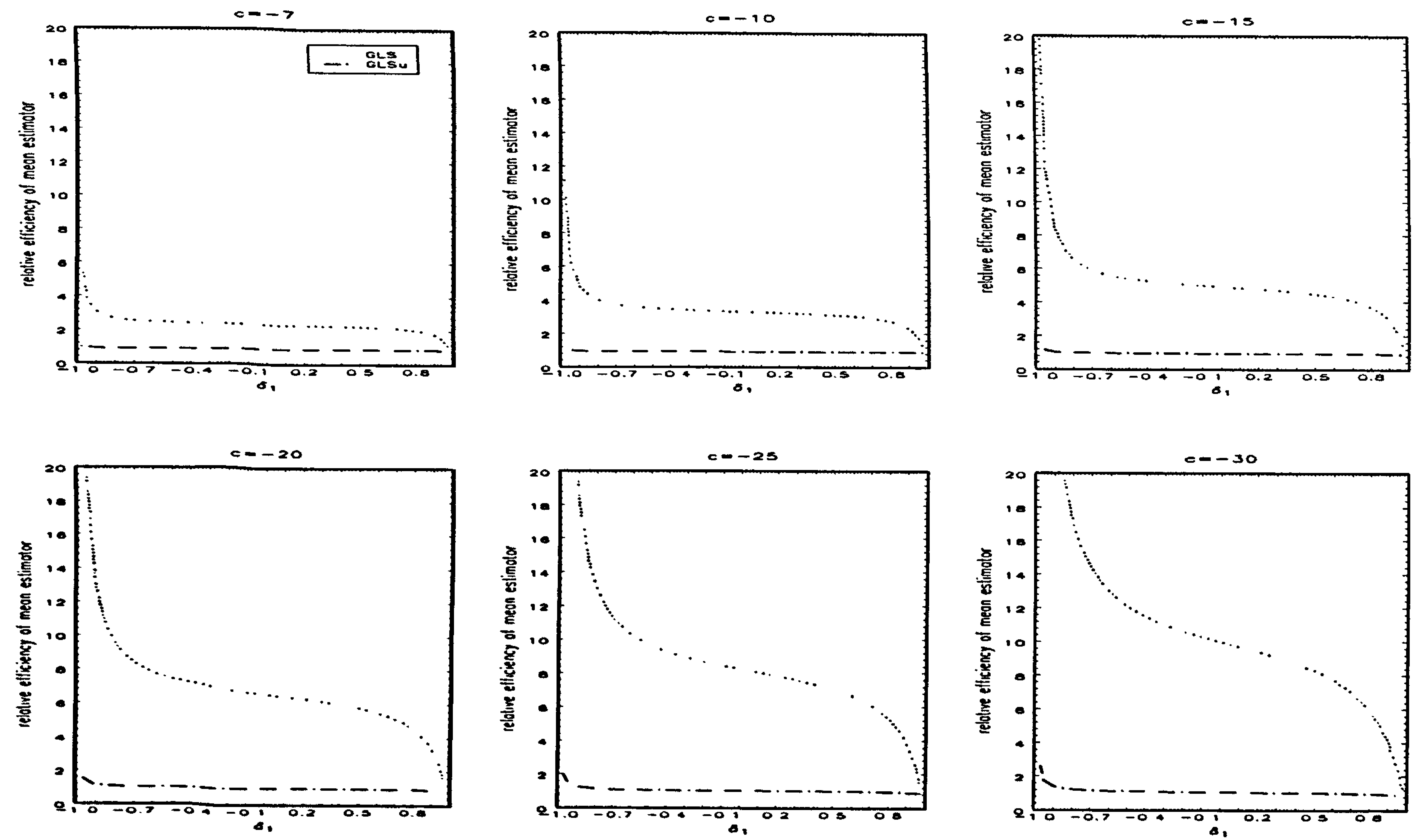


Figure 5.3(b). Relative efficiencies of mean estimators for the *GLS* and *GLSu* tests, $T = 200$.

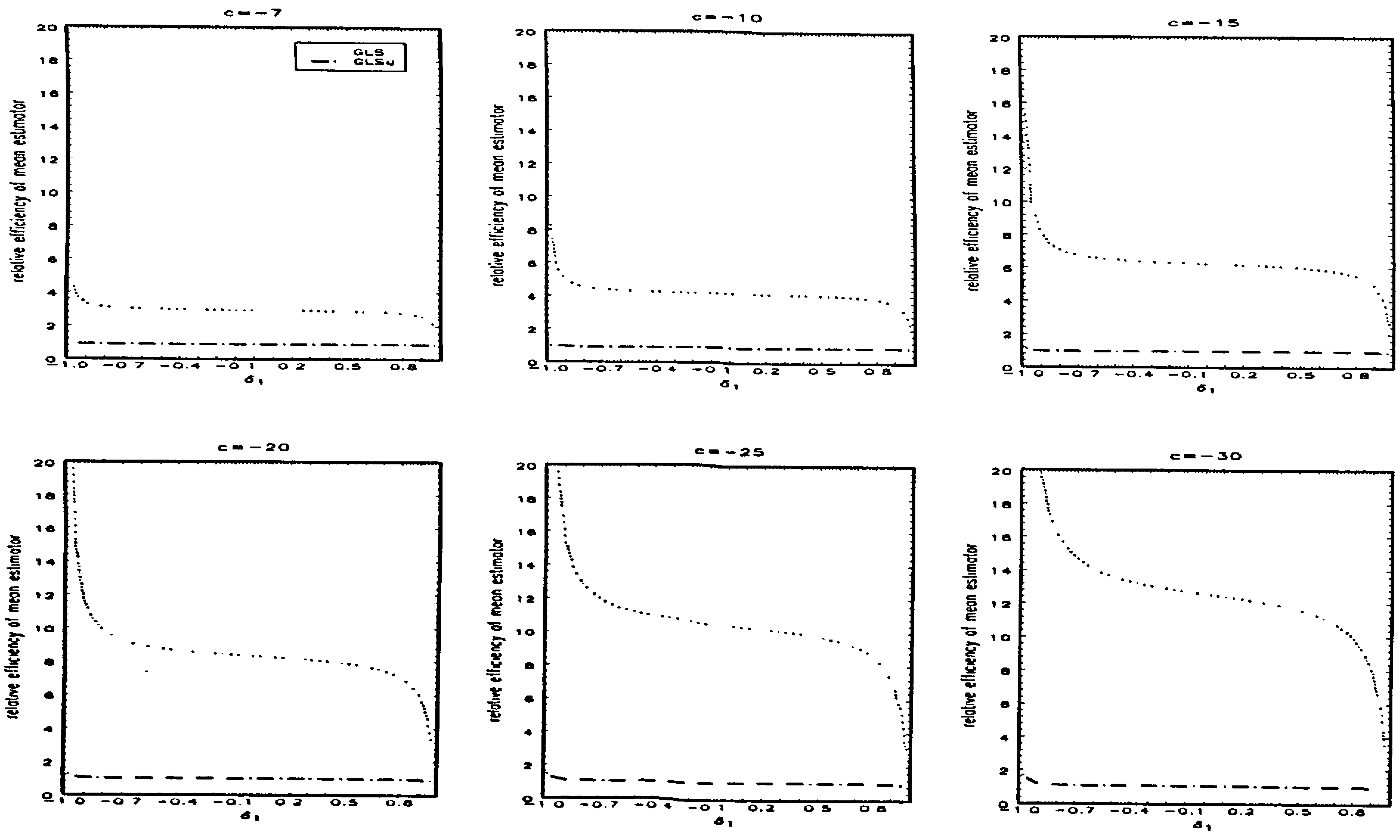


Figure 5.3(c). Relative efficiencies of mean estimators for the GLS and $GLSu$ tests, $T = 500$.

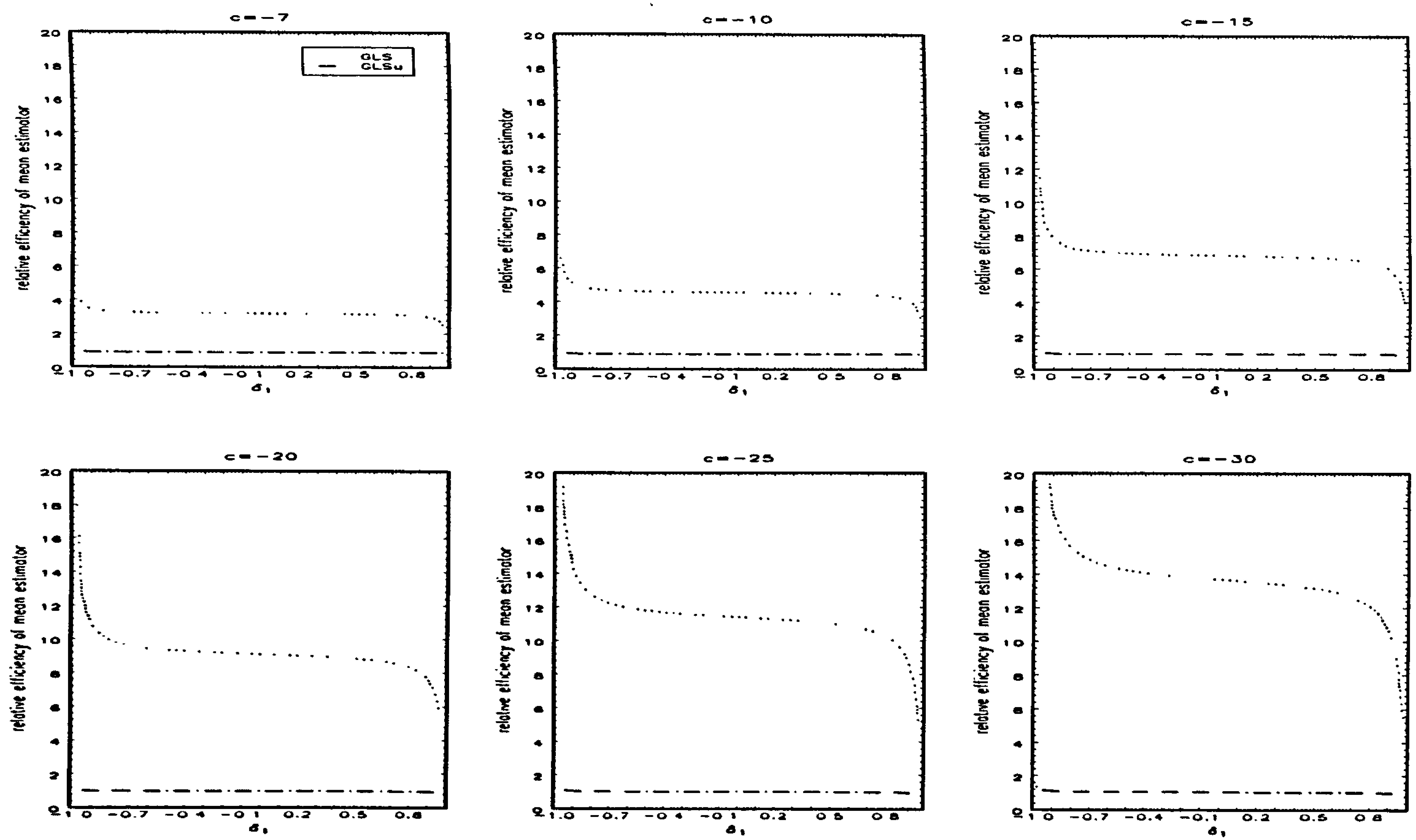


Figure 5.3(d). Relative efficiencies of mean estimators for the GLS and $GLSu$ tests, $T = 1000$.

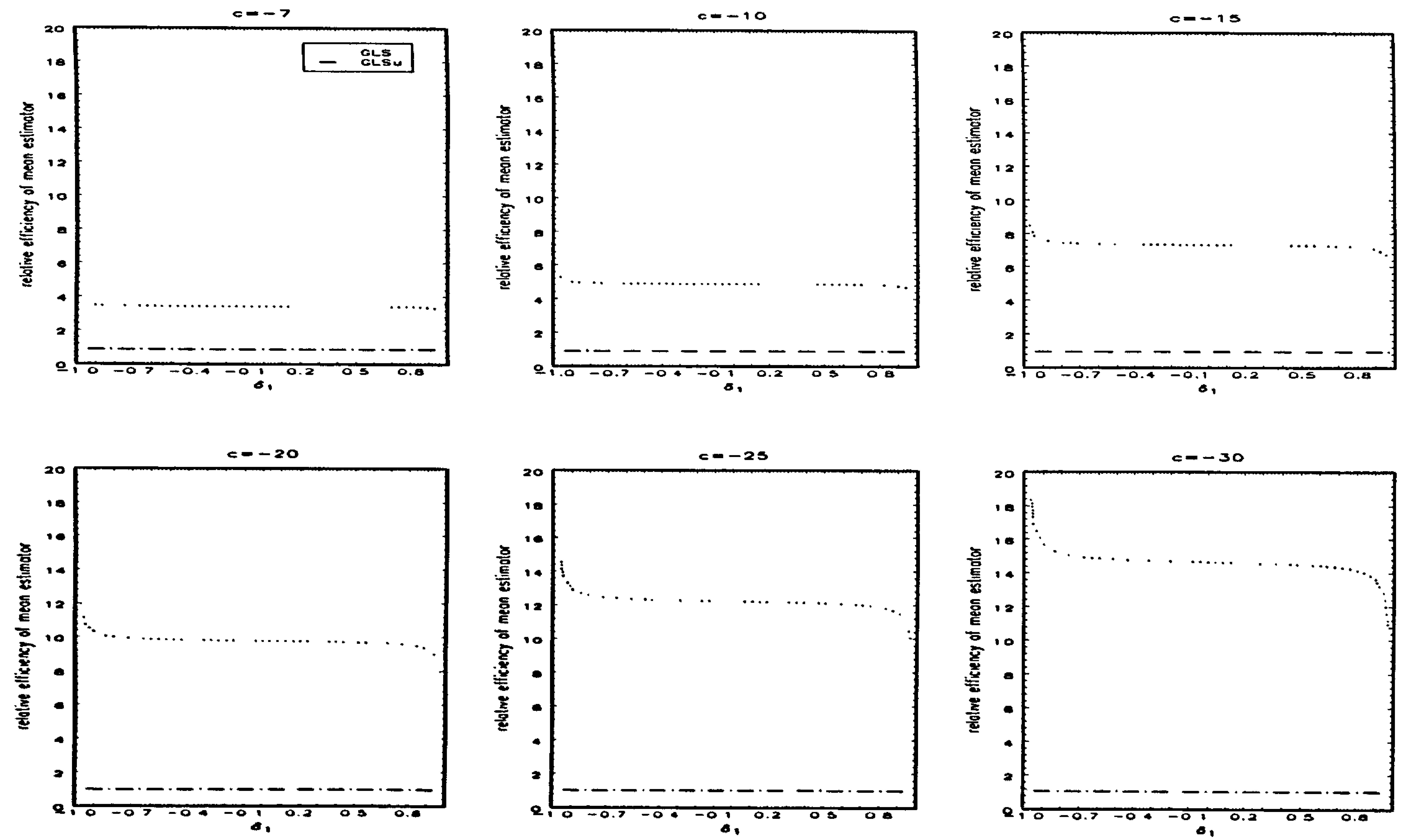


Figure 5.3(e). Relative efficiencies of mean estimators for $T = 4000$.

Table 5.1

Theoretical limits of the relative efficiencies of the mean estimators under *GLS* and *GLSu* demeaning

c	-7	-10	-15	-20	-25	-30
$\lim_{T\to\infty} R_{GLS}$	3.5	5	7.5	10	12.5	15
$\lim_{T\to\infty} R_{GLSu}$	0.881	0.903	0.938	0.972	1.007	1.042

Table 5.2(a)

Size and power of the tests of the $I(1)$ null at the nominal 5% level, $T = 100$

		c						
δ_1		0	−7	−10	−15	−20	−25	−30
−0.98	*WS*	0.056	0.311	0.530	0.839	0.971	0.997	1.000
	MAX	0.054	0.303	0.513	0.834	0.968	0.997	1.000
	DF	0.051	0.193	0.333	0.638	0.877	0.973	0.998
	GLS	0.054	0.183(6.928)	0.263(13.28)	0.360(29.31)	0.407(53.11)	0.429(85.94)	0.443(129.3)
			(7.890)	(14.40)	(30.58)	(54.25)	(86.52)	(128.6)
	GLSu	0.045	0.189(1.077)	0.320(1.284)	0.599(1.807)	0.822(2.584)	0.930(3.654)	0.962(5.068)
			(1.091)	(1.274)	(1.738)	(2.421)	(3.352)	(4.566)
−0.95	*WS*	0.055	0.312	0.531	0.839	0.971	0.996	1.000
	MAX	0.054	0.304	0.513	0.835	0.967	0.998	1.000
	DF	0.051	0.194	0.336	0.640	0.877	0.973	0.998
	GLS	0.048	0.234(3.799)	0.344(6.690)	0.482(13.68)	0.551(23.75)	0.594(37.40)	0.614(55.21)
			(4.381)	(7.366)	(14.53)	(24.75)	(38.51)	(56.32)
	GLSu	0.049	0.212(0.975)	0.358(1.069)	0.658(1.297)	0.879(1.626)	0.971(2.071)	0.993(2.652)
			(0.999)	(1.089)	(1.316)	(1.646)	(2.091)	(2.668)
−0.90	*WS*	0.055	0.311	0.535	0.840	0.970	0.996	1.000
	MAX	0.054	0.304	0.512	0.837	0.967	0.997	1.000
	DF	0.053	0.195	0.335	0.637	0.877	0.972	0.998
	GLS	0.048	0.264(2.755)	0.405(4.492)	0.572(8.460)	0.671(13.95)	0.717(21.20)	0.750(30.49)
			(3.174)	(4.945)	(8.993)	(14.57)	(21.89)	(31.23)
	GLSu	0.049	0.222(0.941)	0.372(0.998)	0.687(1.127)	0.901(1.306)	0.980(1.543)	0.998(1.846)
			(0.960)	(1.011)	(1.138)	(1.317)	(1.555)	(1.858)
−0.85	*WS*	0.056	0.311	0.535	0.839	0.970	0.995	1.000
	MAX	0.056	0.303	0.516	0.837	0.966	0.997	1.000
	DF	0.053	0.195	0.336	0.635	0.874	0.972	0.998
	GLS	0.047	0.277(2.406)	0.430(3.757)	0.622(6.716)	0.736(10.68)	0.783(15.79)	0.816(22.23)
			(2.769)	(4.135)	(7.139)	(11.15)	(16.31)	(22.79)
	GLSu	0.052	0.224(0.929)	0.376(0.974)	0.695(1.070)	0.902(1.199)	0.982(1.366)	0.998(1.576)
			(0.946)	(0.984)	(1.077)	(1.206)	(1.373)	(1.583)
−0.80	*WS*	0.055	0.311	0.536	0.838	0.969	0.995	1.000
	MAX	0.056	0.303	0.519	0.833	0.966	0.997	1.000
	DF	0.053	0.194	0.334	0.633	0.872	0.971	0.998
	GLS	0.050	0.284(2.231)	0.447(3.388)	0.655(5.840)	0.765(9.031)	0.820(13.07)	0.856(18.08)
			(2.566)	(3.727)	(6.208)	(9.433)	(13.51)	(18.55)
	GLSu	0.053	0.225(0.924)	0.378(0.962)	0.695(1.042)	0.904(1.146)	0.982(1.277)	0.998(1.441)
			(0.940)	(0.971)	(1.047)	(1.150)	(1.281)	(1.445)

Note: The results in this and all subsequent tables are based on the data-generating process $y_t = u_t$, $(1-\delta_1 L)(1-\rho_T L)u_t = \epsilon_t$, $\rho_T = 1 + \frac{c}{T}$ The corresponding regression equation is $\Delta\widetilde{y}_t = \varphi\widetilde{y}_{t-1} + \vartheta\Delta\widetilde{y}_{t-1} + e_t$, where $\widetilde{y}_t$ is the demeaned series. The entries for $\delta_1 = 0.0$ are based on $\vartheta = 0$. Values in the upper and lower parentheses denote the approximated and exact relative efficiencies of the mean estimators, respectively.

Table 5.2(a) (continued)

δ_1		c: 0	−7	−10	−15	−20	−25	−30
−0.75	*WS*	0.054	0.311	0.537	0.836	0.967	0.995	1.000
	MAX	0.056	0.303	0.521	0.832	0.964	0.997	1.000
	DF	0.054	0.194	0.332	0.631	0.870	0.970	0.998
	GLS	0.050	0.287(2.125)	0.457(3.165)	0.672(5.311)	0.789(8.038)	0.846(11.43)	0.881(15.58)
			(2.444)	(3.481)	(5.645)	(8.396)	(11.81)	(15.98)
	GLSu	0.053	0.225(0.920)	0.380(0.954)	0.698(1.024)	0.904(1.113)	0.981(1.224)	0.998(1.359)
			(0.936)	(0.963)	(1.029)	(1.116)	(1.226)	(1.361)
−0.70	*WS*	0.054	0.308	0.534	0.835	0.966	0.995	0.999
	MAX	0.056	0.303	0.519	0.829	0.962	0.997	1.000
	DF	0.053	0.194	0.332	0.628	0.868	0.968	0.997
	GLS	0.050	0.288(2.054)	0.463(3.015)	0.681(4.955)	0.804(7.370)	0.866(10.32)	0.901(13.90)
			(2.361)	(3.316)	(5.266)	(7.698)	(10.67)	(14.26)
	GLSu	0.053	0.224(0.918)	0.378(0.949)	0.699(1.013)	0.902(1.091)	0.981(1.188)	0.998(1.304)
			(0.933)	(0.958)	(1.016)	(1.093)	(1.189)	(1.305)
−0.65	*WS*	0.054	0.309	0.533	0.833	0.964	0.994	0.999
	MAX	0.055	0.302	0.518	0.826	0.961	0.997	1.000
	DF	0.053	0.194	0.331	0.629	0.866	0.967	0.996
	GLS	0.050	0.291(2.002)	0.465(2.907)	0.689(4.698)	0.815(6.887)	0.877(9.527)	0.913(12.68)
			(2.301)	(3.196)	(4.993)	(7.194)	(9.848)	(13.01)
	GLSu	0.052	0.224(0.916)	0.379(0.946)	0.696(1.004)	0.900(1.076)	0.980(1.162)	0.998(1.265)
			(0.931)	(0.954)	(1.007)	(1.077)	(1.162)	(1.264)
−0.60	*WS*	0.055	0.307	0.531	0.829	0.961	0.995	0.999
	MAX	0.055	0.301	0.517	0.823	0.959	0.997	1.000
	DF	0.053	0.194	0.329	0.630	0.864	0.966	0.996
	GLS	0.051	0.293(1.963)	0.468(2.824)	0.695(4.503)	0.824(6.520)	0.885(8.922)	0.919(11.76)
			(2.256)	(3.105)	(4.785)	(6.811)	(9.223)	(12.07)
	GLSu	0.053	0.222(0.915)	0.380(0.943)	0.692(0.998)	0.898(1.064)	0.979(1.142)	0.997(1.235)
			(0.930)	(0.951)	(1.001)	(1.064)	(1.142)	(1.234)
0.0	*WS*	0.054	0.310	0.532	0.843	0.974	0.996	1.000
	MAX	0.055	0.307	0.523	0.840	0.972	0.998	1.000
	DF	0.051	0.198	0.344	0.652	0.886	0.976	0.998
	GLS	0.052	0.306(1.769)	0.424(2.418)	0.761(3.545)	0.901(4.735)	0.954(5.993)	0.979(7.325)
			(2.032)	(2.657)	(3.767)	(4.948)	(6.200)	(7.527)
	GLSu	0.052	0.231(0.909)	0.403(0.930)	0.720(0.967)	0.919(1.005)	0.991(1.046)	0.999(1.090)
			(0.922)	(0.936)	(0.968)	(1.004)	(1.043)	(1.086)

Table 5.2(b)

Size and power of the tests of the $I(1)$ null at the nominal 5% level, $T = 200$

		c						
δ_1		0	−7	−10	−15	−20	−25	−30
−0.98	*WS*	0.054	0.333	0.554	0.841	0.972	0.997	1.000
	MAX	0.054	0.323	0.531	0.833	0.970	0.998	1.000
	DF	0.050	0.197	0.336	0.635	0.865	0.971	0.996
	GLS	0.047	0.227(6.137) (7.124)	0.308(11.15) (12.33)	0.411(23.03) (24.57)	0.468(39.63) (41.52)	0.498(61.31) (63.53)	0.514(88.47) (90.99)
	GLSu	0.046	0.205(0.979) (1.014)	0.353(1.091) (1.121)	0.644(1.355) (1.384)	0.862(1.724) (1.753)	0.963(2.206) (2.233)	0.988(2.810) (2.833)
−0.95	*WS*	0.054	0.332	0.556	0.841	0.971	0.997	1.000
	MAX	0.054	0.322	0.531	0.833	0.970	0.998	1.000
	DF	0.051	0.197	0.335	0.635	0.865	0.970	0.996
	GLS	0.048	0.267(3.839) (4.458)	0.387(6.383) (7.066)	0.523(12.03) (12.85)	0.601(19.56) (20.53)	0.646(29.12) (30.22)	0.672(40.85) (42.10)
	GLSu	0.048	0.217(0.928) (0.956)	0.376(0.985) (1.007)	0.681(1.111) (1.131)	0.893(1.278) (1.299)	0.979(1.490) (1.513)	0.997(1.751) (1.776)
−0.90	*WS*	0.052	0.332	0.554	0.841	0.971	0.997	1.000
	MAX	0.054	0.323	0.531	0.833	0.970	0.997	1.000
	DF	0.050	0.197	0.336	0.634	0.862	0.969	0.996
	GLS	0.051	0.287(3.072) (3.566)	0.424(4.794) (5.306)	0.601(8.364) (8.932)	0.694(12.87) (13.50)	0.751(18.37) (19.08)	0.787(24.95) (25.73)
	GLSu	0.051	0.221(0.911) (0.936)	0.385(0.950) (0.968)	0.693(1.029) (1.043)	0.900(1.129) (1.143)	0.983(1.251) (1.265)	0.997(1.398) (1.412)
−0.85	*WS*	0.053	0.332	0.554	0.841	0.971	0.996	1.000
	MAX	0.055	0.323	0.531	0.833	0.969	0.997	1.000
	DF	0.049	0.197	0.334	0.633	0.861	0.968	0.995
	GLS	0.052	0.296(2.816) (3.268)	0.437(4.262) (4.718)	0.631(7.138) (7.622)	0.740(10.63) (11.15)	0.796(14.78) (15.35)	0.834(19.64) (20.25)
	GLSu	0.050	0.226(0.906) (0.930)	0.388(0.938) (0.955)	0.694(1.002) (1.014)	0.902(1.079) (1.090)	0.984(1.172) (1.182)	0.997(1.280) (1.290)
−0.80	*WS*	0.052	0.331	0.554	0.840	0.970	0.996	1.000
	MAX	0.055	0.321	0.532	0.833	0.968	0.997	1.000
	DF	0.050	0.197	0.333	0.633	0.860	0.967	0.995
	GLS	0.051	0.298(2.687) (3.118)	0.448(3.996) (4.422)	0.650(6.521) (6.964)	0.765(9.505) (9.973)	0.826(12.98) (13.48)	0.865(16.97) (17.50)
	GLSu	0.050	0.227(0.903) (0.926)	0.389(0.932) (0.948)	0.695(0.988) (1.000)	0.900(1.054) (1.064)	0.984(1.132) (1.141)	0.998(1.220) (1.229)

Table 5.2(b) (continued)

δ_1		c: 0	−7	−10	−15	−20	−25	−30
−0.75	*WS*	0.054	0.332	0.554	0.839	0.970	0.996	1.000
	MAX	0.056	0.322	0.531	0.833	0.968	0.997	1.000
	DF	0.050	0.197	0.332	0.630	0.860	0.968	0.995
	GLS	0.050	0.300(2.609)	0.455(3.834)	0.664(6.149)	0.778(8.825)	0.845(11.89)	0.884(15.36)
			(3.028)	(4.243)	(6.566)	(9.260)	(12.34)	(15.84)
	GLSu	0.051	0.227(0.901)	0.389(0.928)	0.694(0.980)	0.900(1.039)	0.984(1.107)	0.998(1.185)
			(0.924)	(0.944)	(0.991)	(1.048)	(1.115)	(1.192)
−0.70	*WS*	0.054	0.332	0.551	0.837	0.969	0.996	1.000
	MAX	0.057	0.321	0.530	0.831	0.968	0.997	1.000
	DF	0.050	0.198	0.335	0.630	0.859	0.967	0.995
	GLS	0.051	0.303(2.557)	0.466(3.726)	0.670(5.899)	0.792(8.369)	0.858(11.16)	0.896(14.28)
			(2.967)	(4.123)	(6.299)	(8.781)	(11.58)	(14.72)
	GLSu	0.050	0.227(0.900)	0.389(0.926)	0.692(0.974)	0.899(1.029)	0.983(1.091)	0.997(1.160)
			(0.923)	(0.942)	(0.985)	(1.038)	(1.099)	(1.167)
−0.65	*WS*	0.053	0.333	0.550	0.837	0.969	0.995	1.000
	MAX	0.056	0.322	0.527	0.829	0.967	0.997	1.000
	DF	0.050	0.196	0.333	0.629	0.857	0.966	0.995
	GLS	0.051	0.303(2.519)	0.467(3.648)	0.671(5.718)	0.796(8.039)	0.863(10.63)	0.907(13.50)
			(2.923)	(4.036)	(6.106)	(8.435)	(11.03)	(13.92)
	GLSu	0.050	0.227(0.899)	0.389(0.924)	0.692(0.970)	0.899(1.022)	0.982(1.079)	0.998(1.143)
			(0.922)	(0.940)	(0.981)	(1.030)	(1.086)	(1.149)
−0.60	*WS*	0.054	0.332	0.549	0.837	0.968	0.995	1.000
	MAX	0.056	0.320	0.527	0.829	0.967	0.997	1.000
	DF	0.050	0.196	0.335	0.626	0.855	0.966	0.995
	GLS	0.051	0.303(2.490)	0.471(3.588)	0.672(5.580)	0.801(7.787)	0.868(10.22)	0.911(12.90)
			(2.889)	(3.970)	(5.959)	(8.171)	(10.62)	(13.30)
	GLSu	0.051	0.227(0.898)	0.390(0.923)	0.692(0.967)	0.898(1.016)	0.981(1.070)	0.998(1.130)
			(0.921)	(0.938)	(0.977)	(1.024)	(1.077)	(1.136)
0.0	*WS*	0.052	0.333	0.554	0.848	0.970	0.997	1.000
	MAX	0.055	0.322	0.534	0.837	0.972	0.998	1.000
	DF	0.051	0.205	0.338	0.638	0.870	0.972	0.996
	GLS	0.051	0.303(2.347)	0.483(3.292)	0.713(4.900)	0.839(6.550)	0.911(8.244)	0.950(9.984)
			(3.199)	(3.643)	(5.232)	(6.873)	(8.562)	(10.30)
	GLSu	0.051	0.235(0.895)	0.397(0.916)	0.710(0.952)	0.893(0.989)	0.986(1.026)	0.998(1.065)
			(0.918)	(0.931)	(0.961)	(0.995)	(1.031)	(1.069)

Table 5.2(c)

Size and power of the tests of the $I(1)$ null at the nominal 5% level, $T = 500$

		c						
δ_1		0	−7	−10	−15	−20	−25	−30
−0.98	*WS*	0.055	0.302	0.517	0.832	0.964	0.997	1.000
	MAX	0.055	0.298	0.506	0.827	0.964	0.997	1.000
	DF	0.051	0.196	0.335	0.630	0.867	0.968	0.997
	GLS	0.056	0.243(4.918)	0.360(8.240)	0.492(15.50)	0.568(24.96)	0.611(36.69)	0.640(50.76)
			(5.729)	(9.144)	(16.58)	(26.24)	(38.17)	(52.45)
	GLSu	0.051	0.221(0.921)	0.375(0.977)	0.676(1.101)	0.892(1.261)	0.980(1.460)	0.998(1.699)
			(0.954)	(1.005)	(1.126)	(1.287)	(1.488)	(1.730)
−0.95	*WS*	0.054	0.301	0.518	0.832	0.963	0.997	1.000
	MAX	0.054	0.299	0.506	0.827	0.963	0.997	1.000
	DF	0.052	0.196	0.336	0.630	0.866	0.968	0.997
	GLS	0.054	0.267(3.715)	0.405(5.769)	0.573(9.880)	0.670(14.87)	0.730(20.77)	0.768(27.60)
			(4.327)	(6.402)	(10.57)	(15.64)	(21.61)	(28.52)
	GLSu	0.052	0.229(0.900)	0.387(0.935)	0.690(1.005)	0.904(1.090)	0.983(1.190)	0.999(1.306)
			(0.930)	(0.958)	(1.024)	(1.107)	(1.207)	(1.323)
−0.90	*WS*	0.054	0.301	0.517	0.831	0.962	0.997	1.000
	MAX	0.054	0.300	0.506	0.826	0.964	0.998	1.000
	DF	0.052	0.195	0.336	0.634	0.869	0.967	0.997
	GLS	0.051	0.278(3.314)	0.428(4.945)	0.624(8.006)	0.732(11.51)	0.794(15.46)	0.837(19.87)
			(3.859)	(5.487)	(8.567)	(12.10)	(16.08)	(20.53)
	GLSu	0.053	0.231(0.894)	0.389(0.921)	0.694(0.973)	0.908(1.033)	0.984(1.100)	0.999(1.175)
			(0.922)	(0.943)	(0.990)	(1.047)	(1.113)	(1.188)
−0.85	*WS*	0.054	0.300	0.517	0.833	0.962	0.997	1.000
	MAX	0.054	0.298	0.507	0.827	0.962	0.997	1.000
	DF	0.051	0.195	0.335	0.634	0.868	0.966	0.997
	GLS	0.051	0.278(3.179)	0.437(4.669)	0.641(7.379)	0.755(10.38)	0.819(13.68)	0.860(17.29)
			(3.703)	(5.181)	(7.897)	(10.91)	(14.23)	(17.86)
	GLSu	0.053	0.231(0.891)	0.389(0.917)	0.694(0.963)	0.908(1.014)	0.983(1.070)	0.999(1.131)
			(0.920)	(0.937)	(0.978)	(1.027)	(1.082)	(1.142)
−0.80	*WS*	0.053	0.300	0.518	0.832	0.963	0.997	1.000
	MAX	0.054	0.299	0.508	0.826	0.962	0.997	1.000
	DF	0.051	0.195	0.335	0.633	0.868	0.966	0.997
	GLS	0.051	0.278(3.112)	0.444(4.531)	0.650(7.064)	0.766(9.815)	0.830(12.79)	0.876(15.99)
			(3.624)	(5.027)	(7.560)	(10.32)	(13.31)	(16.52)
	GLSu	0.053	0.231(0.890)	0.389(0.914)	0.696(0.957)	0.909(1.004)	0.983(1.055)	0.998(1.109)
			(0.918)	(0.935)	(0.972)	(1.017)	(1.066)	(1.120)

Table 5.2(c) (continued)

		c						
δ_1		0	−7	−10	−15	−20	−25	−30
−0.75	*WS*	0.053	0.300	0.516	0.831	0.962	0.997	1.000
	MAX	0.054	0.298	0.508	0.826	0.962	0.997	1.000
	DF	0.052	0.195	0.335	0.635	0.866	0.965	0.997
	GLS	0.052	0.279(3.071)	0.445(4.447)	0.656(6.874)	0.773(9.473)	0.836(12.25)	0.882(15.20)
			(3.577)	(4.935)	(7.357)	(9.960)	(12.74)	(15.71)
	GLSu	0.053	0.231(0.890)	0.390(0.913)	0.696(0.954)	0.909(0.998)	0.983(1.045)	0.998(1.096)
			(0.918)	(0.933)	(0.969)	(1.011)	(1.056)	(1.106)
−0.70	*WS*	0.053	0.301	0.514	0.830	0.963	0.997	1.000
	MAX	0.054	0.297	0.507	0.825	0.962	0.997	1.000
	DF	0.052	0.195	0.335	0.633	0.864	0.965	0.997
	GLS	0.052	0.279(3.044)	0.448(4.391)	0.658(6.746)	0.774(9.243)	0.841(11.89)	0.888(14.68)
			(3.545)	(4.872)	(7.220)	(9.719)	(12.37)	(15.17)
	GLSu	0.053	0.232(0.889)	0.390(0.912)	0.694(0.952)	0.909(0.994)	0.983(1.039)	0.998(1.087)
			(0.917)	(0.932)	(0.967)	(1.007)	(1.050)	(1.097)
−0.65	*WS*	0.052	0.301	0.515	0.831	0.962	0.997	1.000
	MAX	0.054	0.297	0.507	0.824	0.961	0.997	1.000
	DF	0.052	0.194	0.334	0.631	0.864	0.964	0.996
	GLS	0.052	0.280(3.024)	0.450(4.350)	0.662(6.654)	0.775(9.077)	0.846(11.62)	0.891(14.30)
			(3.522)	(4.827)	(7.121)	(9.544)	(12.09)	(14.77)
	GLSu	0.054	0.231(0.889)	0.388(0.911)	0.693(0.950)	0.908(0.992)	0.982(1.035)	0.999(1.080)
			(0.917)	(0.931)	(0.965)	(1.004)	(1.045)	(1.090)
−0.60	*WS*	0.054	0.299	0.515	0.831	0.962	0.997	1.000
	MAX	0.054	0.296	0.506	0.824	0.961	0.997	1.000
	DF	0.052	0.193	0.333	0.630	0.864	0.964	0.996
	GLS	0.052	0.281(0.871)	0.450(4.319)	0.663(6.584)	0.779(8.951)	0.850(11.42)	0.895(14.01)
			(0.912)	(4.793)	(7.045)	(9.411)	(11.89)	(14.47)
	GLSu	0.054	0.232(0.888)	0.389(0.911)	0.692(0.949)	0.907(0.989)	0.980(1.031)	0.998(1.075)
			(0.916)	(0.931)	(0.964)	(1.001)	(1.042)	(1.085)
0.0	*WS*	0.054	0.303	0.519	0.833	0.964	0.997	1.000
	MAX	0.054	0.299	0.510	0.827	0.965	0.997	1.000
	DF	0.053	0.193	0.336	0.632	0.867	0.968	0.998
	GLS	0.055	0.282(2.934)	0.457(4.165)	0.676(6.234)	0.798(8.324)	0.870(10.44)	0.912(12.57)
			(3.417)	(4.622)	(6.671)	(8.752)	(10.86)	(12.99)
	GLSu	0.055	0.232(0.887)	0.392(0.908)	0.701(0.943)	0.911(0.979)	0.985(1.015)	0.999(1.051)
			(0.915)	(0.928)	(0.957)	(0.990)	(1.024)	(1.060)

Table 5.2(d)

Size and power of the tests of the $I(1)$ null at the nominal 5% level, $T = 1000$

		c						
δ_1		0	−7	−10	−15	−20	−25	−30
−0.98	*WS*	0.052	0.331	0.537	0.842	0.976	0.997	1.000
	MAX	0.053	0.317	0.520	0.825	0.969	0.997	1.000
	DF	0.047	0.177	0.318	0.609	0.856	0.966	0.996
	GLS	0.048	0.250(4.289)	0.376(6.791)	0.544(11.89)	0.628(18.15)	0.687(25.61)	0.722(34.27)
			(0.500)	(7.541)	(12.73)	(19.10)	(26.66)	(35.43)
	GLSu	0.047	0.213(0.901)	0.373(0.940)	0.677(1.018)	0.895(1.115)	0.982(1.230)	0.998(1.363)
			(0.934)	(0.965)	(1.040)	(1.135)	(1.250)	(1.384)
−0.95	*WS*	0.053	0.331	0.537	0.841	0.976	0.997	1.000
	MAX	0.052	0.317	0.519	0.825	0.968	0.998	1.000
	DF	0.047	0.177	0.317	0.609	0.856	0.966	0.995
	GLS	0.047	0.262(3.625)	0.398(5.432)	0.596(8.814)	0.701(12.66)	0.772(16.99)	0.815(21.79)
			(4.226)	(6.032)	(9.438)	(13.32)	(17.69)	(22.53)
	GLSu	0.048	0.217(0.891)	0.379(0.919)	0.684(0.971)	0.898(1.030)	0.983(1.097)	0.998(1.171)
			(0.922)	(0.942)	(0.989)	(1.046)	(1.112)	(1.185)
−0.90	*WS*	0.053	0.330	0.538	0.841	0.975	0.997	1.000
	MAX	0.052	0.317	0.520	0.825	0.968	0.997	1.000
	DF	0.047	0.177	0.317	0.608	0.855	0.966	0.995
	GLS	0.048	0.270(3.404)	0.411(4.979)	0.620(7.789)	0.736(10.83)	0.814(14.11)	0.864(17.63)
			(3.968)	(5.529)	(8.340)	(11.40)	(14.69)	(18.23)
	GLSu	0.047	0.216(0.888)	0.381(0.912)	0.686(0.955)	0.898(1.002)	0.983(1.053)	0.998(1.107)
			(0.918)	(0.934)	(0.972)	(1.017)	(1.066)	(1.119)
−0.85	*WS*	0.052	0.331	0.539	0.842	0.975	0.997	1.000
	MAX	0.052	0.316	0.520	0.825	0.968	0.997	1.000
	DF	0.048	0.176	0.317	0.608	0.854	0.966	0.996
	GLS	0.048	0.275(3.330)	0.420(4.828)	0.628(7.446)	0.746(10.22)	0.824(13.15)	0.874(16.24)
			(3.882)	(5.361)	(7.973)	(10.75)	(13.69)	(16.79)
	GLSu	0.047	0.218(0.887)	0.382(0.910)	0.688(0.950)	0.898(0.993)	0.983(1.038)	0.998(1.085)
			(0.916)	(0.932)	(0.966)	(1.007)	(1.050)	(1.097)
−0.80	*WS*	0.052	0.330	0.539	0.841	0.974	0.998	1.000
	MAX	0.053	0.316	0.521	0.826	0.968	0.997	1.000
	DF	0.048	0.176	0.317	0.605	0.853	0.965	0.996
	GLS	0.048	0.276(3.293)	0.425(4.752)	0.633(7.274)	0.751(9.912)	0.831(12.67)	0.881(15.54)
			(3.838)	(5.276)	(7.789)	(10.43)	(13.19)	(16.07)
	GLSu	0.047	0.219(0.886)	0.383(0.909)	0.689(0.947)	0.899(0.988)	0.983(1.030)	0.997(1.075)
			(0.916)	(0.930)	(0.964)	(1.002)	(1.043)	(1.086)

Table 5.2(d) (continued)

δ_1		c 0	−7	−10	−15	−20	−25	−30
−0.75	*WS*	0.052	0.332	0.539	0.842	0.974	0.998	1.000
	MAX	0.053	0.316	0.522	0.826	0.968	0.997	1.000
	DF	0.047	0.176	0.317	0.605	0.853	0.965	0.996
	GLS	0.048	0.277(3.270)	0.428(4.706)	0.638(7.170)	0.757(9.726)	0.835(12.38)	0.887(15.12)
			(3.812)	(5.225)	(7.678)	(10.23)	(12.88)	(15.63)
	GLSu	0.047	0.218(0.886)	0.383(0.908)	0.689(0.946)	0.900(0.985)	0.983(1.026)	0.997(1.068)
			(0.915)	(0.930)	(0.962)	(0.999)	(1.038)	(1.079)
−0.70	*WS*	0.052	0.332	0.539	0.842	0.974	0.998	1.000
	MAX	0.053	0.315	0.522	0.826	0.968	0.997	1.000
	DF	0.047	0.175	0.317	0.605	0.854	0.965	0.996
	GLS	0.048	0.277(3.255)	0.429(4.675)	0.640(7.100)	0.761(9.601)	0.840(12.18)	0.889(14.83)
			(3.795)	(5.191)	(7.603)	(10.10)	(12.68)	(15.34)
	GLSu	0.047	0.219(0.886)	0.383(0.907)	0.688(0.945)	0.900(0.983)	0.982(1.023)	0.997(1.064)
			(0.915)	(0.929)	(0.961)	(0.997)	(1.035)	(1.075)
−0.65	*WS*	0.052	0.331	0.540	0.842	0.974	0.998	1.000
	MAX	0.052	0.314	0.521	0.825	0.968	0.997	1.000
	DF	0.047	0.175	0.317	0.607	0.855	0.965	0.995
	GLS	0.048	0.277(3.244)	0.428(4.652)	0.641(7.049)	0.762(9.511)	0.841(12.04)	0.890(14.63)
			(3.782)	(5.166)	(7.549)	(10.01)	(12.53)	(15.12)
	GLSu	0.047	0.218(0.885)	0.384(0.907)	0.687(0.944)	0.900(0.982)	0.982(1.021)	0.997(1.061)
			(0.915)	(0.929)	(0.960)	(0.995)	(1.033)	(1.071)
−0.60	*WS*	0.052	0.331	0.540	0.841	0.974	0.997	1.000
	MAX	0.052	0.314	0.521	0.826	0.968	0.996	1.000
	DF	0.047	0.176	0.317	0.606	0.854	0.965	0.995
	GLS	0.048	0.276(3.236)	0.429(4.635)	0.641(7.011)	0.761(9.442)	0.843(11.93)	0.891(14.47)
			(3.772)	(5.147)	(7.508)	(9.933)	(12.42)	(14.96)
	GLSu	0.047	0.218(0.885)	0.384(0.907)	0.686(0.943)	0.899(0.981)	0.982(1.019)	0.998(1.058)
			(0.915)	(0.928)	(0.959)	(0.994)	(1.031)	(1.069)
0.0	*WS*	0.052	0.328	0.537	0.841	0.976	0.997	1.000
	MAX	0.051	0.314	0.520	0.825	0.969	0.997	1.000
	DF	0.048	0.180	0.320	0.609	0.856	0.966	0.996
	GLS	0.051	0.281(3.195)	0.437(4.551)	0.648(6.819)	0.775(9.100)	0.855(11.39)	0.903(13.70)
			(3.724)	(5.053)	(7.302)	(9.574)	(11.86)	(14.16)
	GLSu	0.049	0.219(0.885)	0.384(0.905)	0.691(0.940)	0.900(0.975)	0.984(1.011)	0.998(1.046)
			(0.914)	(0.927)	(0.956)	(0.989)	(1.022)	(1.057)

Table 5.2(e)

Size and power of the tests of the $I(1)$ null at the nominal 5% level, $T = 4000$

		c						
δ_1		0	−7	−10	−15	−20	−25	−30
−0.98	*WS*	0.053	0.347	0.542	0.840	0.969	0.997	1.000
	MAX	0.053	0.331	0.529	0.831	0.968	0.997	0.999
	DF	0.052	0.196	0.333	0.628	0.865	0.971	0.996
	GLS	0.056	0.322(3.714)	0.454(5.484)	0.636(8.679)	0.747(12.18)	0.806(15.99)	0.847(20.11)
			(4.332)	(6.093)	(9.297)	(12.82)	(16.66)	(20.80)
	GLSu	0.051	0.237(0.887)	0.393(0.912)	0.701(0.958)	0.914(1.008)	0.983(1.062)	0.998(1.121)
			(0.918)	(0.936)	(0.976)	(1.024)	(1.077)	(1.135)
−0.95	*WS*	0.053	0.347	0.543	0.841	0.969	0.997	1.000
	MAX	0.052	0.331	0.529	0.832	0.968	0.997	0.999
	DF	0.051	0.196	0.333	0.628	0.866	0.971	0.996
	GLS	0.056	0.332(3.535)	0.468(5.118)	0.660(7.855)	0.772(10.71)	0.831(13.70)	0.877(16.80)
			(4.123)	(5.686)	(8.415)	(11.28)	(14.27)	(17.38)
	GLSu	0.052	0.237(0.884)	0.394(0.907)	0.701(0.946)	0.914(0.987)	0.984(1.029)	0.998(1.073)
			(0.915)	(0.930)	(0.963)	(1.002)	(1.043)	(1.086)
−0.90	*WS*	0.054	0.347	0.542	0.841	0.969	0.997	1.000
	MAX	0.052	0.331	0.528	0.832	0.968	0.997	0.999
	DF	0.052	0.196	0.333	0.627	0.866	0.971	0.996
	GLS	0.056	0.338(3.476)	0.480(4.996)	0.667(7.580)	0.782(10.23)	0.846(12.93)	0.885(15.70)
			(4.054)	(5.551)	(8.120)	(10.76)	(13.47)	(16.24)
	GLSu	0.051	0.238(0.883)	0.394(0.905)	0.702(0.942)	0.915(0.980)	0.983(1.018)	0.998(1.058)
			(0.914)	(0.928)	(0.959)	(0.994)	(1.031)	(1.070)
−0.85	*WS*	0.054	0.347	0.542	0.840	0.969	0.997	1.000
	MAX	0.052	0.330	0.528	0.832	0.968	0.997	0.999
	DF	0.051	0.196	0.333	0.627	0.865	0.971	0.996
	GLS	0.055	0.335(3.456)	0.485(4.956)	0.672(7.488)	0.783(10.06)	0.848(12.68)	0.887(15.33)
			(4.030)	(5.505)	(8.022)	(10.59)	(13.20)	(15.86)
	GLSu	0.050	0.238(0.883)	0.394(0.905)	0.702(0.941)	0.915(0.977)	0.984(1.015)	0.998(1.052)
			(0.914)	(0.927)	(0.958)	(0.992)	(1.027)	(1.064)
−0.80	*WS*	0.054	0.347	0.543	0.840	0.969	0.997	1.000
	MAX	0.052	0.331	0.528	0.832	0.968	0.997	0.999
	DF	0.051	0.196	0.333	0.628	0.866	0.971	0.996
	GLS	0.056	0.338(3.445)	0.487(4.935)	0.672(7.442)	0.784(9.979)	0.848(12.55)	0.889(15.15)
			(4.019)	(5.483)	(7.972)	(10.50)	(13.07)	(15.67)
	GLSu	0.051	0.238(0.883)	0.394(0.904)	0.702(0.940)	0.916(0.976)	0.984(1.013)	0.998(1.050)
			(0.914)	(0.927)	(0.957)	(0.991)	(1.026)	(1.062)

Table 5.2(e) (continued)

δ_1		c: 0	−7	−10	−15	−20	−25	−30
−0.75	*WS*	0.054	0.349	0.543	0.841	0.969	0.997	1.000
	MAX	0.053	0.331	0.527	0.832	0.967	0.997	0.999
	DF	0.051	0.196	0.333	0.628	0.866	0.971	0.996
	GLS	0.056	0.337(3.439)	0.487(4.923)	0.672(7.414)	0.784(9.929)	0.848(12.47)	0.891(15.03)
			(4.011)	(5.469)	(7.942)	(10.45)	(12.99)	(15.55)
	GLSu	0.051	0.238(0.883)	0.394(0.904)	0.702(0.940)	0.916(0.975)	0.984(1.012)	0.998(1.048)
			(0.914)	(0.927)	(0.957)	(0.990)	(1.024)	(1.060)
−0.70	*WS*	0.053	0.349	0.544	0.841	0.969	0.997	1.000
	MAX	0.052	0.331	0.527	0.832	0.967	0.997	0.999
	DF	0.051	0.196	0.333	0.628	0.867	0.970	0.996
	GLS	0.056	0.336(3.435)	0.488(4.914)	0.674(7.395)	0.784(9.896)	0.850(12.42)	0.893(14.96)
			(4.007)	(5.460)	(7.922)	(10.42)	(12.93)	(15.47)
	GLSu	0.051	0.238(0.883)	0.394(0.904)	0.701(0.939)	0.916(0.975)	0.984(1.011)	0.998(1.047)
			(0.914)	(0.927)	(0.956)	(0.989)	(1.024)	(1.059)
−0.65	*WS*	0.053	0.349	0.544	0.841	0.970	0.997	1.000
	MAX	0.052	0.331	0.527	0.832	0.967	0.997	0.999
	DF	0.051	0.197	0.333	0.629	0.867	0.970	0.995
	GLS	0.056	0.336(3.432)	0.487(4.908)	0.673(7.382)	0.786(9.872)	0.852(12.38)	0.895(14.90)
			(4.003)	(5.453)	(7.908)	(10.39)	(12.89)	(15.42)
	GLSu	0.051	0.238(0.883)	0.395(0.904)	0.702(0.939)	0.916(0.975)	0.984(1.010)	0.998(1.046)
			(0.914)	(0.927)	(0.956)	(0.989)	(1.023)	(1.058)
−0.60	*WS*	0.053	0.349	0.545	0.842	0.970	0.997	1.000
	MAX	0.052	0.332	0.526	0.831	0.967	0.997	0.999
	DF	0.051	0.197	0.333	0.628	0.867	0.970	0.995
	GLS	0.056	0.336(3.430)	0.485(4.904)	0.673(7.371)	0.785(9.853)	0.852(12.35)	0.894(14.86)
			(4.001)	(5.448)	(7.897)	(10.37)	(12.86)	(15.37)
	GLSu	0.051	0.237(0.883)	0.395(0.904)	0.702(0.939)	0.916(0.974)	0.984(1.010)	0.998(1.046)
			(0.913)	(0.927)	(0.956)	(0.989)	(1.023)	(1.057)
0.0	*WS*	0.054	0.347	0.543	0.842	0.970	0.997	1.000
	MAX	0.053	0.331	0.529	0.831	0.967	0.997	0.999
	DF	0.052	0.197	0.334	0.627	0.867	0.971	0.996
	GLS	0.054	0.339(3.419)	0.491(4.881)	0.676(7.320)	0.784(9.762)	0.853(12.21)	0.897(14.66)
			(3.988)	(5.422)	(7.842)	(10.27)	(12.71)	(15.16)
	GLSu	0.051	0.239(0.883)	0.395(0.903)	0.701(0.938)	0.915(0.973)	0.984(1.008)	0.998(1.043)
			(0.913)	(0.926)	(0.955)	(0.987)	(1.021)	(1.054)

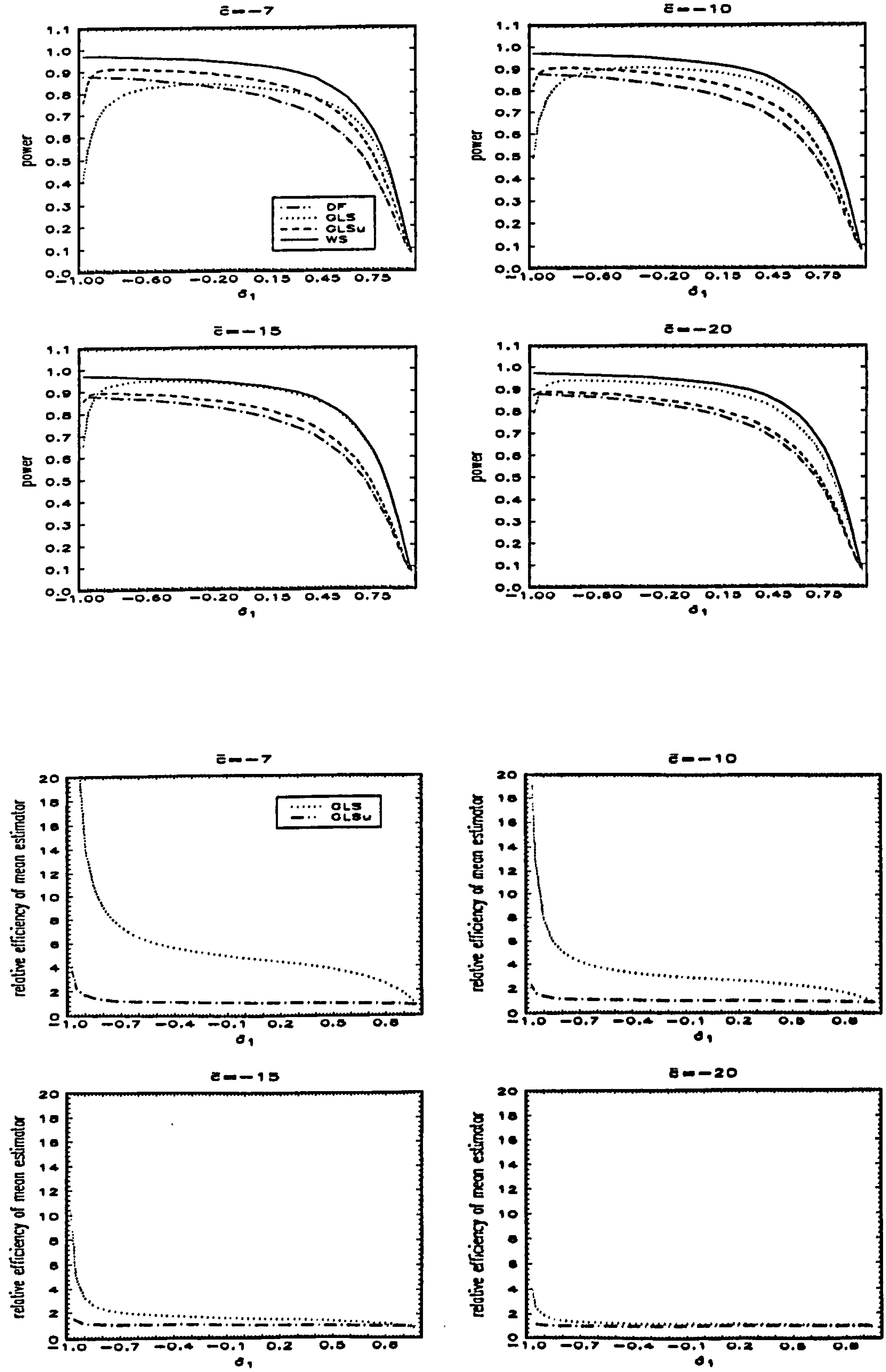


Figure 5.4(a). Power functions of unit root tests and relative efficiencies of the mean estimator for the GLS and $GLSu$ tests, $c = -20$ and $T = 100$.

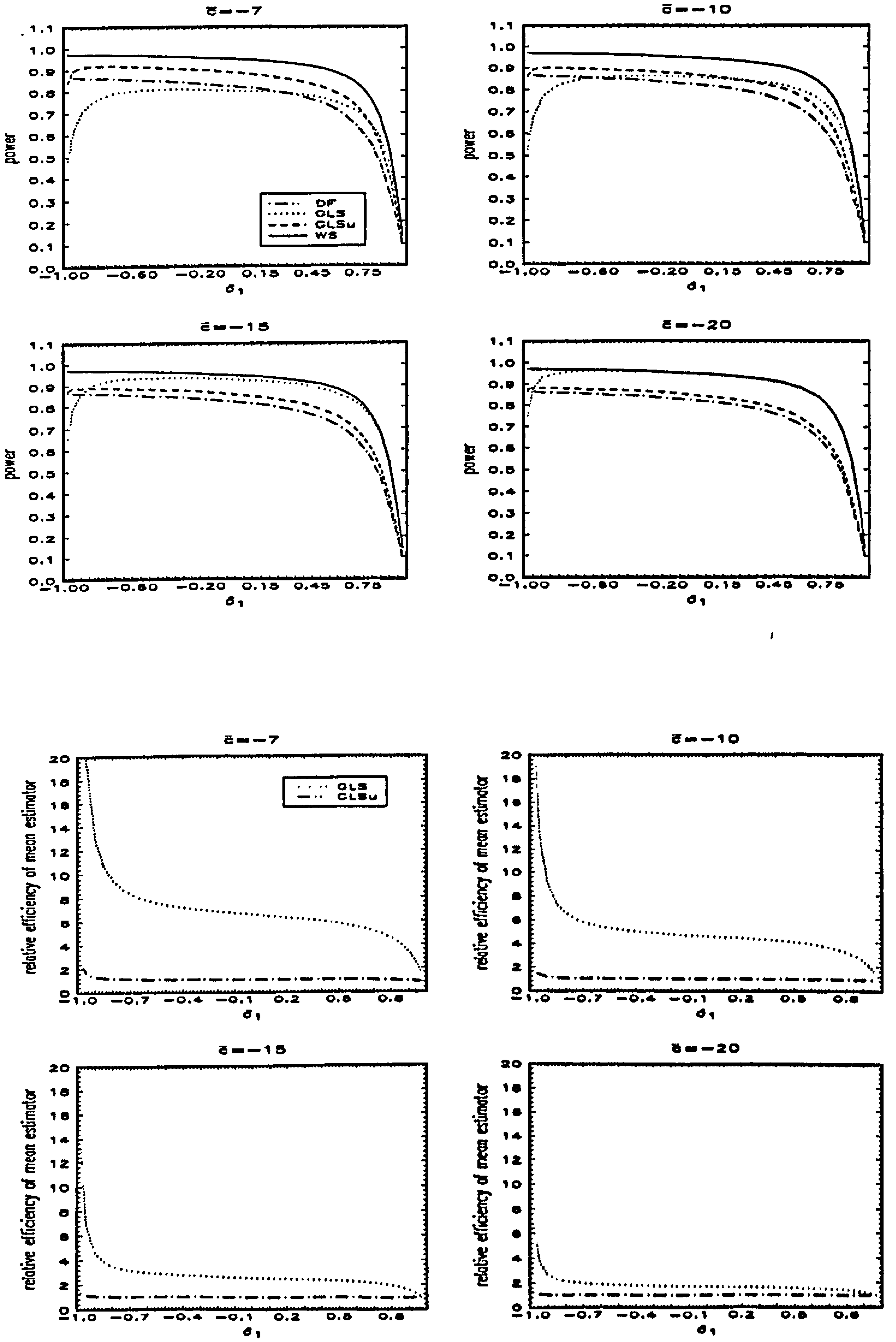


Figure 5.4(b). Power functions of unit root tests and relative efficiencies of the mean estimator for the GLS and $GLSu$ tests, $c = -20$ and $T = 200$.

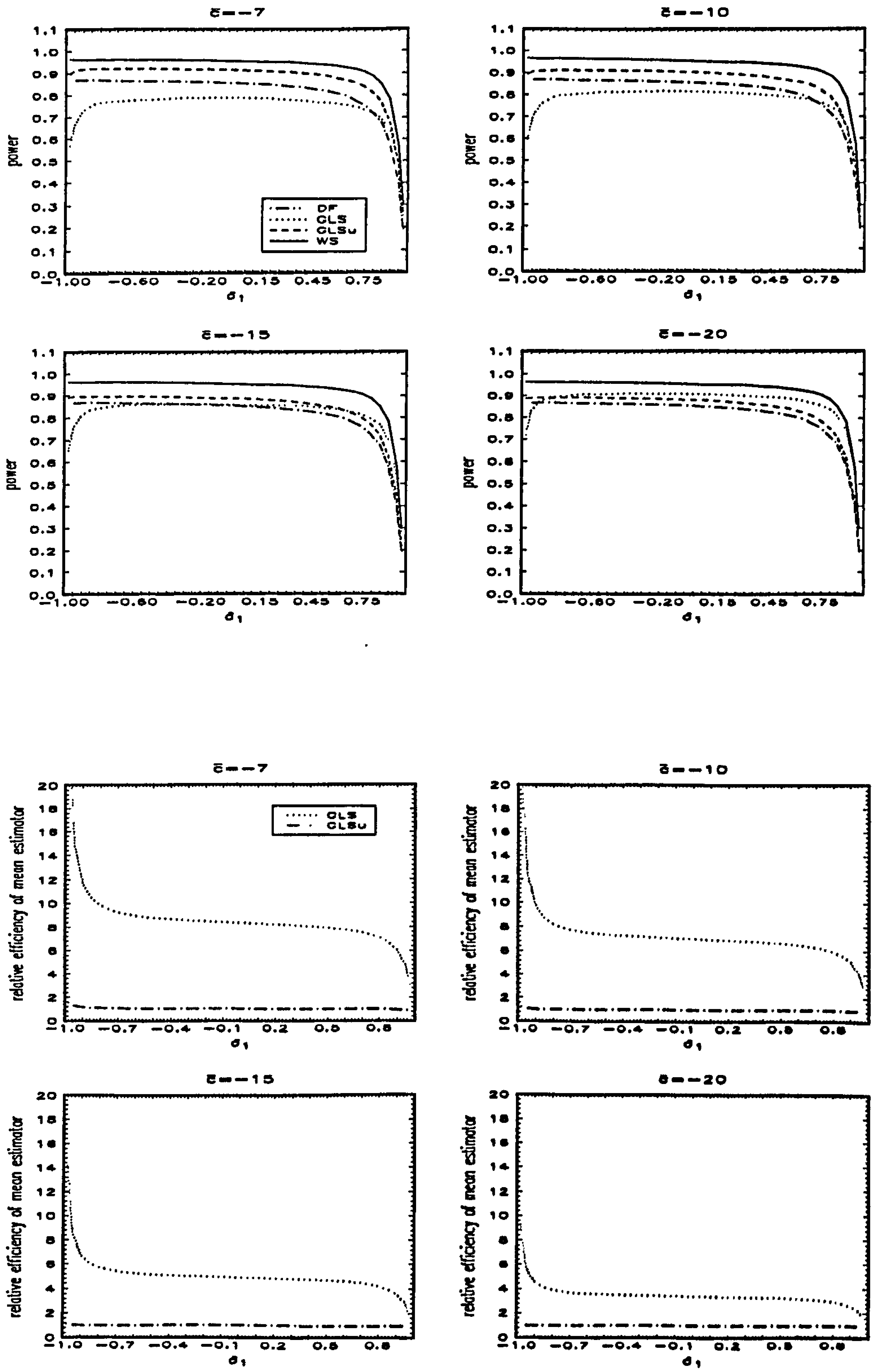


Figure 5.4(c). Power functions of unit root tests and relative efficiencies of the mean estimator for the *GLS* and *GLSu* tests, $c = -20$ and $T = 500$.

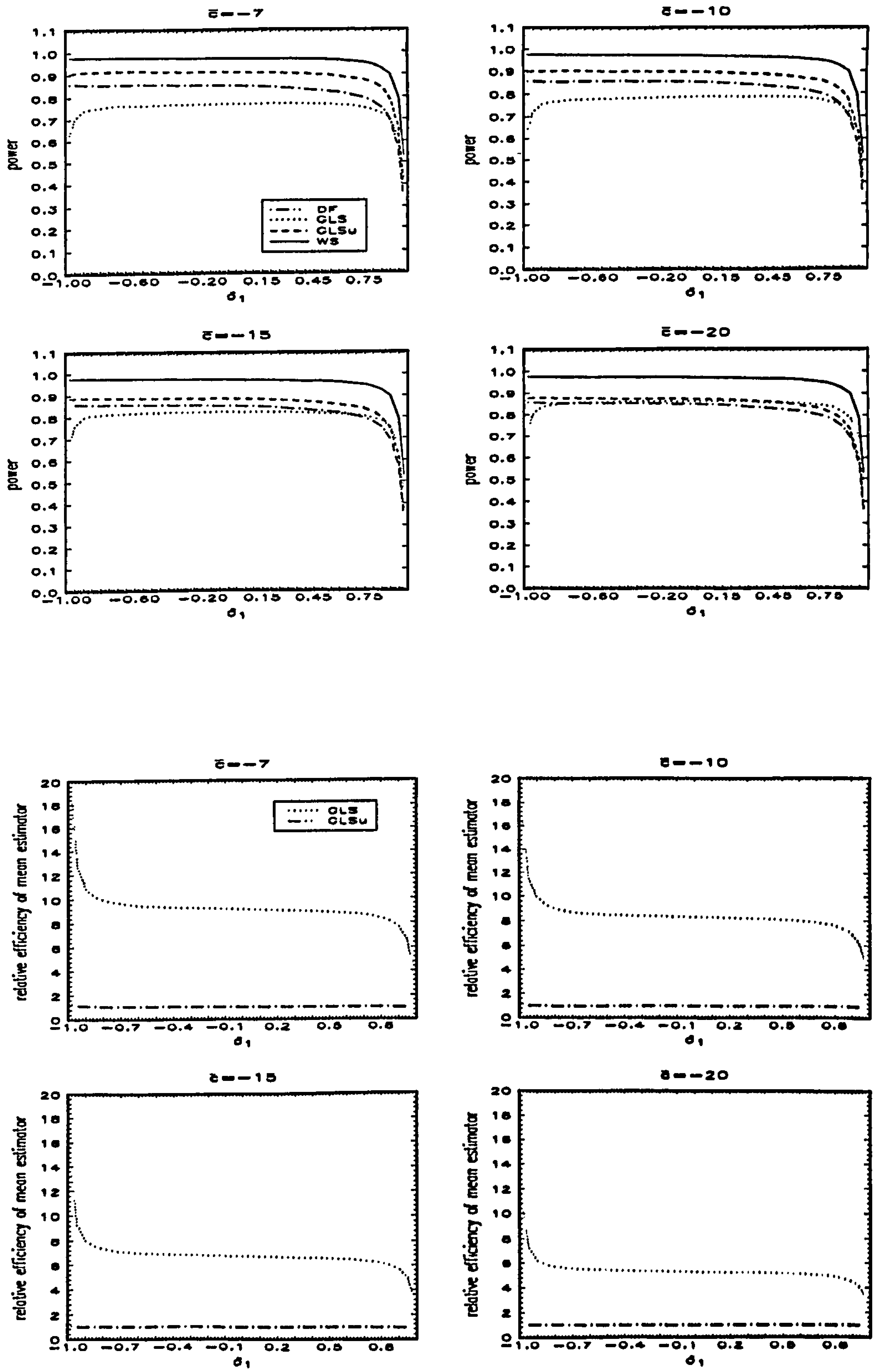


Figure 5.4(d). Power functions of unit root tests and relative efficiencies of the mean estimator for the GLS and $GLSu$ tests, $c = -20$ and $T = 1000$.

Chapter 6

Concluding Remarks and Suggestions for Future Research

This thesis has focused primarily on testing the unit root hypothesis under the phenomenon of structural change and on exploring the behaviour, in alternative settings, of power-enhancing elaborations of commonly applied unit root tests. The latter have emerged as a solution to the well-known problem of low power that traditional unit root tests encounter.

Its main contribution lies in the development of reliable statistical techniques in response to the need for optimally characterising the behaviour of economic series over time which has important implications for forecasting and policy purposes. In particular, unit root tests were developed designed to have power against alternatives involving a shift in persistence. The likelihood of such a shift that refers to a change in structure of a time series from difference stationarity, $I(1)$, to trend stationarity, $I(0)$, or from $I(0)$ to $I(1)$ has been argued by a number of studies. In the most general case, it was not necessary a priori to specify the location or direction of any possible switch under the alternative hypothesis, while consistency of the estimated breakpoint directly followed. Base on simulation evidence the tests appear to perform well in terms of size and power, although some caution was recommended in their use in the case of non-normal errors. Applying the tests to data on consumer price inflation in the G-7 countries uncovered quite strong and consistent evidence of a change from trend-stationarity to difference-stationarity occurring in 1973 for the majority of these countries.

An alternative type of structural change related to the trend function was next revisited, that of a smooth transition type. Given the restrictive nature of the linear trend alternative when testing for a unit root in the standard case, the allowance for a more flexible and intuitive plausible specification of the trend function warrants considerable attention. The aim was to examine whether the additional power generated by the recently emerged unit root tests in the standard case, persists when adopted in this context. In short, the results suggested that it does although the power advantage of the modified smooth transition tests showed up as moderate compared to the standard case. This finding appeared to be linked to the elaborate nature of the trend function in this setting. Application of the modified smooth transition tests to common US macroeconomic series resulted in stronger evidence in favour of stationarity around a smooth transition in linear trend.

The power-enhancing unit root tests in the univariate case were subsequently adopted in the panel setting, where extensions of commonly applied unit root tests have increasingly been employed to tackle issues such as the reversion or otherwise of series to a fixed mean or trend. Such an undertaking led to the development of modified panel data unit root tests that were found to enjoy the same power advantage as the power-inducing tests in the univariate case, both when independence across the units of the panel was assumed and when allowing for a common time-specific component. A panel of series of real exchange rates against the US dollar was analysed, where the bootstrap method was employed to accommodate the heterogeneous nature of cross-section correlation found amongst the innovations generating the individual time series in the panel. Simulation results showed that modified bootstrap tests retain the power gains noted in simpler cases, while the application of these tests yielded appreciably stronger evidence against the unit root null hypothesis for our data than did the unmodified tests. Overall, evidence suggested the value of the Max and WS modifications to the IPS tests. Considerable additional power can be achieved while retaining size reliability.

Finally, the behaviour of the commonly applied GLS power-enhancing unit root test was explored under higher order autoregressive processes, the main issue being the low rejection probabilities observed for this test relative to the

Dickey-Fuller test for certain parameter configurations. A theoretical attempt was made to identify this 'unsatisfactory behaviour' of the *GLS* test, through investigation of the relative efficiencies of the mean estimator, by appealing to local to unity asymptotics. In doing so, simple approximate expressions were obtained for the relative efficiencies under general linear processes which proved fairly precise when compared to exact results. The asymptotic findings showed dependence of the relative efficiencies on the localizing parameter only and not on the underlying model. An extensive simulation study tailored to the AR(2) model both in finite and large samples confirmed the asymptotic predictions while also highlighted the finite sample effects of the second order autoregressive parameter. Comparisons of power across alternative power-enhancing unit root tests in this setting, highlighted the superior and robust performance of the *MAX* and *WS* tests, that did not appear to share any of the shortcomings associated with the *GLS* test. These results raise doubts regarding the use of the *GLS* test in practical applications, particularly if one also considers the dependence of such a test on the value of the localising parameter employed in detrending.

Constructing tests designed to have power against a specific type of structural change known to characterise key economic variables, understanding and evaluating the behaviour of newly existent tests in more flexible settings, introducing more powerful and reliable test statistics in conducting inference, being aware of any shortcomings associated with tests widely applied in the empirical literature; altogether, these were the main issues addressed in this thesis which are of great importance for optimally characterising the behaviour of economic series over time and ultimately for more accurate econometric modelling and policy design. It is hoped that the results will appear useful and provide stimulus for further research.

Regarding issues on the agenda for future research, it would be interesting to extend the tests for a change in persistence proposed in the second chapter to accommodate explosive "phases" as a way of testing for rational bubbles in the stock market. Such an extension would follow the work of Hall *et al.* (1999) who generalize the standard *ADF* testing strategy to allow for the possibility of Markov regime-switching where the regime is employed to make the root of an autoregression move between a unit-root and an explosive root in detecting periodically collapsing bubbles. Furthermore, another important issue that was briefly mentioned at the outset is the extension of the results to allow for changes affecting simultaneously alternative parameters of the model.

The test statistics introduced in chapter four are only valid in the absence of long-run relationships that tie the units of the panel together. Given however the near unit root behavior of many economic series the existence of such relationships is very likely. Recently, Banerjee *et al.* (2001) have demonstrated that in the presence of cross-unit cointegrating relationships quite serious spurious rejections occur, rendering existing tests substantially oversized. A challenging task would therefore be to augment the analysis of chapter four to account for cross-section common stochastic trends. The explicit introduction of cross sectional dependence when working with common driving trends across the units of the panel would mean dealing with both cross-sectional dependence and common trends simultaneously.

Finally, an alternative issue of interest would be to explore the sensitivity of the modified panel data unit root tests to structural breaks and develop further methodology to incorporate such a feature. A first step in this direction was made by Silvestre *et al.* (2001), who in common with the traditional tests designed for the unidimensional case found the panel unit root test proposed by Harris and Tzavalis (1999) to perform poorly when there is a structural break in the time series under the alternative. They proceed by extending such a panel unit root test to account for the existence of a level shift at an unknown time in the deterministic part of the series.[1] In an application to Spanish unemployment they show how the hypothesis of (perfect) hysteresis is rejected in favour of the alternative of the natural unemployment rate, when the possibility of a change in the latter is considered.

Similarly, Im and Lee (1999) extend the LM-bar statistic to take account of the possibility of a structural break. They conclude that their proposed $\overline{LM}_B$ test loses little power by controlling for spurious structural breaks when they do not exist, and thus suggest controlling for breaks even when their presence is only suspected. Such a finding is based on the asymptotic invariance of the individual LM test to the location of the breakpoint.

[1] Their analysis allows for only one structural break to affect all the time series at the same date and therefore accommodates panels that are subjected to similar shocks. Asymptotics are carried out assuming T is fixed and $N \to \infty$.

References

Banerjee, A., Marcellino, M., Osbat, C., 2001. Testing for PPP: Should we use panel methods? Working paper, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

Silvestre, J.L.C., Barrio-Castro,T., Lopez-Bazo, E., 2001. Level shifts in a panel data based unit root test: An application to the rate of unemployment. Working Paper, Departament d' Econometria, Universitat de Barcelona.

Hall, S.G., Psaradakis, Z., Sola, M., 1999. Detecting periodically collapsing bubbles: A Markov-switching unit root test. *Journal of Applied Econometrics*, 14, 143-154.

Harris, R.D.F., Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics*, 91, 201-226.

Im, K.S., Lee, J., 2001. Panel LM unit root test with level shifts. Working Paper, Economics Department, University of Central Florida.