29

TESTING FOR UNIT ROOTS AND COINTEGRATION IN HETEROGENEOUS PANELS

by

Yuthana Sethapramote

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy In Economics

University of Warwick, Department of Economics

January 2005

PAGE NUMBERS CUT OFF IN THE ORIGINAL

To my parents

Table of Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Acknowledgements

I am grateful to numerous people who have contributed to the shaping of this thesis. At the outset, I would like to express my appreciation to my supervisors, Mike Clement, for his guidance, patience and encouragement throughout my PhD study at Warwick, and Jeremy Smith for his encouragement and helpful advice since my MSc study and throughout my PhD research. His MSc time-series analysis class has had a significant influence on my research and future career.

I would also like to acknowledge many other colleagues: P'Mee, P'Pong, May, Prem, Pedro, Tong, Cherry, Kay, Mimi and Pim, for their help and support in many ways during my study at Warwick.

Financial support provided by the Royal Thai Government Scholarship is gratefully acknowledged.

Finally, I owe an invaluable debt to my parents for their unconditional love and support.

Declaration of the Author

I declare that this thesis is my own work and has not been submitted for a degree in another university.

No material contained in this thesis has been previously published.

Yuthana Sethapramote

January 21, 2005

Abstract

This thesis undertakes a Monte Carlo study to investigate the finite sample properties of several panel unit root and cointegration tests. To this end, we consider a number of different experiments which potentially affect the properties of the tests.

We first consider panel unit root tests in heterogenous panels. Application of the panel tests of Im, Pesaran and Shin (2003) (IPS), and Maddala and Wu (1999) (MW) increases their power over the standard ADF test. However, the power of the tests is significantly diminished when the panel is dominated by the non-stationary series. Neglecting the presence of cross-sectional dependence results in serious size distortions. In view of this, a variety of methods are applied to correct the size distortions. However, the power of all tests is diminished as the cross-correlations reduce the amount of independent information in the panel.

The simulation results from the panel cointegration tests extend the findings of the unit root tests to multivariate cases. The likelihood-based panel rank test of Larsson, Lyhagen and Lothgren (2001) is found to be more powerful than the residual-based panel tests of IPS and MW, but slightly over-sized in moderate sample sizes (T). The effects of a mixed panel and of cross-correlations in the errors are similar to those of panel unit root tests. Therefore, we again, use the bootstrap method and the Cross-sectionally augmented IPS test (CIPS) of Pesaran (2003) to correct the size distortions.

The presence of structural breaks affects the size and power properties of any panel unit root tests which fail to cope with it. When the break dates are known, the exogenous break panel LM test is applied, to control the effect of structural shifts. In addition, the endogenous break selection procedures are used to estimate the break points. The endogenous break panel LM test also performs considerably well in terms of the size, power and accuracy with which the true break points are estimated.

Finally, application of the panel unit root and cointegration tests provide some evidence in support of the existence of long-run PPP and the monetary model in Asia Pacific countries. In addition, the presence of structural breaks as the impact of the currency crisis is also detected. However, evidence is found to be sensitive to the choice of deterministic terms (intercepts, trends), the methods used to estimate the panel test statistic (e.g. SUR and CIPS) and the break-point selection criteria.

Chapter 1

Introduction and Overview

1.1 Motivation

Standard unit root (cointegration) tests have limited power to reject the null hypothesis of non-stationarity (no cointegration) when the underlying process is highly persistent. The problem of low power is particularly severe in small samples (see Lothian and Taylor (1997), and Maddala and Kim (1999)). Recently, there has been a surge of interest in the adding of information from the cross-section dimension to form panels, in order to investigate the effect of this additional information on the performance of both unit root and cointegration tests. Various panel unit root and cointegration tests have been proposed, based on macro-panels with both large N (cross-section dimension) and T (length of time-series). In terms of characteristics, these large N, large T panels differ from the traditional large N, small T panels. When T is large, there is an obvious need to address issues of non-stationarity in the data.

Panel unit root tests gained popularity with the tests introduced by Quah (1994), Levin, Lin and Chu (2002) (henceforth, LLC), Im, Persasan and Shin (2003) (henceforth, IPS), Maddala and Wu (1999) (henceforth, MW), and Choi (2001). These panel unit root tests are an extension of standard unit root tests, and offer the opportunity to increase the power of the unit root tests using data from the cross-section dimension. LLC construct a unit root test for homogeneous panels based on the Augmented Dickey-Fuller (ADF) t-statistics constructed from the sample pool estimator with some modifications. However, this homogeneity assumption is very restrictive. The main disadvantage of this approach is that differences in adjustment speeds and dynamics across cross-sectional units are not taken into account. Alternatively, IPS and MW propose panel unit root tests for heterogeneous panels. The panel IPS test is calculated from the average t-statistic of the individual ADF regressions. The t-bar statistic is then adjusted, using its mean and variance. This standardised IPS statistic is asymptotically distributed as a standard normal. Similarly, the panel MW test is calculated from the p-values of the individual unit root tests, and has a standard chi-square distribution. These panel unit root tests provide researchers with the advantage of increasing the dimension from the individual unit root tests, while still allowing for heterogeneity of the individual series in the panel.

The panel data methodology has been further extended to test for cointegration relationships. There are two main approaches in the literature on cointegration analysis within panels. The first approach is a panel version of the Engle and Granger (1987) residual-based two-step cointegration test. In this approach, a long-run relationship is estimated in the first step. In the second step, a test for the existence of unit roots on the residuals obtained from the long-run regressions based on the ADF regressions is constructed. Kao (1999) develops a number of variants of the residual-based panel cointegration tests based on a homogeneity assumption. Pedroni (1999) introduces a number of panel cointegration statistics based on both homogeneity and heterogeneity assumptions. In addition, heterogeneous panel cointegration tests can be estimated, using the panel IPS and MW unit tests on the residuals from the long-run regressions. The second approach in testing for cointegration in panel data applies the likelihood ratio (LR) test for the cointegration rank in a VAR of Johansen (1988). Larsson, Lyhagen and Lothgren (2001) (henceforth, LLL) develop a panel test for determining cointegrating rank in the long-run Π matrix as the average of the individual likelihood-based cointegration rank trace test statistics. This LLL *LR*-bar statistic, defined similarly as the IPS *t*-bar statistic, is also based on heterogeneous panels.

There remains a number of concerns regarding the testing for unit roots and cointegration in panel data. First, MW make reference to the case where there is a mixture of stationary and non-stationary series in the groups as an alternative hypothesis. Theoretically, in heterogeneous panels, the null hypothesis of non-stationarity can be rejected when there is at least one stationary series in the panel. However, the power of any panel test may drop significantly in a mixed panel dominated by non-stationary series. Second, the properties of the panel test statistics are based on the assumption that the error terms in each cross-section are independent. The effect of cross-sectional dependence has been discussed in several papers (see O'Connell (1998) and Cerrato (2001)). In cross-country data, the presence of cross-sectional correlation is likely to arise, due to the existence of inter-economy linkages. However, the presence of cross-sectional dependence in the error terms means that the limit distributions of the panel unit root and cointegration tests are no longer valid. O'Connell (1998) points out that the panel unit root tests that neglect the cross-sectional correlation can be seriously over-sized. In addition, even if the true distribution of the test statistic are available, the power of the test decreases as the total amount of independent information contained in the panel is reduced.

Recently, several methods have been proposed to control for the effect of cross-sectional dependence in the panel unit root tests. IPS suggest removing the effect of the common time-specific components by subtracting the cross-sectional mean from each individual series before applying the tests. However, this demeaning procedure is valid only in the case of homogeneous cross-sectional dependence, and is not robust if the time-specific components vary across the groups. Alternatively, MW recommend a bootstrap procedure to calculate the empirical distribution of the test statistic to compensate for the size distortions of the conventional IPS and MW tests in cross-correlated panels. Finally, Pesaran (2003) introduces a Cross-sectionally augmented IPS test (CIPS), which approximates the structure of error correlation by a factor model. This CIPS test applies the standard ADF regressions augmented with the cross-section average of lagged levels and first-differences of the individual series.

Another issue, which had generated wide-ranging discussion in the unit root literature in the last decade, is the presence of structural changes in time-series data. Testing for unit roots, allowing for possible structural breaks, has received considerable attention since the pioneering work of Perron (1989). A shift in the intercept and/or the trend function of a stationary time-series reduces the power of standard unit root tests (see Perron (1989)). Recently, standard unit root tests have been adjusted to discriminate between the existence of a unit root process and a stationary process with structural instability. Perron (1989, 1990) proposes a modified ADF test to allow for a structural shift, by including a relevant dummy variable in the ADF regression. In these papers, the break point is assumed to be exogenously given. An endogenous break selection method has been developed subsequently by Zivot and Andrews (1992), Banerjee *et al.* (1992), and Perron and Vogelsang (1992), to determine the break point from the data. The most widely used endogenous selection procedure is the minimum test, which selects the break date by minimising the t-statistic for testing unit roots.

In the panel data framework, one would expect the existing panel unit root tests, such as the IPS and MW tests, to suffer from a significant loss of power in the presence of structural breaks in the data. Even though testing for unit roots, allowing for structural breaks, has been widely documented in the literature, panel unit root tests with structural shifts have not received much attention. The main difficulty in the application of structural changes in panel data is that the asymptotic property of Perron-type ADF t-statistics varies according to the location of breaks in the series. The expected values and variances of the ADF t-statistics at all different possible locations of breaks in the sample are then required in computing the IPS-type panel unit root test with structural breaks. Therefore, the calculation of these statistics is practically unmanageable.

Recently, Im, Lee and Tieslau (2002) (henceforth, ILT) have developed a new panel unit root test based on the Lagrangian Multiplier (LM) principle, which is a panel version of the LM unit root test of Amsler and Lee (1995). This LM unit root test has the same asymptotic distribution as that of the LM test without a shift, originally presented by Schmidt and Phillips (1992). The asymptotic distribution of this test does not depend on the nuisance parameters that indicate the position of a structural shift. ILT show that this invariance property of the univariate LM unit root test is still valid in their proposed panel LM unit root test. The panel LM test with a level shift can use the same means and variances that apply to the panel LM test without a shift. Moreover, this invariance property is also useful in constructing the tests based on heterogeneous panels. The panel LM unit root test can be applied when more than one structural shift occurs, or when structural shifts in each cross-section unit occur at different locations.

However, the ILT panel LM unit root test assumes that the number and location of breaks are accepted as a priori. Lee and Strazicich (2003) propose a univariate minimum LM unit root test, which extends the LM unit root test of Amsler and Lee (1995) to allow for the unknown break points that are determined endogenously from the data. The endogenous break LM test of Lee and Strazicich (2003) apply a selection criterion similar to that of the minimum test of Zivot and

Andrews (1992). The break points are selected to minimise the t-statistics used to test for the unit root null hypothesis. Lee and Strazicich (2003) show that the asymptotic property of this minimum LM unit root test does not depend on the location of breaks under the null hypothesis. This endogenous break selection procedure provides the method to determine the presence and the location of breaks from the data, which is practically useful.

Recently, a number of empirical researchers have applied panel unit root and cointegration tests to investigate several key economic issues, for example, growth and convergence (see Evans and Karras (1996) and Lee, Pesaran and Smith (1997)) and international R&D spillovers (see Kao, Chiang and Chen (1999)). However, the empirical study to have generated the greatest attention is in the field of fundamental exchange rate modelling. The standard economic theories that are widely used to explain the exchange rate movements are purchasing power parity (PPP) and the monetary model. Empirical research on exchange rates and their fundamental determination yields controversial results regarding the ability of fundamental economic factors to explain exchange rate movements (see Rogoff (1996) and Taylor (1997)). The failure to find favourable evidence in support of the PPP hypothesis and the monetary model is explained as a result of the low power of standard unit root and cointegration tests. Lothian and Taylor (1997) argue that the standard ADF test has extremely low power in rejecting the unit root null hypothesis for real exchange rates over the post-Bretten Woods sample period. Therefore, panel data analysis has been applied to improve the results over the conventional individual time-series analysis. Recently, several articles, such as Frankel and Rose (1996), Wu (1996), Papell (1997), and Coakley and Fuertes (1997) have found evidence to support PPP with regard to panel data, while Oh (1999) and Groen (2000) find positive results for the long-run relationship according to the monetary model.

Even though a number of empirical studies on exchange rate determination using panel data find evidence to validate PPP and the monetary model, these studies usually focus on industrial OECD countries. For less-developed countries, empirical evidence is still not widely investigated. The importance of Asia Pacific countries has rapidly increased in the last two decades with the rapid economic growth and strong trading ties to the world economy. In 1997, the Asia Pacific region was strongly affected by a severe currency crisis, which forced most countries to change their exchange rate regimes and implement structural economic reforms. The validity of the PPP hypothesis and the monetary model in the region has been applied in studies on the cause and impact of the currency crisis (see, for example, Chin (2000) and Razzaghipour *et al.* (2001)).

The effect of the 1997 East Asian crisis is a major concern for the testing of unit roots in real exchange rates and deviations from monetary fundamental. The crisis is likely to have produced a structural shift, which should be taken into account in testing for unit roots. For this reason, the panel LM unit root test of ILT is useful in modelling exchange rate movements in Asia Pacific countries. In this framework, the impact of the 1997 crisis and the low power of the individual unit root tests can be addressed, using the panel unit root test with structural shifts.

1.2 Objectives of the study

In light of the above discussion, the objectives of the thesis are to build on and extend the research in the field of panel data techniques applied to unit root and cointegration testing, and include:

- Evaluating the finite sample performance of several panel unit root and cointegration tests in terms of the size and power, when the length of time-series is moderate and the speed of adjustment in a mean-reverting process is quite slow.

- Examining the effect of a mixed panel of stationary and non-stationary series and the impact of cross-sectional dependence on the size and power properties of the panel tests.

- Comparing alternative methods for testing unit roots and cointegration in cross-correlated panels.

- Investigating the effect of structural breaks on the size and power properties of the panel unit root tests, both when the shifts are allowed and when they are neglected.

- Applying the panel unit root and cointegration tests in an empirical investigation into the presence of a long-run relationship between exchange rates and their fundamentals in Asia Pacific countries, and the impact of the 1997 East Asian currency crisis.

1.3 Thesis structure

The content of the thesis falls into four overall categories: panel unit root tests, panel cointegration tests, panel unit root tests with structural breaks, and an empirical investigation of exchange rates and their fundamentals in Asia Pacific countries.

In Chapters 2 to 4, the main methodology applied in the studies is Monte Carlo simulations. This method is used to investigate the finite sample properties of several panel unit root and cointegration tests in heterogeneous panels. The simulation results are used to compare the size and power performance of the tests, based on a number of different experiments.

Chapter 2 examines the panel unit root tests of IPS and MW. We focus on the improvement in the power of the panel unit root tests over the standard individual time-series tests. In addition, two concerns in testing for unit roots in heterogeneous panels are addressed. First, the effect of having a mixture of both stationary and non-stationary series in the panel is considered. Second, the impact of cross-sectional dependence is investigated. For cross-sectional dependence, we examine the performance of the three methods for the testing of unit roots in cross-correlated panels: the bootstrap method, the Seeming Unrelated Regression (SUR) and the Cross-sectionally augmented IPS test (CIPS).

Chapter 3 focuses on the panel cointegration tests. The panel unit root tests of IPS and MW are applied to test for cointegration relationships based on the residual-based methodology of Engle and Granger (1987). In addition, we investigate the panel cointegration test of LLL, which applies the method testing for the cointegration rank in a VAR of Johansen (1988) to the panel data framework. We, again, consider the effect of having a mixture of cointegrated and non-cointegrated relationships in the panel and of cross-sectional dependence. The bootstrap method and the CIPS test are applied to control for the effect of cross-correlation in the residual-based and likelihood-based panel cointegration tests.

Chapter 4 investigates the effect of structural breaks in testing for unit roots in panel data. We first examine the impact of level shifts in the series on the size and power properties of the standard panel IPS and LM unit root tests without shifts, as well as the panel LM unit root test of ILT. We evaluate the performance of the tests when the break points are exogenously determined and assumed to be a priori. Next, we apply the endogenous selection procedures to estimate the break points. The finite sample performance of the endogenous break panel LM unit root test is investigated in terms of the size, power and accuracy of selecting the true break dates.

In Chapter 5, we investigate the empirical evidence for a long-run relationship between exchange rates and their fundamentals in Asia Pacific countries. The PPP hypothesis and the monetary model are used as the fundamental determination of exchange rate movements. Several of the panel data methods analysed in Chapters 2 to 4 are then applied to an empirical investigation of long-run PPP and the monetary model. In addition, we address the impact of the 1997 East Asian currency crisis. The presence of a structural break, due to the aftermath of the currency crisis is considered in exchange rate modelling.

Chapter 6 concludes the thesis with a summary of the way in which the research objectives have been investigated. The contributions of the thesis to panel data unit root and cointegration testing are noted, as well as some possible applications of the findings. Some suggestions for future research are also made.

Chapter 2

Unit Root Tests in Heterogeneous Panels

2.1 Introduction

Testing for unit roots in panel data has attracted much attention in recent literature, and various statistics for testing such data have been proposed. Panel unit root tests have gained popularity since the pioneering papers of Quah (1994), and Levin, Lin and Chu (2002) (LLC). However, their tests are based on a homogeneity assumption, in which the autoregressive coefficients are the same across the individual series in the panel. This assumption is quite restrictive, implying identical speeds of mean reversion across series. Heterogeneous panel unit root tests (see, for example, Im, Perasan and Shin (2003) (IPS), Maddala and Wu (1999) (MW) and Choi (2001)) have been introduced to provide a method of increasing data through the cross-section dimension, whilst still preserving the heterogeneity of individual series. Heterogeneity is accommodated by computing unit root tests for each individual series independently. The panel test statistics are then calculated, based on a combination of test statistics across the panel.

IPS propose a t-bar statistic calculated from the t-statistics of the standard ADF test averaged across the panel. This t-bar statistic is then standardised, using its mean and variance, and shown to be asymptotically distributed as a standard normal.

MW propose a Fisher-type statistic calculated from the p-values of the individual unit root tests, and this statistic has a standard chi-square distribution. This Fisher test is also proposed by Choi (2001), who recommends several panel statistics based on combining the p-values from each cross-section unit, which have been often used in meta-analysis. However, Choi (2001) notes that the Fisher test is a more widely used statistic than his other proposed tests. For that reason, in this chapter, we focus on the IPS and MW tests.

Recently, many papers have highlighted several concerns with regard to testing for unit roots in heterogeneous panels. First, MW make reference to a mixed panel that combines both stationary and non-stationary series as an alternative hypothesis. In heterogeneous panels, we can reject the unit root null hypothesis, even though there is only one stationary series in the panel. However, the power of the panel tests would be considerably reduced in a mixed panel dominated by a non-stationary series.

Secondly, most panel unit root tests assume that the disturbance terms of the individual time-series in the panel are cross-sectionally independent. This assumption is acknowledged as being quite restrictive, especially in the context of cross-countries macroeconomics data sets created through strong links across markets. The violation of this assumption may seriously affect the performance of any panel unit root test in terms of size distortion and a loss in power, as suggested by O'Connell (1998).

In this chapter, the finite sample properties of the panel unit root tests of IPS and MW will be examined. The purpose of the chapter is to investigate the size and power performance of the panel IPS and MW tests through Monte Carlo simulations. In addition, we investigate the effect of having a mixture of both stationary and non-stationary series in the panel on the power of the IPS and MW tests. We also consider the effect of cross-sectional dependence in the error terms on both the IPS and MW tests. Finally, we investigate the performance of three alternative methods used to test for unit roots in cross-correlated panels. In particular, the bootstrap method of MW, the Seemingly Unrelated Regressions ADF test (SURADF) of Breuer, McNown and Wallace (2001), and the Cross-sectionally augmented IPS test (CIPS) of Pesaran (2003).

The remainder of this chapter is organised as follows. A literature review on panel unit root tests is carried out in Section 2.2. In Section 2.3, we discuss the effect of cross-sectional dependence on these panel unit root tests and the literature related to this issue. Monte Carlo experiments are carried out in Section 2.4, to evaluate the size and power performance of the IPS and MW tests. Section 2.5 presents the simulation results on a mixed panel. The effect of cross-sectional dependence on the performance of panel unit root tests is investigated in Section 2.6. Section 2.7 considers the performance of the bootstrap panel unit root test, unit root test with Seemingly Unrelated Regression (SUR) and the CIPS test of Pesaran (2003). Finally, Section 2.8 concludes the chapter.

2.2 Literature review

Interest in analysing panel data with non-stationary variables has recently increased. Quah (1994) presents an early development in testing for unit roots based on panel data, and suggests a simple unit root test, using the following regression:

$$y_{i,t} = \phi y_{i,t-1} + \varepsilon_{i,t} \qquad (2.1)$$

where $i = 1,2,\ldots, N$; $t = 1, 2, \ldots, T$ and $\varepsilon_{i,t} \sim iid\,(0,\sigma^2)$.

The asymptotic distribution of test statistics for the unit root null hypothesis ($H_0 : \phi = 1$) is derived as a mixture of standard normal and Dickey-Fuller distribution. However, this test has limited practical application, as it does not accommodate heterogeneity across groups, such as individual fix effects or different patterns of serial correlation in the error terms.

Breitung and Meyer (1994) introduce a panel test, which adjusts for individual specific means by subtracting each time-series with its first observation ($y_{i,1}$), so that the test regression is written as:

$$\Delta y_{i,t} = \rho(y_{i,t-1} - y_{i,1}) + \sum_{j=1}^{p_i} \theta_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t} \qquad (2.2)$$

where $i = 1,2,\ldots, N$; $t = 2, 3, \ldots, T$ and $\varepsilon_{i,t} \sim iid\,(0,\sigma^2)$.

The unit root null hypothesis can be tested by applying a conventional t-statistic of the null hypothesis, $\rho = 0$, using a standard t-distribution. However, this procedure is valid only in a model without trend. In a model with trend, a standard t-distribution is not valid in testing for the null hypothesis.

Levin and Lin (1992) develop a panel unit root test allowing for individual specific intercepts, time trends and serial-correlation in the disturbance terms. This test extends the standard ADF test on individual time-series. The basic equation is given by:

$$\Delta y_{i,t} = \rho y_{i,t-1} + \sum_{j=1}^{p_i} \theta_{i,j} \Delta y_{i,t-j} + \alpha_i + \delta_i t + \varepsilon_{i,t} \tag{2.3}$$

where $i = 1, 2, \ldots, N$; $t = 1, 2, \ldots, T$ and $\varepsilon_{i,t} \sim iid\,(0, \sigma^2)$, the lag order ($p_i$) can vary across i. The asymptotic distributions of the OLS pool panel statistics ($\hat{\rho}$) under models with the different deterministic terms are derived. Details on asymptotic properties of the proposed test are presented in the papers.

LLC extend the work of Levin and Lin (1992), applying it in the case where the error process has a more generally correlated and heteroscedastic structure, and consider three models with different deterministic terms: (i) no intercept, no trend ($\alpha_i = \delta_i = 0$), (ii) intercept, but without trend ($\alpha_i \neq 0, \delta_i = 0$) and (iii) with intercept and trend ($\alpha_i \neq 0, \delta_i \neq 0$). However, the LLC statistics are still based on the assumption that the coefficient ρ is homogenous across i. Therefore, the null and alternative hypotheses of the LLC panel unit root test are:

$H_0 : \rho = 0$ against $H_a : \rho < 0$

LLC propose a multi-step procedure to test for unit roots in panel data:

(1) Apply the ADF test to each individual series, that is:

$$\Delta y_{i,t} = \rho_i y_{i,t-1} + \sum_{j=1}^{p_i} \theta_{i,j} \Delta y_{i,t-j} + \alpha_i + \delta_i t + \varepsilon_{i,t} \tag{2.4}$$

LLC recommend selecting the lag order (p_i) using the method proposed by Hall (1994), which considers the significance of the t-statistic on the last augmented term ($\hat{\theta}_{i,j}$) to determine whether a smaller lag order is preferred. After determining the order of p_i, two auxiliary regressions are estimated, to generate orthogalised residual ($\hat{e}_{i,t}, \hat{v}_{i,t-1}$) by regress $\Delta y_{i,t}$ and $y_{i,t-1}$ against deterministic and augmented terms, respectively, that is:

$$\Delta y_{i,t} = \sum_{j=1}^{p_i} \theta_j^1 \Delta y_{i,t-j} + \alpha_i^1 + \delta_i^1 t + e_{i,t} \tag{2.5}$$

$$\hat{y}_{i,t-1} = \sum_{j=1}^{p_i} \theta_j^2 \Delta y_{i,t-j} + \alpha_i^2 + \delta_i^2 t + v_{i,t-1} \tag{2.6}$$

These partitioned regressions provide estimated residuals: $\hat{e}_{i,t}, \hat{v}_{i,t-1}$. To control for heterogeneity across individuals, $\hat{e}_{i,t}, \hat{v}_{i,t-1}$ are further normalised by the regression standard error ($\hat{\sigma}_{\varepsilon,i}$), i.e.:

$$\tilde{e}_{i,t} = \frac{\hat{e}_{i,t}}{\hat{\sigma}_{\varepsilon,i}}, \quad \tilde{v}_{i,t-1} = \frac{\hat{v}_{i,t-1}}{\hat{\sigma}_{\varepsilon,i}} \tag{2.7}$$

where $\hat{\sigma}_{\varepsilon,i}$ can be calculated from a regression of $\hat{e}_{i,t}$ against $\hat{v}_{i,t-1}$ as:

$$\hat{\sigma}_{\varepsilon,i} = \frac{1}{T - p_i - 1} \sum_{t=p_i+2}^{T} (\hat{e}_{i,t} - \hat{\rho}_i \hat{v}_{i,t-1})^2 \tag{2.8}$$

(2) Estimate the ratio of long-run to short-run standard deviations using the following method.

The long-run variance of equation (2.3) is estimated as:

$$\hat{\sigma}_{y_i}^2 = \frac{1}{T-1} \sum_{t=2}^{T} \Delta y_{i,t}^{*2} + 2 \sum_{L=1}^{\bar{K}} w_{\bar{K}L} \left(\frac{1}{T-1} \sum_{t=L+2}^{T} \Delta y_{i,t}^{*} \Delta y_{i,t-L}^{*} \right) \tag{2.9}$$

where $\Delta y_{i,t}^{*}$ is Δy_{it} adjusted by the cross-section average ($y_{i,t}^{*} = y_{i,t} - \frac{1}{N}\sum_{i=1}^{N} y_{i,t}$). $\bar{K}$ is the truncation lag parameter determined by the Andrew (1991) procedure and $w_{\bar{K}L}$ is a set of sample covariance weights, which depends on the choice of kernal, for example, when the Bartlett kernal is used, $w_{\overline{KL}} = 1 - \frac{k}{\overline{K}+1}$. The ratio of long-run standard deviation to the innovation (short-run) standard deviation (s_i) is then calculated by $\hat{s}_i = \frac{\hat{\sigma}_{y,i}}{\hat{\sigma}_{\varepsilon,i}}$. The average estimated standard deviation ratio is denoted by $\hat{S}_N : \hat{S}_N = \frac{1}{N}\sum_{i=1}^{N}\hat{s}_i$.

(3) Compute the panel test statistic by pooling all cross-sectional and time-series observations to estimate:

$$\tilde{e}_{i,t} = \rho\tilde{v}_{i,t-1} + \tilde{\varepsilon}_{i,t} \qquad (2.10)$$

The regression t-statistics for testing $H_0 : \rho = 0$ against $H_a : \rho < 0$ is then given by:

$$t_\rho = \frac{\hat{\rho}}{STD(\hat{\rho})} \qquad (2.11)$$

where $\hat{\rho} = \frac{\sum_{i=1}^{N}\sum_{t=p_i+2}^{T}\tilde{v}_{i,t-1}\tilde{e}_{i,t}}{\sum_{i=1}^{N}\sum_{t=p_i+2}^{T}\tilde{v}_{i,t-1}}$, $STD(\hat{\rho}) = \hat{\sigma}_{\tilde{\varepsilon}}[\sum_{i=1}^{N}\sum_{t=p_i+2}^{T}\tilde{v}_{i,t-1}^2]^{-\frac{1}{2}}$,

$\hat{\sigma}_{\tilde{\varepsilon}} = \frac{1}{N\tilde{T}}\sum_{i=1}^{N}\sum_{t=p_i+2}^{T}(\tilde{e}_{i,t} - \hat{\rho}\tilde{v}_{i,t-1})^2$, $\tilde{T} = T - p - 1$, $p = \frac{1}{N}\sum_{i=1}^{N}p_i$.

LLC show that this test statistic (t_ρ) has a standard normal distribution in a model without an intercept or trend, but diverges to negative infinity for a model with either an intercept or an intercept and trend. Therefore, they suggest the adjusted t-statistics (t_ρ^*), given by:

$$t_\rho^* = \frac{t_\rho - N\tilde{T}\hat{S}_N\hat{\sigma}_{\tilde{\varepsilon}}^{-2}STD(\hat{\rho})\mu_{m,\tilde{T}}^*}{\sigma_{m,t}^*} \qquad (2.12)$$

where the mean and standard deviation adjustment terms ($\mu_{m,\tilde{T}}^*, \sigma_{m,\tilde{T}}^*$) for a given deterministic specification are calculated by Monte Carlo simulations and given in

Table 2 of LLC. LLC show that, under the null hypothesis, $t_{\rho}^{*} \Rightarrow N(0,1)$ as $T, N \rightarrow \infty$.

The panel unit root tests proposed by Quah (1994), Britieng and Mayer (1994) and LLC are restrictive in assuming the autoregressive coefficients (ρ) to be homogenous across i. Alternative testing procedures, which allow for heterogeneity of the autoregressive coefficients, are proposed by IPS, MW and Choi (2001). Instead of pooling data in the estimation of a single t-statistic, these articles propose panel statistics based on the combining of individual time-series test statistics.

These panel unit root tests can be conducted by estimating separate unit root tests for each individual series in the panel, allowing for heterogeneous autoregressive coefficients in the panel. As an illustration, consider equation (2.4).

The null and alternative hypotheses of these tests are expressed as:

H_0: All series in panel are non-stationary ($\rho_i = 0$).

H_a: There is at least one series in panel, which is stationarity ($\rho_i < 0$).

IPS propose a panel unit root test, which combine the t-statistics from the individual ADF regressions. The IPS standardised t-bar statistic ($\psi_{\bar{t}}$) is defined as:

$$\psi_{\bar{t}} = \frac{\sqrt{N}(\bar{t}_{N,T} - \mu_T)}{\sigma_T} \qquad (2.13)$$

where $\bar{t}_{N,T} = \frac{1}{N}\sum_{i=1}^{N} t_{i,T}$ and $t_{i,T}$ is the t-statistic from the ADF regression for the i^{th} series. The adjustment terms in the IPS t-bar statistic ($\psi_{\bar{t}}$) are for mean (μ_T) and variance (σ_T^2) of $t_{i,T}$, which are tabulated from Monte Carlo simulations and shown in Table 3 of IPS. IPS show that this adjusted t-bar statistic follows a standard normal distribution ($\psi_i \Rightarrow N(0,1)$) as $T \to \infty$, followed by $N \to \infty$.

The value of adjustment terms (μ_t, σ_t) of the standardised IPS statistics depends on both the length of time span (T) and the number of lags in the individual ADF regressions (p_i). In unbalanced panels, when either T or p_i is different across the series, the adjusted t-bar statistic (ψ_i^*) is computed as:

$$\psi_i^* = \frac{\sqrt{N}\{\bar{t}_{N,T} - \frac{1}{N}\sum_{i=1}^{N}\mu_{T_i}(p_i)\}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\sigma_{Ti}^2(p_i)}} \tag{2.14}$$

Maddala and Wu (1999) propose a Fisher-type statistic, which combines the p-values of the test statistics from each cross-sectional unit to form a test statistic in panel data. This MW statistic (p_λ) can be computed as:

$$p_\lambda = -2\sum_{i=1}^{N}\ln(\pi_i) \tag{2.15}$$

where π_i is the p-value from each individual unit root test.

The MW test does not require a balanced panel, and each individual regression can be estimated with different T and p_i. This MW statistic is distributed as a chi-square distribution with $2N$ degrees of freedom as $T_i \to \infty$ for all N. In order to compute the MW test, the p-values of individual unit root tests are derived through simulations.

A Fisher-type statistic is also proposed by Choi (2001), who develops several panel unit root tests based on combining the p-values from individual unit root tests.

In addition to the Fisher test (p_λ), Choi (2001) proposes an inverse normal test (Z) and a logit test (L), which are defined as follows:

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(\pi_i) \quad (2.16)$$

$$L = \sum_{i=1}^{N} \ln\left(\frac{\pi_i}{1-\pi_i}\right) \quad (2.17)$$

where $\Phi(\bullet)$ is the standard normal cumulative distribution function.

Choi (2001) shows that $Z \Rightarrow N(0,1)$ and $L^* \Rightarrow t_{5N+4}$ as $T_i \to \infty$ for all N, where $L^* = \sqrt{\frac{3(5N+4)}{\pi N(5N+2)}} L$.

The concept of the average of the individual test statistics proposed by IPS is extended to calculate panel tests based on different types of unit root tests. Hadri (2000) extends the unit root test of Kwiatkowski *et al.* (1992) (KPSS), and proposes a panel test under the null hypothesis of stationarity. Consider the following model:

$$y_{i,t} = \gamma z'_{i,t} + r_{i,t} + \varepsilon_{i,t} \quad (2.18)$$

where $z'_{i,t}$ is the deterministic component, $r_{i,t}$ is the random walk component ($r_{i,t} = r_{i,t-1} + u_{i,t}$) and $u_{i,t} \sim iid(0, \sigma_u^2)$ and $\varepsilon_{i,t} \sim iid(0, \sigma_\varepsilon^2)$. Using back substitution equation (2.18) can be written as:

$$y_{i,t} = \gamma z'_{i,t} + e_{i,t} \quad (2.19)$$

where $e_{i,t} = \sum_{t=1}^{t} u_{i,t} + \varepsilon_{i,t}$

Let $S_{i,t}$ be a partial sum process of the residual ($S_{i,t} = \sum_{t=1}^{t} \hat{e}_{i,t}$). The panel LM statistic is then calculated as:

$$LM = \frac{(1/N)\sum_{i=1}^{N} \frac{1}{T^2} \sum_{t=1}^{T} S_{i,t}^2}{\hat{\sigma}_e^2} \tag{2.20}$$

where $\hat{\sigma}_e^2$ is the estimate of error variance ($\hat{\sigma}_e^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{e}_{i,t}^2$).

Hadri (2000) shows that the asymptotic distribution of a standardised *LM*-statistic (LM^*) is a normal distribution as:

$$LM^* = \frac{\sqrt{N}(LM - \xi_\mu)}{\zeta_\sigma} \Rightarrow N(0,1) \text{ as } T, N \to \infty \tag{2.21}$$

where ξ_μ and ζ_σ^2 is the mean and variance of the LM statistic, respectively. The values of ξ_μ and ζ_σ^2 are tabulated in the paper by Hadri (2000).

Recently, further development of the panel unit root tests has focused on two major research directions. First, the relaxation of the cross-sectional independence assumption is addressed in many new papers, e.g. Choi (2002) and Phillips and Sul (2003). The literature on panel unit root tests in cross-correlated panels will be presented in Section 2.3.

The second direction is to apply the panel method in testing for cointegration. Several panel cointegration tests have been proposed in many recent papers, e.g. Kao (1999), Pedroni (1999) and Larsson *et al.* (2001). Testing for cointegration in panel data will be discussed in Chapter 3.

2.3 Cross-sectional dependence in panel unit root tests

In the analysis of unit root testing in panel data, the tests discussed in the previous section assume that the individual time-series are cross-sectionally independent. However, this assumption is rather restrictive because of co-movement across the individual units, especially in cross-national data sets. This problem has been pointed out in recent papers, for example, by O'Connell (1998) and Choi (2002). Two main problems arise when the disturbances are cross-sectionally dependent. First, the asymptotic properties of the panel test statistics are no longer valid and the distribution of test statistics becomes unknown. Second, the cross-correlation reduces the total amount of independent information contained in the panel; then, even if the distribution of test statistics is available, the power of the tests may be reduced. Therefore, the reliability of any panel unit root test is affected when the error terms in the panel are cross-correlated.

There are many economic reasons supporting the presence of cross-sectional dependence in the data. First, such dependence can occur through construction of the variables that may include some common component across cross-section units. For example, the definition of real exchange rate includes the value of currency and price index of a numeraire country, which are common across countries. Second, some exogenous shocks can influence the movement of similar economic variables in many countries, simultaneously. For example, the impact of the currency crisis usually spreads across regions, which causes their exchange rates to move together. Finally, model mis-specification, e.g. omission of common variables in the model, may lead to correlation among the error terms.

Consider a panel unit root test (without trend and augmented terms) with cross-sectional dependence in the error terms:

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + \varepsilon_{i,t} \tag{2.22}$$

where $i = 1,\ldots.., N$; $t = 1,\ldots..,$ T, $\varepsilon_t \sim iid\ N(0,\Omega)$, ($\varepsilon_t = [\varepsilon_{1,t}\ \varepsilon_{2,t} \ldots \varepsilon_{N,t}]'$) and Ω is a non-diagonal matrix, such that:

$$E(\varepsilon_{i,t}, \varepsilon_{j,s}) = \begin{cases} \sigma_{i,j} & \text{for } t = s \\ 0 & \text{for } t \neq s \end{cases} \tag{2.23}$$

where the correlations are $|\sigma_{i,j}| < 1$, such that Ω can be expressed as:

$$\Omega = \begin{bmatrix} \sigma_1 & \sigma_{2,1} & \cdots & \sigma_{N,1} \\ \sigma_{1,2} & \sigma_2 & \cdots & \sigma_{N,2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,N} & \sigma_{2,N} & \cdots & \sigma_N \end{bmatrix}_{N\times N} \tag{2.24}$$

Alternatively, the presence of cross-correlated disturbances can be drawn from the time-specific effect component in the errors, i.e.:

$$\varepsilon_{i,t} = r_i\theta_t + \eta_{i,t} \tag{2.25}$$

where θ_t represents the common time effect, which is independently normally distributed across time with variance normalised to unity ($\theta_t \sim iid\ N(0,1)$). r_i is the parameter that measures the impact of the common factor (θ_t) on each individual series, the general error component ($\eta_{i,t}$) is assumed to satisfy $\eta_{i,t} \sim iid\ N(0,\sigma_i^2)$ over t.

O'Connell (1998) shows, through Monte Carlo simulations, that ignoring the contemporaneous correlation can lead to severe size distortions in the LLC panel unit root test.

Several methods have been proposed to control for the effect of cross-sectional dependence in panel unit root tests. In the traditional panel data analysis, the cross-correlation effect is usually accounted for by cross-sectionally demeaning the series. This procedure removes the effect of the common time-specific component by subtracting out cross-sectional means from $y_{i,t}$ before applying the panel unit root test to the demeaned series ($\tilde{y}_{i,t}$), that is:

$$\tilde{y}_{i,t} = y_{i,t} - N^{-1}\sum_{j=1}^{N} y_{j,t} \,. \tag{2.26}$$

This method is equivalent to the inclusion of time dummies (γ_t) in the unit root test equation (2.4), which can be expressed as:

$$\Delta y_{i,t} = \alpha_i - \rho_i y_{i,t} + \gamma_t + \sum_{j=1}^{p_i} \theta_{i,j} \Delta y_{i,t-j} + u_{i,t} \tag{2.27}$$

This procedure assumes that the co-movement in the time-series is due to a common factor, which impacts similarly on all individual series. Therefore, the demeaning method is valid only in the case of homogeneous cross-sectional dependence, i.e. $\varepsilon_{i,t} = r\theta_t + \eta_{i,t}$, in which the off-diagonal elements of Ω are all the same, and will not be robust if the effect of the time-specific component differs across i. In view of this, several recent papers have further developed alternative methods to overcome this deficiency.

O'Connell (1998) suggests applying the GLS method, using the information on the covariance matrix in the estimation. The GLS estimator of the autoregressive coefficient (ρ_i) suggested by O'Connell (1998) is based on a homogeneity assumption ($\rho_i = \rho,\ \forall\, i$). The estimation procedure is displayed as follows.

Let Y is $T \times N$ matrix of first-differences of series $y_{i,t}$ and X is $T \times N$ matrix of lagged series.

$$Y = \begin{bmatrix} \Delta y_{1,1} \cdots \Delta y_{N,1} \\ \Delta y_{1,2} \cdots \Delta y_{N,2} \\ \vdots \quad \ddots \quad \vdots \\ \Delta y_{1,T} \cdots \Delta y_{N,T} \end{bmatrix}_{TxN} \tag{2.28}$$

The GLS estimate of autoregressive coefficient ($\hat{\rho}_{GLS}$) is:

$$\hat{\rho}_{GLS} = tr(X'Y\Omega^{-1}) / tr(X'X\Omega^{-1}) \tag{2.29}$$

In the case of matrix (Ω) being unknown, the feasible GLS estimator is:

$$\hat{\rho}_{GLS} = tr(X'Y\hat{\Omega}^{-1}) / tr(X'X\hat{\Omega}^{-1}) \tag{2.30}$$

where $\hat{\Omega}$ is some consistent estimates of Ω, which is usually obtained from the estimated error terms.

The distribution of the feasible GLS t-statistic is unknown and could be derived, using Monte Carlo simulations under the null hypothesis. However, the reliability of the GLS estimator is based on a consistent estimation of the covariance matrix. Cerrato (2001) mentions that, in the case of equi-correlated error terms, the OLS estimator of $u_{i,t}$ is not a consistent estimator of $u_{i,t}$. Therefore, the covariance matrix is not estimated consistently, as assumed in the GLS procedure.

In the heterogeneous panels framework, Taylor and Sarno (1998) propose a Multivariate Augmented Dickey-Fuller (MADF) test, which applies the GLS method of seemingly unrelated regression in a system of ADF regressions, providing the information with an advantage over the standard ADF test in the presence of cross-correlation in the errors.

$$y_{i,t} = \alpha_i - \sum_{k=1}^{p_i} \rho_{i,k} y_{i,t-k} + u_{i,t} \quad (2.31)$$

The MADF statistic is calculated from the Wald statistic for the null hypothesis of:

$$H_0 : (\sum_{k=1}^{p_i} \rho_{i,k}) - 1 = 0 \quad (2.32)$$

The MADF test allows the sum of autoregressive coefficients to vary across i. Under the alternative hypothesis, at least one of the series in the panel is stationary.

Breuer, McNown and Wallace (2001) (BMW) introduce the Seemingly Unrelated Regressions Augmented Dickey-Fuller test (SURADF), which is estimated as a system of ADF equations across cross-section units, using the GLS estimator of seemingly unrelated regression (SUR). Consider a system of the ADF equations:

$$\begin{aligned}
\Delta y_{1,t} &= \alpha_1 + \phi_1 y_{1,t-1} + \sum_{j=1}^{p_i} \delta_{1,j} \Delta y_{t-j} + u_{1,t} \\
\Delta y_{2,t} &= \alpha_2 + \phi_2 y_{2,t-1} + \sum_{j=1}^{p_i} \delta_{2,j} \Delta y_{t-j} + u_{2,t} \\
&\vdots \\
\Delta y_{N,t} &= \alpha_N + \phi_N y_{N,t-1} + \sum_{j=1}^{p_i} \delta_{N,j} \Delta y_{t-j} + u_{N,t}
\end{aligned} \quad (2.33)$$

The individual ADF equations are estimated as a system of equations, using an iterative SUR method. The null hypothesis of unit roots in each ADF equation is tested separately, using individual critical values, which are calculated through Monte Carlo simulations. Even though the additional information from the contemporaneous covariance matrix of the errors is included in the SURADF test, it is still based on the individual time-series statistics. A panel version of the SURADF test may be calculated by the IPS method as an average of t-statistics from the individual SURADF test.

MW suggest an alternative approach, using a bootstrap method to calculate the empirical distribution of the test statistic. The bootstrap critical values are then applied to correct the size distortions under cross-sectional dependence. The proposed bootstrap procedure applies the sampling scheme from Li and Maddala (1996).

Under the null hypothesis of unit root for $y_{i,t}$, we have:

$$\Delta y_{i,t} = \eta_i \Delta y_{i,t-1} + \varepsilon_{i,t}^0; \tag{2.34}$$

The estimated residuals from these regressions, denoted as $\hat{\varepsilon}_{i,t}^0$, are then re-sampled to get $\varepsilon_{i,t}^*$. To preserve the cross-correlation structure in the error terms ($\hat{\varepsilon}_{i,t}^0$), we resample $\hat{\varepsilon}_{i,t}^0$ indirectly with the cross-sectional index fixed by resampling $\hat{\varepsilon}_t^0 = [\hat{\varepsilon}_{1,t}^0, \hat{\varepsilon}_{2,t}^0, \ldots, \hat{\varepsilon}_{N,t}^0]$, to get $\varepsilon_{i,t}^*$. Next, the bootstrap sample ($y_{i,t}^*$) is constructed as:

$$y_{i,t}^* = y_{i,t-1}^* + u_{i,t}^* \quad \text{with} \quad y_{i,0}^* = 0 \tag{2.35}$$

$$u_{i,t}^* = \hat{\eta}_i u_{i,t-1}^* + \varepsilon_{i,t}^* \quad \text{with} \quad u_{i,0}^* = \sum_{j=0}^{m} \hat{\eta}_i^j \varepsilon_{-j}^* \tag{2.36}$$

where ε_{-j}^* are drawn as an independent bootstrap sample , m is set to be equal to 30, and $\hat{\eta}_i$ is the OLS estimator from equation (2.34).

MW suggest that the size distortion problem, arising from the presence of cross-sectional dependence, would decrease by using this bootstrap method.

Recently, a series of papers have developed panel unit root tests that directly deal with cross-sectional dependence. The cross-sectional independence assumption is usually relaxed by applying a linear factor model to approximate the structure of cross-correlation in the panel. In the tests with a factor model, it is assumed that the structure of cross-correlation can be drawn from common factors (see, for example, equation (2.25)). These proposed procedures basically generalise the traditional cross-section demeaning procedure (see, equation (2.26)). Therefore, the cross-sections are assumed to be independent and conditional on these factors. The majority of these papers suggest de-factoring the data before applying standard procedure on the de-factored data.

Choi (2002) proposes unit root tests for a cross-sectionally correlated error-component model where the non-stochastic trend components and cross-sectional correlations are eliminated from observed data ($y_{i,t}$), using parameters estimated by the GLS method of Elliot *et al.* (1996). In the presence of a linear trend, the model is:

$$y_{i,t} = \beta_0 + \beta_1 t + x_{i,t} \qquad (2.37)$$

where $x_{i,t} = \mu_i + \lambda_t + \gamma_i t + v_{i,t}$; $v_{i,t} = \sum_{j}^{p_i} \alpha_{i,j} v_{i,t-j} + e_{i,t}$; β_0 is a common mean for all i, μ_i is the unobservable individual effect, λ_t denotes the unobservable time effect and $v_{i,t}$ is the remaining random component, which follows an $AR(p_i)$ process.

The cross-sectional means are used to construct a new variable ($y_{i,t}^{\mu}$) as:

$$y_{i,t}^{\mu} = y_{i,t} - \hat{\beta}_{0,i} - \hat{\beta}_{1,i} t - \frac{1}{N}\sum_{i=1}^{N}(y_{i,t} - \hat{\beta}_{0,i} - \hat{\beta}_{1,i} t) \qquad (2.38)$$

where $\hat{\beta}_{0,i}$ and $\hat{\beta}_{1,i}$ is the GLS estimator of $\beta_{0,i}$ and $\beta_{1,i}$, respectively.

A standard ADF test is applied to the new series ($y_{i,t}^{\mu}$). Then, three panel statistics, proposed by Choi (2001), are calculated. This procedure provides a generalisation of the cross-sectional demeaning procedure proposed in other papers.

However, it still does not permit the common factors to have different effects on different cross-section units.

Phillip and Sul (2003) propose an orthogonalisation procedure to eliminate a common factor before applying standard panel unit root tests. In this procedure, a common time effect impacts on each individual series differently. Suppose a single-factor structure in the regression errors ($\varepsilon_{i,t}$) has the form presented by equation (2.25). Therefore, each individual unit of $y_{i,t}$ contains a common random factor (θ_t), generating the correlation among the cross-sectional units. Phillips and Sul (2003) propose a moment-based method to eliminate a common factor. Then, the panel median unbiased estimated is proposed to construct panel statistics based on the de-factored data, which has no cross-sectional dependence.

Moon and Perron (2004) suggest a similar approach to that of Phillips and Sul (2003), in a more general arrangement. This framework assumes that the error terms ($\varepsilon_{i,t}$) follow an approximate *K*-factor model.

$$\varepsilon_{i,t} = \beta_i^{0'} f_t^0 + e_{i,t} \tag{2.39}$$

where f_t^0 is *K*-vector of unobservable random factors, used as a device to generate cross-sectional dependence structure in the error terms. The number of factors (*K*) is unknown.

This method is basically a generalised version of a one-common factor structure proposed by Phillips and Sul (2003) (see equation (2.25)), when there is more than one common factor component. The extent of the correlation is determined by the factor-loading coefficient (β_i^0). Moon and Perron (2004) apply a principal component method (PCM) in estimating β_i^0, which they subsequently apply to generate the de-factored data. A pooled panel unit root test method is proposed to estimate this de-factored data, which can be applied to test for unit roots in a model without a deterministic trend. A procedure for a model with a deterministic trend is presented in a related paper of Moon, Perron and Phillips (2003).

The panel unit root test statistics of Choi (2002), Phillips and Sul (2003) and Moon and Perron (2004) depend on complicated methods to generate the de-factored data. Panel unit root tests are then applied with this de-factored data. By contrast, Pesaran (2003) adopts a different approach to approximate the structure of error covariance matrix. Instead of basing unit root tests on deviations from the estimated factor, Pesaran (2003) introduces the Cross-sectionally Augmented ADF test (CADF), which is an extension of the standard ADF regression augmented with the cross-section averages of lag levels and first-differences of individual series. These averages are used to filter out cross-sectional dependence. This idea is similar to that of augmenting lagged change of the series in dealing with serial-correlation in the standard ADF test. The individual CADF statistics are then used to compute the modified version of the t-bar statistic of IPS.

The CADF regressions can be presented as follows:

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + \delta_i t + \lambda_i \bar{y}_{i,t-1} + \sum_{j=0}^{p_i} \gamma_{i,j} \Delta \bar{y}_{t-j} + \sum_{j=1}^{p_i} \theta_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t} \quad (2.40)$$

The individual CADF statistic ($t_i(N,T)$) is given by the OLS t-statistic of ρ_i. Then, the IPS-type panel unit root test of the CADF regressions (denoted as CIPS) can be constructed as an average of each individual statistic.

$$CIPS(N,T) = N^{-1} \sum_{i=1}^{N} t_i(N,T) \quad (2.41)$$

Pesaran (2003) also considers a truncated version of $t_i(N,T)$ (denoted as $t_i^*(N,T)$) in the construction of the panel test statistics, to avoid undue influences of the extreme outcomes that could arise when T is small ($T \cong 10$ to 20). The $t_i^*(N,T)$ statistic is constructed as:

$$t_i^*(N,T) = \begin{cases} t_i(N,T) & \text{,if } -K_1 < t_i(N,T) < K_2 \\ -K_1 & \text{,if } t_i(N,T) \leq -K_1 \\ K_2 & \text{,if } t_i(N,T) \geq K_2 \end{cases} \tag{2.42}$$

where K_1 and K_2 are positive constants that are sufficiently large, so that $\Pr[-K_1 < t_i(N,T) < K_2]$ is sufficiently large (more than 0.9999).

Pesaran (2003) recommends the value for K_1 and K_2, shown in his paper. The associated truncated panel unit root test, denoted as the CIPS* test, is given by:

$$CIPS^*(N,T) = N^{-1}\sum_{i=1}^{N} t_i^*(N,T) \tag{2.43}$$

However, Pesaran (2003) mentions that the asymptotically correlated between individual series still exist in the CADF regressions, due to their dependence on the common factor. Despite this, he shows that the limit distributions of the $CIPS^*(N,T)$ and $CIPS\ (N,T)$ statistics still exist and are free of nuisance parameters. Then, the critical values can be tabulated and reported in Table 3a to 3c of Pesaran (2003). The finite sample distributions of $CIPS^*(N,T)$ and $CIPS\ (N,T)$ will differ only for very small T, and are indistinguishable for $T > 20$.

This CIPS statistic can be considered as a direct extension of the IPS statistic in the case of cross-sectionally correlated errors.

In Section 2.4 we consider the Monte Carlo simulations.

2.4 A Monte Carlo simulation study

In this section, we report the results of Monte Carlo experiments used to investigate the size and power properties of the panel IPS and MW unit root tests. All simulations are performed in EVIEWS, version 4.1.

2.4.1 Simulation design

In the design of a Monte Carlo simulation, we consider the general model with intercept and trend as deterministic components, that is:

$$\Delta y_{i,t} = \mu_i + \phi_i \mu_i t + \phi_i y_{i,t-1} + u_{i,t} \qquad (2.44)$$

where $i = 1,\ldots, N$ and $t = 1,\ldots., T$;

To investigate the effect of serial-correlation and cross-sectional correlation in the error terms on the performance of the panel unit root tests, we perform simulations with serial-correlated and cross-correlated errors. We consider three cases of Monte Carlo experiments.

Case A: white noise errors

In this case, the error terms ($u_{i,t}$) are generated as *iid* $N(0,1)$.

Case B: serial-correlated errors

In this case, we allow for the presence of 1^{st} order serial correlation in the error terms ($u_{i,t}$), that is:

$$u_{i,t} = \lambda_i u_{i,t-1} + e_{i,t} \tag{2.45}$$

where $\lambda_i \sim U[0.2,0.4]$, and $e_{i,t} \sim iid\ N(0,\delta_i^2)$, $\sigma_i^2 \sim U[0.5,1.5]$.

Case C: serial-correlated and cross-correlated errors

$$u_{i,t} = \lambda_i u_{i,t-1} + e_{i,t} \tag{2.46}$$

where $\lambda_i \sim U[0.2,0.4]$, $E(e_{i,t}) = 0$ and $E(e_{i,t}, e_{j,s}) = \begin{cases} \sigma_{i,j} & \text{for } t = s \\ 0 & \text{for } t \neq s \end{cases}$, such that the error terms has 1st order serial-correlation and cross-correlation. The existence of cross-correlation is represented by a non-diagonal error covariance matrix (Ω), which is equal to $(\sigma_{ij})_{i,j=1}^{N}$. Ω is randomly drawn and then fixed for each panel over experiments. We use the random number process suggested by Chang (2004) to ensure a symmetric positive definite matrix. This procedure is described as follows.

1) Generate $U_{N\times N}$ matrix from Uniform [0,1].

2) Construct an orthogonal matrix H from U, $H = U(U'U)^{-1/2}$.

3) Generate set of eigen values $\lambda_1,\ldots,\lambda_N$, by setting $\lambda_1 = r > 0$, ($r \sim$ U[0,1]) and $\lambda_N = 1$, with $\lambda_2,\ldots,\lambda_{N-1}$ are draw from U[r,1].

4) Construct matrix Λ as a diagonal matrix with ($\lambda_1,\ldots,\lambda_N$) on the diagonal.

5) Covariance matrix, Ω, can be generated as $\Omega = H\Lambda H'$.

The random number generated from this procedure is shown in Table 2.1. The average degree of cross-correlation matrix in Table 2.1 (non-diagonal term) is approximately 0.15.

Table 2.1 Matrix of cross-correlation (Ω) used in the DGP in case C ($N = 5$)

	C1	C2	C3	C4	C5
R1	0.46	0.06	0.10	0.15	0.01
R2		0.60	0.08	0.32	-0.18
R3			0.53	-0.01	0.21
R4				0.54	0.17
R5					0.82

Monte Carlo simulations are performed in the panel with the length of time-series (T) equal to 112, which represents the number of quarterly data in the post-Bretton Woods system from 1973:1 to 2000:4. The number of cross-sectional series in the panel (N) is set equal to 5, 10, 15, 20 and 25 for panels I, II, III, IV and V, respectively. Panel I (N=5) is referred to as the small panel, while panel V (N=25) is referred to as the large panel. All series are generated over 212 observations, with the first hundred observations discarded to achieve randomness of $y_{i,0}$. The value of μ_i is generated randomly from $N(0,1)$ and fixed for each panel over all replications. The number of lags included in the ADF regressions (p_i) to correct for serial-correlation is set equal to 0, 1, 2, in order to investigate the effect of selection of the order of the ADF regressions. Since, in the DGP, the error terms in cases B and C are generated by the $AR(1)$ process, $ADF(1)$ is the appropriate number of lags, while $ADF(0)$ and $ADF(2)$ represent under-selecting and over-selecting, respectively. However, in case A, $ADF(0)$ is an appropriate lag specification, and both $ADF(1)$ and $ADF(2)$ are over-parameterised.

In the analysis of size, ϕ_i is set at zero. In the investigation of power, ϕ_i is set at –0.1, which represents the mean reversion process with a reasonably slow speed of adjustment. This value of autoregressive coefficient corresponds to series with approximately 6.5 quarters (one and a half years) of half-life, which is a reasonable speed of adjustment in the mean reversion process in real exchange rates suggested by Rossi (2004).

The number of replications in Monte Carlo simulations in this section is set to be equal to 10,000. The reported results are based on 5% critical values. Means and variances of the ADF t-statistics used in calculation of the IPS test are extracted from Table 3 of IPS. The p-values of the ADF tests are calculated, using the ADF t-distribution generated by Monte Carlo simulations on the distribution of the ADF regression performed with 100,000 replications.

2.4.2 Simulation results

We first investigate the problem of the low power of the standard ADF test against the near unit root process. Monte Carlo simulations on the performance of the standard ADF test are conducted to address this problem, as mentioned by Maddala and Kim (1999). The DGP is based on equation (2.44) with N=1. The errors are generated according to case A, described in Section 2.4.1. The autoregressive coefficients (ϕ) are set equal to -0.4, -0.3, -0.2, -0.15, -0.1 and -0.05. The results will be used to compare with those of the panel tests to be presented later. Moreover, we perform the simulations on the panel with N=1, T=300 and ϕ = -0.1 to investigate the power property of the standard ADF test with longer time-series span data. The empirical results of the power analysis of the standard ADF test are presented in Table 2.2.

Table 2.2 The empirical power of the standard ADF test

ϕ	$ADF(0)$	$ADF(1)$	$ADF(2)$
-0.05	0.094	0.091	0.088
-0.1	0.234	0.215	0.199
-0.15	0.491	0.433	0.373
-0.2	0.768	0.667	0.574
-0.3	0.989	0.945	0.865
-0.4	1.000	0.995	0.968
-0.1x	0.952	0.923	0.890

Note: The results are based on the standard ADF test, when T=112, with the exception of those of the last line (-0.1x), which are based on T=300. The underlying data are generated by equation (2.44), with N=1 and ϕ = -0.4, -0.3, -0.2, –0.15, –0.1 and –0.05. The error terms are generated from case A.

The simulated results from Table 2.2 show that the empirical power of the standard ADF test is higher than 0.500 only when $\phi < -0.15$. The power results fall markedly as $\phi \rightarrow 0$. The simulated power results are equal to 0.491, 0.234 and 0.094 when ϕ equals to –0.15, –0.1 and –0.05, respectively. These power results show that, when $T = 112$, the ADF test has limited power to reject the unit root null hypothesis, when $\phi_i > -0.2$. To improve the power of the standard ADF test, the larger span of data should be used. The power of the individual ADF test increases dramatically from 0.234 to 0.952 when the data span (T) increases from 112 to 300. Alternatively, combining information from the time-series with that obtained from the cross-sectional dimension by using panel data will be employed to increase the power of unit root tests.

A direct comparison between the panel IPS and MW tests and the standard ADF test is not possible, due to the difference between the null and alternative hypotheses of panel unit root tests and those of the individual ADF test. In heterogeneous panels, the null hypothesis of non-stationarity for all series is tested against the alternative, where at least one series in the panel is stationary. The inclusion of more series in the panel may increase the possibility of rejecting the null hypothesis for at least one series. Therefore, as a point of comparison, we investigate the power of the standard ADF test when applied to each of the N series, based on the same null and alternative hypotheses as those of the panel IPS and MW tests. The null hypothesis of non-stationarity for all N series will be rejected when at least one series (among N individual series) is stationary. The DGP from equation (2.44), with N=5, for cases A, B and C, is used. These simulated size and size-adjusted power results are reported in Table 2.3.

Table 2.3 The empirical size and size-adjusted power of the standard ADF test under the null and alternative hypotheses of the IPS and MW tests

	ϕ_i	Case A	Case B	Case C
Size	0	0.215	0.223	0.216
Size-adjusted Power	-0.05	0.099	0.102	0.101
	-0.1	0.287	0.271	0.252
	-0.15	0.647	0.583	0.509
	-0.2	0.932	0.857	0.768
	-0.3	1.000	0.998	0.986
	-0.4	1.000	1.000	1.000

Note: The results are based on the standard ADF test (T=112). The underlying data are generated by equation (2.44) with N=5 and ϕ = -0.4, -0.3, -0.2, –0.15, –0.1 and –0.05. The error terms are generated from cases A, B and C. The results in case A (white noise errors) are based on $ADF(0)$ specification, while those of case B and C ($AR(1)$ errors) are based on $ADF(1)$ specification.

The results from Table 2.3 show that the ordinary ADF test is over-sized. In case A, with the $ADF(0)$ regression, the empirical size of test is equal to 0.215, for a nominal size of 5%. The size distortion problem renders the power results invalid. In view of this, we report the size-adjusted power results only. In the analysis of power, the autoregressive coefficients (ϕ_i) are set equal to –0.4, -0.3, -0.2, -0.15, -0.1 and -0.05. The size-adjusted power results from Table 2.3 confirm that the power of the standard ADF test remains low, when T=112 and ϕ_i>-0.2. The size-adjusted power results are equal to 0.647, 0.287 and 0.099 when ϕ_i equals to –0.15, –0.1 and –0.05, respectively. These results show that there is only a small improvement in the power from that observed for the ADF test (see Table 2.2) with the ordinary null and alternative hypotheses.

Next, we consider the size and power of the panel IPS and MW unit root tests, using Monte Carlo techniques. The remaining simulations on the performance of panel unit root tests will be performed only in the DGP with $\phi_i = 0$ and -0.1 for all i, for simulations on the size and power properties, respectively.

The simulated results on the size and power performance of the panel IPS and MW tests in the small (N=5) and large (N=25) panels are reported in Tables 2.4 and 2.5, respectively. We first consider case A, in which the error terms are generated as white noise, thus making $ADF(0)$ the appropriate specification. The empirical size of the IPS and MW tests is close to the nominal level of 5%. In panel I (N=5), the size results are equal to 0.050 and 0.051 for the IPS and MW tests, respectively. In panel V (N=25), these size results are equal to 0.051 and 0.057, respectively. Experiments in this section are based on 10,000 replications, implying that the 95% confidence interval of the 0.05 significant level test is between 0.0457 and 0.0543.

In cases B and C, the error terms are generated according to an $AR(1)$ model, rendering $ADF(1)$ an appropriate model. In case B, using the $ADF(1)$ regression, the empirical size of the IPS and MW tests is still reasonably close to the nominal level of 5% in both the small and large panels. The empirical size of the IPS and MW tests is equal to 0.047 (0.051) and 0.049 (0.052), respectively in panel I (panel V).

Table 2.4 The empirical size of the IPS and MW tests

	Number of lags	Panel I ($N = 5$)		Panel V ($N = 25$)	
		IPS	MW	IPS	MW
Case A	*ADF(0)*	0.050	0.051	0.051	0.057
	ADF(1)	0.048	0.051	0.049	0.055
	ADF(2)	0.048	0.051	0.049	0.057
Case B	*ADF(0)*	0.001	0.002	0.000	0.000
	ADF(1)	0.047	0.049	0.051	0.052
	ADF(2)	0.049	0.050	0.048	0.058
Case C	*ADF(0)*	0.002	0.003	0.000	0.000
	ADF(1)	0.065	0.062	0.069	0.070
	ADF(2)	0.065	0.063	0.069	0.073

Note: The results are based on the IPS and MW tests. The underlying data are generated by equation (2.44) with N=5, 25. ϕ_i is set to be 0 and –0.1, in the analysis of size and power, respectively. The error terms are generated from cases A, B and C. In case A (white noise errors), the *ADF*(0) regression represents the correctly chosen order of the ADF regression, while *ADF*(1) and *ADF*(2) are over-fitting. In case B and C (*AR(1)* errors), the *ADF*(1) regression represents the correctly chosen order of the ADF regression, while *ADF*(1) and *ADF*(2) are over-fitting and under-fitting, respectively.

Table 2.5 The empirical power of the IPS and MW tests

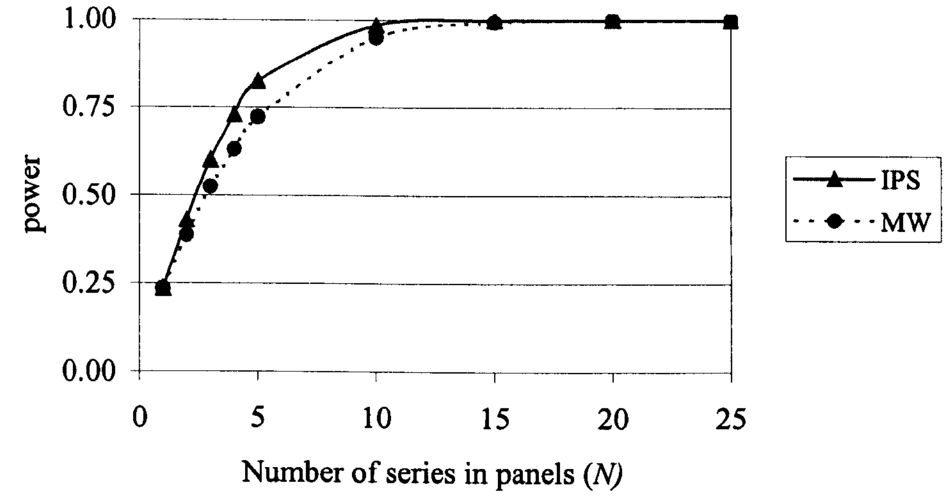
	Number of lags	Panel I ($N = 5$)		Panel V ($N = 25$)	
		IPS	MW	IPS	MW
Case A	*ADF(0)*	0.826	0.723	1.000	1.000
	ADF(1)	0.760	0.660	1.000	1.000
	ADF(2)	0.694	0.600	1.000	0.999
Case B	*ADF(0)*	0.026	0.015	0.045	0.010
	ADF(1)	0.730	0.623	1.000	0.998
	ADF(2)	0.664	0.568	1.000	0.996
Case C	*ADF(0)*	0.037	0.024	0.055	0.014
	ADF(1)	0.686	0.590	1.000	0.997
	ADF(2)	0.623	0.533	0.999	0.992

Note: see notes to Table 2.4.

In cross-correlated panel (case C), the results from Table 2.4 show that both the IPS and MW tests are slightly size-distorted (over-sized). Using the *ADF*(1) regression, the size results of the IPS and MW tests are equal to 0.065 and 0.062, respectively, in panel I, while in panel V, they are equal to 0.069 and 0.070. These size results show evidence of the small size distortions in the panel with cross-correlated errors. The panel unit root tests in the large panel seem to be more severely over-sized than the tests in the small panel.

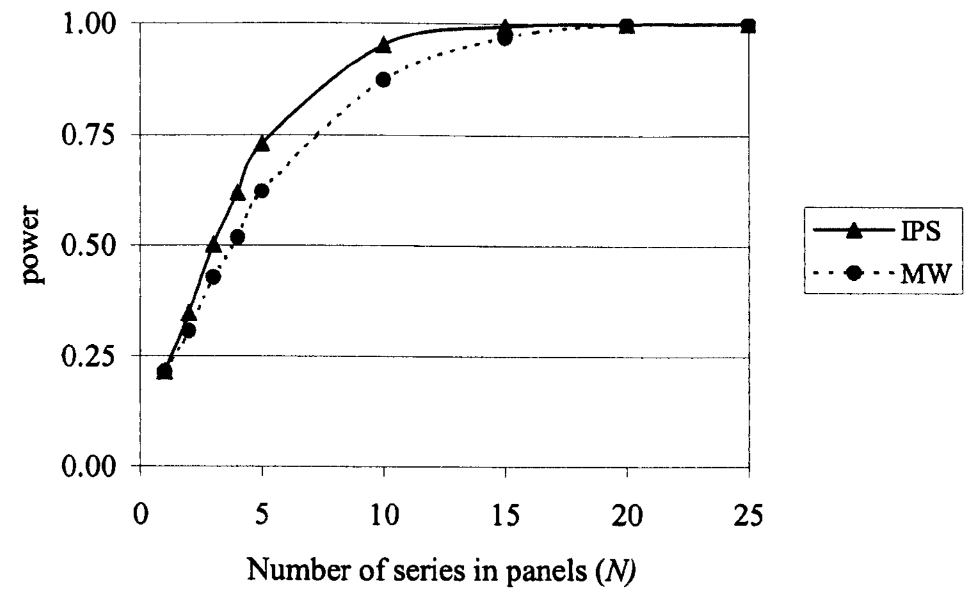
Turning to power performance, the results from Tables 2.5 show that, in cases A and B, the panel IPS and MW tests are significantly more powerful than the standard ADF test in every case. For example, in case A, with *ADF*(0) specification, the simulated results of the IPS and MW tests are equal to 0.826 and 0.723 respectively, when N=5. In addition, an increasing number of series in panel (N) raises the empirical power of both panel tests, as the power results of both the IPS and MW tests are equal to 1.000 when N=25. A pictorial representation of the Monte Carlo results of the empirical power of the IPS and MW tests in case A for the panels with N = 1, 2, 3, 4, 5, 10, 15, 20 and 25 is presented in Figure 2.1. The results from this figure show that the IPS and MW tests produce enough power to distinguish the process from the unit root null hypothesis. The empirical power of both panel tests is higher than 0.500, when $N \geq 3$. The enlarged number of series in the panel increases the empirical power of the IPS and MW tests, monotonically. These power results approach 1.000, when $N>10$. Comparing the IPS and MW tests, the former is slightly more powerful than the latter in every case. However, the empirical power of both tests is approximately the same when N is greater than 10, as these results approach unity.

Figure 2.1 The empirical power of the IPS and MW tests (Case A)



Note: The results are based on the IPS and MW tests (T=112). The underlying data are generated by equation (2.44) with N=1,2,3,4,5,10,15,20,25 and $\phi = -0.1$. The error terms are generated from case A. The results are based on the $ADF(0)$ regression.

Figure 2.2 The empirical power of the IPS and MW tests (Case B)



Note: The results are based on the IPS and MW tests (T = 112). The underlying data are generated by equation (2.44) with N=1,2,3,4,5,10,15,20,25 and $\phi = -0.1$. The error terms are generated from case B. The results are based on the $ADF(1)$ regression.

In serially-correlated panels (case B), the simulated results are similar to those of case A. A pictorial representation of the Monte Carlo results of the empirical power of the IPS and MW tests in case B for the panels with $N = 1, 2, 3, 4, 5, 10, 15, 20$ and 25 is given in Figure 2.2. The results show that the IPS and MW tests still have considerably high power. The empirical power of both panel tests is higher than 0.500, when $N \geq 4$. These power results are close to 1.000, when N is higher than 15. However, in case B, the empirical power of both tests is slightly lower than that of case A. In case C (serial- and cross-correlated errors), the power results are not comparable because both the IPS and MW tests are slightly over-sized.

Next, we consider the effects of incorrect specification of the order of the underlying ADF regressions. The results from Table 2.4 show that under-selecting the order of the ADF regressions results in the remarkably serious size distortions in the panel unit root tests, as the empirical size of both panel tests goes to zero. For example, we consider the results from cases B and C with $ADF(0)$ specification. The errors are generated according to an $AR(1)$ process in both cases. Therefore, in these cases, using $ADF(1)$ is appropriate, while the $ADF(0)$ and $ADF(2)$ regressions both under- and over-select the order of the ADF regressions, respectively. In case B, the simulated size results of the IPS and MW tests using $ADF(0)$ are equal to 0.001 (0.000) and 0.001 (0.000), respectively, in panel I (panel V). Over-fitting is much less damaging of the performance of the panel unit root tests. In case B, the IPS and MW tests still are correctly sized with both $ADF(1)$ and $ADF(2)$ specification. In panel I (panel V), the size results of the IPS and MW tests using $ADF(2)$ are equal to 0.047 (0.051) and 0.049 (0.052), respectively. However, over-selecting the order of the ADF regression slightly affects the power performance of the IPS and MW tests. In panel I, the power of the IPS and MW tests is reduced from 0.730 and 0.623 in the tests based on the $ADF(1)$ regressions to 0.664 and 0.568 when the $ADF(2)$ regressions are used. The simulated results of cases A and C are also similar to those of case B. Under-fitting the order of the ADF regressions severely distorts the size of both the IPS and MW tests, while over-fitting does not affect the empirical size, but slightly reduces the power. IPS and MW also point out the effects of under-selecting and over-selecting the order of ADF regression, which are similar to the results in this chapter.

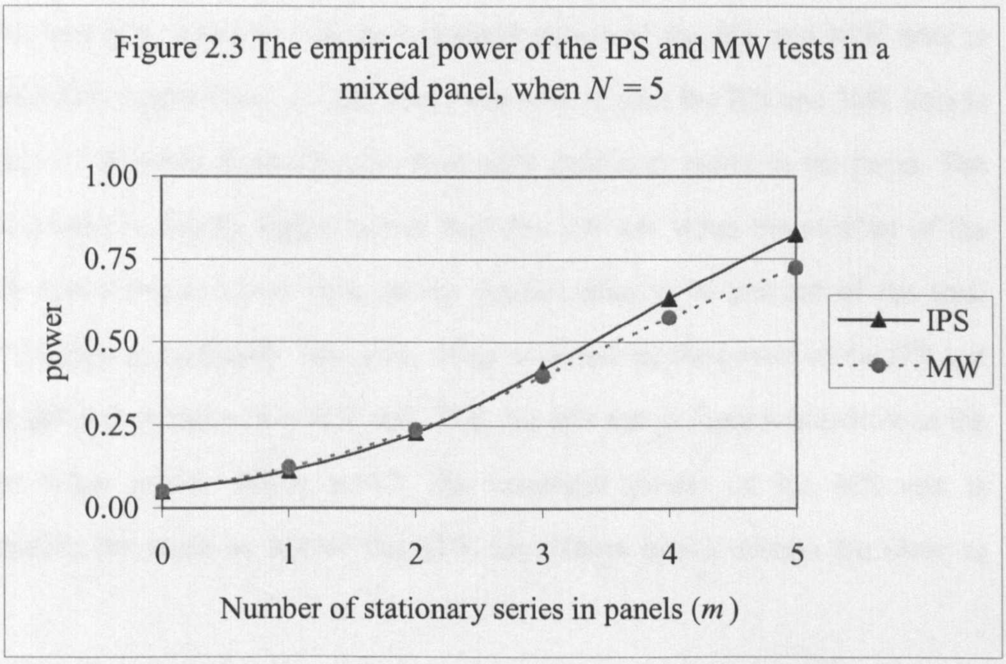
Overall, the simulation results in this section show that, in moderate sample sizes (T), the empirical size of the IPS and MW tests is still reasonably close to the nominal level of 0.05 when the number of lags is correctly specified. By adding the information from the cross-section dimension, the panel IPS and MW tests provide improvement in the power over the standard ADF test. The power of the IPS and MW tests increases with N. The presence of serial-correlation significantly affects the size of the tests when the number of lags is under-specified. These results are similar to those reported in IPS and MW

In the next two sections, two interesting issues in the application of heterogeneous panel unit root tests will be addressed. First, we consider the case in which there is a mixture of stationary and non-stationary series in the panel. Simulations on a mixed panel will be carried out in the next section. Second, the effect of cross-correlation may lead to size distortions in the panel tests, which explicitly assumes that the individual panel tests are independently distributed. The simulation results in this section show some evidence of the size distortions in case C (cross-correlated errors). In Section 2.6, we will further investigate the impact of cross-sectional dependence on the size and power of panel unit root tests.

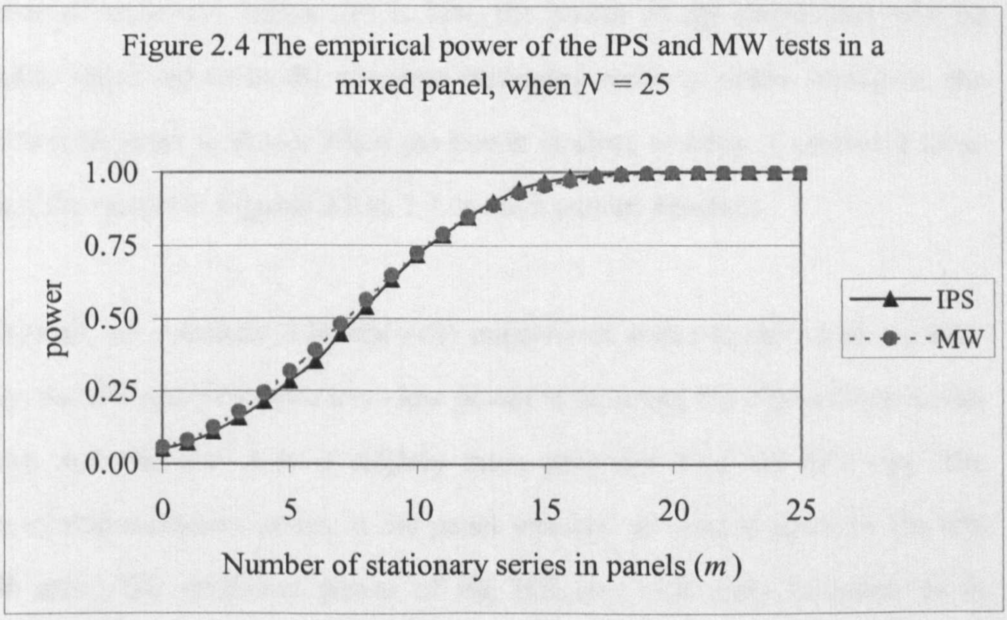
2.5 Unit root tests in a mixed panel of stationary and non-stationary series

In heterogeneous panels, the unit root null hypothesis can be rejected, even though there is only one stationary series in the panel. Consequently, in this section, Monte simulations are conducted to investigate the power performance of the IPS and MW tests in a mixed panel that combine both stationary and non-stationary series. The data is generated according to the DGP of case A, described in Section 2.4. A mixed panel consists of both stationary and non-stationary series. We consider the case of panel I, II, III, IV and V, where $N = 5, 10, 15, 20$ and 25, respectively. In each panel, the number of stationary series included (m) is equal to $1,\ldots, N$, where the autoregressive coefficients (ϕ_i) are equal to -0.1 for $i = 1,\ldots, m$ and 0 for $i = m+1,\ldots,N$. The simulated power results of the small (N=5) and large (N=25) panels are shown in Figures 2.3 and 2.4, respectively.

First, we consider the case of the small panel (N=5). There is no power gain from applying the IPS and MW tests, when there are only one or two stationary series (m = 1,2) in the panel. As shown in Table 2.3, the power of the standard ADF test in case A equates to 0.287 ($ADF(0)$). The empirical power of the IPS (MW) test is equal to 0.114 (0.123) and 0.228 (0.236) when m = 1 and 2, respectively. In the panel with one and two (out of five) stationary series, the MW test has slightly better power than the IPS test. The better performance of the MW test when there are only a few stationary series in the panel may be attributed to the nature of the MW test, which is more flexible in calculation than the IPS test. The empirical power of both panel tests increases considerably with m. The power of the IPS and MW tests exceeds 0.500 when there are at least four stationary series in the panel. Comparing the panel IPS and MW tests, the empirical power of the IPS test grows faster than that of the MW test and exceeds that of the MW test when there are at least three stationary series in the panel.



Note: The results are based on the IPS and MW tests (T = 112). The underlying data are generated by equation (2.44) with N=5 and ϕ_i are equal to -0.1 for i = 1,..., m and 0 for i = m+1,...,5. The error terms are generated from case A. The results are based on the $ADF(0)$ regression.



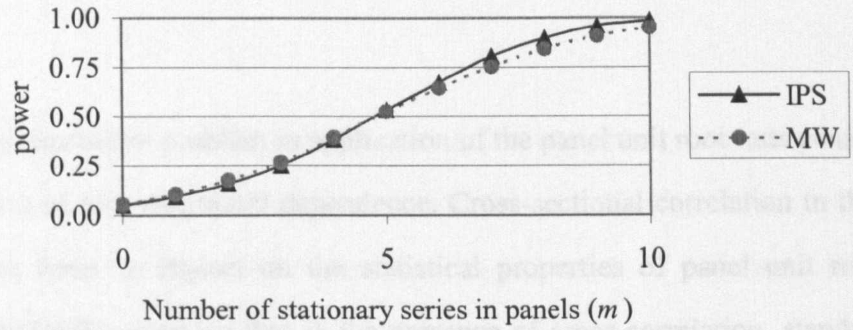
Note: see notes to Figure 2.3, with N=25.

In the large panel ($N = 25$), the simulated results generate a pattern similar to that of the small panel. The panel IPS and MW tests produce no better performance than the individual ADF test when there are fewer than five stationary series in the panel. For example, when $m = 4$, the empirical power of the IPS and MW tests is 0.188 and 0.210, respectively. In addition, the power of both the IPS and MW tests is higher than 0.500 when there are more than eight stationary series in the panel. The MW test produces slightly higher power than the IPS test when the number of the stationary series (m) is lower than eleven (approximately 40 percent of the total number of series in the panel). However, when m increases, the power of the IPS test increases faster than that of the MW test. Then the IPS test is more powerful than the MW test when m>12. When m>17, the empirical power of the IPS test is approximately the same as that of the MW test. These power results are close to 1.000.

The empirical power of the IPS and MW tests in the remaining panels, in which $N = 10$, 15 and 20, are presented in Figures 2.5, 2.6 and 2.7, respectively. All power curves generate a similar pattern, which can be summarised as follows. When the number of stationary series (m) is low, the power of the panel tests will be considerably improved when the marginal stationary series is added. However, the rate of power increase is slower when the power is close to unity. Combining these two stages, the results in Figures 2.3 to 2.7 create a curved J pattern.

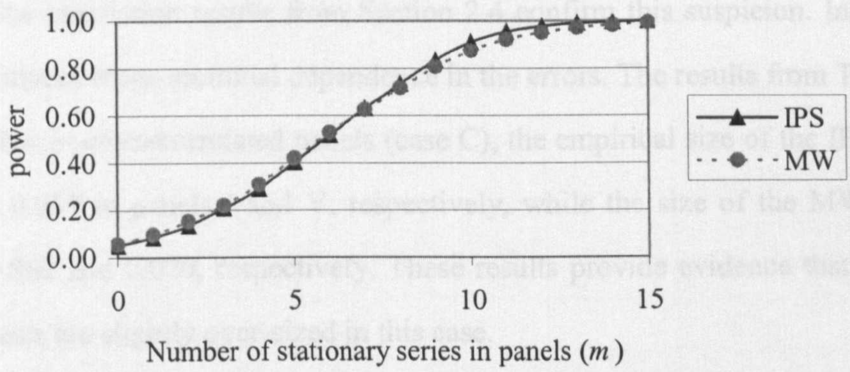
Overall, we conclude that when the majority of series in the panel are non-stationary, the IPS and MW tests have low power in rejecting the non-stationary null hypothesis, and the MW test is slightly more powerful than the IPS test. The inclusion of non-stationary series in the panel worsens the performance of the IPS and MW tests. The empirical power of the IPS and MW tests increases as m increases. The IPS test is slightly more powerful than the MW test when m is large.

Figure 2.5 The empirical power of the IPS and MW tests in a mixed panel, when $N = 10$



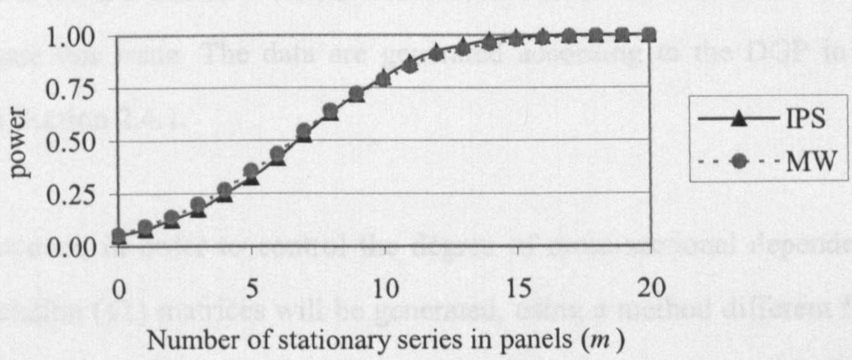
Note: see notes to Figure 2.3, with N=10.

Figure 2.6 The empirical power of the IPS and MW tests in a mixed panel, when $N = 15$



Note: see notes to Figure 2.3, with N=15.

Figure 2.7 The empirical power of the IPS and MW tests in a mixed panel, when $N = 20$



Note: see notes to Figure 2.3, with N=20.

2.6 The effect of cross-sectional dependence

Another major problem in application of the panel unit root tests centres upon the presence of cross-sectional dependence. Cross-sectional correlation in the errors is likely to have an impact on the statistical properties of panel unit root tests. O'Connell (1998) mentions that in the presence of cross-correlation, standard limit distribution of the panel tests will no longer be correct and are not known. Moreover, even if the true distribution of the test statistic is available, the power is likely to decrease, as the total amount of independent information contained in the panel is reduced. The simulation results from Section 2.4 confirm this suspicion. In case C, the DGP induces cross-sectional dependence in the errors. The results from Table 2.4 point out that in cross-correlated panels (case C), the empirical size of the IPS test is 0.065 and 0.069 in panels I and V, respectively, while the size of the MW test is equal to 0.062 and 0.070, respectively. These results provide evidence that the IPS and MW tests are slightly over-sized in this case.

In this section, we further investigate the effect of cross-sectional dependence. The relationship between the values of cross-correlation and size distortions of the IPS and MW tests is considered. Monte Carlo simulations are used to investigate this issue. The data are generated according to the DGP in case C, outlined in Section 2.4.1.

However, in order to control the degree of cross-sectional dependence, the cross-correlation (Ω) matrices will be generated, using a method different from that found in Section 2.4. Here, we apply the cross-correlation matrix used in O'Connell (1998), which takes the form:

$$\Omega = \begin{bmatrix} 1 & \varpi & \cdots & \varpi \\ \varpi & 1 & \cdots & \varpi \\ \vdots & \vdots & \ddots & \vdots \\ \varpi & \varpi & \cdots & 1 \end{bmatrix}_{N \times N} \tag{2.47}$$

The degree of cross-sectional dependence (ϖ) is equal among each series within the panel. We consider the case in which ϖ is set at 0.1, 0.2,..., 0.8 and 0.9, respectively. The equi-correlational error structure is assumed, in order to control the degree of correlation, and to compare the results with different values of ϖ.

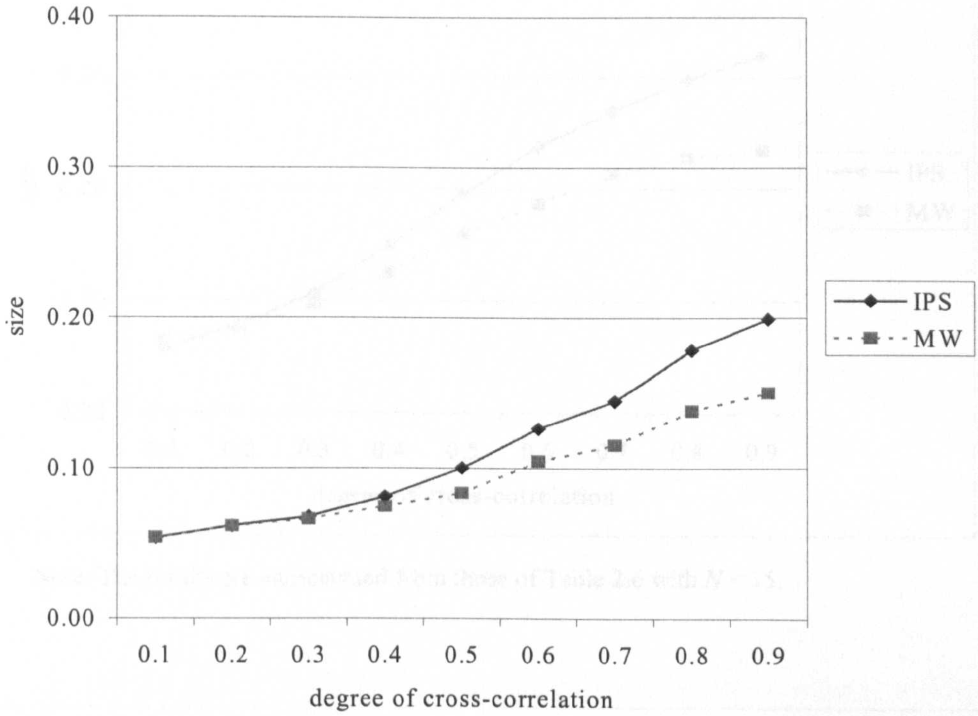
Simulations are conducted on panels I, II, III, IV and V (N = 5, 10, 15, 20 and 25, respectively), using 10,000 replications per experiment. The simulated results of the IPS and MW tests based on *ADF*(1) regression are reported in Table 2.6. A pictorial representation of the empirical size of both panel tests in panels I, II, III, IV and V is presented in Figures 2.8 to 2.12, respectively.

Table 2.6 The empirical size of the IPS and MW tests in cross-correlated panels

	IPS					MW				
ϖ	Panel I	Panel II	Panel III	Panel IV	Panel V	Panel I	Panel II	Panel III	Panel IV	Panel V
0.1	0.054	0.056	0.060	0.062	0.063	0.054	0.061	0.063	0.063	0.067
0.2	0.058	0.066	0.079	0.087	0.101	0.056	0.065	0.075	0.085	0.093
0.3	0.065	0.092	0.107	0.127	0.142	0.061	0.086	0.097	0.112	0.123
0.4	0.081	0.122	0.150	0.171	0.200	0.075	0.106	0.125	0.144	0.168
0.5	0.100	0.154	0.196	0.230	0.255	0.084	0.128	0.161	0.180	0.205
0.6	0.126	0.193	0.238	0.264	0.286	0.105	0.155	0.187	0.209	0.223
0.7	0.145	0.230	0.270	0.298	0.331	0.116	0.177	0.213	0.231	0.256
0.8	0.179	0.250	0.298	0.319	0.352	0.138	0.192	0.228	0.245	0.269
0.9	0.199	0.282	0.320	0.339	0.358	0.150	0.208	0.234	0.251	0.274

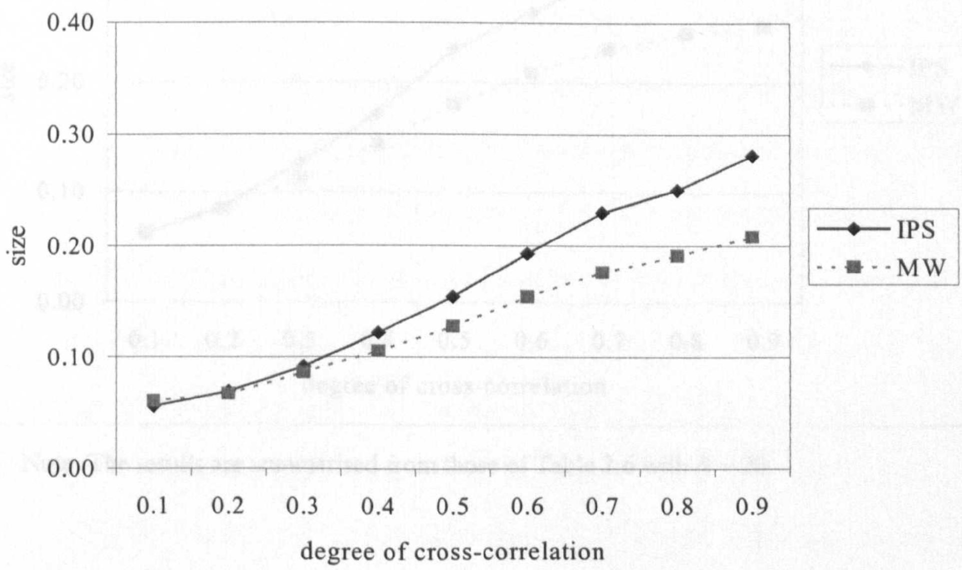
Note: The results are based on the IPS and MW tests (T = 112). The underlying data are generated by equation (2.44) with N=5, 10, 15, 20 and 25. ϕ_i is set to be 0 in the analysis of size. The error terms are generated from case C. The cross-correlation (Ω) matrices are generated as equation (2.47) with ϖ = 0.1, 0.2,... ,0.9. The results are based on the *ADF*(1) regression.

Figure 2.8 The empirical size of the IPS and MW tests in cross-correlated panel, when $N = 5$



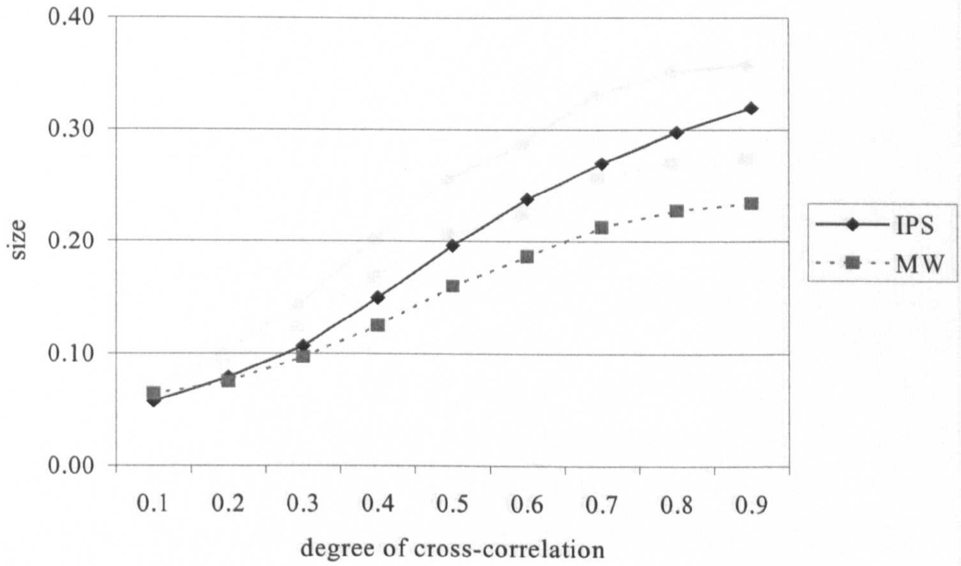
Note: The results are summarised from those of Table 2.6 with $N = 5$.

Figure 2.9 The empirical size of the IPS and MW tests in cross-correlated panel, when $N = 10$



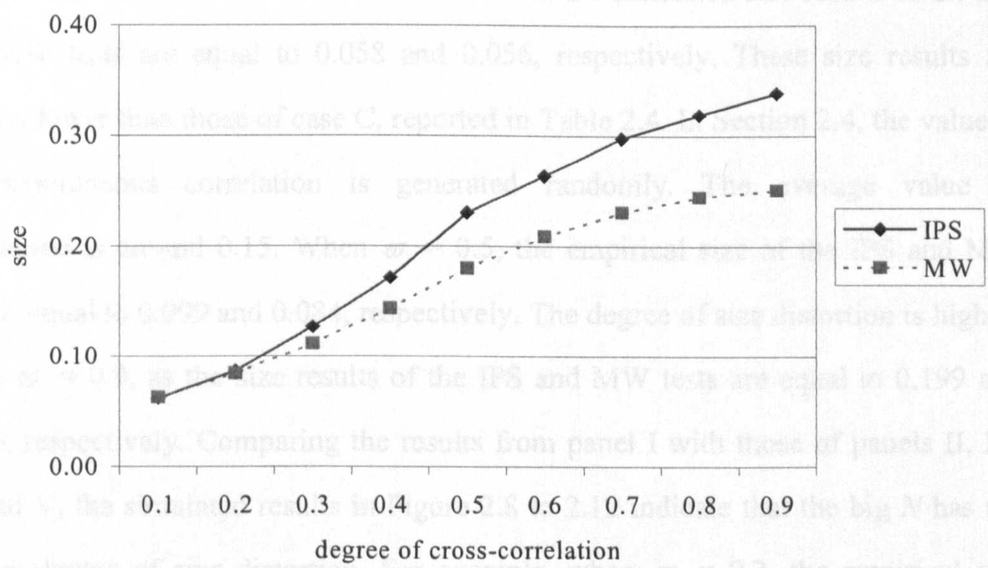
Note: The results are summarised from those of Table 2.6 with $N = 10$.

Figure 2.10 The empirical size of the IPS and MW tests in cross-correlated panel, when N = 15



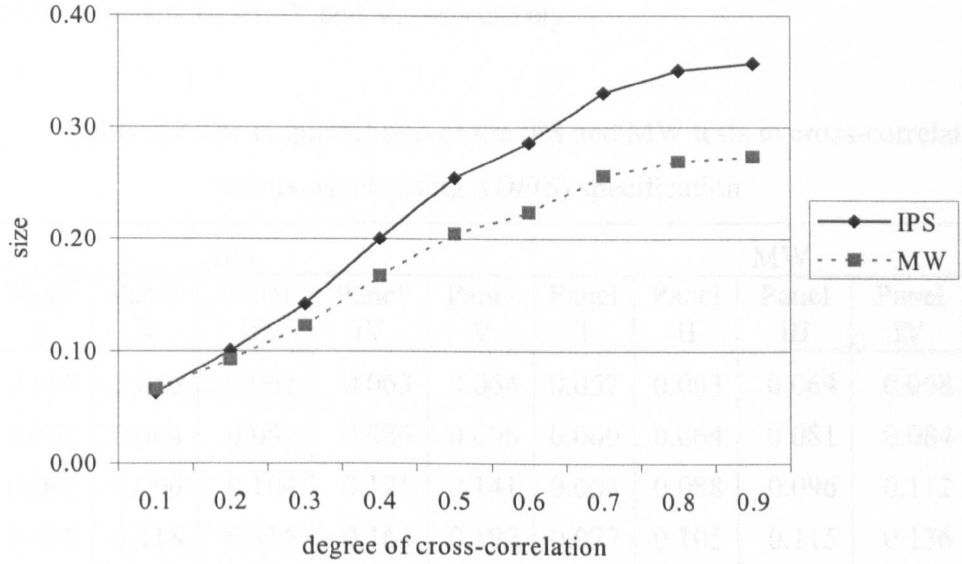
Note: The results are summarised from those of Table 2.6 with N = 15.

Figure 2.11 The empirical size of the IPS and MW tests in cross-correlated panel, when N = 20



Note: The results are summarised from those of Table 2.6 with N = 20.

Figure 2.12 The empirical size of the IPS and MW tests in cross-correlated panel, when N = 25



Note: The results are summarised from those of Table 2.6 with N = 25.

Figures 2.8 to 2.12 show that the degree of size distortion in both panel tests is greater in the panel with higher values of cross-correlation (ϖ). In panel I, the results from Table 2.6 show that when ϖ = 0.2, the simulated size results of the IPS and MW tests are equal to 0.058 and 0.056, respectively. These size results are slightly lower than those of case C, reported in Table 2.4. In Section 2.4, the value of contemporaneous correlation is generated randomly. The average value of correlation is around 0.15. When ϖ = 0.5, the empirical size of the IPS and MW tests is equal to 0.099 and 0.084, respectively. The degree of size distortion is highest when ϖ = 0.9, as the size results of the IPS and MW tests are equal to 0.199 and 0.150, respectively. Comparing the results from panel I with those of panels II, III, IV and V, the simulated results in Figure 2.8 to 2.12 indicate that the big N has the greater degree of size distortion. For example, when ϖ = 0.2, the empirical size results of the IPS test increase from 0.058 for panel I to 0.066, 0.079, 0.087 and 0.101 in panels II, III, IV and V, respectively. The MW test suffers similarly from this problem. However, the degree of size distortion in the MW test is slightly smaller than that of the IPS test in the majority of cases. For example, when ϖ = 0.2, the size results of the MW test increase from 0.056 for panel I to 0.065, 0.075, 0.085, and 0.093 in panels II, III, IV and V, respectively.

Table 2.7 The empirical size of the IPS and MW tests in cross-correlated errors panels using *ADF*(5) specification

	IPS					MW				
ϖ	Panel I	Panel II	Panel III	Panel IV	Panel V	Panel I	Panel II	Panel III	Panel IV	Panel V
0.1	0.053	0.056	0.061	0.063	0.064	0.057	0.063	0.064	0.068	0.068
0.2	0.057	0.064	0.082	0.086	0.096	0.060	0.064	0.081	0.084	0.100
0.3	0.061	0.090	0.104	0.125	0.141	0.063	0.088	0.096	0.112	0.127
0.4	0.080	0.118	0.135	0.163	0.192	0.077	0.105	0.115	0.136	0.164
0.5	0.101	0.142	0.193	0.223	0.239	0.088	0.120	0.159	0.187	0.195
0.6	0.121	0.185	0.232	0.253	0.277	0.105	0.147	0.188	0.204	0.222
0.7	0.144	0.221	0.258	0.288	0.320	0.119	0.173	0.195	0.223	0.249
0.8	0.176	0.254	0.294	0.320	0.339	0.143	0.195	0.222	0.238	0.260
0.9	0.197	0.275	0.314	0.341	0.367	0.150	0.200	0.236	0.255	0.269

Note: The results are based on the IPS and MW tests (T=112). The underlying data is generated by equation (2.44) with N=5, 10, 15, 20 and 25. The error terms are generated from case C. The cross-correlation (Ω) matrices are generated as equation (2.47) with ϖ = 0.1, 0.2,... , 0.9. The results are based on the *ADF*(5) regression.

Table 2.8 The percentage differences between the empirical size of the IPS and MW tests estimated with the *ADF*(1) and *ADF*(5) regressions

	IPS					MW				
ϖ	Panel I	Panel II	Panel III	Panel IV	Panel V	Panel I	Panel II	Panel III	Panel IV	Panel V
0.1	1.9	-1.3	-1.8	-0.5	-1.7	-6.1	-3.6	-2.1	-7.1	-2.9
0.2	4.7	-1.4	-3.7	0.8	4.4	-0.6	-3.7	-7.7	0.9	-7.7
0.3	-1.9	2.3	2.1	1.4	0.9	-4.2	-1.6	1.5	-0.4	-3.2
0.4	2.1	3.2	9.7	4.9	4.3	-3.5	0.3	7.8	5.8	2.2
0.5	-0.8	8.0	1.6	3.2	6.1	-4.8	6.5	1.2	-3.9	4.8
0.6	4.1	4.4	2.6	4.1	3.1	0.0	4.8	-0.8	2.5	0.8
0.7	0.7	4.1	4.4	3.4	3.4	-2.9	2.2	8.3	3.6	3.0
0.8	1.6	-1.6	1.4	-0.3	3.5	-3.1	-1.7	2.6	2.9	3.3
0.9	1.0	2.3	1.8	-0.7	-2.5	0.1	4.0	-0.7	-1.7	1.9
Average	1.4	2.6	2.6	1.9	2.5	-2.4	1.6	1.9	0.9	1.4

Note: The results are the percentage difference between the simulated results based on the *ADF*(1) and *ADF*(5) specification (see notes to Tables 2.6 and 2.7 for details of the DGP).

Next, we consider the results from the *ADF*(5) regression, which represents an over-selecting of the order of the ADF regression. MW suggest that over-fitting may reduce the size distortions of the panel tests. The simulated size results of the IPS and MW tests, using the *ADF*(5) specification are reported in Table 2.7. In addition, the percentage differences between the empirical size of the tests with *ADF*(1) and *ADF*(5) specifications are given in Table 2.8. The results from Tables 2.7 and 2.8 show that over-selecting the order of the ADF regression (*ADF*(5)) generates size results similar to that of the correctly specified regression (*ADF*(1)) in every case. There is no significant evidence that over-fittings alleviate size distortions for both panel tests. The empirical size of the IPS and MW tests based on *ADF*(1) specification is less than 3 percent higher than that of the tests based on *ADF*(5) regressions. Therefore, the differences are virtually negligible.

2.7 Unit root tests in cross-correlated panels

In the previous section, the simulated results show that both the IPS and MW tests are over-sized in cross-correlated panels. The empirical size can be substantially distorted when there is a high degree of cross-correlation between the errors. For example, in the large panel (N=25) with ϖ = 0.9, the size results are 0.358 and 0.274 for the IPS and MW tests, respectively (see Table 2.6). In this section, we investigate the performance of three methods to estimate unit root tests in cross-correlated panels: the bootstrap method of MW, the SURADF test of BMW and the CIPS test of Pesaran (2003).

2.7.1 Bootstrapping panel unit root tests

To correct the size distortions of the panel unit root tests, MW recommend a bootstrap method to calculate the new empirical critical values of the test statistics. In this section, we undertake Monte Carlo simulations to investigate the size and power performance of both the IPS and MW tests with bootstrapped critical values. The data are generated using the DGP, described in Section 2.4, in cases A, B and C for panels I, II and III (N = 5, 10 and 15, respectively).

In case C, the structure of the cross-correlation is generated, using the matrices Ω, outlined in Section 2.6 with ϖ = 0.5 and 0.9. We compare the values of the cross-correlations used in this section with those of an empirical study. Bornhorst (2003) applies several panel unit root statistics in testing the PPP hypothesis, and reports the estimated cross-correlation matrix in the panel of real exchange rates regressions. This matrix is given in Table 2.9. The figures from Table 2.9 show that the degree of cross-correlation is high among EU countries; for example, between Austria and Germany, it is as high as 0.99. The degree of cross-correlation in the errors is lower between some EU and non-EU countries, as the cross-correlation is equal to 0.50 between UK and Japan. Therefore, the selected values of the cross-correlations (0.5 and 0.9) are fairly representative of empirical magnitudes.

The number of replications in each experiment is limited to 500, with 200 bootstrap samples in each replication. We apply the bootstrap method of MW, discussed in Section 2.3, to calculate the empirical distributions. In this chapter, the DGP is set as the trend model (equation (2.44)), in which $y_{i,t}$ is a unit root process with a non-zero drift under the null hypothesis. For this reason, the bootstrap procedure will be slightly different from the process presented in Section 2.3. The equations (2.34) and (2.35) are then adjusted to include an intercept term.

Table 2.9 The estimated cross-sectional correlation matrix in the PPP data reported in Bornhorst (2003)

	UK	At	Be	Dk	Fr	De	Nl	Ca	Jp	Fn	Gr	Es	Au	It	Ch	Ko	Nw	Sw
UK	1.00	0.55	0.59	0.62	0.62	0.59	0.63	0.06	0.50	0.64	0.65	0.63	0.21	0.62	0.53	-0.07	0.60	0.61
At		1.00	0.96	0.92	0.89	0.99	0.97	-0.03	0.62	0.64	0.75	0.74	0.22	0.69	0.84	0.05	0.85	0.71
Be			1.00	0.94	0.89	0.97	0.96	-0.04	0.63	0.65	0.77	0.73	0.24	0.72	0.82	0.03	0.82	0.69
Dk				1.00	0.88	0.94	0.93	0.02	0.64	0.63	0.79	0.73	0.28	0.72	0.82	0.03	0.79	0.68
Fr					1.00	0.90	0.88	-0.01	0.61	0.60	0.79	0.77	0.28	0.72	0.82	0.03	0.85	0.68
De						1.00	0.98	-0.03	0.63	0.65	0.79	0.74	0.22	0.71	0.85	0.05	0.85	0.70
Nl							1.00	-0.04	0.64	0.66	0.79	0.75	0.25	0.71	0.82	0.03	0.81	0.69
Ca								1.00	-0.01	0.11	-0.01	-0.04	0.27	-0.04	0.05	0.22	-0.01	-0.02
Jp									1.00	0.44	0.55	0.51	0.24	0.50	0.64	0.13	0.52	0.42
Fn										1.00	0.58	0.58	0.21	0.59	0.57	0.03	0.64	0.70
Gr											1.00	0.65	0.33	0.69	0.70	0.01	0.73	0.60
Es												1.00	0.22	0.74	0.66	0.10	0.74	0.76
Au													1.00	0.23	0.23	0.29	0.30	0.21
It														1.00	0.70	0.08	0.68	0.68
Ch															1.00	0.09	0.76	0.62
Ko																1.00	0.09	0.05
Nw																	1.00	0.77
Sw																		1.00

Note: The data used are based on a panel of 18 OECD countries, using the US dollar as the base currency. UK, At, Be, Dk, Fr, De, Nl, Ca, Jp, Fn, Gr, Es, Au, It, Ch, Ko, Nw and Sw denote United Kingdom, Austria, Belgium, Denmark, French, Germany, Netherlands, Canada, Japan, Finland, Greece, Spain, Australia, Italy, Switzerland, Korea, Norway and Sweden, respectively.

The simulation results on the finite sample (T=112) size and power properties of the IPS and MW tests with bootstrapped critical values are presented in Tables 2.10 and 2.11. The results from Table 2.10 show that the size of the IPS and MW tests is close to the nominal level (0.05) in most cases when critical values are calculated, using the bootstrap procedure. In the simulation with 500 replications, the 95% confidence interval for a 0.05 test is between 0.0309 and 0.0691. In cross-correlated panels (case C), there is no evidence of the problem of over-sizing in either the IPS or MW tests using bootstrapped critical values. In addition, the effect of under-selecting the order of the ADF regression, discussed in Section 2.4.2, is also corrected. However, the bootstrap IPS and MW tests are slightly under-sized when the $ADF(0)$ regression is used, even in case A, where the $ADF(0)$ regression is appropriate. The empirical size of the IPS (MW) test is equal to 0.032 (0.030), 0.026 (0.024) and 0.016 (0.018), in panels I, II and III, respectively.

Table 2.10 The empirical size of the bootstrap IPS and MW tests

	Number of lags	Panel I (N = 5)		Panel II (N = 10)		Panel III (N = 15)	
		IPS	MW	IPS	MW	IPS	MW
Case A	*ADF (0)*	0.032	0.030	0.026	0.024	0.016	0.018
	ADF (1)	0.046	0.044	0.050	0.056	0.030	0.034
	ADF (2)	0.046	0.040	0.040	0.056	0.040	0.038
Case B	*ADF (0)*	0.028	0.036	0.034	0.030	0.034	0.048
	ADF (1)	0.038	0.044	0.064	0.064	0.050	0.056
	ADF (2)	0.030	0.044	0.044	0.052	0.046	0.052
Case C-1	*ADF (0)*	0.034	0.036	0.038	0.034	0.046	0.048
	ADF (1)	0.040	0.038	0.044	0.056	0.046	0.054
	ADF (2)	0.054	0.050	0.046	0.050	0.042	0.048
Case C-2	*ADF (0)*	0.038	0.040	0.050	0.050	0.054	0.050
	ADF (1)	0.050	0.048	0.056	0.054	0.054	0.046
	ADF (2)	0.040	0.044	0.040	0.040	0.056	0.054

Note: The results are based on the IPS and MW tests. The underlying data are generated by equation (2.44) with N=5, 10 and 15 (see notes to Table 2.6 for details of the DGP). The cross-correlation (Ω) matrices are generated as equation (2.47). ϖ is set to be 0.5 and 0.9 for case C-1 and C-2, respectively. Critical values are obtained from the bootstrap procedure.

Table 2.11 The empirical power of the bootstrap IPS and MW tests

	Number of lags	Panel I ($N = 5$)		Panel II ($N = 10$)		Panel III ($N = 15$)	
		IPS	MW	IPS	MW	IPS	MW
Case A	*ADF (0)*	0.646	0.482	0.902	0.716	0.984	0.896
	ADF (1)	0.748	0.636	0.934	0.842	0.992	0.968
	ADF (2)	0.674	0.566	0.898	0.768	0.986	0.936
Case B	*ADF (0)*	0.370	0.168	0.688	0.316	0.868	0.464
	ADF (1)	0.666	0.564	0.942	0.832	0.992	0.948
	ADF (2)	0.634	0.506	0.886	0.774	0.974	0.920
Case C-1	*ADF (0)*	0.242	0.158	0.318	0.216	0.382	0.242
	ADF (1)	0.496	0.438	0.610	0.526	0.690	0.626
	ADF (2)	0.450	0.380	0.548	0.456	0.628	0.558
Case C-2	*ADF (0)*	0.114	0.112	0.132	0.114	0.126	0.116
	ADF (1)	0.258	0.240	0.278	0.262	0.278	0.262
	ADF (2)	0.210	0.206	0.236	0.212	0.250	0.242

Note: see notes to Table 2.10

The bootstrap method of MW is designed to take care of serial-correlation in the error terms, which may affect the performance of the tests based on the *ADF*(0) regression. The bootstrap sample ($y^*_{i,t}$) is generated, using the estimated residuals and coefficients from equation (2.34), which is based on the *ADF*(1) regression under the null hypothesis (ϕ_i=0). This bootstrap procedure may lead to some size distortion when it is applied with the ADF test with the *ADF*(0) regression. For this reason, it is recommended to include lags in the ADF regression, even though it may not be necessary.

In the analysis of power, the results from Table 2.11 show that, in cases A and B, the bootstrap IPS and MW tests are slightly less powerful than their asymptotic counterparts. However, these power results are still considerably high. For example, in panel I (N=5), the empirical power of the IPS and MW tests is 0.748

and 0.636, respectively in case A, with $ADF(1)$ specification. In case B, these results are 0.666 and 0.564, respectively. In addition, the bootstrap IPS and MW tests are more powerful as N increases

However, in the presence of cross-sectionally correlated errors (case C), the power of the bootstrap IPS and MW tests is significantly lower than their asymptotic counterparts, and is also noticeably lower than that of the tests in cases A and B. The empirical power of the bootstrap IPS and MW tests in panel I (N=5) is 0.496 and 0.438, respectively, when $\varpi = 0.5$ (case C-1) and 0.258 and 0.240, respectively, when $\varpi = 0.9$ (case C-2). These power results offer no improvement over the power of the standard ADF test (see Table 2.3). In the larger panel, the power results of the bootstrap IPS and MW tests are equal to 0.690 and 0.626, respectively in case C-1 (ϖ=0.5) of panel III. However, in case C-2 (ϖ=0.9), the power of the panel unit root tests increases only very slightly with N. The benefit of applying the panel IPS and MW tests is insignificant when the degree of cross-correlation is very high, as the marginal amount of independent information contained in the panel is small.

In summary, the bootstrap method of MW is useful in correcting the critical values of the panel tests when there is cross-sectional correlation in the errors. However, this method should be carefully applied in panel unit root tests. The ADF regressions with different deterministic terms and lag structure require some adjustment from the bootstrap procedure of MW.

2.7.2 Unit root tests with seemingly unrelated regression (SUR)

The Seemingly Unrelated Regression method (SUR) is suggested by O'Connell (1998), Taylor and Sarno (1998), and BMW as an alternative way of estimating unit root tests. The SUR estimator is a multivariate generalised least squares (GLS) method, accounting for the cross-correlations in the errors.

In Section 2.3, we discussed the SURADF test of BMW, which is a system of individual ADF regressions estimated by SUR. Each individual SURADF equation has its own ADF statistic and critical value calculated by the Monte Carlo simulation. Therefore, the SURADF test is directly comparable to the standard ADF test. The null and alternative hypotheses of the SURADF and standard ADF tests are the same. However, BMW note that since the SUR estimation takes account of cross-correlation of the error terms, it should be better than the standard ADF test in the presence of cross-sectional dependence. In addition, the SURADF test supplements the information from panel unit root tests in that rejection of the unit root hypothesis in the panel IPS and MW tests provides information that at least one series in the panel is stationary, but does not indicate how many or which ones are stationary.

Next, we consider the IPS-type t-bar statistic applied with the SURADF regressions (denoted as SURIPS). The SURIPS statistic, calculated in the same way as the IPS statistic, is the average of the ADF t-statistics obtained from the SURADF test. The Fisher-type statistic is not considered in this section because of the problem in calculating p-values of the SURADF t–statistics. The asymptotic ADF distribution cannot be applied to calculate the p-values in this case. The MW test calculated using the p-values from the ordinary ADF distribution suffers from size distortion.

In this section, we perform Monte Carlo simulations to investigate the size and power properties of the SURADF and SURIPS tests. The data are generated according to the DGP outlined in Sections 2.4, 2.6 and 2.8. We consider cases A, B, C-1 and C-2. We set N = 5, 10 and 15 for panels I, II and III, respectively. Critical values of both the SURADF and SURIPS tests are obtained from the bootstrap method discussed in Section 2.6.1. In this section, simulations are carried out with 500 Monte Carlo iterations, each of which uses critical values computed from 200 bootstrap replications. The empirical size and power of the SURADF and SURIPS tests are shown in Tables 2.12 and 2.13, respectively.

The results from Table 2.12 show that the empirical size of the SURADF test approximates the nominal (0.05) in most cases. However, the SURADF test is slightly under-sized in the $ADF(0)$ regression, as in the bootstrap IPS and MW tests discussed in Section 2.7.1. Therefore, we do not consider the power of the tests with the $ADF(0)$ regression. Turning to power performance, we consider the results of the $ADF(1)$ regression. In case A, the empirical power of the SURADF test is 0.186, 0.162 and 0.125 in panels I, II and III, respectively. In case B, these power results are equal to 0.163, 0.145 and 0.116. These power results are worse than those of the standard ADF test (see Table 2.2). However, the power performance of the SURADF test is improved when cross-sectional dependence is presented in the data, as the power of the SURADF test is equal to 0.220 (0.463), 0.190 (0.425) and 0.146 (0.354) for panels I, II and III, respectively, in case C-1 (case C-2). These results show that the SURADF test is markedly better than the standard ADF test only in the presence of strong cross-correlation (case C-2).

Table 2.12 The empirical size and power of the SURADF test

	Number of lags	Panel I ($N = 5$)		Panel II ($N = 10$)		Panel III ($N = 15$)	
		Size	Power	Size	Power	Size	Power
Case A	*ADF(0)*	0.038	0.141	0.035	0.122	0.028	0.094
	ADF(1)	0.046	0.186	0.047	0.162	0.033	0.125
	ADF(2)	0.044	0.168	0.047	0.142	0.036	0.115
Case B	*ADF(0)*	0.043	0.075	0.035	0.055	0.031	0.049
	ADF(1)	0.053	0.163	0.038	0.145	0.038	0.116
	ADF(2)	0.052	0.158	0.040	0.138	0.039	0.106
Case C-1	*ADF(0)*	0.056	0.111	0.042	0.084	0.041	0.070
	ADF(1)	0.053	0.220	0.045	0.190	0.038	0.146
	ADF(2)	0.055	0.201	0.044	0.168	0.040	0.133
Case C-2	*ADF(0)*	0.042	0.215	0.041	0.218	0.044	0.178
	ADF(1)	0.055	0.463	0.048	0.425	0.062	0.354
	ADF(2)	0.042	0.394	0.045	0.374	0.057	0.308

Note: The results are based on the SURADF test (T = 112). The underlying data are generated by equation (2.44) with $N = 5$, 10 and 15. (see notes to Table 2.10 for details of the DGP). Critical values are obtained from the bootstrap procedure.

Table 2.13 The empirical size and power of the SURIPS test

	Number of lags	Panel I ($N = 5$)		Panel II ($N = 10$)		Panel III ($N = 15$)	
		Size	Power	Size	Power	Size	Power
Case A	*ADF(0)*	0.018	0.528	0.014	0.664	0.004	0.650
	ADF(1)	0.038	0.638	0.028	0.786	0.010	0.812
	ADF(2)	0.036	0.552	0.026	0.702	0.010	0.738
Case B	*ADF(0)*	0.026	0.266	0.016	0.350	0.002	0.320
	ADF(1)	0.034	0.578	0.030	0.760	0.010	0.772
	ADF(2)	0.032	0.526	0.026	0.692	0.008	0.668
Case C-1	*ADF(0)*	0.038	0.220	0.032	0.216	0.030	0.162
	ADF(1)	0.048	0.474	0.040	0.450	0.032	0.352
	ADF(2)	0.054	0.420	0.038	0.388	0.038	0.282
Case C-2	*ADF(0)*	0.030	0.290	0.034	0.276	0.040	0.222
	ADF(1)	0.036	0.542	0.032	0.544	0.068	0.440
	ADF(2)	0.036	0.502	0.040	0.466	0.066	0.350

Note: The results are based on the SURIPS test. See notes to Table 2.12 for details of the DGP.

Comparing the power of the SURADF test in the different panel size (N), the results from Table 2.12 show that the power of the SURADF test is reduced when the number of series in the panel (N) increases. The SURADF test is the individual statistic. Therefore, the additional series in the panel do not provide any increase in power. On the other hand, the empirical size of the SURADF test falls as N increases in most cases.

Next, we consider the SURIPS test. The simulated results from Table 2.13 show that the panel SURIPS test is under-sized in many cases. The SURIPS test with the $ADF(0)$ regression is under-sized in every case, a finding similar to those of the IPS, MW, and SURADF tests with bootstrapped critical values; however, the SURIPS test is more severely size-distorted than the other tests. In addition, the SURIPS test is under-sized in cases A and B with $ADF(1)$ and $ADF(2)$ specification. The degree of size distortion increases when N increases. However, in case C, the simulated results are close to the nominal level of 5% in the cases of C-1 and C-2. These results provide evidence that the SURIPS test is not suitable for application when there is no evidence of cross-sectional dependence in the data.

Turning to power performance, the results from Table 2.13 show that by calculating the panel test statistics, the SURIPS test is more powerful than the SURADF test in every case. We do not consider the power from cases A and B because the SURIPS test is massively under-sized. Using the $ADF(1)$ regression, the simulated power results of the SURIPS test are 0.474 (0.542), 0.450 (0.544) and 0.352 (0.440) in panels I, II and III, respectively, in case C-1 (case C-2). These results show that the power of the SURIPS test in the larger panel is lower than that of the smaller panel, which contrasts with the results reported in Sections 2.4 and 2.7. Generally, the power of the panel IPS and MW tests increases when the number of series in panel (N) increases in every case. The possible explanation of this surprising result is that in the presence of cross-correlation, the effect of the size distortion is strong in the large panel, which may lead to an under-estimate of power.

In summary, we recommend the careful application of the SURADF and SURIPS tests only when there is strong evidence of cross-sectional dependence. The SURIPS test is seriously under-sized when it is applied in cross-sectionally independent panels, even though critical values from a bootstrap method of MW are calculated. In addition, application of the SUR method should be used in the small panel. Both the SURADF and SURIPS tests become less powerful as N increases.

2.7.3 Panel unit root tests with a factor model

In this section, we investigate the size and power performance of the CIPS test of Pesaran (2003), using Monte Carlo techniques. The DGP in this section is the same as that of Sections 2.4 and 2.6. The number of cross-section units in the panels (N) is equal to 5, 10, 15, 20, 25 for panels I, II, III, IV and V, respectively. The truncated version of the CIPS test (CIPS*) is applied in this section. Pesaran (2003) notes that the finite sample distribution of the standard CIPS and CIPS* tests are indistinguishable when $T > 20$. Therefore, the CIPS and CIPS* tests provide similar results when the sample size (T) is equal to 112. In addition to the CIPS and CIPS* tests, Pesaran (2003) introduces the cross-sectionally augmented versions of the MW test (denoted as CMW). However, the construction of the CMW test requires the estimation of the individual-specific rejection probabilities by stochastic simulations. Pesaran (2003) shows that the CMW test computed by the standard distribution of the CADF test is over-sized. For this reason, it is necessary that the empirical critical values, obtained by stochastic simulations, compensate for the size distortion. In addition, Pesaran (2003) shows that the CMW test is dominated by the CIPS and $CIPS^*$ tests in terms of higher power. Therefore, we do not consider the CMW test.

The autoregressive coefficient (ϕ_i) is set at 0 and -0.1 in the analysis of size and power, respectively. In this section, the number of replications in each simulation is set to be 1,000. Critical values of the $CIPS^*$ statistic are extracted from Tables 3a to 3c of Pesaran (2003). However, Pesaran (2003) does not provide critical values when $N = 5$ and 25. In view of this, in these cases, the critical values will be obtained from Monte Carlo simulations.

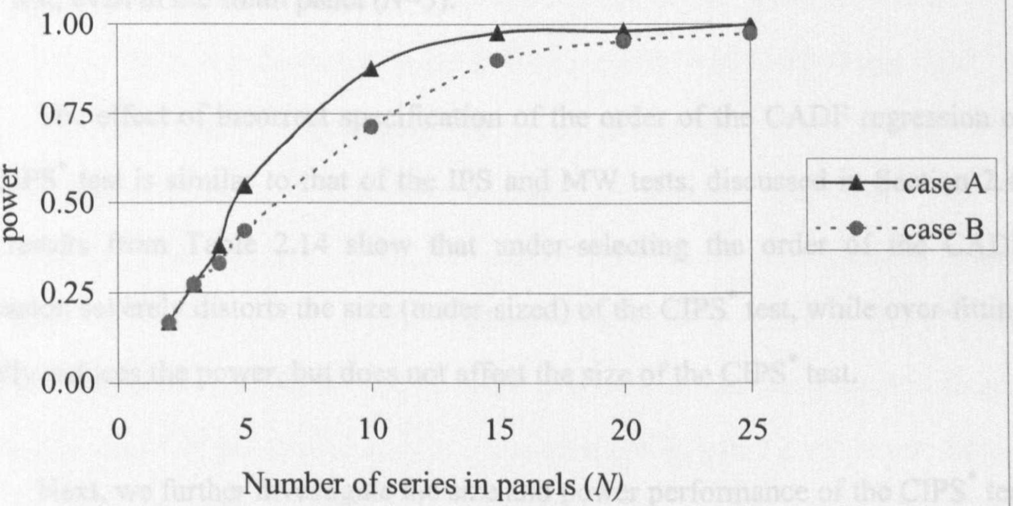
The simulated results pertaining to the size and power of the $CIPS^*$ test in the small (N=5) and large (N=25) panels are presented in Table 2.14. A pictorial representation of the results with regard to the power of the $CIPS^*$ test (cases A and B) in the panels with $N = 2, 3, 4, 5, 10, 15, 20$ and 25 is shown in Figure 2.13.

Table 2.14 The empirical size and power of the $CIPS^*$ test

	Number of lags	Panel I ($N = 5$)		Panel V ($N = 25$)	
		Size	Power	Size	Power
Case A	*CADF(0)*	0.050	0.547	0.060	0.999
	CADF(1)	0.048	0.483	0.061	0.995
	CADF(2)	0.048	0.376	0.052	0.950
Case B	*CADF(0)*	0.001	0.011	0.000	0.007
	CADF(1)	0.047	0.422	0.053	0.975
	ADF(2)	0.049	0.337	0.045	0.898
Case C	*CADF(0)*	0.002	0.018	0.000	0.009
	CADF(1)	0.063	0.441	0.054	0.973
	CADF(2)	0.052	0.363	0.051	0.900

Note: The results are based on the $CIPS^*$ test. The underlying data are generated by equation (2.44) with N=5 and 25. (see notes to Table 2.4 for details of the DGP). Critical values are generated from Monte Carlo simulations. The 95% critical values of the $CIPS^*$ test are equal to −3.04 and −2.68 when $N = 5$ and 25, respectively.

Figure 2.13 The empirical power of the CIPS* test
(Case A and B)



Note: The results are based on the CIPS* test ($T = 112$). The underlying data are generated by equation (2.44) with N=2,3,4,5,10,15,20,25 and $\phi = -0.1$. The error terms are generated from cases A and B. The results are based on the $CADF(0)$ and $CADF(1)$ regressions in cases A and B, respectively.

The results from Table 2.14 show that using the $CADF(1)$ regression, the empirical size of the CIPS* test is close to the nominal level of 5% in every case. For example, in case C, the empirical size of the CIPS* test is equal to 0.063 and 0.054 when N=5 and 25, respectively. The 95% confidence interval of the 0.05 significant level test lies between 0.0365 to 0.0635 when simulations are based on 1,000 replications. These results show that the CIPS* test can be used to correct for the size distortion in cross-correlated panels. Moreover, the CIPS* test does not require the empirical critical values of a bootstrap method.

Considering the power of the CIPS* test in cases A and B, Figure 2.13 shows that power curves of the CIPS* test create a similar pattern to those of the IPS and MW tests (see Figures 2.1 and 2.2). The empirical power of the CIPS* test is higher than 0.500, when $N \geq 5$ and 10, in cases A and B, respectively. The empirical power of the test increases with N. The power is close to 1.000 when N>15 and 20 for cases A and B, respectively. In case B, the power results are slightly lower than those of case A. Comparing the CIPS* test with the IPS and MW tests, the CIPS* test is less powerful than the IPS and MW tests, but is still more powerful than the standard ADF test, even in the small panel (*N*=5).

The effect of incorrect specification of the order of the CADF regression of the CIPS* test is similar to that of the IPS and MW tests, discussed in Section 2.4. The results from Table 2.14 show that under-selecting the order of the CADF regression severely distorts the size (under-sized) of the CIPS* test, while over-fitting slightly reduces the power, but does not affect the size of the CIPS* test.

Next, we further investigate the size and power performance of the CIPS* test in the presence of cross-sectional dependence. The cross-correlation matrices used in Section 2.6 are applied to the DGP in case C. The degree of cross-correlation (ϖ) is set to be 0.1, 0.2, …, 0.9 for panels I, II, II, IV and V (*N*=5, 10, 15, 20 and 25, respectively). The simulated results are reported in Table 2.15.

Table 2.15 The empirical size and power of the CIPS* test in cross-correlated panels

	Size					Power				
ϖ	Panel I	Panel II	Panel III	Panel IV	Panel V	Panel I	Panel II	Panel III	Panel IV	Panel V
0.1	0.046	0.045	0.055	0.051	0.045	0.450	0.689	0.880	0.897	0.933
0.2	0.055	0.054	0.062	0.056	0.049	0.458	0.670	0.903	0.903	0.932
0.3	0.063	0.046	0.049	0.042	0.036	0.428	0.696	0.882	0.891	0.945
0.4	0.053	0.052	0.062	0.050	0.050	0.439	0.680	0.896	0.893	0.956
0.5	0.049	0.053	0.062	0.059	0.052	0.424	0.657	0.903	0.862	0.953
0.6	0.051	0.054	0.050	0.056	0.052	0.453	0.675	0.883	0.893	0.949
0.7	0.062	0.060	0.053	0.048	0.052	0.433	0.657	0.885	0.906	0.939
0.8	0.057	0.048	0.063	0.055	0.059	0.443	0.655	0.906	0.900	0.947
0.9	0.048	0.059	0.055	0.060	0.046	0.438	0.667	0.895	0.897	0.939

Note: The results are based on the CIPS* test with the *CADF*(1) regression. The underlying data are generated by equation (2.44) with *N*=5, 10, 15, 20 and 25 (see notes to Table 2.6 for details of the DGP).

The results from Table 2.15 show that in the presence of cross-sectional dependence, the empirical size of the CIPS* test is still reasonably close to the nominal level (0.05) in every case. For example, in the highly cross-correlated panel (ϖ =0.9), the simulated results are equal to 0.048, 0.059, 0.055, 0.060 and 0.046 in panels I, II, III, IV and V, respectively. The power results of the CIPS* test rise with N. For example, when ϖ =0.5, the power results are equal to 0.424, 0.657, 0.903, 0.862 and 0.953 in panels I, II, III, IV and V, respectively. The value of cross-correlations do not affect the empirical power of the CIPS* test. For fixed N, the empirical power of the CIPS* test is relatively constant in the panels with the different values of cross-correlations (ϖ). For example, in panel I, the power results are equal to 0.450, 0.428, 0.424, 0.433 and 0.438 when ϖ =0.1, 0.3, 0.5, 0.7 and 0.9, respectively.

Overall, the simulation results in this section show that in the finite sample (T=112), the empirical size of the CIPS* test is reasonably close to the nominal level of 0.05 for every panel size (N). The empirical power of the test increases with N. The presence of cross-correlations in the errors does not affect the empirical size and power of the CIPS* test in the finite sample (T=112).

Comparing the power of the CIPS* test, the bootstrap IPS test and the SURIPS test in cross-correlated panels (case C), all three tests have relatively similar power in panel I (N=5) with ϖ =0.5. The simulated power results are equal to 0.496, 0.474 and 0.424 for the bootstrap IPS, SURIPS and CIPS* tests, respectively. However, in the same panel (N=5), the SURIPS test is slightly more powerful than the other tests when ϖ =0.9, as these results are equal to 0.258, 0.542 and 0.438, respectively. For the larger panels ($N \geq 10$), the CIPS* test dominates the other tests in terms of higher power. For example, in panel III, the empirical power of the bootstrap IPS, SURIPS and CIPS* tests is 0.626 (0.262), 0.352 (0.440) and 0.903 (0.895), respectively when ϖ =0.5 (ϖ =0.9).

In summary, the CIPS* test is recommended for use in testing for unit roots in cross-correlated panels because of the advantage of its higher power over the bootstrap IPS and SURIPS tests, when $N \geq 10$. Moreover, the CIPS* test does not require the generation of the empirical critical values of the bootstrap method, which makes it easier to use in empirical work.

2.8 Conclusion

In this chapter, the size and power performance of the IPS and MW tests were investigated using Monte Carlo simulations. We considered the case of a moderate sample size (T) corresponding to quarterly data for the post-Bretton Woods era (1973:1 to 2002:4). The simulated results showed that the standard ADF test has low power to reject the non-stationary null hypothesis when the speed of mean reversion is slow (ϕ_i>-0.2). The panel IPS and MW tests improve the power performance over the standard ADF test and become more powerful as N increases. The empirical power of both panel tests approaches unity when N>10. Comparing the IPS and MW tests, the IPS test is slightly more powerful than the MW test in the majority of cases.

In a mixed panel, the inclusion of non-stationary series in the panel reduces the power of the tests. When N = 5, 10 and 15, the empirical power of the IPS and MW tests is higher than 0.500 when there is more than thirty to forty percent of the stationary series in the panel. When N = 20 and 25, this power figure (0.500) can be achieved when there are more than 7 to 8 stationary series in the panel. Comparing the IPS and MW tests, the MW test is slightly more powerful than the IPS test when the proportion of stationary series in the panel is approximately less than 40 percent. When the number of stationary series in the panel increases, the power of both the IPS and MW tests rises. However, the power of the IPS test increases faster than that of the MW test. Therefore, the IPS test is more powerful than the MW test when there is more than approximately 40 percent of stationary series in the groups.

The presence of cross-sectional dependence in the errors affects the size properties of both the IPS and MW tests. These panel tests are over-sized in cross-correlated panels. The degree of size distortion depends on both the values of correlations (ϖ) and the panel size (N). In the small panel, the size results are less distorted than those of the large panel. The higher size distortions are observed as the values of ϖ increase. The MW test is slightly less size-distorted than the IPS test in most cases. Over-selecting the order of ADF regression does not yield a significant difference in the size distortions from exact-selecting of the order of the ADF regression in our simulation.

To compensate for the size distortion problem, the bootstrap method of MW was then applied to calculate the empirical critical values of the IPS and MW tests. The size results of the bootstrap IPS and MW tests are close to the nominal level. However, in the highly cross-correlated panels (case C-2), the power of the bootstrap panel unit root tests is markedly lower than that of the corresponding tests in case A.

Next, we applied the SUR method to estimate the ADF test. The power of the SURADF and SURIPS tests improves from that of the OLS counterparts when there is a strong contemporaneous correlation in the errors (case C-2). However, in the remaining cases, the SURADF and SURIPS tests do not provide the improvement in the power performance over the tests with standard OLS.

Finally, we considered the CIPS test of Pesaran (2003). This test has the capacity to compensate for the size distortions in cross-correlated panels. The empirical power of the CIPS test depends on the number of series in the panel (N), but does not depend on the degree of cross-correlations (ϖ).

Comparing the three alternative methods used in estimating panel unit root tests in cross-correlated panels, the CIPS test is better than the other two in the majority of cases in terms of its higher power. In addition, the CIPS test does not require the calculation of the critical values from the bootstrap method. However, the SURIPS test is more powerful than the other two tests in the small panel (N=5) with highly correlated errors (ϖ =0.9). The bootstrap IPS test has the highest power in the small panel (N=5) with a moderate degree of cross-correlation (ϖ =0.5).

Chapter 3

Cointegration Tests in Heterogeneous Panels

3.1 Introduction

The analysis of cointegration in panel data has recently received increasing attention. There are two standard approaches in testing for this type of cointegration. The first approach is based on the Engle and Granger (1987) residual-based cointegration test. In this approach, a long-run relationship is estimated in the first step. In the second step, a unit root test on the estimated residuals obtained from the long-run regressions is undertaken. Kao (1999) develops several residual-based panel cointegration statistics based on a homogeneity assumption in both the cointegrating vector in the first step and autoregressive coefficients in the second step. Pedroni (1999) relaxes this homogeneity assumption and proposes several panel tests for cointegration, based on both homogeneity and heterogeneity assumptions.

The second route is a panel version of the likelihood ratio (*LR*) test for the cointegration rank in a VAR of Johansen (1988). Larsson, Lyhagen and Lothgren (2001) (LLL) propose a panel test to estimate the cointegrating rank in the panel as the average of the individual Johansen likelihood-based cointegration rank trace test statistics. This LLL *LR*-bar statistic, defined similarly as the IPS *t*-bar statistic, is based on heterogeneous panels.

Recently, these panel cointegration tests have been used in many empirical studies, mainly focusing on testing for the existence of purchasing power parity (PPP), e.g. Coakley and Fuertes (1997) and Pedroni (2001). In addition, Groen (1999) applies the residual-based panel cointegration test to a panel of fourteen OECD countries for monetary models. LLL also use the *LR*-bar statistic in testing for a consumption function. The empirical results from these papers usually provide more significant evidence of cointegration relationships than that of the standard tests.

In this chapter, we consider panel cointegration tests in heterogeneous panels, using both the residual-based and likelihood-based methods. In the first approach, we consider the direct extension of the panel unit root tests of Im, Pesaran and Shin (2003) (IPS) and Maddala and Wu (1999) (MW) in testing for stationarity of the estimated residuals, allowing for heterogeneity in both the cointegrating vectors and autoregressive coefficients of the residual regressions. In the second framework, the *LR*-bar statistic of LLL is considered in testing for cointegration in heterogeneous panels, based on the likelihood inference for VAR models.

The purpose of this chapter is to compare, by means of Monte Carlo simulations, the size and power properties of the residual-based panel cointegration tests of IPS and MW and the likelihood-based panel LLL rank test. Moreover, we investigate the effect on the cointegration tests of having a mixture of cointegrated and non-cointegrated relationships in the panel, and the effect of cross-sectional dependence in the underlying series. The performance of the bootstrap residual-based and likelihood-based tests and the panel cointegration test of CIPS are also considered to correct the size distortions.

The chapter is outlined as follows. In the next section, we present a review of literature on panel cointegration tests. Section 3.3 outlines the panel residual-based test of IPS and MW, and the panel likelihood-based test of LLL. The results of the Monte Carlo experiments are discussed in Section 3.4. Simulations on the panel with a mixture of cointegrated and non-cointegrated relationships are conducted in Section 3.5. Section 3.6 introduces the bootstrap method to correct the size distortions of the panel cointegration tests and presents some simulation results. Section 3.7 applies the Cross-sectionally augmented IPS (CIPS) panel unit root test of Pesaran (2003) to the residual-based panel cointegration test. Section 3.8 offers some conclusions to this chapter.

3.2 Literature review

The early tests for cointegration in panel data simply apply panel unit root tests to residuals from some long-run regressions, based on the two-step approach of Engle and Granger (1987). The initial applications of panel cointegration tests are developed by Kao (1999), Pedroni (1999), and McCoskey and Kao (1998).

Kao (1999) proposes a residual-based test for cointegration in a panel, under the null hypothesis of no cointegration. The proposed method is based on the spurious LSDV regression model:

$$y_{i,t} = \alpha_i + \beta x_{i,t} + e_{i,t} \tag{3.1}$$

where $i = 1,..., N$; $t = 1, ..., T$ and $x_{i,t}$ and $e_{i,t}$ are $I(1)$ process. The least square dummy variable (LSDV) method is applied to estimate the long-run regression. The LSDV estimator of β is:

$$\hat{\beta} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t}(x_{i,t} - \bar{x}_i)}{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{i,t} - \bar{x}_i)^2} \tag{3.2}$$

The ADF test is then applied to estimated residuals ($\hat{e}_{i,t}$) to test the null hypothesis of no cointegration, that is:

$$\hat{e}_{i,t} = \rho\hat{e}_{i,t-1} + v_{i,t} \tag{3.3}$$

where $\hat{e}_{i,t} = y_{i,t} - \hat{\alpha}_i - \hat{\beta}x_{i,t}$.

The null and alternative hypotheses are:

$H_0 : \rho_i = 0$ for all i against the alternative

$H_a : \rho_i = \rho < 0$ for all i.

The OLS estimator of autoregressive coefficient ($\hat{\rho}$) is:

$$\hat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{i,t} \hat{e}_{i,t-1}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{i,t}^{2}} \tag{3.4}$$

The t-statistic (t_ρ) to test the null hypothesis of no cointegraton is:

$$t_\rho = \frac{(\hat{\rho} - 1)\sqrt{\sum_{i=1}^{N} \sum_{t=2}^{T} \hat{e}_{i,t}^{2}}}{s_e} \tag{3.5}$$

where $s_e^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=2}^{T} (\hat{e}_{i,t} - \hat{\rho}\hat{e}_{i,t-1})^2$.

Kao (1999) proposes five adjusted statistics, which have an asymptotic $N(0,1)$ distribution, the discussion and mathematical exposition of which are contained in his paper. These proposed test statistics are an extension of the statistics of Levin and Lin (1992) (LL) that test for cointegration. Homogeneity is assumed in both the long-run slope coefficient (β) and autoregressive coefficient (ρ) of the estimated residual regressions.

Pedroni (1999) develops panel tests for the null of no cointegration in dynamic panels with multiple regressors. This method utilises residuals from the cointegration regressions given by the general system:

$$y_{i,t} = \alpha_i + x'_{i,t}\beta_i + e_{i,t} \tag{3.6}$$

where $t = 1,...,T$; $i = 1,...,N$ and $x'_{i,t} = \{x_{1,i,t}, x_{2,i,t},...., x_{k-1,i,t}\}'$.

Each equation is estimated independently, allowing for heterogeneity in the long-run cointegrating vectors. Two types of panel statistics are proposed, based on both homogeneity and heterogeneity assumptions in testing for stationarity of the estimated residuals ($\hat{e}_{i,t}$).

The first four statistics are based on estimators that pool autoregressive coefficients for the unit root tests on estimated residuals, referred to as "panel cointegration statistics". The null and alternative hypotheses of this first group of statistics are similar to those of Kao (1999). The second group of statistics is based on estimators that average individual estimated coefficients for each cross-section unit, denoted as "group mean cointegration statistics". While the statistics in the first group have a common autoregressive coefficient ($\rho_i = \rho$), those of the second group are based on heterogeneity of these autoregressive coefficients. Therefore, the null and alternative hypotheses of these statistics are:

$H_0 : \rho_i = 0$ for all i against

$H_a : \rho_i < 0$ for all I

The mathematical exposition of these seven statistics is contained in Table 1 of Pedroni (1999).

Pedroni (1999) shows that the asymptotic distributions for each of seven panel and group mean statistics can be expressed as:

$$\frac{\chi_{N,T} - \mu\sqrt{N}}{\sqrt{v}} \Rightarrow N(0,1)$$

where $\chi_{N,T}$ is a test statistic, appropriately standardised with respect to N and T, and μ and v are the corresponding values of the mean and variance for each of the test statistics, respectively. These values of μ and v are reported in Table 2 of Pedroni (1999).

McCosky and Kao (1998) propose a residual-based panel cointegration test under the null hypothesis of cointegration, which is an extension of the LM unit root test of Kwiatkowski *et al.* (1992) (KPSS) and the panel unit root test of Hadri (2000).

The model allows for varying slopes and intercepts across units:

$$y_{i,t} = \alpha_i + x'_{i,t}\beta_i + v_{i,t} \tag{3.7}$$

where $t = 1,...,T$; $i = 1,...,N$ and $x'_{i,t} = \{x_{1,i,t}, x_{2,i,t},...., x_{k-1,i,t}\}'$. Assume that $x_{i,t}$ are $I(1)$ process for all i, then:

$$x_{i,t} = x_{i,t-1} + \varepsilon_{i,t} \tag{3.8}$$

$$v_{i,t} = \tau_{i,t} + u_{i,t} \quad , \quad \tau_{i,t} = \tau_{i,t-1} + \theta u_{i,t} \tag{3.9}$$

where $u_{i,t} \sim i.i.d.N(0, \sigma_u^2)$.

McCosky and Kao (1998) note that the individual constant forms (α_i) can be extended to include deterministic time trends, such as: $\alpha_{0,i} + \alpha_{1,i}t$.

The null hypothesis of cointegration is tested by:

$H_0 : \theta = 0$ against

$H_a : |\theta| \neq 0$

Then, equation (3.7) can be re-written as:

$$y_{i,t} = \alpha_i + x'_{i,t}\beta_i + e_{i,t} \tag{3.10}$$

where $e_{i,t} = (\theta \sum_{j=1}^{t} u_{i,j}) + u_{i,t}$.

The $L\overline{M}$-statistic is calculated as:

$$L\overline{M} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} S_{i,t}^2}{s^2} \tag{3.11}$$

where $S_{i,t}^2$ is the partial sum process of the estimated residuals, $S_{i,t}^2 = \sum_{j=1}^{T} \hat{e}_{i,j}$; s^2 is a consistent estimator of σ_u^2 under the null hypothesis, $s^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{e}_{i,t}^2$ and estimated residuals from equation (3.10) ($\hat{e}_{i,t}$) can be obtained, using either the dynamic OLS (DOLS) or the fully-modified (FM) estimator presented in Kao and Chiang (2000). The asymptotic result for the test is then shown to follow:

$$\overline{LM}^* = \frac{\sqrt{N}(\overline{LM} - \mu_v)}{\sigma_v} \Rightarrow N(0,1) \tag{3.12}$$

The correction factors (μ_v, σ_v^2) are the mean and variance. The estimation of μ_v, σ_v^2 is discussed in Appendix C of McCosky and Kao (1998).

The second approach in testing for cointegration in panel data is based on the likelihood inference for a VAR model of Johansen (1988). LLL develop a likelihood ratio (*LR*) test for determining the cointegration rank in heterogeneous panels, which is based on the likelihood ratio cointegration rank trace test statistic of Johansen (1988). Further details of this LLL *LR*-bar statistic are presented in Section 3.3.2.

Groen and Kleibergen (1999) also adapt the Vector Error Correction (VEC) framework for cointegration analysis in panel data. Maximum likelihood estimators of the cointegrating vectors are constructed, using the Generalised Method of Moments (GMM), iterated over all parameters. This GMM method can be interpreted as the SUR-type estimator. The maximum likelihood estimates are used to construct likelihood ratio statistics to test for a common cointegration rank based on both heterogeneous and homogenous cointegrating vectors. Groen and Kleibergen (1999) show that the proposed likelihood ratio tests have a limit distribution equal to a summation of limit distributions of Johansen (1988) trace statistics.

The advantage of the Groen and Kleibergen (1999) test over the LLL test is that the LLL panel rank test is based on the average of the individual statistics calculated independently. This statistic is likely to suffer from size distortions when the error terms are cross-sectionally correlated. By contrast, the Groen and Kleibergen (1999) test conducted the likelihood ratio statistics based on the estimation of panel VEC model simultaneously, which allow for the unrestricted disturbance covariance matrix in the panel.

3.3 The panel data cointegration tests

Consider a panel data set $Y_{i,t}$ that consists of a sample of N cross-sections (e.g. industries, countries) observed over T time periods. The number of time-series observations can vary across groups but, for notational convenience, a common T is used. The number of variables in each group is equal to k. Then, $Y_{i,t} = (y_{1,i,t}, y_{2,i,t}, \ldots, y_{k,i,t})'$, where $y_{l,i,t}$ denotes the l^{th} variable for the i^{th} cross-section at time t. The methods of testing for cointegration relationships are explained as follows.

3.3.1 The residual-based panel cointegration tests

We apply the method of the residual-based panel cointegration tests of Pedroni (1999), discussed in Section 3.2. In this chapter, the panel unit tests for the OLS residuals ($\hat{e}_{i,t}$) from the individual cointegration equations (equation (3.6)) are constructed, based on the IPS and MW methods, discussed in Section 2.2. The residual-based panel cointegration test of IPS is similar to one of three group mean cointegration statistics proposed by Pedroni (1999).

The null and alternative hypotheses for these panel statistics are:

$H_0 : \rho_i = 0$ for all $i = 1,\ldots,N$ against the alternative

$H_a : \rho_i < 0$ for some i and $\rho_i = 0$ for the other i.

The null hypothesis implies that there is no cointegration relationship between $y_{i,t}$ and $x'_{i,t}$ for all i systems (e.g. countries, firms). This null is tested against the alternative that there is at least one group in the panel in which the cointegration relationship exists.

In this chapter, we consider the residual-based panel cointegration of IPS and MW tests, in order to extend our study in Chapter 2 into the multivariate case.

3.3.2 The likelihood-based panel cointegration test

The panel LLL rank trace test statistic is given by the average of individual likelihood ratio cointegration rank trace test statistics over the panel individuals. The multivariate cointegration analysis of Johansen (1988) is applied to estimate each individual cross-section system independently, thereby allowing heterogeneity in each cross-sectional unit in the panel. The data generating process (DGP) for each of the groups is characterised by the following heterogeneous VAR(p_i) model.

$$Y_{i,t} = \sum_{j=1}^{p_i} \Lambda_{i,j} Y_{i,t-j} + \varepsilon_{i,t}\,, \tag{3.13}$$

where $i = 1, \dots, N$; $t = 1, \dots, T$.

For each group i, the value of $Y_{i,-j+1}, \dots, Y_{i,0}$ is considered fixed and $\varepsilon_{i,t}$ are independent identically distributed: $\varepsilon_{i,t} \sim N_k(0, \Omega_i)$, where Ω_i is the matrix of cross-correlation in the error terms; $\Omega_i = E(\varepsilon_{i,t}, \varepsilon'_{i,t})$. Equation (3.13) can be re-written in the VECM model as:

$$\Delta Y_{i,t} = \Pi_i Y_{i,t-1} + \sum_{j=1}^{p_i - 1} \Gamma_{i,j} \Delta Y_{i,t-j} + \varepsilon_{i,t} \tag{3.14}$$

where $\Pi_i = \Lambda_{i,1} + \Lambda_{i,2} + \ldots + \Lambda_{p_i} - I$ and $\Gamma_{i,j} = \Lambda_{i,j} - \Lambda_{i,j-1}$, Π_i is of order $k \times k$. If Π_i is of reduce rank, $\text{rank}(\Pi_i) = r_i$, it can be decomposed into $\Pi_i = \alpha_i \beta_i'$, where α_i and β_i are of order $k \times r_i$ and of full column rank, which represent the error correction form.

The null and alternative hypotheses of the panel LLL rank test are:

$H_0 : rank(\Pi_i) = r_i \leq r$ for all $i = 1, \ldots, N$ against

$H_a : rank(\Pi_i) = k$ for all $i = 1, \ldots, N$

The testing procedure is sequential, which is similar to the individual trace test procedure for cointegration rank determination. First, we test for $H_0 : rank(\Pi_i) = r_i \leq r$, $r = 0$. If this hypothesis of no cointegrating vector cannot be rejected, we conclude that there are no cointegration relationships ($\text{rank}(\Pi_i) = r_i = 0$) in all cross-section groups in the panel. If this null hypothesis is rejected, the null hypothesis, $r=1$, is tested. The sequential procedure is continued until the null hypothesis is accepted or the hypothesis, $r = k-1$, is rejected. Rejecting the hypothesis of no cointegration ($r = 0$) and accepting the null of $\text{rank}(\Pi_i) = r_i \leq r$ ($0 < r < k$) implies that there is at least one cross-section unit in the panel that has rank $(\Pi_i) = r > 0$. This procedure is comparable to the residual-based panel cointegration tests of IPS and MW, as heterogeneity in cointegrating vectors in the panel is allowed. Moreover, the possibility of a mixed panel, in which some relationships are cointegrated and others not, is also allowed in these tests.

The likelihood ratio trace test statistic for group i is:

$$LR_{iT}\{H(r) \mid H(k)\} = -2\ln Q_{iT}(H(r) \mid H(k)) = -T \sum_{l=r+1}^{p} \ln(1 - \hat{\lambda}_{li}) \qquad (3.15)$$

where $\hat{\lambda}_l$ is the l^{th} largest eigen value in the i^{th} cross-section unit.

The LR-bar statistic is then calculated as the average of the individual trace statistics:

$$L\overline{R}_{iT}(H(r) \mid H(k)) = \frac{1}{N}\sum_{i=1}^{N} LR_{iT}(H(r) \mid H(k)) \tag{3.16}$$

Finally, the standardised LR-bar statistic is defined by:

$$\gamma_{L\overline{R}}\{H(r) \mid H(k)\} = \frac{\sqrt{N}(L\overline{R}_{NT}\{H(r) \mid H(k)\} - E(Z_k))}{\sqrt{VAR(Z_k)}} \tag{3.17}$$

where $E(Z_k)$ and $Var(Z_k)$ are the mean and variance of the asymptotic trace statistic, which can be obtained from simulations. The relevant values of $E(Z_k)$ and $Var(Z_k)$ are presented in Table 1 of LLL and Table 0, 1,1*, 2 and 2* of Osterwald-Lenum (1992).

LLL prove the central limit theorem for the standardised LR-bar statistic that under the null hypothesis, $\gamma_{L\overline{R}} \Rightarrow N(0,1)$ as N and $T \to \infty$ in such a way that $\sqrt{N}T^{-1} \to 0$, under the assumption that there is no cross-correlation in the error terms, that is:

$$E(\varepsilon_{i,t}) = 0 \quad \text{and} \quad E(\varepsilon_{i,t}\varepsilon_{j,t}) = \begin{cases} \Omega_i & \text{for} \quad i = j \\ 0 & \phantom{\text{for}} \quad i \neq j \end{cases} \tag{3.18}$$

LLL note that $T \to \infty$ is needed for each of the individual test statistics to converge to its asymptotic distribution, while $N \to \infty$ is needed for the central limit theorem. As this panel LLL rank test is one-sided, the null hypothesis is rejected at a significance level of α when $\gamma_{L\overline{R}} > z_{1-\alpha}$.

3.4 A Monte Carlo simulation study

In this section, Monte Carlo experiments are used to investigate the performance of various panel cointegration tests in terms of the size and power. Specifically, we compare the size and power properties of different panel cointegration statistics. The residual-based panel cointegration tests of IPS, MW and the likelihood-based panel LLL rank test are constructed. Simulations are performed in EVIEW, version 4.1.

3.4.1 Simulation design

In this chapter, the DGP is derived as an error correction representation:

$$\Delta Y_{i,t} = \Pi_i Y_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{i,k} \Delta Y_{i,t-j} + \varepsilon_{i,t} \tag{3.19}$$

The simulations are performed when the number of cross-section systems (N) is equal to 5, 10 and 5 for panels I, II and III, respectively. The length of time-series (T) is equal to 112. This is the same combination of N and T as that of Chapter 2. We consider the case of bivariate and trivariate time-series, which may be cointegrated with, at most, a single cointegration relation ($r_i = r = 1$). Therefore, the number of variables in system (k) is equal to 2 and 3, such that $Y_{i,t} = \{y_{i,t}, x_{1,i,t}\}$ or $Y_{i,t} = \{y_{i,t}, x_{1,i,t}, x_{2,i,t}\}$. The cointegrating vectors are assumed to be equal to (1, -1) in the bivariate case (k=2), and (1, -1, 1) in the trivariate case (k=3). The constant term ($\beta_{0,i}$) is restricted to the cointegrated vector, where $\beta_{0,i}$ is generated as a uniform random number ($\beta_{0,i} \sim U[0,10]$) and fixed in simulations. We assume that $x_{1,i,t}, x_{2,i,t}$ are strongly exogenous and are generated as random walks. When the error correction terms (α_i) equal the zero matrix, $Y_{i,t}$ are not cointegrated (r = 0). The cointegration relation exists when r = 1 and α_i vector is equal $(\phi_i,0)'$ and $(\phi_i,0,0)'$ for the bivariate and trivariate systems, respectively. In this chapter, ϕ_i is equal to 0 and –0.1 for size and power experiments, respectively. The parameter values in the DGP are chosen, in order to represent the theoretical coefficients from the purchasing power parity (PPP) hypothesis, which states that the exchange rate should bear a constant proportionate relationship to the relative price level between domestic and foreign countries, expressed as:

$$S_{i,t} = c(\frac{P_{i,t}}{P^*_{i,t}}) \tag{3.20}$$

where $S_{i,t}$ is the nominal exchange rate, $P_{i,t}$ is domestic price level, $P^*_{i,t}$ is foreign price level and c is a constant parameter. Taking logarithms of equation (3.20) gives:

$$s_{i,t} = \gamma + p_{i,t} - p^*_{i,t} \tag{3.21}$$

where $s_{i,t}, p_{i,t}, p^*_{i,t}$ are the logarithms of $S_{i,t}, P_{i,t}, P^*_{i,t}$.

Under the relative PPP hypothesis, $(s_{i,t}, p_{i,t}, p^*_{i,t})$ have a cointegration relationship with (1, -1, 1) cointegrating vector. The equation (3.21) can be rewritten in terms of relative price $(pr_{i,t})$ as:

$$s_{i,t} = \gamma + pr_{i,t} \tag{3.22}$$

where $pr_{i,t} = p_{i,t} - p^*_{i,t}$.

Therefore, in the bivariate system, the relative PPP hypothesis implies that $(s_{i,t}, pr_{i,t})$ have a cointegration relationship with (1, -1) cointegrating vector.

Moreover, the cointegrating vector (1, -1) is also found in other economic relationships, e.g. the theory of term structure and the test for market efficiency.

The error correction term (ϕ_i), set at –0.1, represents a mean reversion process with the slow speed of adjustment towards a long-run equilibrium. This value of the autoregressive coefficient corresponds to approximately 6.5 quarters (1.5 years) of half-life, which was used in Chapter 2. Therefore, the DGP is given by:

<u>$k = 2$ (bivariate system)</u>

$$\Delta y_{i,t} = \phi_i (y_{i,t-1} - x_{1,i,t-1} - \beta_{0,i}) + \varepsilon_{y,i,t} \tag{3.23}$$

$$\Delta x_{1,i,t} = \varepsilon_{x_1,i,t} \tag{3.24}$$

where $\varepsilon_{x_1,i,t} \sim N_p(0, \Omega_{x_1,i})$, $\varepsilon_{y,i,t} \sim N_p(0, \Omega_{y,i})$.

<u>$k = 3$ (trivariate system)</u>

$$\Delta y_{i,t} = \phi_i (y_{i,t-1} - x_{1,i,t-1} + x_{2,i,t-1} - \beta_{0,i}) + \varepsilon_{y,i,t} \tag{3.25}$$

$$\Delta x_{1,i,t} = \varepsilon_{x_1,i,t} \tag{3.26}$$

$$\Delta x_{2,i,t} = \varepsilon_{x_2,i,t} \tag{3.27}$$

where $\varepsilon_{x_1,i,t} \sim N_p(0, \Omega_{x_1,i})$, $\varepsilon_{x_2,i,t} \sim N_p(0, \Omega_{x_2,i})$, $\varepsilon_{y,i,t} \sim N_p(0, \Omega_{y,i})$.

This DGP implies that these systems have a single long-run relationship (r = 1). The short-run relationship does not exist in this DGP. Isard (1995) mentions that various economic forces may cause large and prolonged fluctuations in real exchange rate over time. Such arguments imply that the PPP hypothesis is not valid in the short-run, but should not necessarily be rejected over the long-run. Therefore, we set the DGP to characterise this argument.

In Chapter 2, we demonstrated the effect of cross-correlated errors on the size of the panel unit root tests. In light of this, the presence of cross-sectional dependence should affect the performance of panel cointegration tests. To investigate this issue, we perform simulations with cross-correlated errors in the DGP. Therefore, the error terms ($\varepsilon_{x_1,i,t}, \varepsilon_{x_2,i,t}, \varepsilon_{y,i,t}$) are generated as follows.

Case 1: no cross-sectional correlation in the error terms:

$$E(\varepsilon_{x_1,i,t};\varepsilon_{x_1,j,t}) = \begin{cases} \sigma_{x_1,i} & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases} \tag{3.28}$$

$$E(\varepsilon_{x_2,i,t};\varepsilon_{x_2,j,t}) = \begin{cases} \sigma_{x_2,i} & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases} \tag{3.29}$$

$$E(\varepsilon_{y,i,t};\varepsilon_{y,j,t}) = \begin{cases} \sigma_{y,i} & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases} \tag{3.30}$$

$$E(\varepsilon_{x_1,i,t};\varepsilon_{x_2,j,t}) = 0, \forall i, j. \tag{3.31}$$

Case 2: The error terms are cross-correlated:

$$E(\varepsilon_{x_1,i,t};\varepsilon_{x_1,j,t}) = \begin{cases} \sigma_{x_1,i} & \text{for } i=j \\ \sigma_{x_1,i,j} & \text{for } i \neq j \end{cases} \tag{3.32}$$

$$E(\varepsilon_{x_2,i,t};\varepsilon_{x_2,j,t}) = \begin{cases} \sigma_{x_2,i} & \text{for } i=j \\ \sigma_{x_2,i,j} & \text{for } i \neq j \end{cases} \tag{3.33}$$

$$E(\varepsilon_{y,i,t};\varepsilon_{y,j,t}) = \begin{cases} \sigma_{y,i} & \text{for } i=j \\ \sigma_{y,i,j} & \text{for } i \neq j \end{cases} \tag{3.34}$$

$$E(\varepsilon_{x_1,i,t};\varepsilon_{x_2,j,t}) = 0, \forall i,j. \tag{3.35}$$

We consider two case of cross-sectional dependence.

(2.1) The moderate degree of cross-correlation, $\sigma_{x_1,i,j}$, $\sigma_{x_2,i,j}$ and $\sigma_{y,i,j}$ is set to equal 0.5.

(2.2) The high degree of cross-correlation, $\sigma_{x_1,i,j}$, $\sigma_{x_2,i,j}$ and $\sigma_{y,i,j}$ is set to equal 0.9.

$\sigma_{x_1,i}$, $\sigma_{x_2,i}$ and $\sigma_{y,i}$ are generated as uniform random number generators, i.e. $\sigma_{x_1,i}$, $\sigma_{x_2,i}, \sigma_{y,i}$ ~ $U[0.5,1.5]$, in both the bivariate and trivariate systems, and then fixed over each replication.

Therefore, cross-correlation matrices (Ω_i) in case 2 can be shown as:

$$\Omega_{x_1,i} = \Omega_{x_2,i} = \Omega_{y,i} = \begin{bmatrix} 1 & \varpi & \dots & \varpi \\ \varpi & 1 & \dots & \varpi \\ \vdots & \vdots & \ddots & \vdots \\ \varpi & \varpi & \dots & 1 \end{bmatrix}_{NxN} ;\varpi = 0.5,0.9 \tag{3.36}$$

The values of the cross-correlations used in this chapter represent moderate and high degrees of correlation when ϖ is 0.5 and 0.9, respectively. The equi-correlational error structure is assumed, in order to control the degree of correlation and compare the results with the different values of ϖ.

Let us compare the cross-correlated matrix in our study with the matrix reported in an empirical study. Groen (2000) applies the residual-based panel cointegration tests in the study of a monetary approach in exchange rate determination, and reports the cross-correlation matrix of the OLS residuals of the ADF regression of the estimated residuals from the cointegration regressions of a monetary exchange rate model. This matrix is presented in Table 3.1

Table 3.1 The estimated cross-correlation matrix of the OLS residuals from a monetary exchange rate model reported in Groen (2000)

	Austr	Aut	Can	Fin	Fr	Ger	Ita	Jap	Nl	Nor	Sp	Swe	Switz	UK
Austr	1.00	0.17	0.21	0.19	0.24	0.23	0.19	0.31	0.11	0.32	0.19	0.16	0.17	0.30
Aut		1.00	0.04	0.58	0.86	0.91	0.71	0.65	0.91	0.68	0.63	0.59	0.81	0.56
Can			1.00	0.06	0.05	0.04	0.06	0.18	0.03	0.20	0.03	0.10	0.00	0.16
Fin				1.00	0.61	0.62	0.61	0.39	0.54	0.59	0.62	0.63	0.51	0.61
Fr					1.00	0.86	0.80	0.51	0.83	0.60	0.66	0.57	0.80	0.61
Ger						1.00	0.70	0.57	0.89	0.69	0.60	0.55	0.79	0.56
Ita							1.00	0.43	0.70	0.54	0.64	0.57	0.59	0.67
Jap								1.00	0.56	0.46	0.44	0.30	0.62	0.47
Nl									1.00	0.71	0.63	0.57	0.78	0.56
Nor										1.00	0.62	0.51	0.56	0.59
Sp											1.00	0.56	0.49	0.58
Swe												1.00	0.53	0.52
Switz													1.00	0.54
UK														1.00

Note: Austr, Aut, Can, Fin, Fr, Ger, Ita, Jap, Ni, Nor, Sp, Swe, Switz and UK denote Australia, Austria, Canada, Finland, France, Germany, Italy, Japan, Netherlands, Spain, Sweden, Switzerland and United Kingdom, respectively.

In Table 3.1, the degree of cross-correlation is high between EU countries. This correlation is as high as 0.91 between Austria and Germany. The degree of crosss-correlation is lower between EU and non-EU countries, as correlation is equal to 0.47 between UK and Japan. The degree of cross-sectional dependence in cases 2.1 and 2.2 represents these numbers of cross-correlated errors.

In this section, the simulation results are based on 10,000 replications. The nominal size for the simulation results is set at a significance level of 0.05. The *p*-values of the ADF tests for stationarity of the estimated residuals from the long-run regressions are calculated, using the ADF *t*-distributions generated by Monte Carlo simulations for the corresponding ADF *t*-test statistics in the bivariate and trivariate systems. These simulations also provide the means and variances of the ADF *t*-statistics used in the construction of the IPS statistic in the residual-based panel cointegration test. These mean (variance) values are equal to –2.049 (0.691) and

–2.486 (0.659), when $k = 2$ and 3, respectively. The means and variances of trace statistics used in calculation of the standardised LR-bar statistic are extracted from the simulation results in Osterwald-Lenum (1992). These mean (variance) values are equal to 4.03 (7.07), 11.91 (18.94) and 23.84 (37.98), when $k - r = 1$, 2 and 3, respectively.

3.4.2 Simulation results

Table 3.2 presents the estimates of empirical size and power for the standard Engle and Granger (1987) two-step test (E-G), and the Johansen (1988) trace test. The results are computed, using a DGP with no cross-correlation in the error terms (case 1) for the standard E-G (1987) and Johansen (1988) tests procedure ($N = 1$). Serially-correlated errors are not included in any of the DGP used in this chapter (cases 1, 2.1 and 2.2). Therefore, the empirical results of the residual-based test are based on the DF test (no lags) of the residuals of long-run estimation. The results of trace statistics are also based on tests with no lags included in VECM models.

From Table 3.2, the reported results show that the empirical size of these standard tests is close to the nominal size of 5% in both the bivariate and trivariate systems. However, the standard two-step test has low power in rejecting the null hypothesis of no cointegration, with a power of 0.183 and 0.103 in the bivariate ($k = 2$) and trivariate cases ($k = 3$), respectively. The standard trace test has markedly more power than the standard two-step test to distinguish the cointegrated system from the null hypothesis in both systems, with a power of 0.516 and 0.571 for k=2 and 3, respectively. For the trace test, the power results of the trivariate system are similar to those of the bivariate system. By contrast, the residual-based test is less powerful when k increases.

Table 3.2 The empirical size and power of the Engle and Granger (1987) test and the Johansen (1988) trace test under the null hypothesis of no cointegration (N=1)

	Bivariate sytem (k = 2)		Trivariate system (k = 3)	
	Size	Power	Size	Power
2-step test	0.051	0.183	0.053	0.107
Trace test	0.054	0.516	0.057	0.571

Note: The results are based on the standard E-G two-step test and the Johansen (1988) trace test, when T = 112. The underlying data are generated by equation (3.23) – (3.27), with N=1. The error terms are generated from case A (equation (3.28) – (3.31)).

In order to compare the power of the standard tests with that of the panel tests to be presented later, under the same null and alternative hypotheses, in each replication, the standard E-G and Johansen (1988) tests are applied to each N system. The DGP with white noise errors (case A) is used to generate data. We set N=5. Each system is estimated and tested for the null hypothesis independently. The null hypothesis of no cointegration relationship for all cross-section units will be rejected if we can reject the null of no cointegration relationship for at least one out of N individual systems in each replication. These size, power and size-adjusted power results are reported in Table 3.3. The size and power results for the residual-based panel cointegration test of IPS and MW and the panel LLL rank test in panel I (N=5) are presented in Tables 3.4 and 3.5, respectively.

The simulated results from Table 3.3 show that, under the null of no cointegration for all N individuals of the panel, these standard tests are massively over-sized. Therefore, the size-adjusted power is computed. From Table 3.3, the size-adjusted power results of the standard trace test in the panel are 0.812 and 0.907 in bivariate and trivariate systems, respectively. These results imply that more information from the cross-section dimension increases the chance of rejecting the null hypothesis. However, the size-adjusted power of the standard two-step test in the panel yields little improvement over its individual counterpart reported in Table 3.2. The simulated size-adjusted power results in this case are 0.200 and 0.115 for k = 2,3, respectively.

Table 3.3 The empirical size, power and size-adjusted power of the standard Engle and Granger (1987) two-step test and the Johansen (1988) trace test when the same null and alternative hypotheses as those of the panel cointegration tests (N =5) are applied

	Bivariate system ($k = 2$)			Trivariate system ($k = 3$)		
	Size	Power	Size-adjusted Power	Size	Power	Size-adjusted Power
2-step test	0.224	0.620	0.200	0.236	0.427	0.115
Trace test	0.245	0.968	0.812	0.249	0.987	0.907

Note: The results are based on the standard E-G two-step test and the Johansen (1988) trace test (T = 112). The underlying data are generated by equation (3.23) – (3.27) with N=5. The error terms are generated from case A (equation (3.28) – (3.31)).

From Tables 3.4 and 3.5, the simulated results of the panel cointegration tests can be noted as follows: first, we consider the benchmark case, in which the errors are generated as white noises (case 1). For the residual-based panel cointegration tests of IPS and MW, the empirical size of these tests is reasonably close to the nominal size (0.05) in all cases. In the bivariate (trivariate) case, the size results are equal to 0.046 (0.048) and 0.050 (0.050) for the residual-based panel test of IPS and MW, respectively. However, the size results of the panel LLL rank test are equal to 0.082 and 0.080, for k=2 and 3, respectively. These results are slightly over-sized. Simulations in this section are based on 10,000 replications. This number of replications implies that the 95% confidence interval of a test at the 0.05 significant level is between 0.0456 to 0.0543. The empirical size results, reported in Tables 3.4 and 3.5, are similar to those of LLL for the panel LLL rank test, and to those of

McCoskey and Kao (1999) for the IPS panel two-step cointegration test. LLL perform Monte Carlo simulations on the different combination of T and N, and show that the standard trace test has better size than the panel LLL rank test, but that the empirical power of the panel rank test is markedly higher than that of the standard trace test. Moreover, in the panel with a large cross-sectional dimension (N), the panel LLL rank test will be over-sized if the time-series dimension (T) is small. A large T relative to N is required, to avoid the size distortion problem. For fixed N, the empirical size of the panel rank test approaches the nominal 5% level as T increases, and for fixed T, the size increases with increased N. When N = 5 and k = 3, The empirical size of the panel LLL rank test for the null of no cointegration ($H_0 : rank(\Pi_i) = 0$), reported in Table 3 of LLL, is 0.106, 0.081 and 0.075, when T = 50, 100 and 200, respectively.

Table 3.4 The empirical size and power of the panel cointegration tests in the bivariate system, when N=5

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.046	0.049	0.082	0.649	0.536	0.979
Case 2.1 (cross-correlated errors)	0.077	0.070	0.102	0.642	0.534	0.935
Case 2.2 (cross-correlated errors)	0.189	0.142	0.194	0.572	0.460	0.822

Note: The results are based on the residual-based panel cointegration tests of IPS and MW and the likelihood-based panel LLL rank test. The underlying data are generated by equation (3.23) – (3.27) with N=5. In case A, the error terms are generated from equation (3.28) – (3.31). In cases 2.1 and 2.2, the error terms are generated from equation (3.32) – (3.35). $\sigma_{x_1,i,j}$, $\sigma_{x_2,i,j}$ and $\sigma_{y,i,j}$ are equal to 0.5 and 0.9 for case 2.1 and 2.2, respectively.

Table 3.5 The empirical size and power of the panel cointegration tests in the trivariate system, when N=5

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.048	0.050	0.080	0.298	0.240	0.990
Case 2.1 (cross-correlated errors)	0.071	0.071	0.112	0.346	0.275	0.973
Case 2.2 (cross-correlated errors)	0.178	0.135	0.203	0.388	0.301	0.878

Note: see notes to Table 3.4.

Even though these results provide evidence that the empirical size of the panel rank tests approaches the nominal 5% level as T increases (for fixed N), the empirical size of the panel rank test is still slightly over-sized when T is relatively large (T = 200, T/N = 40). There are some differences between simulation design in our study and that of LLL. In this chapter, the Johansen trace test is used with a constant restricted to the cointegrating vector. LLL considers the trace test without a constant term. Nevertheless, our results are still similar to those of LLL.

McCoskey and Kao (1999) also perform Monte Carlo simulations to compare the size and power performance of several residual-based panel cointegration tests. The panel test of the average t-statistics from the ADF test (the IPS test) is also considered in their study. For fixed N, the empirical size of this t-bar statistic approaches the nominal 5% level as T increases and, for fixed T, the size increases with increased N. The empirical size of the panel two-step test, reported in Table 1 of McCoskey and Kao (1999), is 0.057, when T = 100, N = 15 and k = 2. Comparing the size results of the panel two-step test in McCosKey and Kao (1999) with those of the panel LLL rank test reported in LLL, the size of the panel LLL rank test approaches the nominal level more slowly than that of the panel two-step test. The empirical size of the residual-based panel test of McCoskey and Kao (1999) is closer to the nominal size than the likelihood-based test of LLL in the same combination of T and N. These results imply that the panel rank test has a slower rate of convergence to its asymptotic result than that of the residual-based panel tests. In terms of size, the residual-based panel cointegration tests perform better than the likelihood-based panel rank test.

Turning to power performance, the power results of the panel tests increase significantly over the individual counterparts (see Tables 3.2 and 3.3). The power results of the panel LLL rank tests are 0.979 and 0.990 for k = 2 and 3, respectively. When k=2 (k=3), the empirical power is equal to 0.649 (0.298) and 0.536 (0.240) for the panel two-step tests of IPS and MW, respectively. The panel LLL rank test is considerably more powerful than the residual-based tests of IPS and MW, especially in the trivariate system. The additional variable in the system affects the empirical power of the panel two-step tests, where the simulated power results of the trivariate case are more significantly reduced than those of the bivariate case. The additional variable in the cointegrated system has no impact on the power performance of the panel LLL rank test. The empirical power of the panel rank test of the bivariate and trivariate systems is quite similar. These findings are not inconsistent with those of the standard individual tests (N=1). The standard likelihood-based trace test is significantly more powerful than the standard residual-based two-step test. Comparing the two residual-based panel cointegration tests, the IPS panel cointegration test is more powerful than the MW panel cointegration test, although the difference in power is small.

Next, we consider the effect of cross-sectional dependence on the error terms. The underlying DGP of case 2 contains cross-sectionally correlated errors. The results from Table 3.4 show evidence of size distortions in both the panel two-step tests and panel rank test. The size distortions are not large when the degree of cross-sectional dependence is moderate (case 2.1, $\sigma_{x_1,i,j}$, $\sigma_{x_2,i,j}$, $\sigma_{y,i,j}$ = 0.5 in a DGP). In the bivariate case, the empirical size of the panel residual-based test with IPS and MW and the panel rank test is 0.077, 0.070 and 0.102, respectively. The additional variables in the cointegrated system do not affect these results. In the trivariate case, these results are equal to 0.071, 0.071 and 0.112, respectively. All panel cointegration tests are slightly over-sized in the presence of cross-correlation.

In case 2.2 ($\sigma_{x_i,i,j}$, $\sigma_{x_2,i,j}$, $\sigma_{y,i,j}$ = 0.9 in a DGP), the error terms are strongly cross-correlated. In general, all of the panel tests are severely over-sized. The degree of size distortion in all tests is considerably higher than that of case 2.1. For k=2, the simulated size results of the residual-based panel tests of IPS and MW and the panel LLL rank test are equal to 0.189, 0.142 and 0.194, respectively. The size results of the trivariate system are close to those of the bivariate system, as these results are equal to 0.178, 0.135 and 0.203, respectively. Comparing the three panel cointegration tests, the panel LLL rank test has the highest degree of size distortion, followed by the residual-based tests of IPS and MW, respectively. Turning to power performance in case 2, we find that the simulated power results of panel cointegration tests are lower than those of case 1. Even though these numbers are not significantly different from those of the corresponding tests in case 1, the over-sized property of the tests will make the size-adjusted power drop considerably. This issue will be discussed again in Section 3.6, when the bootstrap method is applied to correct the size distortion.

Next, we consider the results from panels II and III (N=10,15). The simulated size and power results of the panel cointegration tests of IPS and MW and LLL, when N=10, for k=2 and 3, are presented in Tables 3.6 and 3.7, respectively. When N=15, the results are given in Tables 3.8 and 3.9.

Table 3.6 The empirical size and power of the panel cointegration tests in the bivariate system, when N=10

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.050	0.048	0.084	0.936	0.816	1.000
Case 2.1 (cross-correlated errors)	0.122	0.099	0.148	0.821	0.709	0.992
Case 2.2 (cross-correlated errors)	0.268	0.200	0.255	0.695	0.565	0.931

Note: see notes to Table 3.4 with N=10.

Table 3.7 The empirical size and power of the panel cointegration tests in the trivariate system, when N=10

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.049	0.045	0.090	0.543	0.387	1.000
Case 2.1 (cross-correlated errors)	0.111	0.088	0.157	0.541	0.412	0.996
Case 2.2 (cross-correlated errors)	0.254	0.190	0.280	0.511	0.397	0.945

Note: see notes to Table 3.4 with N=10.

Table 3.8 The empirical size and power of the panel cointegration tests in the bivariate system, when N=15

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.048	0.045	0.087	0.989	0.935	1.000
Case 2.1 (cross-correlated errors)	0.143	0.114	0.164	0.885	0.786	0.997
Case 2.2 (cross-correlated errors)	0.304	0.226	0.296	0.747	0.617	0.955

Note: see notes to Table 3.4 with N=15.

Table 3.9 The empirical size and power of the panel cointegration tests in the trivariate system, when N=15

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.046	0.045	0.100	0.726	0.530	1.000
Case 2.1 (cross-correlated errors)	0.071	0.071	0.112	0.649	0.514	0.999
Case 2.2 (cross-correlated errors)	0.178	0.135	0.203	0.558	0.432	0.962

Note: see notes to Table 3.4 with N=15.

The results from Tables 3.6 to 3.9 show that the empirical size of the residual-based tests of IPS and MW is close to the nominal level (0.05) in case 1, but shows some size distortions (over-sized) in the presence of cross-sectional dependence (cases 2.1 and 2.2). The panel LLL rank test is over-sized in every case. The size distortions tend to get greater as N increases in every case. With regard to power performance, the residual-based tests of IPS and MW become more powerful as N increases. However, the improvement in power is small in the panel LLL rank test. The power of the LLL test has already been close to 1.000, when N=5. Therefore, the power is not significantly improved as N increases.

In summary, the panel LLL rank test is markedly more powerful than the residual-based panel cointegration tests of IPS and MW. However, the panel LLL rank test suffers from the problem of size bias when T is not large enough. Moreover, in the presence of cross-correlation, all panel cointegration tests are size-distorted, and are severely over-sized when the values of the cross-correlations are high (case 2.2). The power of the residual-based panel cointegration tests of IPS and MW increases with N, while the size of the tests is close to the nominal size in every case. However, the likelihood-based cointegration test of LLL performs better in the small panel because the size distortions are more severe for big N, and the power is near unity for N≥5.

3.5 Cointegration tests in a mixed panel of cointegrated and non-cointegrated relationships

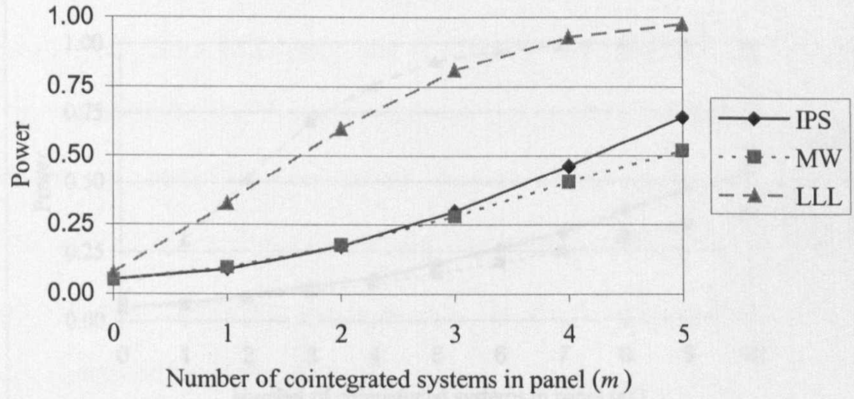
Both the residual-based panel cointegration tests of IPS and MW and the likelihood-based panel LLL rank tests are based on the assumption of heterogeneity. These panel cointegration tests allow for the possibility of a mixture of cointegrated and non-cointegrated groups in the panel. In this section, Monte Carlo experiments are performed on a mixed panel. The data are generated according to the DGP in case 1, outlined in Section 3.4, with the exception of our setting of $\phi_i = 0$ for some of the individual groups, and $\phi_i = -0.1$ for others. We consider the case where there are 1, 2, ..., N cointegrated systems (m) in panels of N=5, 10 and 15, T=112. The simulation results in this section are based on 10,000 replications. When N=5, the power results in the bivariate and trivariate systems are given in Table 3.10. The pictorial representation of the power results in panels I, II III are presented in Figures 3.1 to 3.6.

Table 3.10 The empirical power of panel cointegration tests in a mixed panel (N=5)

Number of	Bivariate case	($k = 2$)		Trivariate case	($k = 3$)	
Cointegrated groups (m)	IPS	MW	LLL	IPS	MW	LLL
1	0.089	0.096	0.328	0.080	0.080	0.511
2	0.173	0.177	0.598	0.110	0.106	0.783
3	0.299	0.281	0.811	0.155	0.145	0.925
4	0.461	0.405	0.930	0.215	0.182	0.979
5	0.649	0.536	0.979	0.298	0.240	0.994

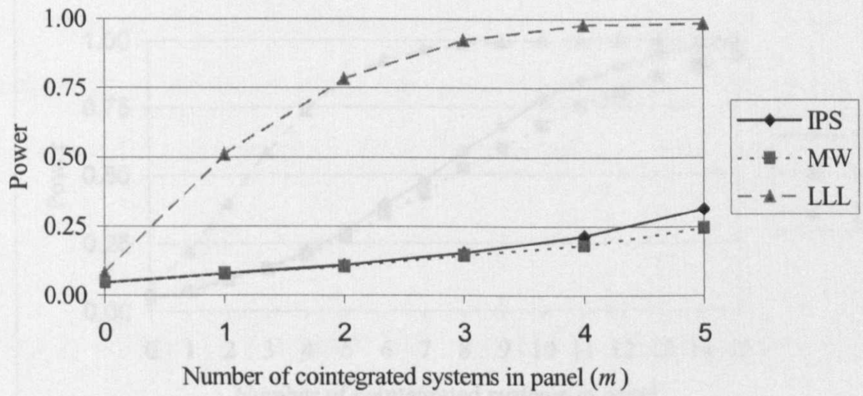
Note: The results are based on the residual-based panel cointegration tests of IPS and MW and the likelihood-based panel LLL rank test. The underlying data are generated by equation (3.23) – (3.27), with N=5. The error terms are generated from case A (equation (3.28) – (3.31)). The error correction term (ϕ_i) is set to be –0.1 for i =1,..,m and 0 for $i = m$+1,...,N).

Figure 3.1 The empirical power of panel cointegration tests in a mixed panel, when k=2, N=5



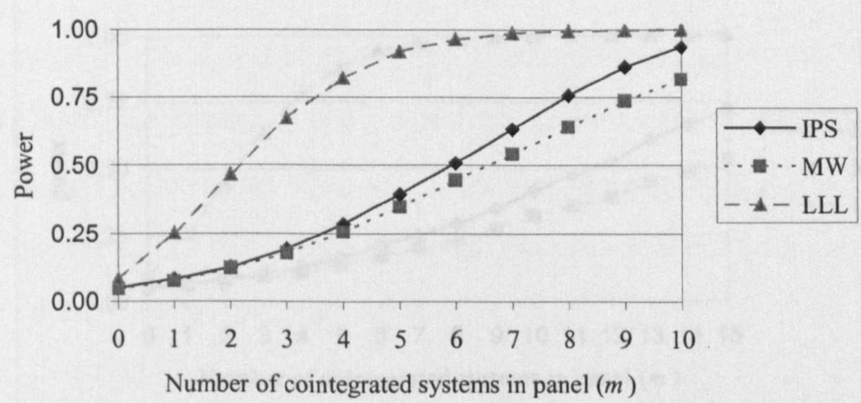
Note: see notes to Table 3.10 with N=5.

Figure 3.2 The empirical power of panel cointegration tests in a mixed panel, when k=3, N=5



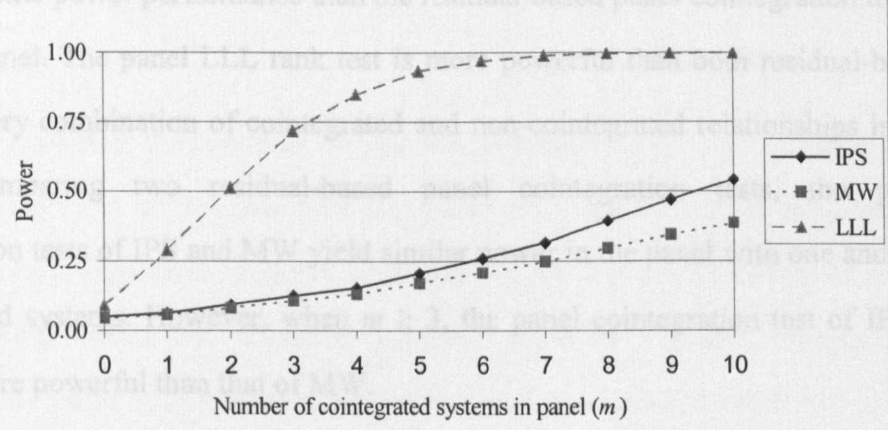
Note: see notes to Table 3.10 with N=5.

Figure 3.3 The empirical power of panel cointegration tests in a mixed panel, when k=2, N=10



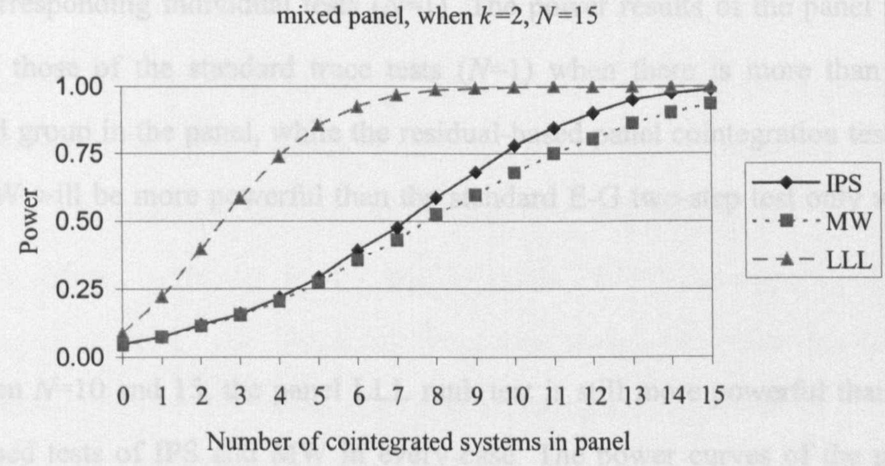
Note: see notes to Table 3.10 with N=10.

Figure 3.4 The empirical power of panel cointegration tests in a mixed panel, when k=3, N=10



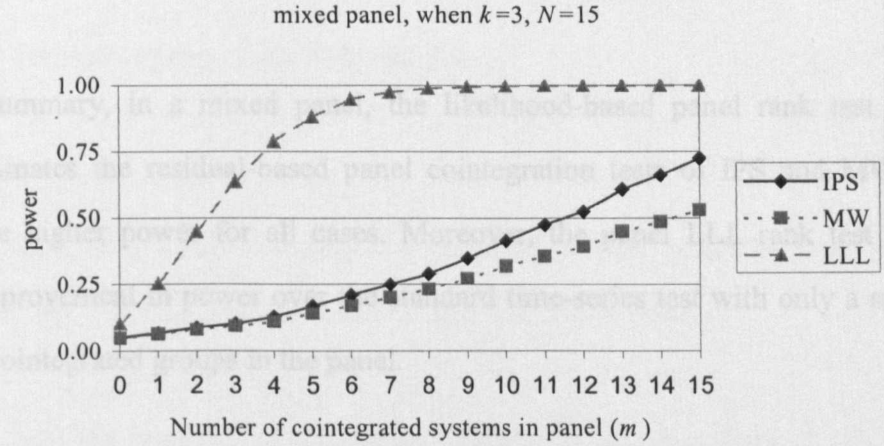
Note: see notes to Table 3.10 with N=10.

Figure 3.5 The empirical power of panel cointegration tests in a mixed panel, when k=2, N=15



Note: see notes to Table 3.10 with N=15.

Figure 3.6 The empirical power of panel cointegration tests in a mixed panel, when k=3, N=15



Note: see notes to Table 3.10 with N=15.

The simulated results in this section show that the panel LLL rank test still provides better power performance than the residual-based panel cointegration test in a mixed panel. The panel LLL rank test is more powerful than both residual-based tests in every combination of cointegrated and non-cointegrated relationships in the panel. Comparing two residual-based panel cointegration tests, the panel cointegration tests of IPS and MW yield similar power in the panel with one and two cointegrated systems. However, when $m \geq 3$, the panel cointegration test of IPS is slightly more powerful than that of MW.

Next, we compare the power results of these panel cointegration tests with those of corresponding individual tests (N=1). The power results of the panel rank test exceed those of the standard trace tests (N=1) when there is more than one cointegrated group in the panel, while the residual-based panel cointegration tests of IPS and MW will be more powerful than the standard E-G two-step test only when $m \geq 2$.

When N=10 and 15, the panel LLL rank test is still more powerful than the residual-based tests of IPS and MW in every case. The power curves of the panel cointegration tests of IPS and MW are similar to those of the panel IPS and MW unit root tests reported in Chapter 2 (see Figures 2.3, 2.5 and 2.6).

In summary, in a mixed panel, the likelihood-based panel rank test still clearly dominates the residual-based panel cointegration tests of IPS and MW in terms of the higher power for all cases. Moreover, the panel LLL rank test also provides improvement in power over the standard time-series test with only a small number of cointegrated groups in the panel.

3.6 Bootstrapping panel cointegration tests

In this section, the bootstrap method is introduced to panel cointegration tests. The main purpose in utilising the bootstrap panel cointegration tests is to attempt to correct the size distortions of panel cointegration tests in the presence of cross-sectional dependence. The bootstrap approach provides a feasible method for estimating the finite sample distribution of test statistics under the null hypothesis. Therefore, it provides the empirical distribution that can be used to calculate critical value when the error terms are cross-sectionally correlated.

However, the ordinary bootstrap method assumes that the underlying disturbances are independent across i cross-sections. This assumption is not appropriate in the presence of cross-correlation. For this reason, the stationary bootstrap guideline, explained in Li and Maddala (1996), is applied in our bootstrapping procedure. The stationary bootstrap resamples entire blocks of adjacent residuals, which preserves the cross-correlation structure among cross-sectional units.

3.6.1 Bootstrapping residual-based test

We apply the stationary bootstrap method to the residual-based panel cointegration test. The bootstrap procedure is presented as follows.

1. Apply the standard residual-based panel cointegration tests of IPS and MW, discussed in Section 3.3.1, to obtain the panel statistics.

2. Estimate the ADF regression for $\hat{e}_{i,t}$ under the null hypothesis that $y_{i,t}$ and $x'_{i,t}$ are not cointegrated ($\rho_i = 0$); that is:

$$\Delta\hat{e}_{i,t} = \sum_{q=1}^{r_i} v_{i,q}\Delta\hat{e}_{i,t-q} + \varepsilon_{i,t} \tag{3.37}$$

These estimated regressions provide the estimated parameters ($\tilde{v}_{i,1},\ldots,\tilde{v}_{i,r}$) and residuals ($\tilde{\varepsilon}_{i,t}$) under $H'_0(\rho_i = \rho = 0)$. The number of augmented terms ($\Delta\hat{e}_{i,t-q}$) included in equation (3.37) is the same as those of the regressions on the estimated residuals in step 1.

3. Generate bootstrap disturbances ($\varepsilon^*_{i,t}$) by resampling blocks of adjacent residuals ($\tilde{\varepsilon}_t = \{\tilde{\varepsilon}_{1,t},\ldots,\tilde{\varepsilon}_{N,t}\}$) structure, according to the stationary bootstrap method, to preserve the cross-correlation structure across the series in the panel. Next, the bootstrap sample of $e_{i,t}$ ($e^*_{i,t}$) is generated from $\varepsilon^*_{i,t}$, using the estimated coefficients ($\tilde{v}_{i,1},\ldots,\tilde{v}_{i,r}$) from equation (3.37).

$$\Delta e^*_{i,t} = \sum_{q=1}^{r_i} \tilde{v}_{i,q}\Delta e^*_{i,t-q} + \varepsilon^*_{i,t} \tag{3.38}$$

where $\{e^*_{i,1},\ldots,e^*_{i,1+r_i}\} = \{\hat{e}_{i,1},\ldots,\hat{e}_{i,1+r_i}\}$

4. The bootstrap sample of $y_{i,t}$ ($y^*_{i,t}$) is generated, using the estimated parameters from the cointegrating regressions in step 1.

$$y^*_{i,t} = \hat{\alpha}_i + \hat{\beta}_i x_{i,t} + e^*_{i,t} \tag{3.39}$$

5. Re-estimate the two-step residual-based panel cointegration tests with the bootstrap sample of $y_{i,t}$ ($y^*_{i,t}$).

$$y^{*}_{i,t} = \widetilde{\alpha}_i + \widetilde{\beta}_i x'_{i,t} + \widetilde{\eta}_{i,t} \qquad (3.40)$$

The estimated residuals are extracted. The ADF test is applied to the $\widetilde{\eta}_{i,t}$ to compute the bootstrap ADF t-statistics, which will be used to compute the bootstrap t-bar statistic.

6. Repeat the process numerous times to generate the bootstrap distribution of the residual-based tests of IPS and MW. The bootstrap critical value is then computed from the bootstrap distribution.

3.6.2 Bootstrapping likelihood-based test

The stationary bootstrap algorithm for testing the existence of the cointegration rank, using the panel trace statistics, can be explained as follows.

1. Obtain the individual trace test statistics $LR(r|k)$, using the Johansen (1988) procedure, for each individual system of equations in the panel, based on the assumption of a restricted intercept lying only in a cointegrating vector. The standardised LR-bar statistics is then calculated from the individual statistics.

2. Obtain the estimated parameters ($\hat{\alpha}_i, \hat{\beta}_i, \hat{\Gamma}_{i,1}, ..., \hat{\Gamma}_{i,p_i-1}$) and residuals ($\hat{\varepsilon}_{i,t}$) from the vector error correction model (VECM) under the null hypothesis $H_0 : rank(\Pi_i) = r_i \leq r$ by estimating the model:

$$\Delta Y_{i,t} = \alpha_i \beta'_i Y_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{i,j} \Delta Y_{i,t-j} + \varepsilon_{i,t} \qquad (3.41)$$

The estimated parameters $\hat{\alpha}_i, \hat{\beta}_i$ are $k \times r$ matrices and $\hat{\Gamma}_{i,1},\ldots,\hat{\Gamma}_{i,k-1}$ are $k \times k$ matrices. The adjusted residuals ($\tilde{\varepsilon}_{i,t}$), which are scaled and centred residuals, i.e. $\tilde{\varepsilon}_{i,t} = \sqrt{T/\{T-(p_i-1)k\}}(\hat{\varepsilon}_{i,t} - \hat{\mu}_{i,T})$, where $\hat{\mu}_{i,T} = T^{-1}\sum_{i=1}^{T}\hat{\varepsilon}_{i,t}$, are then obtained. $\tilde{\varepsilon}_{i,t}$ are used in the resampling process to get bootstrap disturbances ($\varepsilon^{*}_{i,t}$).

3. Generate bootstrap disturbances ($\varepsilon^{*}_{i,t}$) by resampling blocks of adjacent adjusted residuals ($\tilde{\varepsilon}_t$) according to the stationary bootstrap. $\tilde{\varepsilon}_t$ are defined as:

$$\tilde{\varepsilon}_t = \{\tilde{\varepsilon}'_{1,t},\ldots,\tilde{\varepsilon}'_{N,t}\} \tag{3.42}$$

where $\tilde{\varepsilon}_{i,t} = \{\tilde{\varepsilon}_{1,i,t},\ldots,\tilde{\varepsilon}_{k,i,t}\}$

Then, a bootstrap sample of $Y_{i,t}$ ($Y^{*}_{i,t}$) is generated from $\varepsilon^{*}_{i,t}$, using the estimated parameters from the VECM model in step 1, as:

$$\Delta Y^{*}_{i,t} = \hat{\alpha}\hat{\beta}' Y^{*}_{i,t-1} + \hat{\Gamma}_{i,1}\Delta Y^{*}_{i,t-1} + \ldots + \hat{\Gamma}_{i,p_i-1}\Delta Y^{*}_{i,,t-p_i+1} + \varepsilon^{*}_{i,t} \tag{3.43}$$

where $\{Y^{*}_{i,1},\ldots,Y^{*}_{i,p_i-1}\} = \{Y_{i,1},\ldots,Y_{i,p_i-1}\}$

4. Apply the Johansen (1988) trace test procedure to the bootstrap sample ($Y^{*}_{i,t}$), and then compute the bootstrap panel LLL rank test statistic.

5. Repeat the process numerous times to generate the bootstrap distribution of the panel LLL rank test statistic, which will be used to compute the bootstrap critical value.

3.6.3 Monte Carlo results

In this section, the finite sample size and power properties of the bootstrap residual-based and likelihood-based panel cointegration tests are investigated. Monte Carlo experiments are carried out with 500 iterations, each of which uses the bootstrap critical values computed from 200 bootstrap replications. The data are generated according to the DGP outlined in Section 3.4.1 in the panel where N=5 and T=112.

Tables 3.11 and 3.12 present the simulated size and power results of the bootstrap residual-based panel tests of IPS and MW and the bootstrap panel LLL rank test in the bivariate and trivariate systems, respectively. In case 1, the size results of these panel tests are close to the nominal level (0.05), with the sizes of 0.052, 0.058 and 0.054 for the residual-based panel tests of IPS and MW and the panel LLL rank test, respectively, in the bivariate case. Similar results are found in the trivariate system, with the sizes of 0.058, 0.042 and 0.044, respectively. In the simulation with 500 replications, the 95% confidence interval of the test at the 0.05 significantly level lies between 0.0309 and 0.0691. These simulated size results show that the bootstrap method can correct the size distortions of the panel LLL rank test in our panel size (N=5, T=112).

Table 3.11 The empirical size and power of the bootstrap panel cointegration tests in the bivariate system

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.052	0.058	0.054	0.720	0.564	0.964
Case 2.1 (cross-correlated errors)	0.060	0.052	0.064	0.600	0.494	0.878
Case 2.2 (cross-correlated errors)	0.076	0.070	0.040	0.314	0.286	0.536

Note: The results are based on the residual-based panel cointegration tests of IPS and MW and the likelihood-based panel LLL rank test. The underlying data are generated by equation (3.23) – (3.27) with N=5. In case A, the error terms are generated from equation (3.28) – (3.31). In cases 2.1 and 2.2, the error terms are generated from equation (3.32) – (3.35). $\sigma_{x_1,i,j}$, $\sigma_{x_2,i,j}$ and $\sigma_{y,i,j}$ are set to be 0.5 and 0.9 for case 2.1 and 2.2, respectively. Critical values are obtained from the bootstrap procedure.

Table 3.12 The empirical size and power of the bootstrap panel cointegration tests in the trivariate system

	Size			Power		
	IPS	MW	LLL	IPS	MW	LLL
Case 1 (white noise errors)	0.058	0.042	0.044	0.350	0.272	0.980
Case 2.1 (cross-correlated errors)	0.054	0.054	0.064	0.310	0.258	0.882
Case 2.2 (cross-correlated errors)	0.072	0.064	0.056	0.158	0.152	0.588

Note: see notes to Table 3.11.

In case 2, the empirical size of all bootstrap panel tests is closer to the nominal size (0.05) than that of the corresponding tests with the asymptotic critical values reported in Tables 3.4 and 3.5. In case 2.1, the size results are reasonably close to the nominal size (0.05) in all three tests. For k=2 (k=3), the size results are 0.060 (0.054), 0.052 (0.054) and 0.064 (0.064) for the bootstrap panel cointegration tests of IPS and MW and the panel LLL rank test, respectively. In case 2.2, the empirical size of the bootstrap panel LLL rank test also approximates the nominal level (0.05), as the size results are 0.056 and 0.040 when k = 2 and 3, respectively. However, the simulated size results are 0.076 (0.074) and 0.070 (0.064), for the panel cointegration tests of IPS and MW, respectively, when k = 2 (k = 3). These size results are closer to the nominal size (0.05) than those of the tests with asymptotic critical values (see Tables 3.4 and 3.5). However, in case 2.2, the empirical size lies just outside the 95% confidence interval. MW note that using the bootstrap method results reduces these size distortions, although it does not eliminate them entirely.

The results in terms of the power performance of the bootstrap panel cointegration tests show that the bootstrap panel LLL rank test is more powerful than the bootstrap residual-based tests of IPS and MW in all cases. Let us compare the power of the bootstrap panel cointegration tests with those of the corresponding tests with the asymptotic critical values reported in Tables 3.4 and 3.5. In case 1, the empirical power results of bootstrapping tests remain similar to those of the corresponding tests with the asymptotic critical values. However, in case 2, the empirical power of the bootstrap tests is lower than that of the asymptotic tests. In the bivariate system, the simulated power results of the bootstrap panel cointegration test of IPS, MW and the bootstrap panel LLL rank test are equal to 0.600 (0.314), 0.494 (0.286) and 0.878 (0.536), respectively, in case 2.1 (case 2.2). Moreover, in the trivariate system, the simulated power results of these bootstrap tests are equal to 0.310 (0.158), 0.258 (0.152) and 0.882 (0.588), respectively, in case 2.1 (case 2.2).

These results show that when the empirical critical values are available from the bootstrap method, the presence of cross-sectional dependence reduces the power of all panel tests, findings similar to those of Chapter 2. The arguments from Chapter 2 are still valid when testing for cointegration; the higher the values of the cross-correlations, the greater the reduction in the empirical power. The bootstrap method can correct the size distortions. However, the power to reject the null hypothesis of no cointegration is diminished, as the total amount of independent information contained in the panel is reduced in cross-correlated panels.

In summary, the bootstrap panel LLL rank test is superior to the bootstrap residual-based cointegration test of IPS and MW in terms of the size and power. The empirical size of all bootstrap panel cointegration tests is close to the nominal level of 0.05. The bootstrap panel LLL rank test is more powerful than the bootstrap panel residual-based tests of IPS and MW in every case. In addition, the empirical size of the bootstrap panel LLL rank test is clearly better than that of the test with the asymptotic critical values. The bootstrap method can correct the size distortions, which occur from either the presence of cross-correlated errors or the short span of the time-series (T).

3.7 Panel cointegration test with a factor model

In this section, we investigate the size and power performance of the residual-based panel cointegration test of CIPS, which applies the CIPS panel unit root test of Pesaran (2003), to test for unit roots of the estimated residuals in the panel two-step cointegration test. The data are generated according to a DGP, described in Section 3.4, for N = 5, 10 and 15. A total of 10,000 trials are used in computing the Monte Carlo results. The critical values are obtained from Monte Carlo simulations, similar to those of the CIPS panel unit root test in Chapter 2. The simulated size and power results are presented in Table 3.13.

Table 3.13 The empirical size and power of the residual-based panel cointegration test of CIPS in the bivariate and trivariate systems

	Panel size (N)	Bivariate system (k = 2)		Trivariate system (k = 3)	
		Size	Power	Size	Power
Case 1 (white noise errors)	N=5	0.054	0.484	0.050	0.238
	N=10	0.047	0.825	0.049	0.429
	N=15	0.052	0.947	0.050	0.625
Case 2-1 (cross-correlated errors)	N=5	0.049	0.426	0.048	0.182
	N=10	0.041	0.710	0.043	0.317
	N=15	0.045	0.839	0.043	0.442
Case 2-2 (cross-correlated errors)	N=5	0.048	0.385	0.047	0.160
	N=10	0.059	0.615	0.051	0.270
	N=15	0.064	0.714	0.063	0.359

Note: The results are based on the residual-based panel cointegration test of CIPS. The underlying data are generated by equation (3.23) – (3.27) with N=5, 10 and 15. In case A, the error terms are generated from equation (3.28) – (3.31). In cases 2.1 and 2.2, the error terms are generated from equation (3.32) – (3.35). $\sigma_{x_1,i,j}$, $\sigma_{x_2,i,j}$ and $\sigma_{y,i,j}$ are set to be 0.5 and 0.9 for cases 2.1 and 2.2, respectively. When N=5, the 5% critical values are equal to –2.871 and –3.245, for k=2 and 3, respectively. These critical values are –2.662 (-2.577) and –3.039 (-2.947), when N=10 (N=15).

The simulated results from Table 3.13 show that the empirical size of the residual-based panel cointegration test of CIPS is reasonably close to the nominal level (0.05) in every case. The application of the CIPS test can correct the size distortion problem in the presence of cross-sectional dependence. However, in cases 1 and 2.1, the simulated power results decrease from those of the panel residual-based test of IPS with both asymptotic and bootstrap critical values reported in Sections 3.4 and 3.6. For example, when N=5, the empirical power of the panel cointegration test of CIPS is 0.484 (0.426) and 0.238 (0.182), when k = 2 and 3, respectively, in case 1 (case 2.1). By contrast, in case 2.2, the residual-based panel test of CIPS is slightly more powerful than the bootstrap panel test of IPS when k=2. When N=5, these power results are equal to 0.385 and 0.310 for the panel test of CIPS and the bootstrap panel test of IPS, respectively. In the trivariate case, these two panel tests have similar power, as the simulated results are equal to 0.160 and 0.158, respectively. When N=15, the empirical power of the panel cointegration test of CIPS is equal to 0.947 (0.625), 0.839 (0.442) and 0.714 (0.359) in cases 1, 2.1 and 2.2, respectively for k=2 (k=3). These power results show that increasing the number of series in the panel (N) improves the power results of the test. The residual-based panel cointegration test of CIPS of the bivariate system (k=2) is more powerful than that of the trivariate system (k=3). In addition, the presence of cross-sectional dependence reduces the power of the test.

3.8 Conclusion

In this chapter, we investigated the finite sample performance of several panel cointegration tests in terms of the size and power. We compared the residual-based panel cointegration tests with the likelihood-based panel rank test in heterogeneous panels, using Monte Carlo simulations. The main conclusion from our experiments was that the likelihood-based panel LLL rank test outperforms the residual-based panel tests of IPS and MW in terms of higher power. Moreover, the panel LLL rank test also has the highest power in the panel with a mixture of cointegrated and non-cointegrated relationships. However, the panel LLL rank test is slightly over-sized if the time-series dimension is not large enough. In addition, all panel cointegration tests are over-sized in the presence of cross-sectional dependence in the data. The degree of size distortion is high in the strongly cross-correlated panel (case 2.2).

The bootstrap method was applied to correct the size distortion problem. The empirical size of the bootstrap panel LLL rank test and the residual-based panel cointegration test of IPS and MW are clearly better than those of the corresponding asymptotic tests, and reasonably close to the nominal level of 0.05. The bootstrap panel LLL rank test is still more powerful than both the bootstrap residual-based panel cointegration tests.

The residual-based panel cointegration test of CIPS was also applied to correct the size distortions in cross-correlated panels. The empirical power of the CIPS test is slightly higher than that of the bootstrap panel test of IPS only when k=2 and the errors are highly cross-correlated (case 2.2). In other cases, the panel cointegration test of CIPS does not provide any better power results than those of the bootstrap panel cointegration test of IPS.

Overall, we conclude that the panel LLL rank test is better than the residual-based panel cointegration tests of IPS and MW. However, we recommend applying the bootstrap method to correct the size distortions of the panel LLL rank test.

Chapter 4

Panel Unit Root Tests with Structural Breaks

4.1 Introduction

The presence of structural breaks in time-series data can induce behaviour similar to that of an integrated process, making it difficult to differentiate between a unit root and a stationary process with regime shift. Perron (1989) shows that the ADF test suffers from a loss of power when there is a shift in the intercept and/or slope of the trend function of a stationary time series. Recently, standard unit root tests have been modified to discriminate between structural break and unit root processes. Perron (1989, 1990) proposes a modified ADF test that allows for a structural shift by including a relevant dummy variable in the ADF test, assuming that the break point is exogenously given. Subsequent researchers have adopted an endogenous selection method to determine the break date (see, for example, Zivot and Andrews (1992), Banerjee *et al.*(1992) and Perron and Vogelsang (1992)). A widely used procedure selects the break point where the *t*-statistic for testing the null hypothesis of unit roots is minimised. Lumsdcaine and Papell (1997), and Clemente *et al.* (1998) extend these tests in the presence of multiple breaks.

Recently, testing for unit roots in panel data has attracted increasing attention. Heterogeneous panel unit root tests have been introduced by Im, Perasan and Shin

(2003) (IPS), Maddala and Wu (1999) (MW) and Choi (2001). However, panel unit root tests allowing for structural breaks have, to date, not widely been researched. Existing panel unit root tests, such as the IPS and MW tests, may potentially suffer from a significant loss of power under the presence of structural breaks in the data. However, Im, Lee and Tieslau (2002) (ILT) mention that constructing a valid panel unit root test allowing for structural shifts is complicated. The asymptotic property of the Perron-type t-statistic varies according to the location of break. Therefore, computing a modified IPS-type panel unit root test to include a dummy variable in each ADF regression to control for the effect of structural changes is practically unmanageable. The IPS procedure of standardising the t-bar statistic of the ADF test with structural breaks requires the expected values and variances of the ADF t-statistics at all different possible break points for each cross-section unit in the panel. In addition, the asymptotic validity of these test statistics under the null hypothesis is also affected by the incorrect placement of a structural break, by allowing for a break when there is no break, and by not allowing a break when there is one. Nunes, Newbold and Kuan (1997), and Lee and Strazicich (2001) demonstrate that the assumption of no break under the null hypothesis in the modified ADF-type tests when the DGP has a unit root with a break, causes the test statistic to diverge from its asymptotic property, leading to size distortions.

ILT propose a panel unit root test that can allow for structural shifts in level based on the Lagrangian Multiplier (LM) principle, and provide the relevant asymptotic results. An important feature of this panel LM unit root test is that its asymptotic distribution does not depend on the nuisance parameters that indicate the position of structural shifts. The panel unit root test of ILT uses the work of Schmidt and Phillips (1992) (SP), who propose a univariate LM unit root test whose asymptotic distribution is independent of the nuisance parameters of the deterministic components (intercept, trend). Amsler and Lee (1995) (AL) extend the LM unit root test to allow for a shift in level of a series, and show that the asymptotic distribution of the LM unit root test is invariant to the presence and location of a shift. ILT note that this invariance property of the univariate LM unit root test is useful in constructing heterogeneous panel unit root tests when either the number or the location of break is different in each cross-section unit. This important implication makes the panel LM unit root test very practical.

Lee and Strazicich (2003) (LS) propose a LM unit root test with an endogenous break selection procedure. The break dates are selected where the test statistic for the unit root null hypothesis is minimised. LS show that an asymptotic property of this LM unit root test is invariant to the location of shifts under the null hypothesis. This endogenous break LM unit root test provides greater flexibility in determining the location of breaks. For this reason, it is interesting to apply an endogenous break selection procedure to estimate the panel LM unit root test. The important points for the endogenous break test are the accuracy with which the break point is estimated and the way this affects the property of the panel unit root test. The performance of the tests using alternative break point selection criteria, such as the maximised values of the statistics for testing the significance of the shift dummy variable, is also of interest.

Even though the asymptotic property of the LM unit root test is invariant under the presence of level shifts, this invariance property does not hold in the presence of a change in trend slope. LS show that, in this case, the asymptotic distribution of the LM unit root test depends on the location of breaks. Therefore, in this chapter, we focus on the model with level shifts alone and do not consider the model with a change in trend slope or the model with a change in both level and trend.

The purpose of this chapter is to investigate, by means of a Monte Carlo simulation, the size and power performance of the panel LM unit root test. First, the performance of the panel LM unit root test without shifts will be investigated. The simulation results will be compared with those of the panel unit root tests of IPS and MW, reported in Chapter 2. Next, we consider the exogenous break panel LM unit root test of ILT. Finally, we examine the endogenous break panel LM unit root test in terms of the size, power and break point estimation.

The chapter is outlined as follows. The following section provides a review of literature pertaining to unit root testing allowing for structural breaks. The procedures of the panel LM unit root tests are presented in Section 4.3. Monte Carlo experiments are conducted in Section 4.4 to evaluate the size and power performance of the panel LM unit root test without shifts. Section 4.5 investigates the size and power properties of the exogenous break panel LM unit root test. Section 4.6 examines the performance of the endogenous break panel LM unit root test. Section 4.7 concludes this chapter.

4.2 Literature review

Since the influential paper of Perron (1989), the importance of allowing for a structural break in testing for unit roots is well recognised. Perron (1989, 1990) presents several unit root tests within a framework of structural shift when a break occurs at a known date. Perron (1989) considers three structural change models. The crash model (model A) allows for a one-time change in an intercept. The changing growth model (model B) allows for a break in a trend with the two segments joined at the break point, and the mixed model (model C) includes a one-time change in both level and trend. He develops a procedure for testing the unit root (with drift) null hypothesis, when an exogenous break occurs at time T_B ($1 < T_B < T$) against the alternative hypothesis that the series is stationary about a deterministic time trend.

Under the null hypothesis, models A, B and C can be described as follows:

Model A: $$y_t = \mu_0 + \delta D(T_B)_t + y_{t-1} + \upsilon_t \quad (4.1)$$

Model B: $$y_t = \mu_0 + (\mu_1 - \mu_0)DU_t + y_{t-1} + \upsilon_t \quad (4.2)$$

Model C: $$y_t = \mu_0 + \delta D(T_B)_t + (\mu_1 - \mu_0)DU_t + y_{t-1} + \upsilon_t \quad (4.3)$$

Under the trend-stationary alternative hypothesis, the models are expressed as follows:

Model A: $$y_t = \mu_0 + \alpha_1 t + (\mu_1 - \mu_0)DU_t + \upsilon_t \quad (4.4)$$

Model B: $$y_t = \mu_0 + \alpha_1 t + (\alpha_1 - \alpha_0)DT_t^* + \upsilon_t \quad (4.5)$$

Model C: $$y_t = \mu_0 + \alpha_1 t + (\alpha_1 - \alpha_0)DT_t + (\mu_1 - \mu_0)DU_t + \upsilon_t \quad (4.6)$$

where $DU_t = \begin{cases} 1 & \text{for } t \geq T_B + 1 \\ 0 & \text{otherwise} \end{cases}$; $D(T_B)_t = \begin{cases} 1 & \text{for } t = T_B + 1 \\ 0 & \text{otherwise} \end{cases}$

$$DT_t = \begin{cases} t - T_B & \text{for } t \geq T_B + 1 \\ 0 & \text{otherwise} \end{cases} ; \; DT_t^* = \begin{cases} t & \text{for } t \geq T_B + 1 \\ 0 & \text{otherwise} \end{cases}.$$

In addition, Perron (1990) considers a test that allows for a change in the mean of the series under both the null and alternative hypotheses, which complements the study of Perron (1989). This null hypothesis is stated as follows:

Model D: $y_t = \delta D(T_B)_t + y_{t-1} + \upsilon_t$ (4.7)

Under the alternative hypothesis, the model is given by:

Model D: $y_t = \mu_0 + \delta DU_t + \upsilon_t$ (4.8)

The proposed method in testing for unit roots with a structural shift is to specify a point of shift, and then to estimate a regression that nests the random walk null hypothesis and the alternative hypothesis of trend break stationarity.

Perron (1989, 1990) proposes two approaches in estimating the unit root tests with a structural break, in particular: the Additive-Outlier (AO) model and the Innovative-Outlier (IO) model.

The AO model assumes that a break occurs instantly and is not affected by the dynamics of the series. Testing for unit roots in the AO framework consists of a two-step procedure. The first step involves detrending the series using the following regressions:

Model A: $y_t = \mu + \alpha t + \gamma DU_t + \tilde{y}_t$ (4.9)

Model B: $y_t = \mu + \alpha t + \gamma DT_t^* + \tilde{y}_t$ (4.10)

Model C: $y_t = \mu + \alpha t + \gamma DT_t + \delta DU_t + \tilde{y}_t$ (4.11)

Model D: $y_t = \mu + \gamma DU_t + \tilde{y}_t$ (4.12)

In the second step, the unit root hypothesis is tested using the following regression:

$$\tilde{y}_t = \beta\tilde{y}_{t-1} + \sum_{j=1}^{p}\theta_j \Delta\tilde{y}_{t-j} + u_{i,t} \tag{4.13}$$

The IO model is useful to demonstrate a gradual shift occurring more slowly over time. The null hypothesis can be tested in the following ADF-type regressions:

Model A: $$y_t = \mu_0 + \alpha_1 t + \gamma DU_t + dD(T_B)_t + \beta y_{t-1} + \sum_{j=1}^{p}\theta_j \Delta y_{t-j} + u_t \tag{4.14}$$

Model B: $$y_t = \mu_0 + \alpha_1 t + \phi DT_t^* + \beta y_{t-1} + \sum_{j=1}^{p}\theta_j \Delta y_{t-j} + u_t \tag{4.15}$$

Model C: $$y_t = \mu_0 + \alpha_1 t + \gamma DU_t + dD(T_B)_t + \phi DT_t + \beta y_{t-1} + \sum_{j=1}^{p}\theta_j \Delta y_{t-j} + u_t \tag{4.16}$$

Model D: $$y_t = \mu + \gamma DU_t + dD(T_B)_t + \beta y_{t-1} + \sum_{j=1}^{p}\theta_j \Delta y_{t-j} + u_t \tag{4.17}$$

Perron (1989, 1990) derives the limit distributions of the t-statistics for testing unit roots for each model, which depend on the location and form of break under the alternative hypothesis. The key assumption of the Perron (1989, 1990) test is that the break date is fixed and chosen independently of the data. This assumption has been widely criticised in subsequent papers. The important argument is that the break date is often chosen after looking at the data, the choice of the break point should be correlated with the data, leaving room for data mining.

Subsequent papers have proposed a procedure to endogenise the choice of the break point and make it data dependent. Zivot and Andrews (1992) extend the test of Perron (1989) to allow for an endogenous selection procedure, and propose a minimum test, which determines the break point by utilising a grid search. To eliminate the end point, the model is estimated for each possible point of break over the time interval [.1T, .9T]. The break date is selected where the t-statistic for testing unit roots is minimised, which gives the least favourable result for the null hypothesis.

$$t_{\hat{\beta}}[\hat{T}_{B,\inf}] = \inf_{T_B \in \Lambda} t_{\hat{\beta}}(T_B) \tag{4.18}$$

where $\hat{T}_{B,\inf}$ denotes the minimised value.

Banerjee *et al.* (1992), and Perron and Vogelsang (1992) also propose unit root tests with several methods to estimate the break point. Banerjee *et al.* (1992) consider both mean and trend shift models:

$$y_t = \mu + \gamma\phi_{1,t} + \beta y_{t-1} + \upsilon_{i,t} \tag{4.19}$$

where $\phi_{1,t} = \begin{cases} t - T_B & \text{for } t > T_B \\ 0 & \text{otherwise} \end{cases}$, then equation (4.19) represents the trend shift model; when $\phi_{1,t} = \begin{cases} 1 & \text{for } t > T_B \\ 0 & \text{otherwise} \end{cases}$, the model is for a shift in mean.

The break point is also selected to minimise the ADF-type unit root test statistic. In addition, alternative criteria are also proposed, i.e. the maximised values of the Wald test for the significance of the break point and the Quandt likelihood ratio test for a break in any coefficients.

Perron and Vogelsang (1992) suggest a sequential procedure for testing the unit root null hypothesis when the point of break is unknown, based on the model D of Perron (1990). This break point selection method is also based on the minimum t-statistic procedure. The estimated model is the same as that of Perron (1990) for the IO model. However, in the AO model, the model can be presented as follows:

$$\tilde{y}_t = \sum_{l=1}^{p} \omega_l D(TB)_{t-l} + \beta\tilde{y}_{t-1} + \sum_{j=1}^{p} \theta_j \Delta\tilde{y}_{t-j} + u_t \tag{4.20}$$

The dummy variables, $D(TB)_{t-l}$, are included to ensure that the t-statistic on β converges to its asymptotic distribution.

Vogelsang and Perron (1998) propose an endogenous break unit root test for the three models (A, B and C) of Perron (1989), using the minimum t-statistic criterion. Moreover, they suggest an alternative procedure that selects the point of the break where the absolute value of the t-statistic for the shift dummy is maximised.

All of these studies consider the tests for those cases in which there is only one structural shift in the data. Later studies extend the tests to the case of multiple breaks. Clemente *et al.* (1998) extend the Perron and Vogelsang (1992) procedure to the case of two changes in mean, and derive the critical values for this test. The two-dimensional grid search for the two break points ($T_{B,1}$ and $T_{B,2}$) is used for either the AO or the IO model. In the IO model, the unit root null hypothesis is tested by estimating the following model:

$$y_t = \mu_0 + \gamma_1 DU_{1,t} + \gamma_2 DU_{2,t} + d_1 D(T_{B,1})_t + d_2 D(T_{B,2})_t + \beta y_{t-1} + \sum_{j=1}^{p} \theta_j \Delta y_{t-j} + u_t \quad (4.21)$$

where $DU_{1,t} = \begin{cases} 1 & \text{for } t \geq T_{B,1}+1 \\ 0 & \text{otherwise} \end{cases}$; $DU_{2,t} = \begin{cases} 1 & \text{for } t \geq T_{B,2}+1 \\ 0 & \text{otherwise} \end{cases}$;

$D(T_{B,1})_t = \begin{cases} 1 & \text{for } t = T_{B,1}+1 \\ 0 & \text{otherwise} \end{cases}$; $D(T_{B,2})_t = \begin{cases} 1 & \text{for } t = T_{B,2}+1 \\ 0 & \text{otherwise} \end{cases}$.

The AO model can be tested through a two-step procedure. First, the deterministic part of the variables is removed by estimating the following equation:

$$y_t = \mu + \gamma_1 DU_{1,t} + \gamma_2 DU_{2,t} + \tilde{y}_t \quad (4.22)$$

In the second step, the unit root hypothesis is tested, using the equation:

$$\tilde{y}_t = \beta \tilde{y}_{t-1} + \sum_{j=1}^{p} \omega_{1,j} D(T_{B,1})_{t-j} + \sum_{j=1}^{p} \omega_{2,j} D(T_{B,2})_{t-j} + \sum_{j=1}^{p} \theta_j \Delta \tilde{y}_{t-j} + u_t \quad (4.23)$$

The break points are selected by minimising the t-statistic on the autoregressive coefficient ($t_{\hat{\beta}}$) across all the possible break time combinations adjusted for end points. The restriction that $T_{B,2} > T_{B,1} + 1$ is also imposed, to eliminate those cases where the breaks occur in consecutive periods.

Lumsdaine and Papell (1997) also propose a unit root test that allows for the possibility of two endogenous break points, which basically extends the test of Banarjee *et al.* (1992) to take account of the multiple breaks.

While most of the unit root tests allowing for structural breaks are based on the ADF parameterisation, AL propose an alternative test based on the LM (score) principle. This test extends the LM unit root test of SP, to allow for a shift in mean. In this approach, all nuisance parameters of the process are estimated in the first step in such a way that the limit distributions of the subsequent unit root tests do not depend on these parameters. Therefore, the advantages of the LM unit root test are that the same critical value can be applied in the tests with different deterministic terms, and that the meaning of the parameters is the same under the null and alternative hypotheses. A one-time structural break in an intercept is allowed, based on the assumption that a level shift occurs at a known time. The LM unit root test allowing for a shift is estimated from the following regression using OLS:

$$\Delta y_t = \delta' \Delta Z_t + \beta \widetilde{S}_{t-1} + \sum_{j=1}^{p} \theta_j \Delta \widetilde{S}_{t-j} + u_t \qquad (4.24)$$

where $\widetilde{S}_t = y_t - Z_t \widetilde{\delta}'$; Z_t are the set of deterministic terms, e.g. in the crash model of Perron (1989), $Z_t = [1, t, DU_t]$; $\widetilde{\delta}'$ are the coefficients in the regression of Δy_t on ΔZ_t. The LM statistic (LM_T) is given by the standard t–statistic for testing $\beta = 0$.

An important feature of this LM unit root test is that its asymptotic distribution is not influenced by the presence of a level shift, and is invariant to the location of break in the data. Therefore, the asymptotic distribution of the LM unit root test is not affected by incorrect placement of a structural break or by allowing a break when it does not occur (or vice versa). This property provides an advantage over the ADF-type unit root test with shift.

The LM unit root test of AL is based on the exogenous selection of the break point. The number and location of breaks are taken as a priori. LS propose a univariate minimum LM unit root test, which adopts an endogenous break selection procedure to the LM unit root test (with shifts) of AL. The break point is selected to minimise the LM statistic (LM_T) for testing the unit root null hypothesis. Therefore, the LM test statistics of LS (LM_T^*) is given by:

$$LM_T^* = \inf_{T_B \in \Lambda}(LM_T(T_B)) \tag{4.25}$$

In the panel data framework, the modified IPS test, which includes a dummy variable in each ADF regression to allow for a break in mean, has a problem as its asymptotic property depends on the location of break in the data. For this reason, there were no significant developments in panel unit root tests with structural shifts until ILT proposed a panel LM unit root test with level shifts. The invariance property of the LM unit root test is useful when applying the test to the panel data framework. Due to this invariance property, the same adjustment values as those in the baseline case (without shifts) can be used to standardise the panel LM unit root test of ILT. The details of estimation procedures of the LM unit root test without shifts, with exogenous and endogenous shifts, and the calculation of the panel LM unit root test will be presented in the next section.

4.3 Panel LM unit root tests allowing for structural breaks

In this section, we present the procedure of the panel LM unit root test allowing for the presence of structural shifts in the data proposed by ILT. We demonstrate the case of a one-time shift in mean. However, ILT note that the asymptotic results can be applied to the tests with multiple level shifts.

Suppose a structural shift in mean occurs at time period $T_{B,i}$ in the i^{th} series. Then, the data generating process (DGP) is given as:

$$y_{i,t} = Z_{i,t} + x_{i,t} \tag{4.26}$$

where $Z_{i,t} = \gamma_{1,i} + \gamma_{2,i}t + \delta_i D_{i,t}$; $x_{i,t} = \phi_i x_{i,t-1} + \varepsilon_{i,t}$;

$$i = 1,\dots,N;\ t = 0,1,\dots,T\ ;\ \varepsilon_{i,t} \sim iid(0,\sigma_i^2);\ D_{i,t} = \begin{cases} 1 & \text{for } t \geq T_{B,i}+1 \\ 0 & \text{for } t < T_{B,i}+1 \end{cases}.$$

The null hypothesis of unit roots implies that $\phi_i = 1$ for all i. Rearranging equation (4.26) yields:

$$\Delta y_{i,t} = \beta_i y_{i,t-1} + \beta_i \gamma_{1,i} + [1-(\beta_i+1)(t-1)]\gamma_{2,i} + (\Delta D_{i,t} - \beta_i D_{i,t-1})\delta_i + \varepsilon_{i,t} \tag{4.27}$$

where $i = 1,\dots,N$; $t=1,\dots,T$; $\beta_i = -(1-\phi_i)$ and $\Delta D_{i,t} = D_{i,t} - D_{i,t-1}$, i.e.

$$\Delta D_{i,t} = \begin{cases} 1 & \text{for } t = T_{B,i}+1 \\ 0 & \text{otherwise} \end{cases}.$$

The panel LM unit root test is conducted by estimating separate LM unit root tests for each of the N individual series in the panel over T time periods, allowing for heterogeneity.

The null and alternative hypotheses for the panel LM unit root test are:

$H_0 : \beta_i = 0$ for all $i = 1,\ldots,N$ against

$H_a : \beta_i < 0$ for at least one i

We first consider the panel LM statistic without any structural shifts, which is a panel version of the univariate LM unit root test of SP. Next, is considered the panel LM unit root test, where a level shift occurs at a known date. This test is essentially a panel version of the LM unit root test with structural change proposed by AL. Finally, we present the panel LM unit root test with endogenous break selection procedures.

4.3.1 The panel LM unit root test without breaks

The panel LM unit root test statistic is derived from the results of the univariate LM unit root test. The LM-type test statistic is obtained by estimating the following regression:

$$\Delta y_{i,t} = \alpha_i + \beta_i \widetilde{S}_{i,t-1} + \sum_{j=1}^{p_i} \theta_{i,j} \Delta \widetilde{S}_{i,t-j} + u_{i,t} \tag{4.28}$$

According to the LM (score) principle, the residual variable ($\widetilde{S}_{i,t-1}$) is defined as follows:

$$\widetilde{S}_{i,t-1} = y_{i,t-1} - \widetilde{\gamma}_{2i}(t-1) \tag{4.29}$$

where $\widetilde{\gamma}_{2,i}$ is obtained as the OLS estimator of $\gamma_{2,i}$ in the restricted regression under the null hypothesis:

$$\Delta y_{i,t} = \gamma_{2,i} + \varepsilon_{i,t} \tag{4.30}$$

The augmented terms ($\sum_{j=1}^{p_i} \Delta\widetilde{S}_{i,t-j}$) are included to correct for serial-correlation. The LM statistic for the i^{th} series ($LM_{i,T}$) is given by the standard t-statistic for testing $\beta_i = 0$; that is:

$$LM_{i,T} = \frac{\hat{\beta}_i}{\hat{\sigma}_{\beta_i}} \tag{4.31}$$

where $\hat{\sigma}^2_{\beta_i}$ is the estimated variance of $\hat{\beta}_i$

Following ILT, the average of the individual LM test statistics is denoted as:

$$L\overline{M}_{NT} = \frac{1}{N}\sum_{i=1}^{N} LM_{i,T} \tag{4.32}$$

Then, the standardised panel LM unit root test statistic (Γ_{LM}) can be calculated as follows:

$$\Gamma_{LM} = \frac{\sqrt{N}[L\overline{M}_{NT} - E(L_T)]}{\sqrt{V(L_T)}} \tag{4.33}$$

where $E(L_T)$ and $V(L_T)$ denote the expected value and variance of $LM_{i,T}$ under the null hypothesis.

ILT show that this panel LM statistic follows a normal distribution: $\Gamma_{LM} \Rightarrow N(0,1)$, as N increase (for finite T).

The length of time-series (T) and number of augmentation terms (p_i) can be varied among each cross-section unit in the panel. Therefore, when T and p_i are not the same across cross-section units, Γ_{LM} is given by:

$$\Gamma_{LM} = \frac{\sqrt{N}[L\overline{M}_{NT} - \frac{1}{N}\sum_{i=1}^{N} E(L_T(p_i))]}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} V(L_T(p_i))}} \tag{4.34}$$

4.3.2 The exogenous break panel LM unit root test

In this section, we consider the panel LM unit root test where a level shift occurs in each individual time series. The LM statistic ($LM_{i,T}^{B}$) is derived from the results of the univariate LM unit root test of AL, obtained as the *t*-statistic for testing the null hypothesis that $\beta_i = 0$ in the following regression:

$$\Delta y_{i,t} = \alpha_i + \delta_i \Delta D_{i,t} + \beta_i \widetilde{S}_{i,t-1} + \sum_{j=1}^{p_i} \theta_{i,j} \Delta \widetilde{S}_{i,t-j} + u_{i,t} \tag{4.35}$$

where $\widetilde{S}_{i,t-1} = y_{i,t-1} - \widetilde{\gamma}_{2,i}(t-1) - \widetilde{\delta}_i D_{i,t-1}$ and $\widetilde{\gamma}_{2,i}$ and $\widetilde{\delta}_i$ are obtained as the OLS estimators of $\gamma_{2,i}$ and δ_i in the restricted regression:

$$\Delta y_{i,t} = \gamma_{2,i} + \delta_i \Delta D_{i,t} + \varepsilon_{i,t} \tag{4.36}$$

The average of the individual *LM* test (with a shift) statistic is denoted as:

$$\overline{LM}_{NT}^{B} = \frac{1}{N} \sum_{i=1}^{N} LM_{i,T}^{B} \tag{4.37}$$

ILT show that the limit distribution of $LM_{i,T}^{B}$ depends on *T* and p_i, but does not depend on the parameter indicating the location of shift point (λ_i; $\lambda_i = \frac{TB_i}{T}$). The difference between the LM statistic with a shift and that without one is asymptotically negligible. Therefore, the standardised panel LM unit root test (Γ_{LM}^{B}) can be expressed as follows:

$$\Gamma_{LM}^{B} = \frac{\sqrt{N}[\overline{LM}_{NT}^{B} - E(L_T)]}{\sqrt{V(L_T)}} \tag{4.38}$$

where $E(L_T)$ and $V(L_T)$ denote the expected value and the variance of $LM_{i,T}$ under the null hypothesis.

Again, ILT show that the panel LM statistic with a shift also follows a normal distribution: $\Gamma_{LM}^B \Rightarrow N(0,1)$, as $N,T \to \infty$.

When T and p_i differ across cross-section units, Γ_{LM}^B is given by:

$$\Gamma_{LM}^B = \frac{\sqrt{N}[\overline{LM}_{NT}^B - \frac{1}{N}\sum_{i=1}^N E(L_T(p_i))]}{\sqrt{\frac{1}{N}\sum_{i=1}^N V(L_T(p_i))}} \tag{4.39}$$

4.3.3 The endogenous break panel LM unit root test

The ILT panel LM unit root test with a level shift is based on the assumption that the number and the location of the breaks are accepted as a priori. However, this assumption is quite restrictive. In this section, we consider the panel LM unit root test when the location of shift is endogenously selected from the data. Three endogenous break selection methods are considered. The first method is the minimum t-statistic procedure suggested by LS. This approach involves selecting the break point to minimise the t-statistic for testing the null hypothesis of unit roots across all possible regressions. The chosen break date corresponds to the point that is most likely to reject the null hypothesis. The second method is to apply a procedure that maximises the statistic, testing for the significance of the shift dummy variable suggested by Vogelsang and Perron (1998). Finally, we apply the minimum Schwarz Bayesian Criterion (*SBC*) procedure used by Nunes, Newbold and Kuan (1997).

We first consider a panel version of the minimum LM unit root test (denoted as min-t_β test). LS propose a minimum LM unit root test, which is the LM unit root test of AL with an endogenous selection procedure to determine the number of augmentation terms (p_i) and the location of the break points (λ_i). This procedure is presented as follows.

First, the number of augmentation terms (p_i) is determined at each break point. The optimal value of p_i for each time-series is selected, using the general to specific procedure suggested by Ng and Perron (1995). Beginning with a maximum number of lagged terms ($k_{\max}$), if the last augmented term ($\Delta\widetilde{S}_{i,t-k_{\max}}$) is insignificantly different from zero at 10% significant level using the asymptotic critical values (± 1.645), the term is dropped from the regression. Then, the model is re-estimated using $k_{\max}$ - 1 lagged terms, and the significance of the last augmented term is tested. The process continues until the optimal number of lags is found or p_i=0.

Next, we use a grid search to determine the break at the location where the LM statistic ($LM_{i,T}^{B}$) is minimised. The LM statistics ($LM_{i,T}^{B}$) with the optimal number of lags are calculated at each possible break date over the time interval [.1T, .9T] (to eliminate end points). Therefore, the minimum LM unit root test statistics ($LM_{i,T}^{B_1*}$) are given by:

$$LM_{i,T}^{B_1*} = \inf_{\lambda_i \in \Lambda}(LM_{i,T}^{B}(\lambda_i)) \qquad (4.40)$$

The point of break (TB_i) chosen from this procedure is denoted as TB_i^1. LS show that the invariance property of the LM unit root test suggested by AL carries over to the endogenous break LM unit root test. Therefore, the asymptotic distribution of the endogenous break LM unit root test will not diverge in the presence of breaks under the null hypothesis and is robust against mis-specification.

The second approach considers the statistic used for testing the significance of the one or more break parameters as a criterion for selecting the break points (denoted as max-$|t_{\delta}|$ test). In the one-break test, the break point is chosen to maximise the absolute value of the t-statistic for the shift dummy variable in the LM unit root test ($|\hat{t}_{\delta_i}|$). In the two-break test, the location of break is obtained by maximising the F-statistic on the joint significance of the two dummy variables. The resulting break point from this procedure is denoted as TB_i^2.

The third approach utilises the Schwarz Bayesian Criterion (*SBC*) to determine the break dates (denoted as min-*SBC* test). Nunes, Newbold and Kuan (1997) have previously employed the *SBC*-based test, where the break date is selected to minimise the *SBC* statistic. The break point is chosen as argmin of *SBC* from the LM regressions, denoted as TB_i^3.

The standardised endogenous break panel LM unit root test statistic is obtained as follows:

$$\Gamma_{LM}^{B_k^*} = \frac{\sqrt{N}[\overline{LM}_{NT}^{B_k^*} - \frac{1}{N}\sum_{i=1}^{N} E(L_T^{B_k}(p_i))]}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} V(L_T^{B_k}(p_i))}} \qquad (4.41)$$

where $\overline{LM}_{NT}^{B_k^*} = \frac{1}{N}\sum_{i=1}^{N} LM_{i,T}^{B_k^*}$, $k = 1, 2$ and 3 denote each selection criterion for TB_i^1, TB_i^2 and TB_i^3.

4.4. Monte Carlo experiments on the panel LM unit root test without shifts

To investigate the performance of the panel LM unit root test in terms of the size and power, Monte Carlo experiments are carried out. Three experiments are conducted, based on the three different procedures described in Section 4.3. The first experiment investigates the panel LM unit root test in the benchmark case without structural breaks in the DGP. In the second experiment, we allow for structural shifts in the DGP. The exogenous break panel LM unit root test is considered. Finally, the third experiment investigates the case of the panel LM unit test with endogenous break selection procedures.

Simulations are performed when the number of cross-section series in the panel (N) is equal to 5 and 25, to represent the case of small and large panels, respectively. The length of time-series (T) is equal to 112, which is similar to that of Chapters 2 and 3. We generate additional 100 pre-sample values, which are then discarded. Simulations are performed using EVIEWS, version 4.1.

In this section, we consider the first experiment of the panel LM unit root test without shifts. The second and third experiments, which examine the exogenous and endogenous break panel LM unit root test, will be considered in Sections 4.5 and 4.6, respectively. The simulation results of the panel LM unit root test without shifts will be compared with those of the IPS and MW tests reported in Chapter 2. Therefore, we conduct experiments similar to those of Chapter 2, i.e. the effect of cross-correlation in the error terms. A mixed panel of stationary and non-stationary series is also considered.

In order to compare the size and power results of the panel LM unit root test with those of the IPS and MW tests, we apply the same DGP as that used in Chapter 2. Therefore, in this section, the DGP is given by:

$$\Delta y_{i,t} = \mu_i + \phi_i \mu_i t + \phi_i y_{i,t-1} + u_{i,t} \qquad (4.42)$$

We consider three cases of the DGP, similar to those of Chapter 2. Simulations with white noise errors (case A), serial-correlated errors (case B) and serial- and cross-correlated errors (case C) are conducted. The details of the DGP in each case were presented in Chapter 2, Section 2.4.1.

In the analysis of size, ϕ_i is set to be zero. In the investigation of power, ϕ_i is set to be –0.1. We consider the 0.05 significant level. The number of replications in the Monte Carlo simulation is equal to 10,000. Therefore, the 95% confidence interval lies between 0.0457 and 0.0543. The means and variances of the *LM* statistics used in the calculation of the panel LM test are extracted from Table 1 of ILT. The number of lags included in the LM regressions (p_i) to correct for serial-correlation is set to be 0, 1, 2 for the *LM*(0), *LM*(1) and *LM*(2) regressions, respectively.

4.4.1 The finite sample size and power

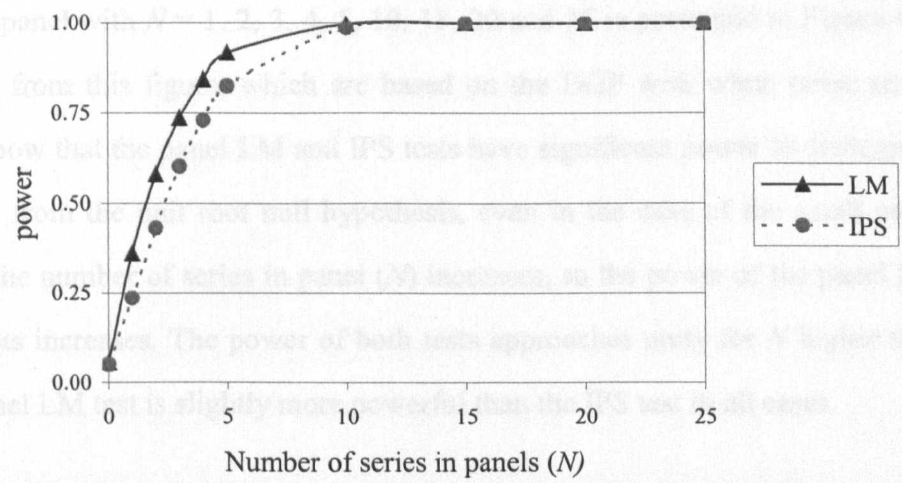
We first consider the size and power results of the panel LM unit root test. The empirical size and power results of the test in the small (*N*=5) and large panels (*N*=25) are reported in Table 4.1.

Table 4.1 The empirical size and power of the panel LM unit root test

	Number of lags	Small panel ($N = 5$)		Large panel ($N = 25$)	
		Size	Power	Size	Power
Case A	*LM(0)*	0.057	0.914	0.050	1.000
	LM(1)	0.056	0.861	0.053	1.000
	LM(2)	0.056	0.799	0.054	1.000
Case B	*LM(0)*	0.000	0.048	0.000	0.107
	LM(1)	0.059	0.837	0.057	1.000
	LM(2)	0.060	0.777	0.058	1.000
Case C	*LM(0)*	0.000	0.067	0.000	0.127
	LM(1)	0.075	0.814	0.074	1.000
	LM(2)	0.080	0.756	0.073	1.000

Note: The results are based on the panel LM unit toot test. The underlying data are generated by equation (4.42) with N=5 and 25. ϕ_i is set to be 0 and –0.1, in the analysis of size and power, respectively. In case A (white noise errors), the *LM*(0) regression represents the correctly chosen order of the LM regression, while *LM*(1) and *LM*(2) are over-fitting. In cases B and C (*AR*(1) errors), the *LM*(1) regression represents the correctly chosen order of the LM regression, while *LM*(1) and *LM*(2) are over-fitting and under-fitting, respectively.

Figure 4.1 The empirical power of the panel LM and IPS tests.



Note: The results are based on the panel LM and IPS tests (T=112). The underlying data are generated by equation (4.42) with N=1,2,3,4,5,10,15,20,25 and $\phi = -0.1$. The error terms are generated as white noises (case A). The results are based on the *LM*(0) and *ADF*(0) regressions.

The results show that, in cases A and B, when the optimal number of lags is correctly specified, the empirical size of the test is reasonably close to the nominal level (0.05). In the large panel (N=25), the size results are closer to the nominal size than those in the small panel (N=5), indicating that the size approaches the nominal level as N increase. In case C, the panel LM test is slightly size-distorted (over-sized), due to the cross-correlation, which is similar to the IPS and MW tests discussed in Chapter 2. This issue will be addressed when we consider the effect of cross-sectional dependence. Turning to the power performance, the results from Table 4.1 show that the empirical power of the panel LM test is slightly higher than that of the IPS and MW tests. For example, in case A, when N=5, the empirical power of the panel LM test with the $LM(0)$ regression is equal to 0.914, while the power of the IPS and MW tests reported in Chapter 2 is equal to 0.827 and 0.773, respectively. When N=25, these three tests have the same power, as their power results are equal to 1.000.

A pictorial representation of the empirical power of the panel LM and IPS tests in the panel with N = 1, 2, 3, 4, 5, 10, 15, 20 and 25 is presented in Figure 4.1. The results from this figure, which are based on the DGP with white noise errors (case A), show that the panel LM and IPS tests have significant power to distinguish the process from the unit root null hypothesis, even in the case of the small panel (N=5). As the number of series in panel (N) increases, so the power of the panel LM and IPS tests increases. The power of both tests approaches unity for N higher than 10. The panel LM test is slightly more powerful than the IPS test in all cases.

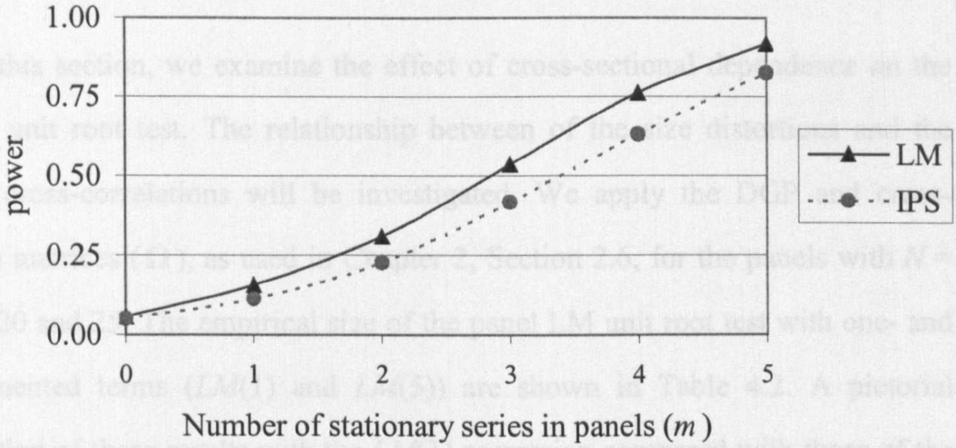
These results are consistent with the simulated results reported by SP, who compared the power of the LM unit root test with that of the standard ADF test for the time-series data. SP argue that the LM test is more powerful than the ADF test for length of time-series (T) and autoregressive coefficient (ϕ), such that the power is low, but less powerful than the ADF test for T and ϕ, such that the power is high. The difference between the LM and ADF tests is the way in which the deterministic terms (intercept, trend) are estimated. The LM and ADF tests estimate these parameters from regression, in first differences and in levels, respectively. SP note that estimation in differences is superior when the null hypothesis is true, or close to being true.

4.4.2 Simulations with a mixture of stationary and non-stationary series in the panel

In this section, we consider a mixed panel with stationary and non-stationary series. The simulated power results of the panel LM and IPS tests with N=5 and 25 are shown in Figures 4.2 and 4.3, respectively. In the case of a small panel (N=5), the panel LM test remains slightly more powerful than the IPS test in every case. The power of the panel LM and IPS tests exceeds 0.500 when there are more than three and four stationary series in the panel, respectively. In addition, the power of the panel LM test grows faster than that of the IPS test when the number of stationary series in panel (m) increases.

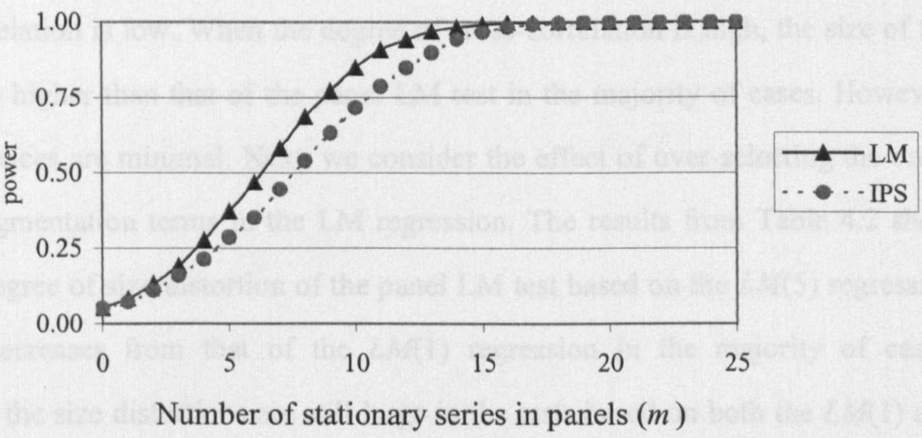
In the case of the large panel (N=25), the results still yield a pattern common to that of the small panel (N=5). The panel LM test still produces slightly better power performance than the IPS test in a mixed panel. The gap between the power results is higher than 0.100 when the number of stationary series (m) is between 6 to 11. This gap is narrowed when m>12, as the power of both tests approaches unity. In addition, the power of the tests is higher than 0.500 when there are more than seven and eight stationary series for the panel LM and IPS tests, respectively.

Figure 4.2 The empirical power of the panel LM and IPS tests in a mixed panel of stationary and non-stationary series in the case of small panel (N=5)



Note: The results are based on the panel LM and IPS tests (T=112). The underlying data are generated by equation (4.42) with N=5 and ϕ_i are set to be equal to -0.1 for i = 1…, m and 0 for i = m+1,…,N. The errors are generated as white noises (case A). The results are based on the LM(0) and ADF(0) regression for the panel LM and IPS tests, respectively.

Figure 4.3 The empirical power of the panel LM and IPS tests in a mixed panel of stationary and non-stationary series in the case of large panel (N=25)



Note: see notes to Figure 4.2, with N=25.

4.4.3 The effect of cross-sectional dependence

In this section, we examine the effect of cross-sectional dependence on the panel LM unit root test. The relationship between of the size distortions and the values of cross-correlations will be investigated. We apply the DGP and cross-correlation matrices (Ω), as used in Chapter 2, Section 2.6, for the panels with N = 5, 10, 15, 20 and 25. The empirical size of the panel LM unit root test with one- and five- augmented terms ($LM(1)$ and $LM(5)$) are shown in Table 4.2. A pictorial representation of these results with the $LM(1)$ regression compared with those of the IPS test is also presented in Figures 4.4 to 4.8.

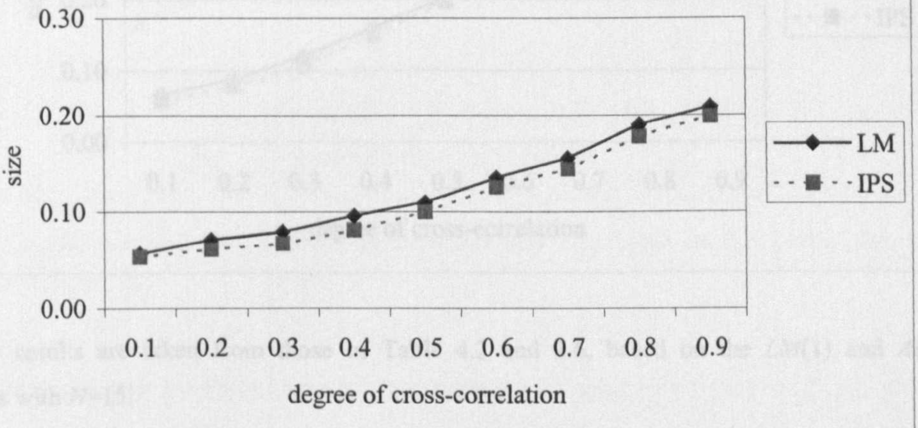
Figures 4.4 to 4.8 show that the degree of size distortion in the panel LM unit root test is similar to that of the IPS test reported in Chapter 2. The panel LM unit root test will be more severely size-distorted when either degree of cross-correlation (ω) or panel size (N) increases. Comparing the empirical size of the panel LM and IPS tests, the size of the former is higher than that of the latter when the degree of cross-correlation is low. When the degree of cross-correlation is high, the size of the IPS test is higher than that of the panel LM test in the majority of cases. However, the differences are minimal. Next, we consider the effect of over-selecting the order on the augmentation terms in the LM regression. The results from Table 4.2 show that the degree of size distortion of the panel LM test based on the $LM(5)$ regression slightly decreases from that of the $LM(1)$ regression in the majority of cases. However, the size distortions are still large in the tests based on both the $LM(1)$ and $LM(5)$ regressions. Over-fitting does not significantly affect the degree of size distortion of the panel LM test.

Table 4.2 The empirical size of panel LM test in the panel with cross-correlated errors estimated using the *LM*(1) and *LM*(5) regression

ω	*LM*(1)					*LM*(5)				
	$N=5$	$N=10$	$N=15$	$N=20$	$N=25$	$N=5$	$N=10$	$N=15$	$N=20$	$N=25$
0.1	0.057	0.065	0.066	0.068	0.069	0.054	0.061	0.063	0.063	0.067
0.2	0.072	0.075	0.088	0.097	0.102	0.056	0.065	0.075	0.085	0.093
0.3	0.080	0.102	0.120	0.136	0.153	0.061	0.086	0.097	0.112	0.123
0.4	0.096	0.125	0.161	0.182	0.207	0.075	0.106	0.125	0.144	0.168
0.5	0.110	0.166	0.202	0.232	0.258	0.084	0.128	0.161	0.180	0.205
0.6	0.136	0.188	0.236	0.270	0.288	0.105	0.155	0.187	0.209	0.223
0.7	0.156	0.227	0.269	0.301	0.317	0.116	0.177	0.213	0.231	0.256
0.8	0.190	0.254	0.280	0.318	0.331	0.138	0.192	0.228	0.245	0.269
0.9	0.209	0.266	0.323	0.313	0.342	0.150	0.208	0.234	0.251	0.274

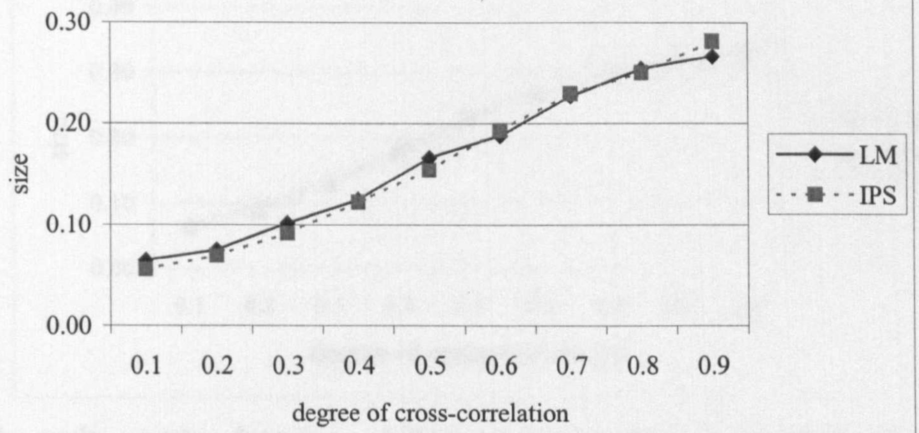
Note: The results are based on the panel LM unit root test (T=112). The underlying data are generated by equation (4.42) with N=5, 10, 15, 20 and 25. The error terms are generated according to case C. The cross-correlation (Ω) matrices are generated as equation (2.47) with ω = 0.1, 0.2,… ,0.9. The results are based on the *LM*(1) and *LM*(5) regressions.

Figure 4.4 The empirical size of the panel LM and IPS tests in cross-correlated panel, when N = 5



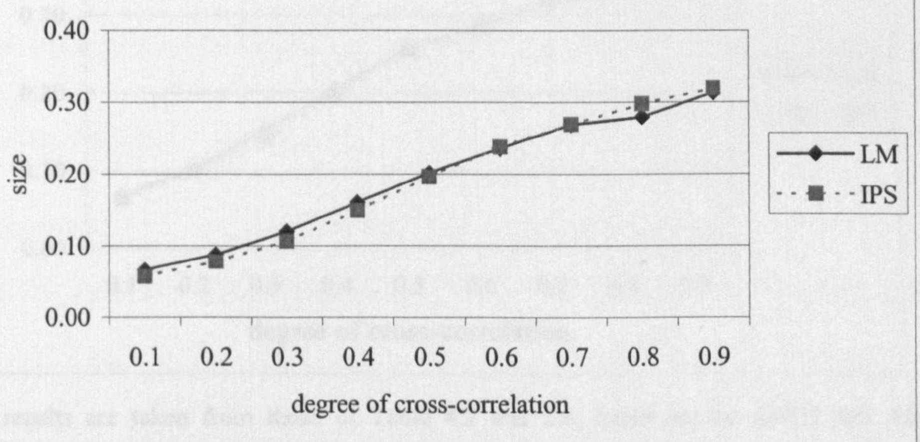
Note: The results are taken from those of Table 4.2 and 2.6, based on the *LM*(1) and *ADF*(1) regressions with N=5.

Figure 4.5 The empirical size of the panel LM and IPS tests in cross-correlated panel, when N = 10



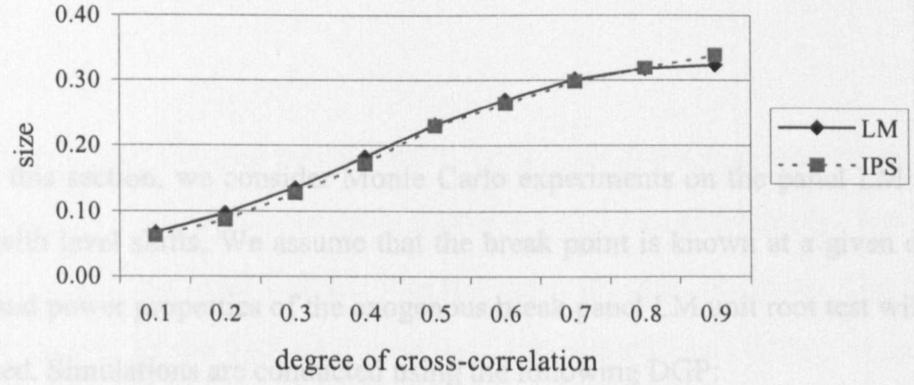
Note: The results are taken from those of Table 4.2 and 2.6, based on the $LM(1)$ and $ADF(1)$ regressions with N=10.

Figure 4.6 The empirical size of the panel LM and IPS tests in cross-correlated panel, when N = 15



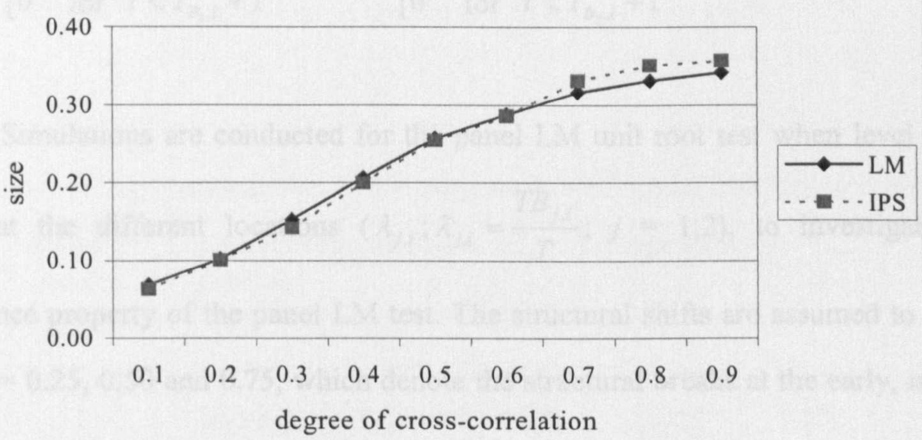
Note: The results are taken from those of Table 4.2 and 2.6, based on the $LM(1)$ and $ADF(1)$ regressions with N=15.

Figure 4.7 The empirical size of the panel LM and IPS tests in cross-correlated panel, when $N = 20$



Note: The results are taken from those of Table 4.2 and 2.6, based on the *LM*(1) and *ADF*(1) regressions with *N*=20.

Figure 4.8 The empirical size of the panel LM and IPS tests in cross-correlated panel, when $N = 25$



Note: The results are taken from those of Table 4.2 and 2.6, based on the *LM*(1) and *ADF*(1) regressions with *N*=25.

Overall, the performance of the panel LM unit root test without shifts is similar to that of the IPS and MW tests, although the panel LM test is slightly more powerful than the IPS and MW tests. The presence of cross-sectional dependence in the errors and a mixture of stationary and non-stationary series in the panel affect the panel LM, IPS and MW tests similarly.

4.5 Monte Carlo experiments on the exogenous break panel LM unit root test

In this section, we consider Monte Carlo experiments on the panel LM unit root test with level shifts. We assume that the break point is known at a given date. The size and power properties of the exogenous break panel LM unit root test will be investigated. Simulations are conducted using the following DGP:

$$y_{i,t} = Z_{i,t} + x_{i,t} \tag{4.43}$$

where $Z_{i,t} = \delta_{1,i} D_{1,i,t}$ in the case of a shift in level and $Z_{i,t} = \delta_{1,i} D_{1,i,t} + \delta_{2,i} D_{2,i,t}$ in the case of two level shifts; $x_{i,t} = \phi_i x_{i,t-1} + \varepsilon_{i,t}$; i = 1,...,N; t = 1,...,T; $\varepsilon_{i,t} \sim iid\ N(0,1)$;

$$D_{1,i,t} = \begin{cases} 1 & \text{for } t \geq T_{B_1,i} + 1 \\ 0 & \text{for } t < T_{B_1,i} + 1 \end{cases}, \quad D_{2,i,t} = \begin{cases} 1 & \text{for } t \geq T_{B_2,i} + 1 \\ 0 & \text{for } t < T_{B_2,i} + 1 \end{cases}.$$

Simulations are conducted for the panel LM unit root test when level shifts occur at the different locations ($\lambda_{j,i}$; $\lambda_{j,i} = \frac{TB_{j,i}}{T}$; j = 1,2), to investigate the invariance property of the panel LM test. The structural shifts are assumed to occur at $\lambda_{j,i}$ = 0.25, 0.50 and 0.75, which denote the structural breaks at the early, middle and late stage of the series, respectively. The magnitude of breaks ($\delta_{j,i}$) applied in our simulation is equal to 5 and 10, representing the moderate and large scale shifts in the data, respectively. We set ϕ_i=1 in computing size and ϕ_i= 0.9 in the case of power for all i. In this section, the simulated results are based on 10,000 replications.

4.5.1 Experiments on the exogenous break test when the break points are correctly specified

The empirical size and power results for the exogenous one- and two-break panel LM unit root test are displayed in Tables 4.3 and 4.4, respectively. We first consider the size of the panel LM unit root test under the null hypothesis. The 95% confidence interval for the 0.05 significant level test is 0.0457 and 0.0543 when the number of replications is equal to 10,000. The results from Table 4.3 show that the empirical size of the LM test is marginally over-sized in the small panel (N=5). However, the empirical size gets closer to the nominal level as N increases. When N=25, the size results are close to the nominal level of 5%. These results show that the size of the panel LM test with shifts still approaches the nominal level as N increases, which is similar to that of the panel LM test without shifts

Table 4.3 The empirical size of the exogenous break panel LM unit root test

Number of breaks	Location of breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.057	0.056	0.050	0.051
1	0.50	0.061	0.058	0.056	0.053
1	0.75	0.057	0.057	0.053	0.055
2	0.25, 0.50	0.059	0.064	0.056	0.054
2	0.25, 0.75	0.060	0.054	0.054	0.054
2	0.50, 0.75	0.056	0.059	0.053	0.054

Note: The results are based on the exogenous one- and two-break panel LM tests with the $LM(0)$ regression. The underlying data are generated by equation (4.43) with N=5 and 25. ϕ_i is set to be 1 and 0.9, in the analysis of size and power, respectively.

Table 4.4 The empirical power of the exogenous break panel LM unit root test

Number of breaks	Location of breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.900	0.905	1.000	1.000
1	0.50	0.903	0.903	1.000	1.000
1	0.75	0.900	0.906	1.000	1.000
2	0.25, 0.50	0.893	0.894	1.000	1.000
2	0.25, 0.75	0.893	0.894	1.000	1.000
2	0.50, 0.75	0.889	0.893	1.000	1.000

Note: see notes to Table 4.3.

The simulated power results of the panel LM unit root test controlling for level shifts approximate those of the corresponding results of the benchmark case without shifts, reported in Table 4.1. In addition, the size and power results are similar for any location and magnitude of breaks. For N=5, the empirical power is equal to 0.91, 0.90 and 0.89 for the panel LM tests without breaks, with one break ($\delta_{1,i}$=5) and two breaks ($\delta_{1,i}=\delta_{2,i}$=5), respectively. In the large panel (N=25), these power results are equal to 1.000 in every case. These results clearly demonstrate that the invariance property of the panel LM test is still valid in the finite sample (T=112). The presence of structural shifts does not affect the property of the test in any value of $\lambda_{j,i}$ and $\delta_{j,i}$, if the number and location of breaks are correctly specified.

4.5.2 Experiments on the exogenous break test when the number of breaks is over- and under-specified

First, we examine the exogenous break panel LM unit root test when the number of breaks in the data is over-specified. We consider the case in which the panel LM test is estimated on the assumption of one or two breaks when there is, in fact, no break in the DGP. In addition, we also consider the case of assuming two breaks when there is actually only one break in the DGP. The simulation results are presented in Tables 4.5 and 4.6.

The size and power figures of Tables 4.5 and 4.6 remain similar to the corresponding figures reported in Tables 4.1, 4.3 and 4.4. There is no significant evidence of size distortion or a loss of power in the panel LM test when it mistakenly assumes a shift, if, in fact, there is no shift in the DGP, or if more shifts are allowed for than actually occurred. The important implication of these results is that allowing for a break when it does not exist does not affect the size and power properties of the panel LM test in the finite sample (T=112).

Table 4.5 The empirical size and power of the exogenous one- and two-break panel unit root test when there is no break in the DGP

Specified number of break	Specified location of break	Small panel (N = 5)		Large panel (N = 25)	
		Size	Power	Size	Power
1	0.25	0.058	0.905	0.052	1.000
1	0.50	0.056	0.903	0.053	1.000
1	0.75	0.058	0.906	0.053	1.000
2	0.25, 0.50	0.061	0.895	0.055	1.000
2	0.25, 0.75	0.061	0.892	0.054	1.000
2	0.50, 0.75	0.057	0.887	0.050	1.000

Note: The results are based on the exogenous one- and two-breaks panel LM tests with the $LM(0)$ regression. The underlying data are generated by equation (4.43) with N=5 and 25, when δ_i =0.

Table 4.6 The empirical size and power of the exogenous two-break panel unit root test when there is one break in the DGP

Specified number of Break	Actual location of break	Specified location of break	Small panel (N = 5)		Large panel (N = 25)	
			Size	Power	Size	Power
2	0.50	0.25, 0.50	0.056	0.894	0.052	1.000
2	0.50	0.50, 0.75	0.058	0.898	0.052	1.000

Note: The results are based on the exogenous two-break panel LM unit root test with the $LM(0)$ regression. The underlying data are generated by equation (4.43) with N=5 and 25, when δ_i =5.

Next, we consider the panel LM unit root test when the number of breaks is under-specified. The finite sample properties of the panel LM unit root test (without allowing shifts in the series) when there are one or two breaks in the DGP will be investigated. In this case, the IPS test will be computed for a direct comparison. In addition, we examine the case of the panel LM test with a one-time shift in the series that suffers from two shifts. The size and power results are shown in Tables 4.7 to 4.10.

Table 4.7 The empirical size of the panel LM test (without shifts) and the IPS test (without shifts) when there are one or two breaks in the DGP

Actual Number of Break	Actual Location Of break	Small panel ($N = 5$)				Large panel ($N = 25$)			
		$\delta_{j,i} = 5$		$\delta_{j,i} = 10$		$\delta_{j,i} = 5$		$\delta_{j,i} = 10$	
		IPS	*LM*	*IPS*	*LM*	*IPS*	*LM*	*IPS*	*LM*
1	0.25	0.035	0.056	0.010	0.032	0.021	0.044	0.001	0.017
1	0.50	0.042	0.053	0.020	0.032	0.043	0.047	0.010	0.018
1	0.75	0.042	0.056	0.015	0.032	0.035	0.047	0.006	0.016
2	0.25, 0.50	0.046	0.077	0.019	0.052	0.035	0.089	0.005	0.056
2	0.25, 0.75	0.065	0.107	0.068	0.190	0.127	0.232	0.347	0.783
2	0.50, 0.75	0.067	0.069	0.073	0.049	0.160	0.094	0.395	0.057

Note: The results are based on the panel LM and IPS test without shifts with the *LM*(0) and *ADF*(0) regressions. See notes to Table 4.3 for details of the DGP.

Table 4.8 The empirical power of the panel LM test (without shifts) and the IPS test (without shifts) when there are one or two breaks in the DGP

Actual Number of Break	Actual Location Of break	Small panel ($N = 5$)				Large panel ($N = 25$)			
		$\delta_{j,i} = 5$		$\delta_{j,i} = 10$		$\delta_{j,i} = 5$		$\delta_{j,i} = 10$	
		IPS	*LM*	*IPS*	*LM*	*IPS*	*LM*	*IPS*	*LM*
1	0.25	0.352	0.641	0.005	0.095	0.949	0.999	0.005	0.473
1	0.50	0.538	0.665	0.091	0.148	0.998	1.000	0.595	0.608
1	0.75	0.416	0.639	0.019	0.094	0.985	0.999	0.076	0.469
2	0.25, 0.50	0.271	0.552	0.004	0.064	0.881	0.997	0.002	0.332
2	0.25, 0.75	0.699	0.890	0.437	0.857	1.000	1.000	1.000	1.000
2	0.50, 0.75	0.589	0.549	0.210	0.059	1.000	0.996	0.992	0.330

Note: See notes to Table 4.7.

The results from Tables 4.7 and 4.8 show that the ignoring of shifts in the data affects the size and power results in both the panel LM and IPS tests. Assuming too few breaks leads to incorrect size. When the DGP has one break, the size of the panel LM test (allowing no shifts) is only slightly under-sized when the magnitude of break is moderate ($\delta_{j,i}$= 5). However, the panel LM test is markedly under-sized when the magnitude of break is high ($\delta_{j,i}$= 10). The degree of size distortion increases when either the panel size (*N*) or magnitude of break ($\delta_{j,i}$) is large. Comparing the panel LM and IPS tests, the IPS test is more size-distorted than the panel LM test in the majority of cases. Moreover, the size results of the IPS test also vary according to the location of a break. The size distortions in the IPS test are most severe when a break occurs in the early stage of the series, followed by the late and middle stages of the series, respectively. However, this variation is not found in the panel LM unit root test, as the empirical size of the test is similar, regardless of the location of break, when there is one break in the DGP.

Table 4.9 The empirical size of the exogenous one-break panel LM unit root test when there are two breaks in the DGP

Actual number of Breaks	Actual location of break	Specified location of break	Small panel (*N* = 5)		Large panel (*N* = 25)	
			$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
2	0.25, 0.50	0.50	0.055	0.037	0.050	0.017
2	0.50, 0.75	0.50	0.055	0.031	0.050	0.019

Note: The results are based on the exogenous one-break panel LM test with the *LM*(0) regression. See notes to Table 4.3 for details of the DGP.

Table 4.10 The empirical power of the exogenous one-break panel LM unit root test when there are two breaks in the DGP

Actual number of Breaks	Actual location of break	Specified location of break	Small panel ($N = 5$)		Large panel ($N = 25$)	
			$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
2	0.25, 0.50	0.50	0.598	0.091	0.999	0.436
2	0.50, 0.75	0.50	0.618	0.088	0.999	0.443

Note: See notes to Table 4.9.

When there are two breaks in the DGP, we still observe the size distortion in both the panel LM and IPS tests ignoring the breaks, where the pattern of distortion varies across N, $\lambda_{j,i}$ and $\delta_{j,i}$ for both the LM and IPS tests. Surprisingly, the panel LM test is over-sized when the breaks occur at the early and late stages of the series ($\lambda_{j,i} = 0.25, 0.75$). In addition, when the empirical size of the tests is close to nominal, there is a loss of power when too few breaks are specified.

The simulated results from Tables 4.9 and 4.10 show that the size and power results of the exogenous one-break panel LM test, when two shifts actually occur, are similar to those of the panel LM test when ignoring one structural shift, reported in Tables 4.7 and 4.8.

Overall, in finite samples, the size distortions are serious when we ignore existing structural shifts in estimating the panel LM unit root test. Even though AP show that both the LM and ADF tests ignoring existing structural shifts are still valid under the null hypothesis, LS and ILT provide evidence that these results hold only asymptotically. For moderate sample sizes, the tests which ignore structural shifts may result in notable size distortions, depending on the values of $\lambda_{j,i}$, N and T. The size divergence is magnified when either N or $\delta_{j,i}$ increases. The size distortions in our simulations are more serious than those reported in LS, who investigate the performance of the exogenous two-break LM unit root test in the univariate framework. ILT note that any small size distortion in individual time-series accumulates in the panel data framework as N increases. In addition, our simulated results exhibit more serious size distortions than those of ILT, who consider only the case of one-break under the DGP. Another interesting aspect of our results is that the problem of over-sizing is found in some cases, a problem not reported by either LS or ILT. A possible explanation for this finding is the effect of spurious cross-sectional dependence in the errors, due to the mis-specification of the model by ignoring the break points. The simulated results in Section 4.4.3 show that the panel LM test is over-sized in the presence of cross-correlation in the errors. The degree of size-distortion in this section is close to that reported in Table 4.2 with cross-correlation similar values. For example, in the small panel (N=5), when the presence of breaks is ignored, the average degree of cross-correlation between the residuals of each LM regression in the panel is approximately equal to 0.26 (0.59) when $\lambda_{j,i}$ = 0.25,0.75 and $\delta_{j,i}$=5 ($\delta_{j,i}$=10). In this case, the empirical size of the panel LM test is equal to 0.107 (0.190). The problem of cross-correlation does not affect the individual time-series test and hence, is not found in LS. These results clearly demonstrate the importance of controlling for possible structural shifts. Comparing the panel LM and IPS tests, the performance of the panel LM test is clearly superior to that of the IPS test. The size distortions in the panel LM test are less severe than in the IPS test.

4.5.3 Experiments on the exogenous break test when the location of breaks is mis-specified

Finally, we examine the effect of incorrectly specifying the break date. When there is one mis-specified break point, the location of break ($\lambda_{j,i}$) of 0.25, 0.50 and

0.75 is used. For the two-break test, when both of them are mis-specified, we consider the panel when the location of break ($\lambda_{j,i}$) is specified at 0.20, 0.40, 0.60 and 0.80. The simulated results are presented in Tables 4.11 and 4.12.

Table 4.11 The empirical size of the panel LM test when the break points are incorrectly specified

Number of Breaks	Actual location of break	Specified location of Break	Small panel ($N = 5$)		Large panel ($N = 25$)	
			$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.50	0.057	0.052	0.047	0.017
1	0.25	0.75	0.058	0.034	0.051	0.020
1	0.50	0.25	0.054	0.035	0.043	0.020
1	0.50	0.75	0.058	0.031	0.050	0.019
1	0.75	0.25	0.049	0.033	0.048	0.020
1	0.75	0.50	0.053	0.031	0.051	0.017
2	0.25, 0.50	0.25, 0.75	0.058	0.030	0.049	0.019
2	0.25, 0.50	0.50, 0.75	0.056	0.034	0.047	0.016
2	0.25, 0.75	0.25, 0.50	0.056	0.033	0.042	0.016
2	0.25, 0.75	0.50, 0.75	0.055	0.031	0.049	0.017
2	0.50, 0.75	0.25, 0.50	0.054	0.030	0.051	0.019
2	0.50, 0.75	0.25, 0.75	0.056	0.034	0.048	0.018
2	0.20, 0.40	0.60, 0.80	0.062	0.005	0.056	0.007
2	0.20, 0.60	0.40, 0.80	0.108	0.169	0.211	0.635
2	0.20, 0.80	0.40, 0.60	0.095	0.156	0.196	0.594
2	0.40, 0.60	0.20, 0.80	0.058	0.018	0.049	0.005
2	0.40, 0.80	0.20, 0.60	0.098	0.160	0.211	0.632
2	0.60, 0.80	0.20, 0.40	0.063	0.026	0.056	0.010

Note: See notes to Table 4.3.

Table 4.12 The empirical power of the panel LM test when the break points are incorrectly specified

Number of Breaks	Actual location of break	Specified location of Break	Small panel ($N = 5$) $\delta_{j,i} = 5$	Small panel ($N = 5$) $\delta_{j,i} = 10$	Large panel ($N = 25$) $\delta_{j,i} = 5$	Large panel ($N = 25$) $\delta_{j,i} = 10$
1	0.25	0.50	0.619	0.094	0.999	0.442
1	0.25	0.75	0.632	0.101	0.999	0.458
1	0.50	0.25	0.655	0.149	0.999	0.585
1	0.50	0.75	0.652	0.150	0.999	0.576
1	0.75	0.25	0.630	0.099	0.999	0.455
1	0.75	0.50	0.622	0.091	1.000	0.442
2	0.25, 0.50	0.25, 0.75	0.636	0.134	0.999	0.546
2	0.25, 0.50	0.50, 0.75	0.603	0.092	0.999	0.426
2	0.25, 0.75	0.25, 0.50	0.609	0.088	0.999	0.424
2	0.25, 0.75	0.50, 0.75	0.623	0.088	0.999	0.424
2	0.50, 0.75	0.25, 0.50	0.609	0.091	0.999	0.422
2	0.50, 0.75	0.25, 0.75	0.643	0.133	0.999	0.540
2	0.20, 0.40	0.60, 0.80	0.374	0.005	0.954	0.008
2	0.20, 0.60	0.40, 0.80	0.831	0.677	1.000	1.000
2	0.20, 0.80	0.40, 0.60	0.815	0.650	1.000	1.000
2	0.40, 0.60	0.20, 0.80	0.407	0.015	0.955	0.013
2	0.40, 0.80	0.20, 0.60	0.824	0.697	1.000	1.000
2	0.60, 0.80	0.20, 0.40	0.364	0.001	0.957	0.006

Note: See notes to Table 4.3.

The size and power results of the panel LM unit root test which incorrectly specifies the points of break are similar to those of the test when the number of breaks are under-specified. When there is only a mis-specified break point, the size of the panel LM test can approximate the nominal level (0.05). However, the size results diverge from the nominal level when both $\delta_{j,i}$ and N are large. The downward size distortion is found when $\delta_{j,i}$ is large ($\delta_{j,i}$ =10). Therefore, a loss of power is observed in most cases. In addition, a loss of power is also found in the case in which the size of the test approaches the nominal level. Both the size distortion and power loss problems are magnified when there are two mis-specified break points, similar results to those of the test omitting two shifts. The upward size distortion is observed when the gap between each actual break is large ($\lambda_{j,i}$ = 0.20,0.60; 0.20,0.80 and 0.40,0.80). For example, when the breaks actually occur at $\lambda_{j,i}$ = 0.20,0.80, the size results of the panel LM unit root test are equal to 0.095, and 0.156 when the breaks are mis-specified at $\lambda_{j,i}$ = 0.40,0.60 and $\delta_{j,i}$ is equal to 5 and 10, respectively. In this case, the average degree of cross-correlation between the residuals of each LM regression in the panel is approximately equal to 0.29 and 0.66. These figures are close to those reported in Table 4.2 for the corresponding values of cross-correlations reported in Table 4.2.

Evidence of size distortion and power loss is also reported by LS, who consider the univariate two-break LM unit root test. However, our findings are far more significant than those of LS, especially when N is large. Again, any size distortions in individual time-series accumulate in the panel data framework as N increases. In addition, we observe the upward size distortion due to cross-sectional dependence in the errors.

In summary, the simulated results in this section show the importance of controlling for structural shifts in the data. The presence of structural breaks in the data does not affect the size and power performance of the panel LM unit root test when the number and location of breaks are correctly specified. Ignoring existing structural changes may lead either to size distortions or power loss. However, it is also important for the number and location of shifts to be correctly specified. Incorrect specification of break point locations does not improve the results over the panel test without shifts. These results show the importance of applying a method to endogenously select the points of break from the data. In the next section, the performance of the panel LM unit root test with endogenous break selection procedures will be examined.

4.6 Monte Carlo experiments on the endogenous break panel LM unit root test

In this section, the primary goal is to evaluate the performance of the endogenous break panel LM unit root test. Strazicich, Tieslau and Lee (2001) applied the min-t_β test procedure in testing for hysteresis in unemployment, using panel data, although the performance of this endogenous break LM unit root test has not been fully investigated in the panel data framework. We also compare the performance of the panel min-t_β test with that of the endogenous break panel LM unit root tests with different estimators of the break points, i.e. the max-$|t_\delta|$ test and the min-*SBC* test, in terms of the size, power and the accuracy of estimating the break date. This accuracy is considered by calculating the frequency of correctly estimating the true break point ($T_{B,i}$) in each test. The frequencies of estimating the break date at $T_{B,i} \pm 10$, $T_{B,i} \pm 20$ and $T_{B,i} \pm 30$ are also calculated. Lee and Strazicich (2001) note that estimators of the unit root *t*-test statistics become biased when the incorrect break point is used. Th is bias is maximised at $T_{B,i}$-1. Therefore, the modified ADF tests with endogenous selection procedures tend incorrectly to choose the break point at $T_{B,i}$-1. Consequently, we also report the frequency of incorrectly selecting the point of break at $T_{B,i}$-1, to investigate this concern. Simulations are undertaken, using 1,000 replications in the one-break test and 500 replications in the two-break test. The underlying data is generated by equation (4.43), which is similar to that of Section 4.5.

4.6.1 The finite sample means and variances

Before we investigate the size and power properties of the endogenous break panel LM unit root tests, their finite sample distributions should, first, be addressed. The means and variances of the endogenous break LM unit root tests when the points of break are determined by various procedures must be available before we compute the standardised *LM*-bar statistics. The finite sample distribution of the exogenous break LM unit root test may not be valid in the endogenous break test.

LS mention that the invariance property of the exogenous break LM unit root test is still carried over to the endogenous break min-t_β test. The asymptotic null distribution of the test is still invariant to the location of structural break. However, some recent studies raise concerns about the finite sample distribution of the unit root test with endogenous break selection procedures. First, LS show that critical values of the endogenous break LM min-t_β test differ from those of the exogenous break test. Therefore, the finite sample means and variances of the standard LM unit root test reported in Table 1 of ILT are no longer valid for computing a panel statistic of the min-t_β test. Second, Nunes, Newbold and Kuan (1997) point out that critical values of the endogenous break test depend on the method used for break date estimation. In view of this, the different mean and variance figures should be applied to the tests with different selection criteria. Finally, Lee and Strazicich (2001) note that, in general, the distribution of any endogenous break test depends on the accuracy with which the break point is estimated. This accuracy should depend on the magnitude of break, implying that the finite sample property of the endogenous break LM unit root test should not be invariant to the magnitude of break under the null hypothesis. For this reason, it is necessary to investigate the finite sample means and variances of the endogenous break LM unit root tests before we apply these tests in panel data. We use Monte Carlo simulations to calculate the finite sample means and variances of the endogenous break LM unit root test, using different methods in estimating the location of break under the null hypothesis with different values of magnitude ($\delta_{j,i}$) and location of break ($\lambda_{j,i}$). The results are reported in Tables 4.13 to 4.16, and 4.17 to 4.20 for the one- and two-break tests, respectively. These means and variances are derived, using 10,000 replications in the one-break test and 5,000 replications in the two-break test in sample T=112. In all cases, we report the means and variances of the tests with the $LM(0)$, $LM(1)$ and $LM(2)$ regressions

Table 4.13 Means of the endogenous one-break LM unit root test with different magnitudes of break (δ_i)

Size Of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\delta_i = 0$	-2.226	-2.236	-2.221	-1.983	-1.986	-1.967	-2.085	-2.085	-2.061
$\delta_i = 2$	-2.235	-2.241	-2.226	-1.982	-1.985	-1.965	-2.077	-2.078	-2.055
$\delta_i = 4$	-2.256	-2.264	-2.248	-1.974	-1.975	-1.956	-2.006	-2.007	-1.988
$\delta_i = 5$	-2.270	-2.277	-2.262	-1.974	-1.974	-1.955	-1.983	-1.983	-1.965
$\delta_i = 6$	-2.281	-2.289	-2.272	-1.969	-1.970	-1.950	-1.971	-1.971	-1.952
$\delta_i = 8$	-2.292	-2.299	-2.282	-1.969	-1.970	-1.950	-1.969	-1.970	-1.950
$\delta_i = 10$	-2.290	-2.295	-2.279	-1.969	-1.970	-1.950	-1.969	-1.970	-1.950

Note: The results are the means of the LM statistic of the endogenous one-break test. The underlying data are generated by equation (4.43), with N=1 and $\lambda_i = 0.50$.

Table 4.14 Means of the endogenous one-break LM unit root test with different locations of break (λ_i)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\lambda_i = 0.25$	-2.270	-2.277	-2.262	-1.974	-1.974	-1.955	-1.983	-1.983	-1.965
$\lambda_i = 0.50$	-2.273	-2.280	-2.264	-1.973	-1.973	-1.954	-1.982	-1.983	-1.964
$\lambda_i = 0.75$	-2.269	-2.276	-2.259	-1.974	-1.974	-1.954	-1.982	-1.983	-1.964

Note: The results are the means of the LM statistic of the endogenous one-break test. The underlying data are generated by equation (4.43), with N=1 and $\delta_i = 5$.

Table 4.15 Variances of the endogenous one-break LM unit root test with different magnitudes of break (δ_i)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\delta_i = 0$	0.471	0.470	0.461	0.358	0.362	0.356	0.462	0.465	0.455
$\delta_i = 2$	0.472	0.472	0.460	0.356	0.361	0.354	0.458	0.461	0.450
$\delta_i = 4$	0.474	0.472	0.459	0.348	0.352	0.346	0.391	0.394	0.388
$\delta_i = 5$	0.472	0.471	0.458	0.345	0.348	0.342	0.359	0.363	0.360
$\delta_i = 6$	0.465	0.462	0.449	0.342	0.346	0.341	0.346	0.349	0.344
$\delta_i = 8$	0.444	0.439	0.425	0.342	0.345	0.340	0.342	0.345	0.340
$\delta_i = 10$	0.414	0.408	0.394	0.342	0.345	0.340	0.342	0.345	0.340

Note: The results are the variances of the LM statistic of the endogenous one-break test. The underlying data are generated by equation (4.43), with N=1 and $\lambda_i = 0.50$.

Table 4.16 Variances of the endogenous one-break LM unit root test with different locations of break (λ_i)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\lambda_i = 0.25$	0.472	0.471	0.458	0.345	0.348	0.342	0.359	0.363	0.360
$\lambda_i = 0.50$	0.475	0.472	0.459	0.342	0.348	0.344	0.359	0.364	0.360
$\lambda_i = 0.75$	0.474	0.470	0.458	0.344	0.350	0.345	0.358	0.365	0.361

Note: The results are the variances of the LM statistic of the endogenous one-break test. The underlying data are generated by equation (4.43), with N=1 and $\delta_i = 5$.

Table 4.17 Means of the endogenous two-break LM unit root test with different magnitudes of break ($\delta_{j,i}$)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\delta_i = 0$	-2.483	-2.497	-2.484	-2.036	-2.037	-2.015	-2.213	-2.216	-2.185
$\delta_i = 2$	-2.510	-2.523	-2.510	-2.039	-2.037	-2.014	-2.202	-2.198	-2.170
$\delta_i = 4$	-2.584	-2.592	-2.576	-2.003	-2.000	-1.976	-2.055	-2.056	-2.037
$\delta_i = 5$	-2.572	-2.577	-2.564	-1.958	-1.951	-1.935	-1.972	-1.969	-1.951
$\delta_i = 6$	-2.638	-2.641	-2.622	-1.983	-1.979	-1.958	-1.985	-1.981	-1.960
$\delta_i = 8$	-2.650	-2.650	-2.631	-1.982	-1.978	-1.957	-1.982	-1.978	-1.957
$\delta_i = 10$	-2.631	-2.629	-2.611	-1.982	-1.978	-1.957	-1.982	-1.978	-1.957

Note: The results are the means of the LM statistic of the endogenous two-break test. The underlying data are generated by equation (4.43), with N=1 and $\lambda_{j,i} = 0.25, 0.50$.

Table 4.18 Means of the endogenous two-break LM unit root test with different locations of break ($\lambda_{j,i}$)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\lambda_{2,i} - \lambda_{1,i} = 0.25$									
$\lambda_{j,i} = 0.25, 0.50$	-2.572	-2.577	-2.564	-1.958	-1.951	-1.935	-1.972	-1.969	-1.951
$\lambda_{j,i} = 0.35, 0.60$	-2.580	-2.590	-2.569	-1.973	-1.973	-1.948	-1.988	-1.988	-1.967
$\lambda_{j,i} = 0.45, 0.70$	-2.577	-2.584	-2.567	-1.971	-1.970	-1.948	-1.982	-1.985	-1.963
$\lambda_{j,i} = 0.50, 0.75$	-2.578	-2.583	-2.567	-1.959	-1.952	-1.934	-1.974	-1.969	-1.951
$\lambda_{2,i} - \lambda_{1,i} = 0.50$									
$\lambda_{j,i} = 0.20, 0.70$	-2.641	-2.652	-2.632	-1.971	-1.970	-1.946	-1.985	-1.984	-1.961
$\lambda_{j,i} = 0.25, 0.75$	-2.633	-2.639	-2.622	-1.962	-1.956	-1.937	-1.975	-1.969	-1.950
$\lambda_{j,i} = 0.30, 0.80$	-2.638	-2.645	-2.625	-1.966	-1.965	-1.942	-1.977	-1.977	-1.956

Note: The results are the means of the LM statistic of the endogenous two-break test. The underlying data are generated by equation (4.43), with N=1 and $\delta_{j,i} = 5$.

Table 4.19 Variances of the endogenous two-break LM unit root test with different magnitudes of break ($\delta_{j,i}$)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\delta_i = 0$	0.608	0.611	0.587	0.412	0.411	0.399	0.613	0.622	0.598
$\delta_i = 2$	0.604	0.615	0.592	0.413	0.411	0.398	0.591	0.605	0.583
$\delta_i = 4$	0.612	0.618	0.598	0.378	0.378	0.367	0.457	0.472	0.468
$\delta_i = 5$	0.593	0.574	0.558	0.330	0.327	0.331	0.356	0.355	0.355
$\delta_i = 6$	0.585	0.581	0.563	0.352	0.357	0.349	0.357	0.362	0.354
$\delta_i = 8$	0.531	0.521	0.506	0.351	0.356	0.348	0.351	0.356	0.348
$\delta_i = 10$	0.469	0.457	0.442	0.351	0.356	0.348	0.351	0.356	0.348

Note: The results are the variances of the LM statistic of the endogenous two-break test. The underlying data are generated by equation (4.43), with N=1 and $\lambda_{j,i} = 0.25, 0.50$.

Table 4.20 Variances of the endogenous two-break LM unit root test with different locations of break ($\lambda_{j,i}$)

Size of break	Min-t_β test			Max-$\lvert t_\delta \rvert$ test			Min-*SBC* test		
	LM(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)	*LM*(0)	*LM*(1)	*LM*(2)
$\lambda_{2,i} - \lambda_{1,i} = 0.25$									
$\lambda_{j,i} = 0.25, 0.50$	0.593	0.574	0.558	0.330	0.327	0.331	0.356	0.355	0.355
$\lambda_{j,i} = 0.35, 0.60$	0.593	0.593	0.577	0.347	0.352	0.343	0.371	0.377	0.376
$\lambda_{j,i} = 0.45, 0.70$	0.585	0.585	0.570	0.345	0.351	0.346	0.367	0.374	0.371
$\lambda_{j,i} = 0.50, 0.75$	0.584	0.583	0.572	0.332	0.328	0.328	0.357	0.355	0.353
$\lambda_{2,i} - \lambda_{1,i} = 0.50$									
$\lambda_{j,i} = 0.20, 0.70$	0.508	0.510	0.489	0.346	0.349	0.342	0.367	0.370	0.363
$\lambda_{j,i} = 0.25, 0.75$	0.515	0.506	0.489	0.337	0.333	0.332	0.362	0.354	0.354
$\lambda_{j,i} = 0.30, 0.80$	0.515	0.518	.0505	0.346	0.344	0.338	0.355	0.361	0.361

Note: The results are the variances of the LM statistic of the endogenous two-break test. The underlying data are generated by equation (4.43), with N=1 and $\delta_{j,i} = 5$.

The interesting observations of the results can be summarised as follows. First, the finite sample means and variances of the endogenous break LM test unit root are larger (in absolute value) than those of the exogenous break test reported in Table 1 of ILT. These differences are largest for the min-t_β test, but are markedly smaller for the max-$|t_\delta|$ and min-*SBC* tests. Under the null of no break ($\delta_i = 0$), means of the min-t_β test with *LM(0)* specification are equal to -2.226 and -2.483 for the one- and two-break tests, respectively. In this case (p=0 and T=112), mean of the LM unit root test reported in Table 1 of ILT is equal to -1.973. Means of the one- and two-break min-*SBC* tests with the *LM*(0) regression are equal to -2.085 and -2.213, respectively. Finally, the results of the max-$|t_\delta|$ test are equal to –1.983 and –2.036 (see Table 4.13 and 4.17).

Second, the results show that in moderate samples, the endogenous one- and two-break LM tests are not invariant to the magnitude of break. For example, means (variances) of the one-break min-t_β test with the *LM*(0) regression are equal to –2.226 (0.471), -2.270 (0.472) and –2.290 (0.414) when $\delta_i = 0$, 5 and 10, respectively (see Table 4.13 (4.15)). In addition, means and variances of the endogenous two-break LM test are also larger than those of the corresponding one-break test in every case. For example, means (variances) of the two-break min-t_β test with the *LM*(0) regression are equal to –2.463 (0.608), -2.572 (0.593) and –2.631 (0.469) when $\delta_{j,i} =$ 0, 5 and 10, respectively (see Table 4.17 (4.19)). In the max-$|t_\delta|$ and min-*SBC* tests, the differences between means and variances of the tests with different sizes of break ($\delta_{j,i}$) are smaller than those of the min-t_β test. The mean and variance figures of the max-$|t_\delta|$ and min-*SBC* tests approach those of the exogenous break test as $\delta_{j,i}$ increases.

Third, in the presence of breaks under the null hypothesis, the mean and variance results of the endogenous one-break tests with different locations of break are similar. For instance, means (variances) of the one-break min-t_β test, with δ_i=5 and the *LM*(0) regression, are equal to –2.270 (0.472), -2.273 (0.475) and –2.269 (0.474), when λ_i=0.25, 0.50 and 0.75, respectively (see Table 4.14 (4.16)). These results confirm that the finite sample properties of the endogenous one-break tests are invariant to the location of break. Means and variances of the two-break test are still invariant to the location of break when the gap between each break point is similar. For example, means (variances) of the two-break min-t_β test, with δ_i=5 and the *LM*(0) regression, are equal to –2.641 (0.508), -2.633 (0.515) and –2.638 (0.515), when λ_i = 0.25,0.70; 0.25,0.75 and 0.30,0.80, respectively (see Table 4.18 (4.20)).

Finally, means and variances of the two-break min-t_β test vary according to the gap between each break point. For example, when $\lambda_{j,i}$ = 0.25,0.50 and 0.50,0.75, means (variances) of the two-break min-t_β test (*p*=0) are equal to –2.572 (0.593) and –2.578 (0.584), respectively (see Table 4.18 (4.20)). The mean (variance) values are –2.633 (0.515) when the gap is relatively large ($\lambda_{j,i}$= 0.25,0.75) (see Table 4.18 (4.20)). However, this variation is not observed in the max-| t_δ | and min-*SBC* tests.

In summary, the reported results confirm our suspicion about the finite sample properties of the endogenous break LM unit root tests, discussed earlier in this section. The finite sample means and variances of the endogenous break LM unit root tests are different from those of the exogenous break test, and also depend on the methods used to select the location of breaks. In addition, in the finite sample (*T*=112), the invariance property of the endogenous break test should be applied only to the location of breaks, and cannot be applied to the magnitude of breaks and the gap between each location of breaks under the null hypothesis.

4.6.2 The accuracy of estimating the true break points

Next, we examine the accuracy of estimating the true break points of the endogenous one- and two-break LM unit root tests. The frequencies of correctly estimating the true break point ($T_{B,i}$) and estimating the break point at $T_{B,i}$-1 and in the range of $T_{B,i} \pm 10$, $T_{B,i} \pm 20$ and $T_{B,i} \pm 30$ on the one-break LM test under the null (ϕ_i=1) and alternative (ϕ_i=0.9) hypotheses are reported in Tables 4.21 and 4.22, respectively. The results of the two-break LM test are given in Tables 4.23 and 4.24.

The results of the break point selection accuracy from Tables 4.21 to 4.24 can be summarised as follows. First, comparing the accuracy results, the max-$|t_\delta|$ and min-*SBC* tests determine the break date more accurately than the min-t_β test. Both the one- and two-break max-$|t_\delta|$ and min-*SBC* tests have a very high chance of correctly choosing the break points under both null and alternative hypotheses. For example, when δ_i=5 and λ_i= 0.25 (0.25,0.50), frequencies of correctly choosing the break dates at $T_{B,i}$ of the one-break (two-break) max-$|t_\delta|$ and min-*SBC* tests are equal to 0.975 (0.961) and 0.976 (0.956), respectively, under the null hypothesis (see Table 4.21 (4.23)). Under the alternative hypothesis, these results are equal to 0.971 (0.944) and 0.978 (0.942) for the one-break (two-break) max-$|t_\delta|$ and min-*SBC* tests, respectively (see Table 4.22 (4.24)). Similar results are obtained from other locations of shift under both null and alternative hypotheses. In addition, both the max-$|t_\delta|$ and min-*SBC* tests can accurately estimate the true break points 100 per cent of the time, when the size of breaks is relatively large ($\delta_{j,i}$=10) in every case. These results show that the accuracy of estimating the break points of both the max-$|t_\delta|$ and min-*SBC* tests is invariant to the location of breaks. This accuracy increases as the size of break ($\delta_{j,i}$) increases. The one-break test estimates the break point slightly more accurately than the two-break test.

Table 4.21 The accuracy of estimating the true break point of the endogenous one-break LM unit root test (under the null hypothesis)

	Location of breaks	Min-t_β test		Max-| t_δ | test		Min-*SBC* test	
		δ_i=5	δ_i=10	δ_i=5	δ_i=10	δ_i=5	δ_i=10
$T_{B,i}$	0.25	0.348	0.477	0.975	1.000	0.976	1.000
	0.50	0.361	0.482	0.978	1.000	0.975	1.000
	0.75	0.351	0.463	0.977	1.000	0.973	1.000
$T_{B,i}\pm 10$	0.25	0.444	0.666	0.981	1.000	0.980	1.000
	0.50	0.453	0.658	0.983	1.000	0.980	1.000
	0.75	0.439	0.652	0.984	1.000	0.980	1.000
$T_{B,i}\pm 20$	0.25	0.586	0.745	0.986	1.000	0.985	1.000
	0.50	0.600	0.740	0.989	1.000	0.987	1.000
	0.75	0.591	0.732	0.988	1.000	0.984	1.000
$T_{B,i}\pm 30$	0.25	0.683	0.798	0.990	1.000	0.990	1.000
	0.50	0.759	0.836	0.992	1.000	0.992	1.000
	0.75	0.683	0.788	0.990	1.000	0.987	1.000
$T_{B,i}$-1	0.25	0.007	0.022	0.000	0.000	0.000	0.000
	0.50	0.006	0.024	0.000	0.000	0.000	0.000
	0.75	0.007	0.023	0.001	0.000	0.001	0.000

Note: The figures are frequencies of estimating the true break point in the range using the endogenous one-break LM unit root test. The data are generated under the null hypothesis (ϕ_i=1). See notes to Table 4.3 for details of the DGP.

Table 4.22 The accuracy of estimating the true break point of the endogenous one-break LM unit root test (under the alternative hypothesis)

	Location of breaks	Min-t_β test		Max-$\lvert t_\delta \rvert$ test		Min-*SBC* test	
		δ_i=5	δ_i=10	δ_i=5	δ_i=10	δ_i=5	δ_i=10
$T_{B,i}$	0.25	0.482	0.744	0.971	1.000	0.978	1.000
	0.50	0.512	0.742	0.969	1.000	0.972	1.000
	0.75	0.485	0.731	0.968	1.000	0.976	1.000
$T_{B,i}\pm 10$	0.25	0.603	0.888	0.979	1.000	0.983	1.000
	0.50	0.647	0.912	0.976	1.000	0.977	1.000
	0.75	0.617	0.874	0.975	1.000	0.980	1.000
$T_{B,i}\pm 20$	0.25	0.684	0.904	0.982	1.000	0.985	1.000
	0.50	0.755	0.934	0.983	1.000	0.983	1.000
	0.75	0.705	0.893	0.981	1.000	0.985	1.000
$T_{B,i}\pm 30$	0.25	0.741	0.914	0.987	1.000	0.987	1.000
	0.50	0.853	0.959	0.988	1.000	0.989	1.000
	0.75	0.760	0.907	0.986	1.000	0.988	1.000
$T_{B,i}$-1	0.25	0.012	0.019	0.000	0.000	0.000	0.000
	0.50	0.018	0.022	0.000	0.000	0.000	0.000
	0.75	0.014	0.021	0.000	0.000	0.000	0.000

Note: The figures are frequencies of estimating the true break point in the range using the endogenous one-break LM unit root test. The data are generated under the alternative hypothesis (ϕ_i =0.9). See notes to Table 4.3 for details of the DGP.

Table 4.23 The accuracy of estimating the true break points of the endogenous two-breaks LM unit root test (under the null hypothesis)

	Location of breaks	Min-t_β test		Max-$\lvert t_\delta \rvert$ test		Min-*SBC* test	
		δ_i=5	δ_i=10	δ_i=5	δ_i=10	δ_i=5	δ_i=10
$T_{B,i}$	0.25, 0.50	0.120	0.243	0.961	1.000	0.956	1.000
	0.25, 0.75	0.035	0.095	0.954	1.000	0.946	1.000
	0.50, 0.75	0.113	0.249	0.955	1.000	0.953	1.000
$T_{B,i} \pm 10$	0.25, 0.50	0.228	0.475	0.970	1.000	0.965	1.000
	0.25, 0.75	0.149	0.295	0.964	1.000	0.960	1.000
	0.50, 0.75	0.206	0.461	0.965	1.000	0.963	1.000
$T_{B,i} \pm 20$	0.25, 0.50	0.448	0.593	0.979	1.000	0.972	1.000
	0.25, 0.75	0.346	0.476	0.974	1.000	0.974	1.000
	0.50, 0.75	0.423	0.591	0.971	1.000	0.971	1.000
$T_{B,i} \pm 30$	0.25, 0.50	0.699	0.758	0.992	1.000	0.990	1.000
	0.25, 0.75	0.542	0.653	0.978	1.000	0.977	1.000
	0.50, 0.75	0.684	0.663	0.992	1.000	0.992	1.000
$T_{B,i}$-1	0.25, 0.50	0.000	0.000	0.000	0.000	0.000	0.000
	0.25, 0.75	0.000	0.001	0.000	0.000	0.000	0.000
	0.50, 0.75	0.000	0.000	0.000	0.000	0.000	0.000

Note: The figures are frequencies of estimating the true break point in the range using the endogenous two-break LM unit root test. The data are generated under the null hypothesis (ϕ_i =1). See notes to Table 4.3 for details of the DGP.

Table 4.24 The accuracy of estimating the true break points of the endogenous two-break LM unit root test (under the alternative hypothesis)

	Location of breaks	Min-t_β test		Max-$\lvert t_\delta \rvert$ test		Min-*SBC* test	
		δ_i=5	δ_i=10	δ_i=5	δ_i=10	δ_i=5	δ_i=10
$T_{B,i}$	0.25, 0.50	0.226	0.541	0.944	1.000	0.942	1.000
	0.25, 0.75	0.149	0.350	0.948	1.000	0.954	1.000
	0.50, 0.75	0.227	0.558	0.940	1.000	0.950	1.000
$T_{B,i} \pm 10$	0.25, 0.50	0.382	0.776	0.959	1.000	0.953	1.000
	0.25, 0.75	0.305	0.567	0.961	1.000	0.962	1.000
	0.50, 0.75	0.408	0.784	0.952	1.000	0.959	1.000
$T_{B,i} \pm 20$	0.25, 0.50	0.521	0.812	0.968	1.000	0.967	1.000
	0.25, 0.75	0.490	0.660	0.972	1.000	0.971	1.000
	0.50, 0.75	0.560	0.882	0.964	1.000	0.971	1.000
$T_{B,i} \pm 30$	0.25, 0.50	0.742	0.877	0.990	1.000	0.991	1.000
	0.25, 0.75	0.653	0.780	0.975	1.000	0.974	1.000
	0.50, 0.75	0.747	0.882	0.988	1.000	0.992	1.000
$T_{B,i}$-1	0.25, 0.50	0.000	0.000	0.000	0.000	0.000	0.000
	0.25, 0.75	0.000	0.002	0.000	0.000	0.000	0.000
	0.50, 0.75	0.000	0.000	0.000	0.000	0.000	0.000

Note: The figures are frequencies of estimating the true break point in the range using the endogenous two-break LM unit root test. The data are generated under the alternative hypothesis (ϕ_i=0.9). See notes to Table 4.3 for details of the DGP.

Second, even though the frequencies of correctly estimating the break dates of the min-t_β test are less than the max-$\lvert t_\delta \rvert$ and min-*SBC* tests, they still estimate the break points reasonably well. Under the alternative hypothesis, the accuracy performance of the min-t_β test is better than that of the test under the null hypothesis. For example, when δ_i=5, frequencies of the correct choice of the break date at $T_{B,i}$ of the one-break min-t_β test under the null (alternative) hypothesis are equal to 0.348 (0.482), 0.361 (0.512) and 0.351 (0.485), when λ_i= 0.25, 0.50 and

0.75, respectively (see Table 4.21 (4.22)). Again, these accuracy results increase in the relatively large size of breaks ($\delta_{j,i}$ =10), and decline in the two-break test. In the two-break test, the min-t_β test determines the break point correctly at 12.0 (22.6), 3.5 (13.7) and 11.3 (22.7) percent of time, under the null (alternative) hypothesis, when $\delta_{j,i}$=5 and $\lambda_{j,i}$= 0.25,0.50; 0.25,0.75 and 0.50,0.75, respectively (see Table 4.23 (4.24)). These results show that the accuracy of the two-break min-t_β test depends on the location of breaks, and decreases when these gaps are large.

Third, our accuracy results of the two-break min-t_β test are similar to those reported by LS. Frequencies of estimating the true break point ($T_{B,i}$) of the two-break min-t_β test, when $\delta_{j,i}$=10, under the alternative hypothesis (ϕ_i=0.9) reported in Table 4 of LS, are equal to 0.226 and 0.101 when λ_i=0.20,0.50 and 0.25,0.75, respectively. The accuracy increases when the gap between each break is small. The frequency result, under the alternative hypothesis (ϕ_i=0.9), is equal to 0.325, when δ_i=5 and λ_i=0.20,0.30.

Finally, comparing the break point estimation accuracy of the endogenous break LM unit root test with that of the endogenous one-break ADF-type test, reported in Table 1 of Lee and Strazicich (2001), the min-t_β LM and max-$| t_\delta |$ LM tests determine the break date more accurately than the corresponding ADF-type tests. However, results from the min-*SBC* tests are similar in terms of the LM- and ADF-type unit root tests. In addition, Lee and Strazicich (2001) report that the min-t_β and max-$| t_\delta |$ ADF-type tests tend to determine the break point incorrectly at $T_{B,i}$-1. By contrast, the results from Tables 4.21 to 4.24 show that the chance of incorrectly estimating the break point at $T_{B,i}$-1 is negligible for the endogenous break LM unit root test in every case.

4.6.3 The finite sample size and power

Next, we investigate the size and power performance of the endogenous break panel LM unit root test. The simulated size and power results of the endogenous one- and two-break panel LM unit root tests, using three methods in estimating the break dates, are presented in Tables 4.25 to 4.30. The results are based on the standardised panel LM statistics with correct adjustment parameters (means, variances) for the tests with different sizes of break ($\delta_{j,i}$) and gaps between each break point. These parameters are obtained from Tables 4.13 to 4.20.

Table 4.25 The empirical size of the endogenous one- and two-break min-t_β panel LM unit root test

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.049	0.061	0.049	0.047
1	0.50	0.065	0.058	0.046	0.055
1	0.75	0.065	0.045	0.045	0.044
2	0.25, 0.50	0.060	0.054	0.086	0.050
2	0.25, 0.75	0.060	0.058	0.062	0.056
2	0.50, 0.75	0.046	0.052	0.056	0.036

Note: The results are based on the endogenous one- and two-break min-t_β panel LM unit root tests with the $LM(0)$ regression. See notes to Table 4.3 for details of the DGP.

Table 4.26 The empirical size of the endogenous one- and two-break max-$|t_{\delta}|$ panel LM unit root test

Number of breaks	Location of Breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.062	0.057	0.048	0.043
1	0.50	0.056	0.064	0.041	0.049
1	0.75	0.060	0.057	0.044	0.052
2	0.25, 0.50	0.086	0.054	0.060	0.044
2	0.25, 0.75	0.060	0.066	0.078	0.048
2	0.50, 0.75	0.064	0.050	0.064	0.042

Note: The results are based on the endogenous one- and two-break max-$|t_{\delta}|$ panel LM unit root tests with the $LM(0)$ regression. See notes to Table 4.3 for details of the DGP.

Table 4.27 The empirical size of the endogenous one- and two-break min-*SBC* panel LM unit root tests

Number of breaks	Location of Breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.057	0.057	0.048	0.043
1	0.50	0.056	0.064	0.041	0.049
1	0.75	0.062	0.057	0.044	0.052
2	0.25, 0.50	0.086	0.054	0.078	0.044
2	0.25, 0.75	0.064	0.066	0.078	0.048
2	0.50, 0.75	0.068	0.050	0.070	0.042

Note: The results are based on the endogenous one- and two-break min-*SBC* panel LM tests with the $LM(0)$ regression. See notes to Table 4.3 for details of the DGP.

Table 4.28 The empirical power of the endogenous one- and two-break min-t_{β} panel LM unit root tests

Number of breaks	Location of breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.745	0.644	1.000	1.000
1	0.50	0.780	0.706	1.000	1.000
1	0.75	0.755	0.606	0.999	1.000
2	0.25, 0.50	0.672	0.292	1.000	0.942
2	0.25, 0.75	0.726	0.432	1.000	0.984
2	0.50, 0.75	0.636	0.308	0.946	0.912

Note: See notes to Table 4.25.

Table 4.29 The empirical power of the endogenous one- and two-break max-$|t_\delta|$ panel LM unit root tests

Number of breaks	Location of breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.904	0.906	1.000	1.000
1	0.50	0.907	0.904	1.000	1.000
1	0.75	0.901	0.909	1.000	1.000
2	0.25, 0.50	0.888	0.870	1.000	1.000
2	0.25, 0.75	0.906	0.898	1.000	1.000
2	0.50, 0.75	0.884	0.888	1.000	1.000

Note: See notes to Table 4.26.

Table 4.30 The empirical power of the endogenous one- and two-breaks min-*SBC* panel LM unit root test

Number of Breaks	Location of Breaks	Small panel ($N=5$)		Large panel ($N=25$)	
		$\delta_{j,i}=5$	$\delta_{j,i}=10$	$\delta_{j,i}=5$	$\delta_{j,i}=10$
1	0.25	0.891	0.906	1.000	1.000
1	0.50	0.903	0.904	1.000	1.000
1	0.75	0.896	0.909	1.000	1.000
2	0.25, 0.50	0.882	0.870	1.000	1.000
2	0.25, 0.75	0.884	0.898	1.000	1.000
2	0.50, 0.75	0.882	0.888	1.000	1.000

Note: See notes to Table 4.27.

The results from Tables 4.25, 4.26 and 4.27 show that the size results of the endogenous one- and two-break tests approach the nominal level of 0.05 as N increases. The 95% confidence interval of the 0.05 significant level test lies between 0.0365 to 0.0635 in the one-break test, in which simulations are based on 1,000 replications. In the two-break test, the results are based on 500 trials. In this case, the 95% confidence interval is between 0.0309 and 0.0691. Under the alternative hypothesis, in the case of a small panel (N=5), the panel max-$|t_\delta|$ and min-*SBC* tests are more powerful than the panel min-t_β test. The power results of the one- and two-break panel max-$|t_\delta|$ and min-*SBC* tests are close to the corresponding figures of the exogenous-break test reported in Table 4.4. For example, when $\lambda_i = 0.5$ and $\delta_i = 5$, the empirical power of the one-break panel max-$|t_\delta|$ and min-*SBC* tests is equal to 0.907 and 0.903, respectively; and the power results for the two-break panel max-$|t_\delta|$ and min-*SBC* tests test are 0.888 and 0.882, respectively, when $\lambda_{j,i} = 0.25,0.50$ and $\delta_{j,i} = 5$ (see Tables 4.29 and 4.30). The power of the min-t_β LM test decreases from that of the exogenous break test. However, the panel min-t_β test still has significant power to reject the null hypothesis when δ_i=5, as the power of the one-break test is equal to 0.745, 0.780 and 0.755 when λ_i= 0.25, 0.50 and 0.75, respectively (see Table 4.28). For the two-break min-t_β test, the power results are 0.672, 0.744 and 0.636, when $\lambda_{j,i}$= 0.25,0.50; 0.25,0.75 and 0.50,0.75, respectively (see Table 4.28). A loss of power in the min-t_β test tends to increase with the size of break (δ_i). For example, the power results of the two-break min-t_β test are 0.292, 0.506 and 0.308, when δ_i=10 and λ_i= 0.25,0.50; 0.25,0.75 and 0.50,0.75, respectively (see Table 4.28).

The results from Table 4.28 also show that the location of breaks does slightly affect the power of the min-t_β test. The power of the one-break test when the break occurs in the middle of the series (λ_i=0.50) is slightly higher than that of the test when the break occurs in either the early (λ_i= 0.25) or late (λ_i= 0.75) stages of the series. In the two-break test, the power increases when the gap between each break point is large (λ_i=0.25,0.75). In the large panel (*N*=25), the power of the endogenous break tests is equal, or close to, 1.000 in every case. These findings are similar to those of the univariate min-t_β test reported in LS. The power of the univariate min-t_β test also decreases with δ_i and slightly drops when the gap between each break point is small.

Next, we consider the effect of incorrectly applying the means and variances in standardising the endogenous break panel test. The results from Tables 4.13 to 4.20 show that the means and variances of the endogenous break LM unit root tests are different from those of the exogenous break LM test, and depend on the methods of break date estimation and magnitude of break (δ_i) under the DGP. The simulated size and power results of the endogenous break test calculated from the means and variances corresponding to the no break case are presented in Tables 4.31 to 4.36. The results of the standardised panel LM test calculated from the means and variances of the exogenous break test are shown in Tables 4.37 to 4.42.

The results from Tables 4.31, 4.32 and 4.33 show that the one- and two-break panel min-t_β and max-$|t_\delta|$ tests are slightly size-distorted in this case. For example, the size of the two-break panel min-t_β (max-$|t_\delta|$) test is 0.088 (0.026) and 0.080 (0.030), when λ_i= 0.25,0.50 and δ_i=5 and 10, respectively (see Table 4.31 (4.32)) However, the panel one- and two-break min-*SBC* test is substantially under-sized. The empirical size of the one-break (two-break) panel min-*SBC* test is 0.016 (0.002) and 0.018 (0.004) when λ_i=0.50 ($\lambda_{j,i}$= 0.25,0.50) and δ_i= 5 and 10, respectively (see Table 4.33). In the case of the large panel (*N*=25), the size distortions are slightly larger than those of the small panel (*N*=5) in every case. For instance, when λ_i= 0.50, the empirical size of the one-break panel min-*SBC* test decreases to 0.004 and 0.005 with δ_i= 5 and 10, respectively. The empirical power of the tests approximates that of the corresponding tests with correctly standardised parameters when the size is close to the nominal level. For example, when λ_i= 0.50, the power of the one-break max-$|t_\delta|$ test is equal to 0.893 and 0.884 with δ_i= 5 and 10, respectively (see Table 4.35).

Table 4.31 The empirical size of the endogenous break min-t_β panel LM unit root test, using adjustment parameters from the endogenous-break test without shifts

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.075	0.072	0.090	0.102
1	0.50	0.086	0.069	0.076	0.117
1	0.75	0.079	0.057	0.096	0.103
2	0.25, 0.50	0.088	0.080	0.192	0.158
2	0.25, 0.75	0.100	0.112	0.300	0.470
2	0.50, 0.75	0.070	0.090	0.148	0.128

Note: The results are based on the endogenous one- and two-break min-t_β panel LM unit root tests with the $LM(0)$ regression adjusted by means and variances of the endogenous break test without shifts. See notes to Table 4.3 for details of the DGP.

Table 4.32 The empirical size of the endogenous break max-$|t_\delta|$ panel LM unit root test, using adjustment parameters from the endogenous-break test without shifts

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.049	0.056	0.042	0.031
1	0.50	0.052	0.055	0.034	0.036
1	0.75	0.056	0.048	0.034	0.037
2	0.25, 0.50	0.026	0.030	0.012	0.008
2	0.25, 0.75	0.024	0.030	0.024	0.008
2	0.50, 0.75	0.026	0.032	0.006	0.012

Note: The results are based on the endogenous one- and two-break max-$|t_\delta|$ panel LM unit root tests with the $LM(0)$ regression adjusted by means and variances of the endogenous break test without shifts. See notes to Table 4.3 for details of the DGP.

Table 4.33 The empirical size of the endogenous break min-*SBC* panel LM unit root test, using adjustment parameters from the endogenous-break test without shifts

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.015	0.018	0.004	0.002
1	0.50	0.016	0.018	0.004	0.005
1	0.75	0.020	0.010	0.002	0.003
2	0.25, 0.50	0.002	0.004	0.000	0.000
2	0.25, 0.75	0.004	0.000	0.000	0.000
2	0.50, 0.75	0.006	0.006	0.000	0.000

Note: The results are based on the endogenous one- and two-break min-*SBC* panel LM unit root tests with the *LM*(0) regression adjusted by means and variances of the endogenous break test without shifts. See notes to Table 4.3 for details of the DGP.

Table 4.34 The empirical power of the endogenous break min-t_β panel LM unit root test, using adjustment parameters from the endogenous-break test without shifts

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.803	0.689	1.000	1.000
1	0.50	0.837	0.742	1.000	1.000
1	0.75	0.811	0.659	1.000	1.000
2	0.25, 0.50	0.736	0.416	1.000	1.000
2	0.25, 0.75	0.844	0.618	1.000	1.000
2	0.50, 0.75	0.738	0.434	0.990	0.994

Note: See notes to Table 4.31.

Table 4.35 The empirical power of the endogenous break max-$|t_\delta|$ panel LM unit root test, using adjustment parameters from the endogenous-break test without shifts

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.885	0.888	1.000	1.000
1	0.50	0.893	0.884	1.000	1.000
1	0.75	0.888	0.892	1.000	1.000
2	0.25, 0.50	0.760	0.782	1.000	1.000
2	0.25, 0.75	0.790	0.798	1.000	1.000
2	0.50, 0.75	0.770	0.800	1.000	1.000

Note: See notes to Table 4.32.

Table 4.36 The empirical power of the endogenous break min-*SBC* panel LM unit root test, using adjustment parameters from the endogenous-break test without shifts

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.730	0.727	1.000	1.000
1	0.50	0.730	0.720	1.000	1.000
1	0.75	0.722	0.685	1.000	1.000
2	0.25, 0.50	0.400	0.400	0.992	0.984
2	0.25, 0.75	0.450	0.398	0.992	0.982
2	0.50, 0.75	0.424	0.388	0.984	0.982

Note: See notes to Table 4.33.

Next, we consider the results of the panel LM statistics when they are incorrectly adjusted, using the means and variances of the exogenous-break test presented in Table 1 of ILT. The simulated results from Tables 4.37, 4.38 and 4.39 show that the one- and two-break panel min-t_β tests are massively over-sized in both the small (N=5) and large (N=25) panels. The size of the one- and two-break panel min-*SBC* tests is slightly over-sized when $\delta_{j,i} = 5$. When $\delta_{j,i} = 10$, the size of the one- and two-break min-*SBC* tests approaches the nominal level. Finally, there is no significant size distortions in the one- and two-break panel max-$|t_\delta|$ tests.

Table 4.37 The empirical size of the endogenous break min-t_β panel LM unit root test, using adjustment parameters from the exogenous-break test

Number of breaks	Location of Breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.332	0.332	0.762	0.826
1	0.50	0.328	0.319	0.753	0.842
1	0.75	0.330	0.323	0.793	0.829
2	0.25, 0.50	0.746	0.758	1.000	1.000
2	0.25, 0.75	0.756	0.902	1.000	1.000
2	0.50, 0.75	0.660	0.734	0.998	1.000

Note: The results are based on the endogenous one- and two-break max-$|t_\delta|$ panel LM unit root tests with the *LM*(0) regression adjusted by means and variances of the exogenous break test. See notes to Table 4.3 for details of the DGP.

Table 4.38 The empirical size of the endogenous break max-| t_δ | panel LM unit root test, using adjustment parameters from the exogenous-break test

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.062	0.056	0.049	0.040
1	0.50	0.056	0.060	0.041	0.046
1	0.75	0.060	0.056	0.044	0.048
2	0.25, 0.50	0.080	0.054	0.048	0.052
2	0.25, 0.75	0.054	0.064	0.062	0.044
2	0.50, 0.75	0.050	0.054	0.050	0.046

Note: The results are based on the endogenous one- and two-break max-| t_δ | panel LM unit root tests with the $LM(0)$ regression adjusted by means and variances of the exogenous break test. See notes to Table 4.3 for details of the DGP.

Table 4.39 The empirical size of the endogenous break min-*SBC* panel LM unit root test, using adjustment parameters from the exogenous-break test

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.069	0.056	0.066	0.040
1	0.50	0.061	0.060	0.059	0.046
1	0.75	0.067	0.056	0.060	0.048
2	0.25, 0.50	0.088	0.054	0.080	0.052
2	0.25, 0.75	0.070	0.064	0.090	0.044
2	0.50, 0.75	0.068	0.054	0.074	0.046

Note: The results are based on the endogenous one- and two-break min-*SBC* panel LM unit root tests with the $LM(0)$ regression adjusted by means and variances of the exogenous break test. See notes to Table 4.3 for details of the DGP.

Table 4.40 The empirical power of the endogenous break min-t_β panel LM unit root test, using adjustment parameters from the exogenous-break test

Number of breaks	Location of breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.984	0.974	1.000	1.000
1	0.50	0.987	0.977	1.000	1.000
1	0.75	0.987	0.970	1.000	1.000
2	0.25, 0.50	1.000	0.998	1.000	1.000
2	0.25, 0.75	1.000	1.000	1.000	1.000
2	0.50, 0.75	0.998	0.998	1.000	1.000

Note: See notes to Table 4.37.

Table 4.41 The empirical power of the endogenous break max-$|t_\delta|$ panel LM unit root test, using adjustment parameters from the exogenous-break test

Number of breaks	Location of Breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.904	0.903	1.000	1.000
1	0.50	0.907	0.900	1.000	1.000
1	0.75	0.901	0.903	1.000	1.000
2	0.25, 0.50	0.868	0.878	1.000	1.000
2	0.25, 0.75	0.888	0.896	1.000	1.000
2	0.50, 0.75	0.870	0.896	1.000	1.000

Note: See notes to Table 4.38.

Table 4.42 The empirical power of the endogenous break min-*SBC* panel LM unit root test, using adjustment parameters from the exogenous-break test

Number of breaks	Location of Breaks	Small panel ($N = 5$)		Large panel ($N = 25$)	
		$\delta_{j,i} = 5$	$\delta_{j,i} = 10$	$\delta_{j,i} = 5$	$\delta_{j,i} = 10$
1	0.25	0.912	0.903	1.000	1.000
1	0.50	0.917	0.900	1.000	1.000
1	0.75	0.910	0.903	1.000	1.000
2	0.25, 0.50	0.886	0.878	1.000	1.000
2	0.25, 0.75	0.898	0.896	1.000	1.000
2	0.50, 0.75	0.882	0.896	1.000	1.000

Note: See notes to Table 4.39.

In summary, the panel max-$|t_{\delta}|$ and min-*SBC* tests outperform the panel min-t_{β} test in terms of the size, power and accuracy of break date estimation. The values of standardised parameters (means, variances) should be carefully applied, to compute the panel statistics. Applying incorrectly standardised parameters may lead to a distortion in size. The size results of the panel max-$|t_{\delta}|$ and min-*SBC* tests are less sensitive to the incorrectly standardised use of means and variances of the exogenous break LM unit root test than the panel min-t_{β} test. Comparing the panel max-$|t_{\delta}|$ and min-*SBC* tests, the panel max-$|t_{\delta}|$ test is less sensitive than the panel min-*SBC* tests to the incorrectly adjusted use of parameters of the incorrect magnitude of breaks ($\delta_{j,i}$). When the tests are calculated, using the correct adjustment parameters, the empirical power of the panel max-$|t_{\delta}|$ and min-*SBC* tests is higher than that of the panel min-t_{β} test. In addition, the max-$|t_{\delta}|$ and min-*SBC* procedures select the true break dates more accurately than the min-t_{β} test.

4.7 Conclusion

In this chapter, we studied the panel LM unit root tests both with and without structural shifts. We first examined the panel LM test without shifts, showing that the empirical size and power of the test is reasonably close to that of the IPS and MW tests. The power of the panel LM test reduces in a mixed panel of stationary and non-stationary series. The panel LM test is over-sized in the presence of cross-sectional dependence in the errors. These results are similar to those of the panel IPS and MW tests reported in Chapter 2.

Second, we applied the panel LM unit test when the location of breaks (λ_i) is exogenously given. The simulated results show that the exogenous break panel LM unit root test performs well when the break points are correctly specified. The size and power performance of the exogenous break test is similar to that of the test without shifts. However, incorrectly specifying the number and/or location of breaks results in size-distortions.

Finally, the performance of the endogenous break LM unit root tests was investigated in terms of the size, power and frequencies of estimating the true break points. The means and variances of the endogenous break tests vary between the tests with different procedures of break point estimation and magnitudes of break ($\delta_{j,i}$) under the null hypothesis. We computed the finite sample means and variances and reported the results in Tables 4.13 to 4.20. When the panel statistics are standardised, using the correct parameters (means, variances), the endogenous break max-$| t_\delta |$ and min-*SBC* tests perform well in terms of the size, power and accuracy in selecting the true break points. In this case, the size and power results approximate those of the exogenous break test, and the accuracy of estimating the break date is very high. The performance of the min-t_β test is worse than the max-$|t_\delta|$ and min-*SBC* tests. However, the min-t_β test still has the powerful capacity to reject the null hypothesis in the majority of cases. When the panel LM statistics are incorrectly adjusted with invalid mean and variance values, the panel LM tests will be seriously size-distorted. Therefore, application of the endogenous break test requires the careful application of the correct adjustment parameters. Comparing the endogenous break tests with various break point selection procedures, the max-$|t_\delta|$ test performs better than the other tests in terms of the size, power and accuracy in estimating the break point. In addition, the max-$|t_\delta|$ test is also less sensitive to a choice of adjustment parameters (means and variances) than the other tests.

Chapter 5

Panel Evidence on Fundamental Exchange Rate Modelling from Asia Pacific Countries

5.1 Introduction

For many years, the empirical analysis of fundamental exchange rate modelling has been an active area of research in the field of international macroeconomics. The standard economic theories that underlie the analysis of exchange rate movements are purchasing power parity (PPP) and the monetary model. The PPP hypothesis is usually considered as a critical assumption, in modern theories, of open economy macroeconomic models. However, empirical research on the validity of a PPP relationship yields one of the most puzzling results in international macroeconomics. Using long-horizon data sets, Kim (1990), and Lothian and Taylor (1996) provide some evidence for the existence of PPP. However, a number of empirical studies (see Taylor (1995) and Rogoff (1996)) find no evidence to support the validity of PPP as a long-run relationship under the current floating period. Rogoff (1996) points out that the adjustment toward a long-run equilibrium of PPP is quite slow, with a three to five year half-life. This persistence cannot be explained by monetary factors, which should have a faster adjustment rate. In addition, the volatility of exchange rate movements is too large for the persistence to be related to real factors, which should not yield such a high volatility.

For the monetary model, MacDonald and Taylor (1993, 1994) find a long-run relationship between exchange rates and monetary fundamentals, using the likelihood-based cointegration test of Johansen (1988) for both the Deutsche mark/US dollar and pound sterling/US dollar; however, Sarantis (1994) fails to find any significant evidence of a cointegrating relationship, using the monetary model in a study utilising pound sterling based exchange rates.

The low power of unit root and cointegration tests in short spans of data is often mentioned as one explanation for the failure to find a long-run relationship between exchange rates and fundamentals. Lothian and Taylor (1997) argue that the standard unit root tests, such as the ADF test, have extremely low power in testing for mean reversion under PPP for small sample sizes corresponding to the post-Bretton Woods era. The recent development of panel data techniques has offered an alternative approach to increasing the power of the unit root and cointegration tests over the conventional time-series tests. Several panel unit root tests have been developed by Levin, Lin and Chu (2002) (LLC), Im, Persaran and Shin (2003) (IPS) and Maddala and Wu (1999) (MW), based on both homogeneity and heterogeneity assumptions. Panel cointegration tests have been proposed as an extension of the residual-based test of Engle and Granger (1987) and the likelihood-based test of Johansen (1988) (see Chapters 2 and 3). These panel methods have been used to examine the validity of PPP and the monetary model as a long-run relationship in recent studies. A number of researchers find evidence for the existence of PPP (see, for example, Frankel and Rose (1996), Oh (1996), Wu (1996) and Sarno and Taylor (1998)), while Oh (1999) and Groen (2000) find positive evidence of a cointegration relationship, using monetary models.

Most of the empirical studies on exchange rates and their fundamentals usually focus on the industrial OECD countries. For developing countries, the empirical evidence is relatively scarce. However, numerous countries in the Asia Pacific region have experienced rapid economic growth with strong trading tied to the world economy - the so-called Asian miracle (see Krugman (1994) for discussion of the economic growth in East Asian countries). However, the 1997, the East Asian currency crisis affected most countries in the Asia Pacific region, as a consequence of which, many countries implemented changes to both the exchange rate regimes and their structural economic programmes. In light of this, it is important to examine whether or not the traditional PPP and monetary model can explain exchange rate movements in the region. In addition, we must allow for the possibility of a structural shift, due to the effect of the currency crisis. The effect of structural changes and the development of panel unit root tests with shifts were discussed in Chapter 4. These methods are useful to investigate a long-run relationship between exchange rates and fundamentals in the presence of a structural shift due to the crisis.

Consequently, the objective of this chapter is to examine a long-run relationship between exchange rates and fundamentals in Asia Pacific countries. PPP and the monetary approach are used as the fundamental determinants of exchange rate movements. We apply various panel unit root and cointegration tests and compare the empirical results from these tests with the simulation results reported in Chapters 2,3 and 4. The panel IPS and MW unit root tests are applied in testing for the unit root null hypothesis of real exchange rates, according to the PPP hypothesis and the monetary model. The bootstrap method is also used to correct the size distortions, which potentially occur due to the presence of cross-sectional dependence. In addition, the alternative methods, i.e. the Seeming Unrelated Regression (SUR) method and Cross-sectionally augmented IPS test (CIPS) of Pesaran (2003), are also applied to test for unit roots accounting for the cross-correlations. In addition, the residual-based panel cointegration tests of IPS, MW and

CIPS and the likelihood-based panel rank test of Larsson, Lyhagen and Lothgren (2001) (LLL) are used to investigate a long-run relationship between exchange rates and fundamentals in the multivariate framework. Next, we use the panel LM unit root test with structural shifts to account for a level shift in real exchange rates due to the effect of the currency crisis.

The chapter is organised as follows. The following section provides a literature review of the empirical studies on PPP and the monetary model. Section 5.3 outlines a short description of the underlying economic theories of PPP and the monetary approach to exchange rate behaviour. Section 5.4 sets out the data sources and presents the empirical results from the unit root and cointegration tests. Section 5.5 investigates the impact of the 1997 currency crisis, using the panel unit root tests allowing for level shifts. Finally, Section 5.6 provides conclusions.

5.2 Literature review

Many empirical studies of exchange rate behaviour use variants of PPP and the monetary approach as the fundamental determinants of exchange rate movements. However, the power of the fundamental factors to explain exchange rate behaviour is still one of the key controversies in the area of international macroeconomics (see Messe and Rogoff (1983) and Rogoff (1996)). Unit root and cointegration tests are usually employed to test for the existence of a long-run relationship between exchange rates and fundamentals. However, the empirical evidence is mixed. For example, Mark (1990) uses the Engle and Granger (1987) two-step cointegration test to investigate a long-run PPP relationship, using the data from eight OECD countries during the post-Bretton Woods era. The results produce little support for long-run PPP. However, Lothian and Taylor (1996) apply the ADF test for stationarity of real exchange rates as evidence of mean reversion under the PPP hypothesis. Using long-horizon data spanning over two centuries for US dollar/sterling and franc/sterling real exchange rates, they find strong evidence of mean reversion in real exchange rates. Lothian and Taylor (1997) use Monte Carlo simulations to demonstrate that the standard unit root tests have extremely low power in rejecting the unit root null hypothesis in real exchange rates over sample sizes corresponding to the post-Bretton Woods data set. This argument is usually mentioned in many empirical studies as an explanation for the failure to find support for PPP.

In the empirical studies of the monetary model, MacDonald and Taylor (1993) examine a relationship between exchange rates and monetary factors, using the likelihood-based Johansen cointegration test for monthly data on the Deutsche mark/US dollar exchange rate over the period 1976:1 to 1990:12. Empirical results indicate the presence of significant cointegrating vectors among the series of exchange rates, domestic and foreign money supplies, domestic and foreign real incomes and domestic and foreign interest rates. These findings indicate that the monetary model is valid when it is considered as a long-run equilibrium. In addition, MacDonald and Taylor (1994) use the same technique in a study of the pound sterling/US dollar exchange rate, and find a long-run relationship in the monetary model. By contrast, Sarantis (1994) finds no evidence to support the monetary model, when using the four pound-sterling based exchange rates (US dollar, Deutsche mark, yen, French franc).

A number of empirical studies on fundamental exchange rate determination have applied recent developments in panel unit root and cointegration tests to improve the power over the conventional time-series tests. These studies generally provide encouraging results with regard to a relationship between exchange rates and fundamentals. Wu (1996) applies panel data set on the US dollar based real exchange rates against eighteen OECD countries. Using the Levin and Lin (1992) (LL) panel unit root test, Wu (1996) rejects the null hypothesis of non-stationarity in the panel of real exchange rates during the post-Bretton Woods period. These findings support the validity of long-run PPP. Oh (1996) also uses the panel LL unit root test in testing for unit roots during the floating exchange rate period, and finds evidence to support the PPP hypothesis in most panels and sub-panels of twenty-three OECD and eighty-eight developing countries. Frankel and Rose (1996), Coakley and Fuertes (1997), and Pedroni (2001) provide additional evidence in support of PPP, using various panel unit root and cointegration tests.

However, O'Connell (1998) mentions that the presence of cross-sectional dependence in the error terms, which arises from the co-movement pattern in macroeconomics data, affects the properties of the panel tests. He shows that the panel unit root tests that neglect this cross-correlation suffer from serious size distortions. Therefore, the importance of controlling for the effect of the cross-correlations has been well recognised in recent studies. Sarno and Taylor (1998), and Taylor and Sarno (1998) apply the Multivariate Augmented Dickey-Fuller test (MADF) and the Johansen Likelihood Ratio test (*JLR*) in testing for unit roots of real exchange rates among G-5 countries during the post-Bretton Woods period. They find strong evidence for mean reversion in real exchange rates. Wu and Wu (2001) extend the IPS and MW tests, allowing for a general serial correlation structure and contemporaneous correlation in the errors across countries, using a bootstrap procedure. The results reject the null hypothesis of non-stationarity in real exchange rates for the panel of twenty industrial countries under the current floating period. However, Breuer, McNown and Wallace (2001) apply the SURADF test in testing for unit roots in the system of fourteen industrial countries, and find only little improvement in the results for the panel over those of the standard ADF test.

The sensitivity of empirical results to the choice of base currency has been investigated in some recent studies. Coakley and Fuertes (2000) compare the panel unit root tests based on the Deutsche mark real exchange rates with those of the US dollar based real exchange rates. They consider the panel LL test, the panel LL test of SUR method, the panel IPS test and the *JLR* test. The empirical findings support the PPP hypothesis in most panels and sub-panels for the nineteen OECD currencies, and also find some evidence of a base currency effect when the presence of cross-correlations in the data is ignored. However, when the contemporaneous correlations are controlled, similar results between the panel based on the Deutsche mark and US dollar are found. Papell and Theodoridis (2001) investigate the implication of the choice of numeraire currency on the panel tests for PPP in twenty-one industrial countries under the post-Bretton Woods period. The panel LL unit root test is applied in the panels of real exchange rates based on twenty-one different base currencies, using SUR estimation to account for the contemporaneous correlations. In contrast to the findings of Coakley and Fuertes (2000), the results show that evidence of PPP is stronger when European currencies are used as the base currency, instead of non-European currencies.

Turning to the monetary model, Oh (1999) applies the panel LL unit root test to the residual-based panel cointegration test, to examine exchange rate determination in seven industrial countries during the post-Bretton Woods period, and finds favourable results for the monetary model. Oh (1999) argues that evidence of a long-run relationship in the monetary model is stronger when the number of countries in the panel increases and when low frequency data (e.g. annual data) is used. Groen (2000) uses the same method as Oh (1999) in testing for the monetary model, using the SUR estimation to account for the contemporaneous correlations across countries, by means of which he finds evidence of cointegration in the panel of fourteen OECD countries.

The validity of PPP and the monetary model in exchange rate determination has been extensively tested for industrialised countries, especially for European economies. However, empirical studies on less-developed countries in the Asia Pacific region are relatively limited. Moreover, the majority of them are based on standard time-series analysis. Oh (1996) mentions that the empirical evidence of PPP in developing countries is weaker than in OECD countries. Kim and Enders (1991) find little evidence in support of stationarity of real exchange rates for six Pacific rim countries. The results show that the unit root null hypothesis can be rejected only for Thailand. Bahmani-Oskooee (1993) reports weak evidence for PPP for five Asian developing countries, using the residual-based Engle-Granger cointegration method during the post-Bretton Woods era, in which the PPP hypothesis is accepted only for the Philippines. Moreover, Baharumshah and Ariff (1997) apply both the residual-based and likelihood-based cointegration tests to investigate PPP in five South-East Asian countries during the same period. The results fail to support a long-run PPP relationship in most countries. The null hypothesis of no cointegration can be rejected only for Indonesia, using the *JLR* test.

In contrast, results in support of PPP in Asia Pacific countries have been found in some recent studies. Phylaktis and Kassimatis (1994) conduct the unit root tests in the system of equations estimated by the SUR method, and find evidence to support PPP in eight Asia Pacific currencies, using the exchange rate data from black markets rather than the official rates. Lee (1999) investigates the validity of PPP in thirteen Asia Pacific currencies, using the generalised error correction model and the ADF test for stationarity of real exchange rates. The results from the ADF test support PPP only for Mexico. However, the results support PPP in most countries when the generalised dynamic error correction model is applied. Wang (2000) uses the Johansen cointegration approach to examine the validity of PPP in seven Asian countries during the flexible exchange rate periods. A long-run relationship between exchange rates and prices is found, but the cointegrating vector implied by PPP does not exist. Fujii (2002) uses the likelihood-based cointegration test with pre-specified cointegrating vectors to investigate the behaviour of real exchange rates in five East Asian countries and the effect of the 1997 currency crisis. The results support the existence of long-run PPP for all of the US dollar-based currencies, excluding Indonesia. However, the results from the yen-based currencies are much weaker, and PPP is found only in the case of the Philippines.

In his study of the monetary approach in Asia Pacific countries, Chin (1999) employs the *JLR* test, to estimate the monetary model in eight East Asian countries. The results show that a cointegration relationship is found for various types of monetary models. In addition, Chin (2000) investigates a long-run relationship between exchange rates and fundamental factors, based on PPP and the monetary model, in assessing the overvaluation of eight East Asian currencies before the 1997 crisis. Using the Johansen method, the results suggest the existence of PPP in most countries. However, when the cointegrating vector according to PPP are pre-specified, the results are much weaker, as the PPP hypothesis is found only for five US dollar-based and two yen-based currencies, respectively.

The application of panel unit root and cointegration tests in testing for PPP and the monetary model in Asia Pacific countries can be found in some recent studies. Wu and Chen (1999) use the panel IPS and MW tests to test for stationarity of real exchange rates in eight Pacific basin countries during the post-Bretton Woods period. Using critical values from the bootstrap method, they fail to find empirical support for PPP in both US dollar and Singapore dollar based panels. Basher and Mohsin (2001) apply the panel cointegration test of Pedroni (1999) in testing for PPP. Their results support the existence of PPP in the panel of ten Asian developing economies. Choong *et al.* (2000) apply the panel two-step cointegration test and find evidence to support PPP. Using the monetary model, Husted and MacDonald (1999) employ a panel of data from nine Asia Pacific counties, producing results that support the monetary model in the panel estimates, but not in the individual country estimates.

An important concern in the studies of exchange rates and fundamentals in Asia Pacific countries is the existence of the currency crisis. Breitung and Candelon (2003) apply the panel unit root test of Breitung and Meyer (1994) to account for the effect of structural breaks. Stationarity of real exchange rate is investigated in the panel of ten Asian and South American countries. The panel unit root test with shifts is applied when the break points are exogenously given. The results suggest that PPP holds for Asian countries, but does not exist for South American counties, which have fixed exchange rate regimes.

Overall, empirical studies on PPP and the monetary model in Asia Pacific countries, which include data after the 1997 currency crisis, tend to show evidence of

PPP in some countries. However, these results are not consistent across all papers. In addition, the empirical evidence from the panel data framework is also mixed. Hence, we are extending the panel unit root and cointegration tests discussed in Chapters 2, 3 and 4 and also look more specifically at the role of the 1997 currency crisis on PPP and the monetary models using data up to 2002 quarter 4.

5.3 Purchasing power parity and the monetary approach to exchange rate modelling

In this chapter, two major approaches are implemented to determine a long-run equilibrium between exchange rates and fundamentals: PPP and the monetary model. A brief summary of the underlying theories is presented, as follows.

5.3.1 Purchasing power parity

A plethora of theoretical and empirical models of exchange rate behaviour has been built around purchasing power parity (PPP). Under PPP, the equilibrium value of an exchange rate is determined by the change in the relative price level.

The PPP hypothesis has two major variants: the absolute PPP and the relative PPP hypotheses. According to the absolute PPP hypothesis, the nominal exchange rate between the currencies of two countries is proportional to a ratio of the foreign and domestic price levels; specifically:

$$S_t = \frac{P_t}{P_t^*} \qquad (5.1)$$

where S_t is the nominal exchange rate defined as the domestic currency price of foreign currency, P_t is the domestic price level, P_t^* is the foreign price level. The logarithmic transformation of equation (5.1) has the form:

$$s_t = p_t - p_t^* \qquad (5.2)$$

where s_t, p_t, p_t^* are the logarithms of S_t, P_t, P_t^*.

Next, the relative PPP hypothesis states that the exchange rate should bear a constant proportionate relationship to the ratio of national price level; in particular:

$$S_t = c(\frac{P_t}{P_t^*}) \qquad (5.3)$$

where c is a constant parameter. The logarithmic transformation has the form:

$$s_t = \alpha + p_t - p_t^* \qquad (5.4)$$

where s_t, p_t, p_t^* are the logarithms of S_t, P_t, P_t^*.

Under the relative PPP hypothesis, (s_t, p_t, p_t^*) have a cointegration relationship with (1, -1, 1) cointegrating vector. We can re-write the equation (5.4) in the relative price term (pr_t) as:

$$s_t = \alpha + pr_t \qquad (5.5)$$

where $pr_t = p_t - p_t^*$. In the bivariate system, the relative PPP hypothesis implies that (s_t, pr_t) have a cointegration relationship with (1, -1) cointegrating vector.

On the majority of occasions, the PPP hypothesis is restated in terms of real exchange rate (Q_t). The nature of deviations from PPP can be examined through real exchange rate because its logarithm, q_t, can be defined as deviations from PPP:

$$q_t \equiv s_t - p_t + p_t^* \qquad (5.6)$$

where q_t denotes logarithm of real exchange rate.

In this case, the rejection of the unit root null hypothesis in real exchange rate has been taken as evidence for the existence of PPP.

5.3.2 The monetary approach to exchange rate modelling

The monetary model is often used as the fundamental structural theory underlying exchange rate movements. In the monetary approach, the exchange rate is viewed as the relative price of two monies. Variants of the monetary model have been employed in many empirical studies in international macroeconomics, depending on the assumption with regard to the flexibility of price and exchange rate expectation. In this chapter, we focus on the flexible price monetary model with rational expectation.

The model relies on relative money market conditions based on the quantity theory of money, where fully flexible prices are determined by a monetary equilibrium between a stable real money demand and real money supply. Demand for log real balances is static and linearly related to log real income and the nominal interest rate. We suppose that PPP, which links the exchange rate to home and foreign price levels, and uncovered interest rate parity (UIP), which links home and foreign interest rates and the expected rate of exchange rate change, hold continuously. The above conditions can be expressed as follows:

$$m_t - p_t = \lambda y_t - \phi r_t \tag{5.7}$$

$$m_t^* - p_t^* = \lambda y_t^* - \phi r_t^* \tag{5.8}$$

$$m_t' - p_t' = \lambda y_t' - \phi r_t' \tag{5.9}$$

$$s_t = p_t - p_t^* \tag{5.10}$$

$$r_t - r_t^* = E_t(s_{t+1} - s_t) \tag{5.11}$$

where m_t denotes the logarithm of the money supply at time t; y denotes the logarithm of domestic real income; r denotes the domestic interest rate. The asterisks denote the foreign variables. The money demand income elasticity and interest rate semi-elasticity are denoted by λ and ϕ, respectively, which are assumed to be equal across countries. The prime denotes the domestic minus foreign variables. E_t denotes the expectation operator, conditional on information available at time t.

Solving equations (5.7) and (5.8) with respect to p_t and p_t^*, and then substituting into equation (5.10) yields:

$$s_t = m' - \lambda y_t' + \phi r_t' \tag{5.12}$$

Equation (5.12) is the basic equation of the flexible price monetary model, in which nominal exchange rate movements are driven by the relative excess supply of money. The changes in output levels or interest rates have their effects on the exchange rate indirectly through the effect on money demand.

The uncovered interest rate condition and the assumption that expectation about expected future spot rates is formed rationally, is introduced. Rearranging equations (5.11) and (5.12), we obtain:

$$s_t = \frac{1}{1+\phi} f_t + \frac{\phi}{1+\phi} E_t(s_{t+1}) \tag{5.13}$$

The simple flexible price monetary fundamental, denoted as f_t, is defined as:

$$f_t = m_t' - \lambda y_t' \tag{5.14}$$

By solving equation (5.13) forward, we obtain the forward-looking monetary exchange rate equation:

$$s_t = \frac{1}{1+\phi} \sum_{j=1}^{\infty} (\frac{\phi}{1+\phi})^j E_t(f_{t+j}) \tag{5.15}$$

where the transversality condition has been imposed:

$$\lim_{j \to \infty} \left[\frac{\phi}{1+\phi} \right]^j E_t(s_{t+j}) = 0 \tag{5.16}$$

Subtracting f_t from both side and rearranging yields:

$$s_t - f_t = \sum_{j=1}^{\infty} (\frac{\phi}{1+\phi})^j E_t(\Delta f_{t+j}) \tag{5.17}$$

Under the assumption that $m_t, y_t, m_t^*, y_t^*, f_t$ are first difference stationary ($I(1)$) variables, from equation (5.15), s_t is a nonstationary $I(1)$ series, and the right hand side of equation (5.17) is also stationary. Hence, a relationship between the log exchange rate (s_t) and fundamental variable (f_t) can be expressed as follows:

$$s_t = f_t = (m_t - m_t^*) - \lambda(y_t - y_t^*) \tag{5.18}$$

Equation (5.18) will be used as the basic equation for the flexible price monetary model with rational expectations. f_t may be interpreted as the long-run equilibrium value of the log of nominal exchange rate. According to an error-correction mechanism, the exchange rate might be expected to react to deviation from its fundamental value (z_t), defined as:

$$z_t = s_t - f_t \tag{5.19}$$

This framework will be used as one of the underlying economic theories in the remainder of this chapter.

We apply both the univariate and multivariate methods in testing for the validity of a long-run relationship between exchange rates and monetary fundamentals. In the univariate framework, the unit root tests for stationarity of deviations of exchange rate from its monetary fundamental value (z_t) are used. In the multivariate framework, the monetary model implies that exchange rates (s_t) and fundamental variables (f_t) are cointegrated with a cointegration vector equal to (1,-1) in the bivariate system, and (s_t, m_t', y_t') are cointegrated with a $(1,-1,\lambda)$ cointegrating vector in the trivariate system.

5.4 Empirical results

5.4.1 Data

In this chapter, the data are quarterly for the period between 1980:1 and 2002:4. We consider ten Asia Pacific countries, i.e. Australia (AU), Indonesia (ID), Japan (JP), Korea (KR), Malaysia (ML), New Zealand (NZ), the Philippines (PH), Singapore (SG), Thailand (TH) and Taiwan (TW). We have grouped the countries into two categories. The first group, denoted as the AS5 sub-panel, consists of five South-East Asian countries: Indonesia, Malaysia, the Philippines, Singapore and Thailand. The second group contains the remaining five Pacific rim countries: Australia, Japan, Korea, New Zealand and Taiwan, denoted as the AP5 sub-panel. We use the US dollar as the base currency for every country in the panels. This sample period is suggested by a number of recent studies, e.g. Wu and Chen (1999), Wang (2000), and Esaka (2003), to represent the relatively flexible exchange rate regimes in most countries in the panels (see Razzaghipour *et al.* (2001) and Esaka (2003) for the details of the exchange rate regimes in each country). The data are obtained from the International Monetary Fund (IMF)'s *International Financial Statistics* (IFS) CD-ROM, Datastream and the central bank of China. The nominal exchange rates ($s_{i,t}$) are end of period data expressed as price of domestic currencies per unit of US dollar obtained from the IFS line *ae*. However, the new Taiwan dollar exchange rate is extracted from the central bank of China database, due to the unavailability of the data from the IFS CD-ROM. The consumer price index (CPI), used as the proxy of price level ($p_{i,t}$), is also taken from the IFS line 64. The money supply data ($m_{i,t}$) is represented by the seasonally adjusted narrow money supply (M1), taken from the IFS line 34*s*. The IFS CD-ROM does not provide the CPI and

M1 data for Taiwan. Datastream is then used as the data source of these variables. However, as Taiwan's M1 data from Datastream is seasonal unadjusted, the X12 procedure is used to remove the seasonality effect. Finally, the real income variable ($y_{i,t}$) is measured by the quarterly real gross domestic product (GDP), taken from Datastream. As the real income data are seasonally unadjusted in a number of countries, we also apply the X12 procedure to adjust for the seasonal effect. In addition, the sample period of the quarter real GDP series varies across counties. The details of the sample spans and seasonal adjustment property of the real income data ($y_{i,t}$) in each country are presented in Table 5.1. All variables ($s_{i,t}$, $p_{i,t}$, $m_{i,t}$, $y_{i,t}$) are presented in log form.

Table 5.1 The sample spans and seasonal adjustment property of the real income data in Asia Pacific countries

Countries	Sample Period	Seasonal Adjustment
AS5		
Indonesia	1990:1-2002:4	No
Malaysia	1991:1-2002:4	No
Philippines	1981:1-2002:4	No
Singapore	1980:1-2002:4	Yes
Thailand	1993:1-2002:4	No
AP5		
Australia	1980:1-2002:4	Yes
New Zealand	1987:2-2002:4	Yes
Japan	1980:1-2002:4	Yes
Korea	1980:1-2002:4	Yes
Taiwan	1980:1-2002:4	No

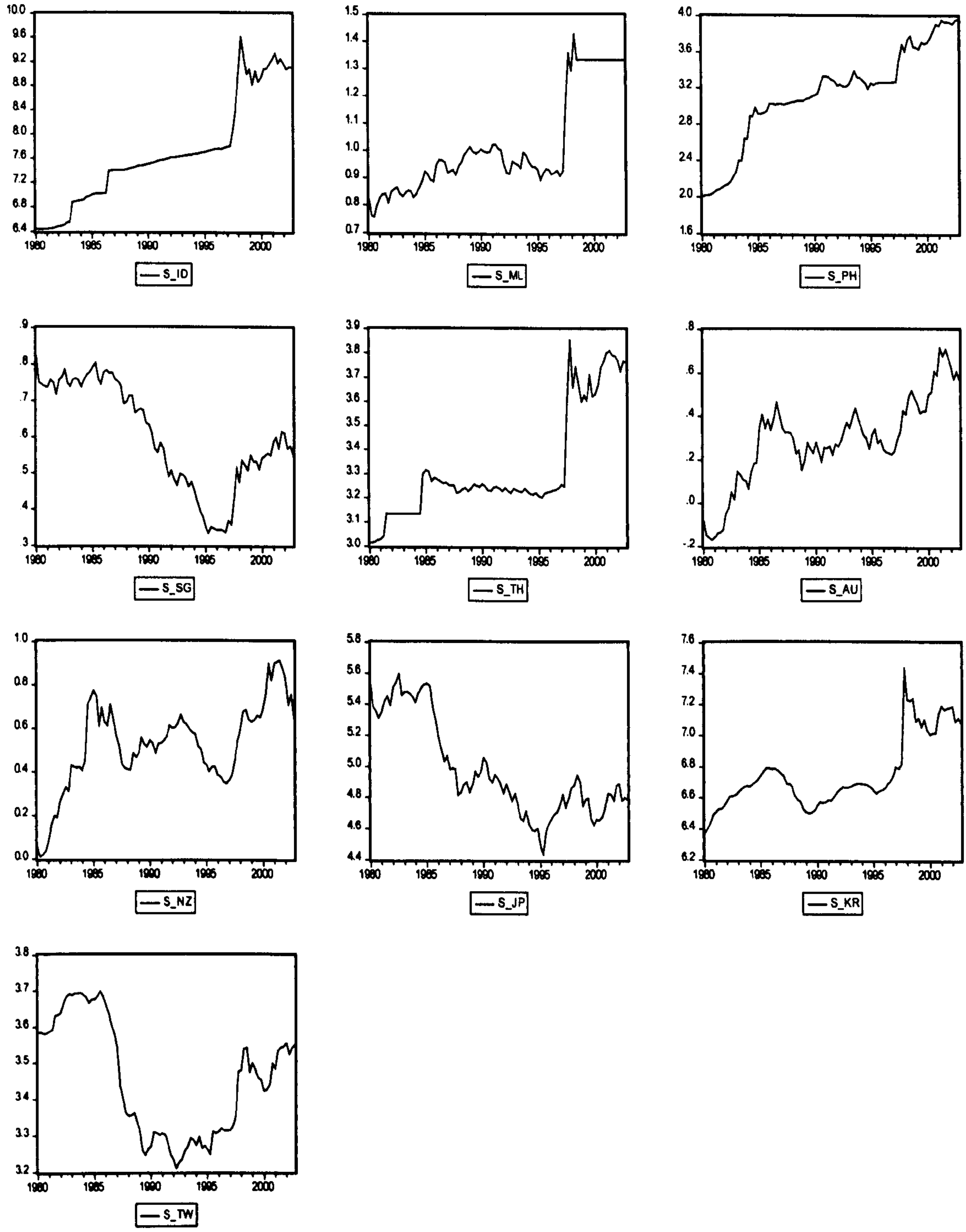
The monetary fundamental variable ($f_{i,t} = (m_{i,t} - m_{i,t}^{*}) - \lambda(y_{i,t} - y_{t}^{*})$) is constructed on the assumption that λ = 1. This value is used to calculate monetary fundamental in Mark (1995), Kilian (1999) and Taylor and Peel (2000). In addition, the empirical results from Dekle and Pradhan (1999) shows that the income elasticity of money demand in South-East Asian countries is close to unity. The real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$) are calculated as equation (5.6)($q_{i,t} \equiv s_{i,t} + p^{*}_{i,t} - p_{t}$) and (5.19)($z_{i,t} = s_{i,t} - (m_{i,t} - m^{*}_{i,t}) - \lambda(y_{i,t} - y^{*}_{i,t})$), respectively.

5.4.2 Empirical results of the single country time-series data

As a preliminary analysis, Figures 5.1 to 5.4 provide a graphical representation of nominal exchange rates ($s_{i,t}$), price levels ($p_{i,t}$), relative money supplies ($m'_{i,t} = m_{i,t} - m^{*}_{i,t}$) and relative real incomes ($y'_{i,t} = y_{i,t} - y^{*}_{i,t}$) for each country in our panel.

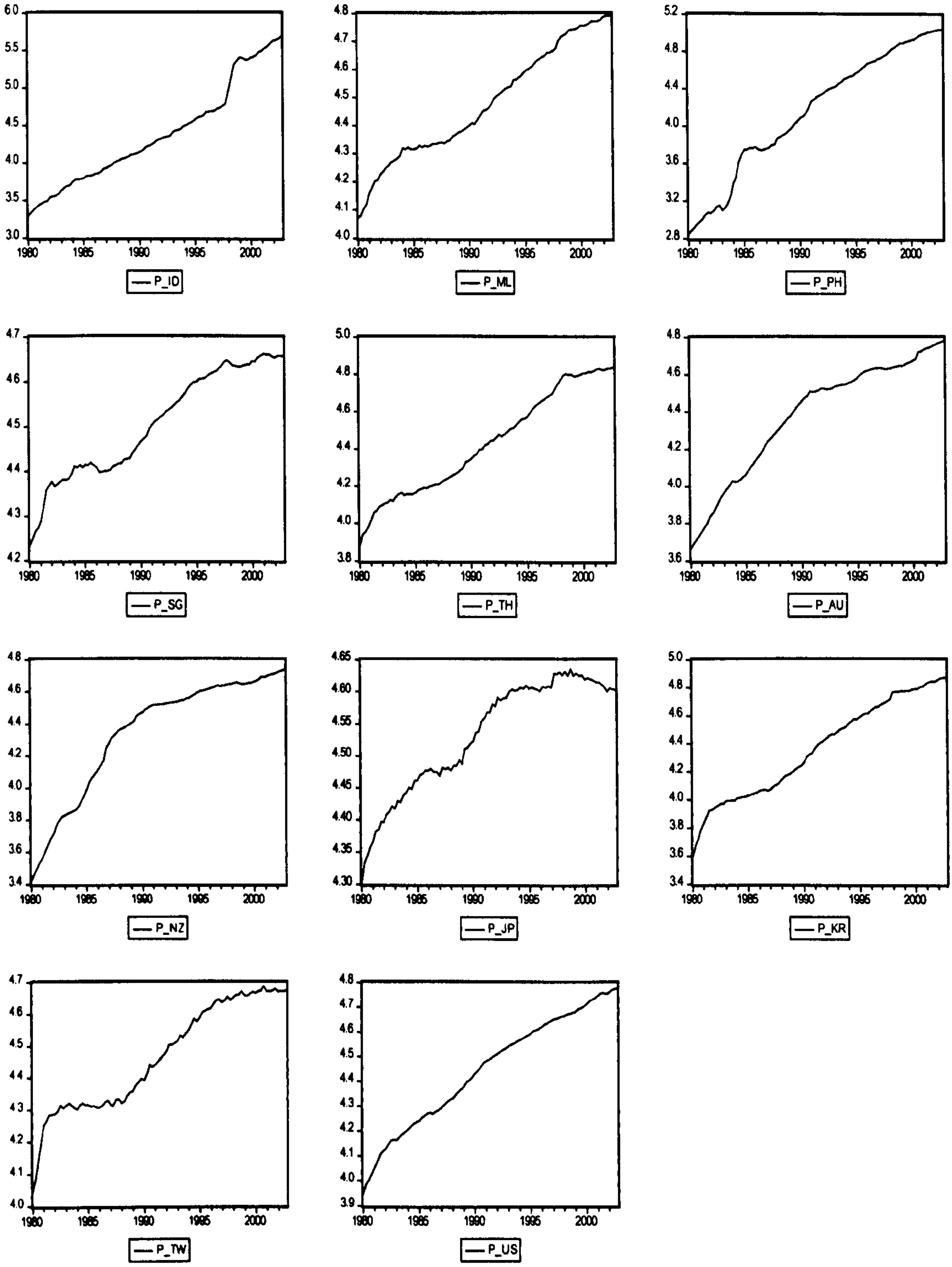
From Figures 5.1 to 5.4, evidence of the non-stationary property for each country is quite obvious, as all data ($s_{i,t}, p_{i,t}, m'_{i,t}, y'_{i,t}$) look persistent. However, nominal exchange rates ($s_{i,t}$) demonstrate greater volatility compared with the their fundamental determinants ($p_{i,t}, m'_{i,t}, y'_{i,t}$). The devaluation and change in exchange rate regimes, which regularly occur in the region, coupled with the effect of the 1997 currency crisis, resulted in the sudden changes in nominal exchange rates for Thai baht, Indonesian rupiah Malaysian ringgit after 1997:2 at the beginning of the 1997 East Asian currency; and for the Philippine peso, Singaporean dollar, Korean won and new Taiwan dollar after 1997:3 when the effect of the crisis spread throughout the region. For Australia, New Zealand and Japan, which have a stronger economic fundamental, the large scale of change in nominal exchange rates in 1997 is not observed (see Radelet and Sachs (1998) for discussion of the 1997 crisis). For price levels ($p_{i,t}$), the data look very persistent, and the presence of time trends is obvious across countries.

Figure 5.1 Nominal exchange rates ($s_{i,t}$) in Asia Pacific countries



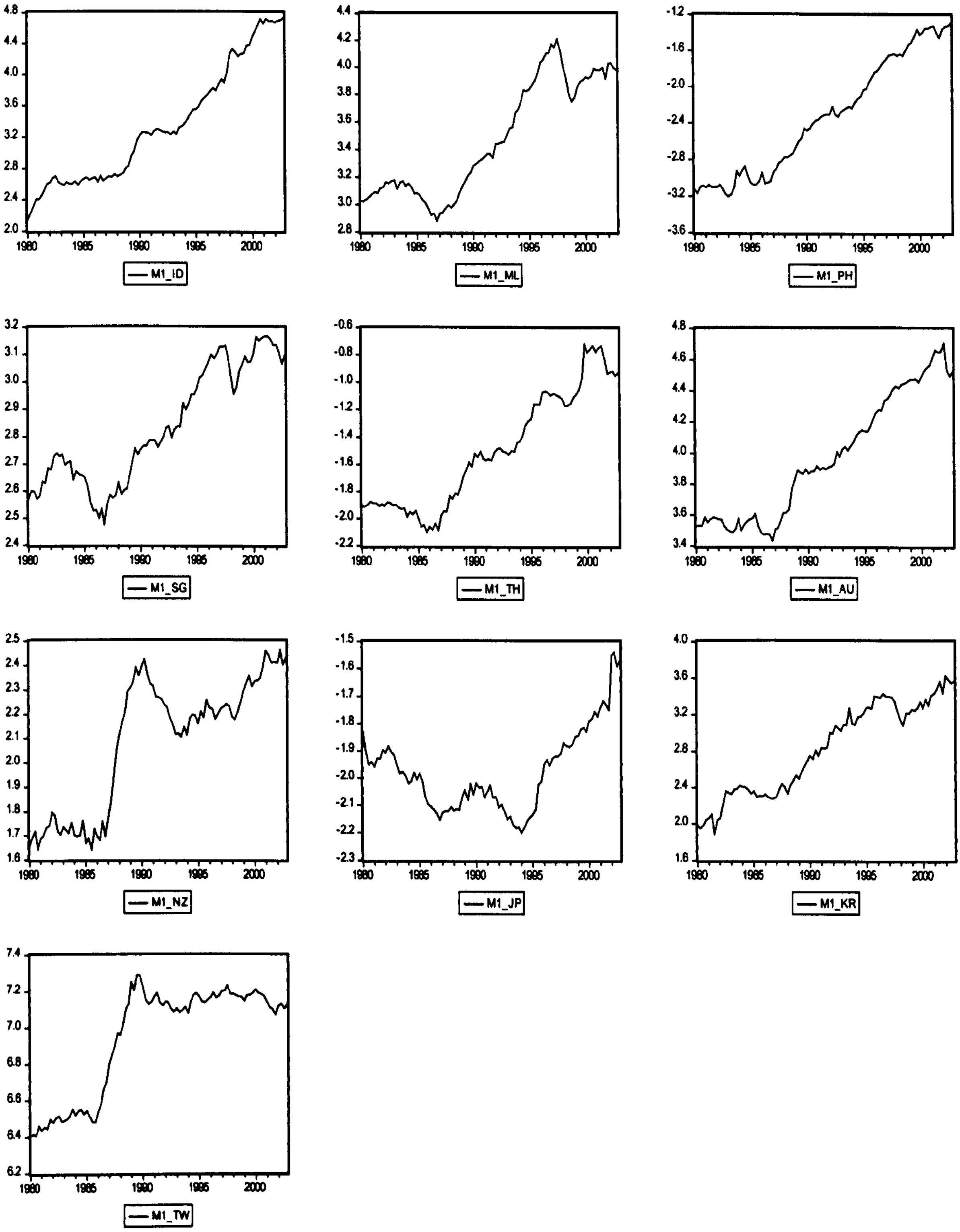
Note: ID, ML, PH, SG, TH, AU, NZ, JP, KR, TW denote Indonesia, Malaysia, the Philippines, Singapore, Thailand, Australia, New Zealand, Japan, Korea and Taiwan, respectively.

Figure 5.2 Price levels ($p_{i,t}$) in Asia Pacific countries



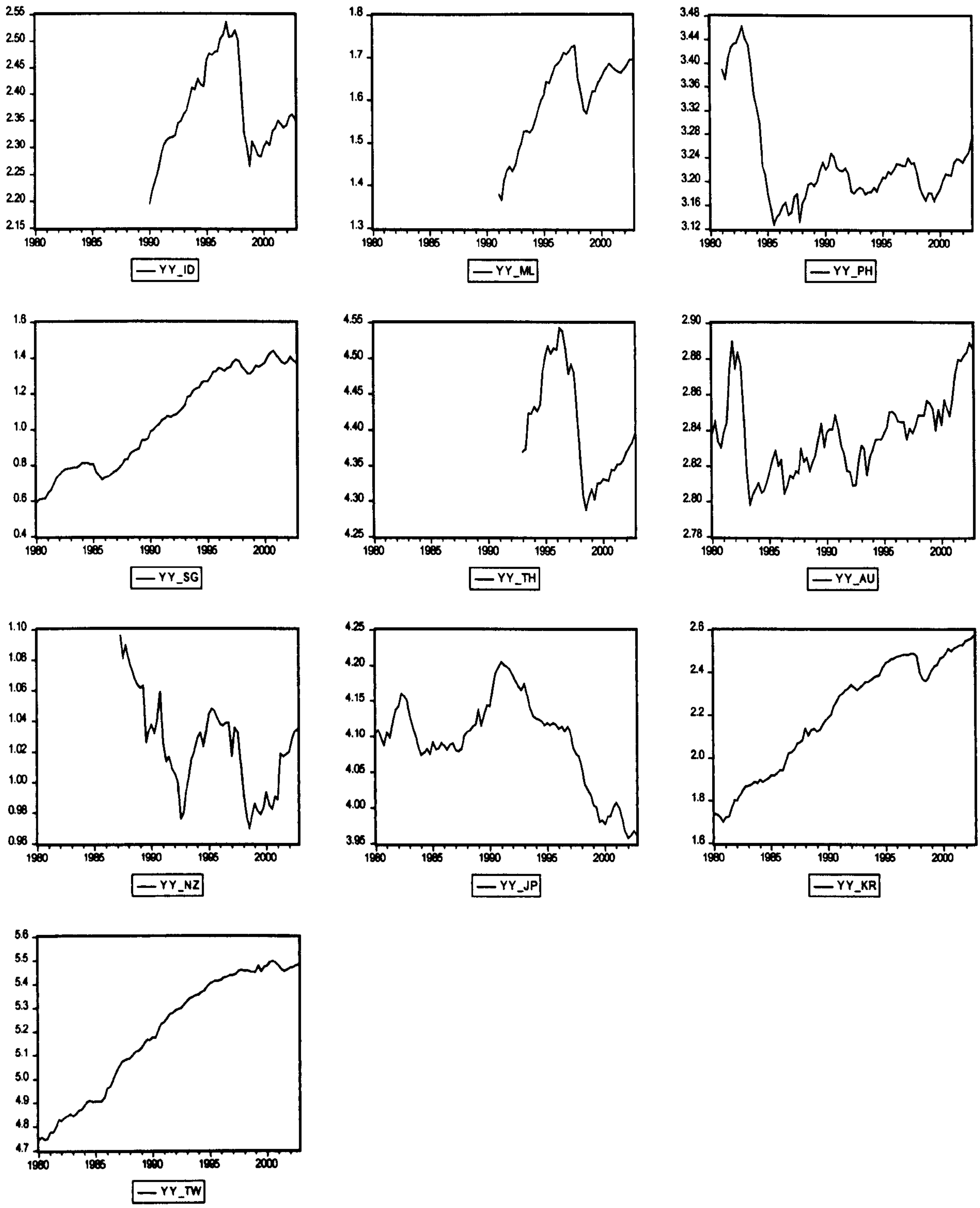
Note: See notes to Figure 5.1. US denotes the United States.

Figure 5.3 Relative money supplies ($m'_{i,t}$) in Asia Pacific countries



Note: See notes to Figure 5.1.

Figure 5.4 Relative real incomes ($y'_{i,t}$) in Asia Pacific countries

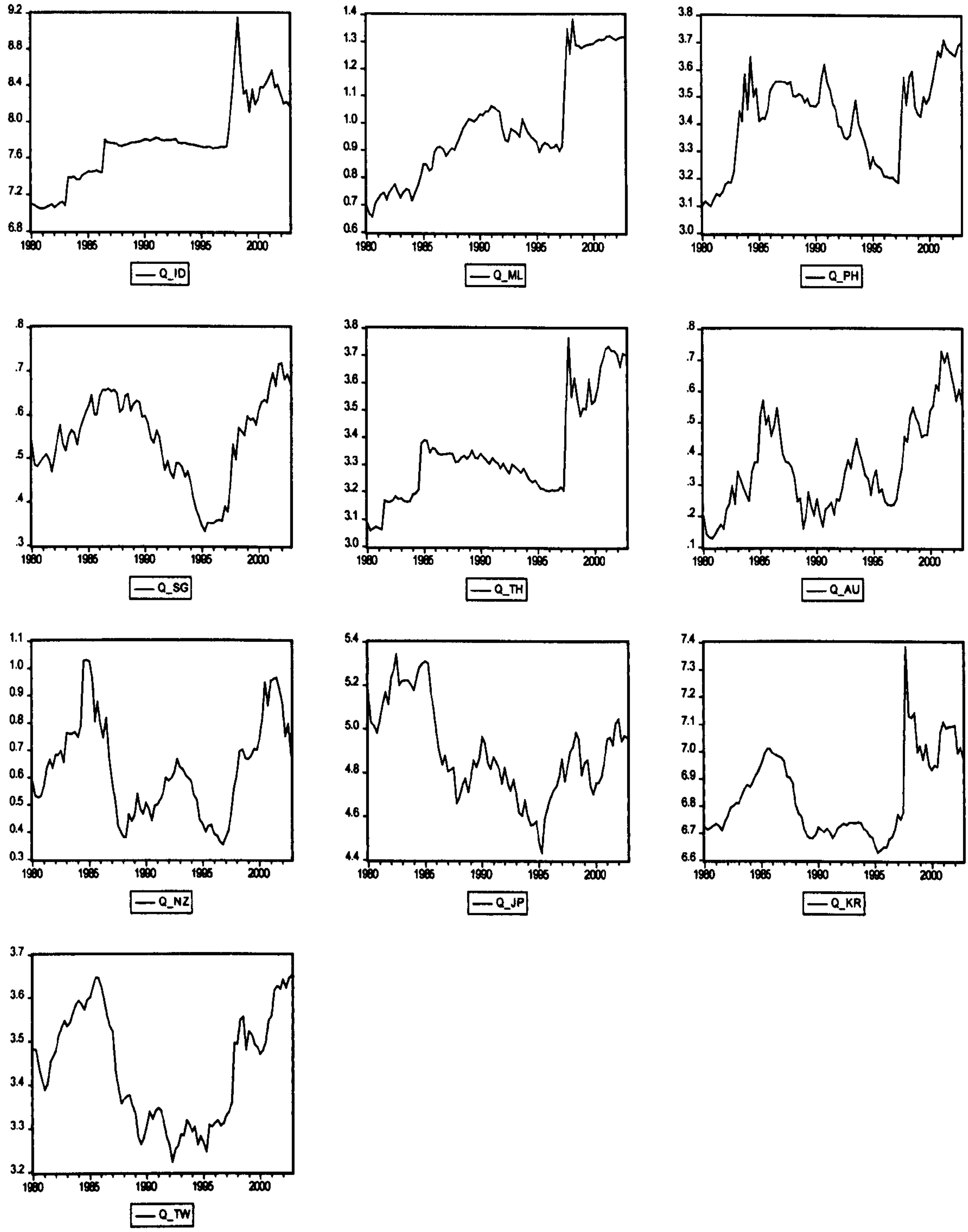


Note: See notes to Figure 5.1

For relative money supplies ($m'_{i,t}$), the presence of an upward trend is observable in the majority of countries. A sharp increase in $m'_{i,t}$ is found in Taiwan and New Zealand in 1986 and 1987, respectively, which may arise from the effects of the financial liberalisation in Taiwan and structural economic reform in New Zealand during those periods. For relative real incomes ($y'_{i,t}$), the upward trend is obvious for Singapore, Korea and Taiwan, which represent the fast-growing economies (compared with the US) in the period of study. For Indonesia, Malaysia and Thailand, the upward trend is also observable before 1997, when the economies also experienced a period of rapid growth. However, the effect of the 1997 crisis was particularly severe in these countries, causing them to go into recession after the crisis. The downward trend is observed in Japan after 1991, when the economy was in recession after the collapse of the bubble economy.

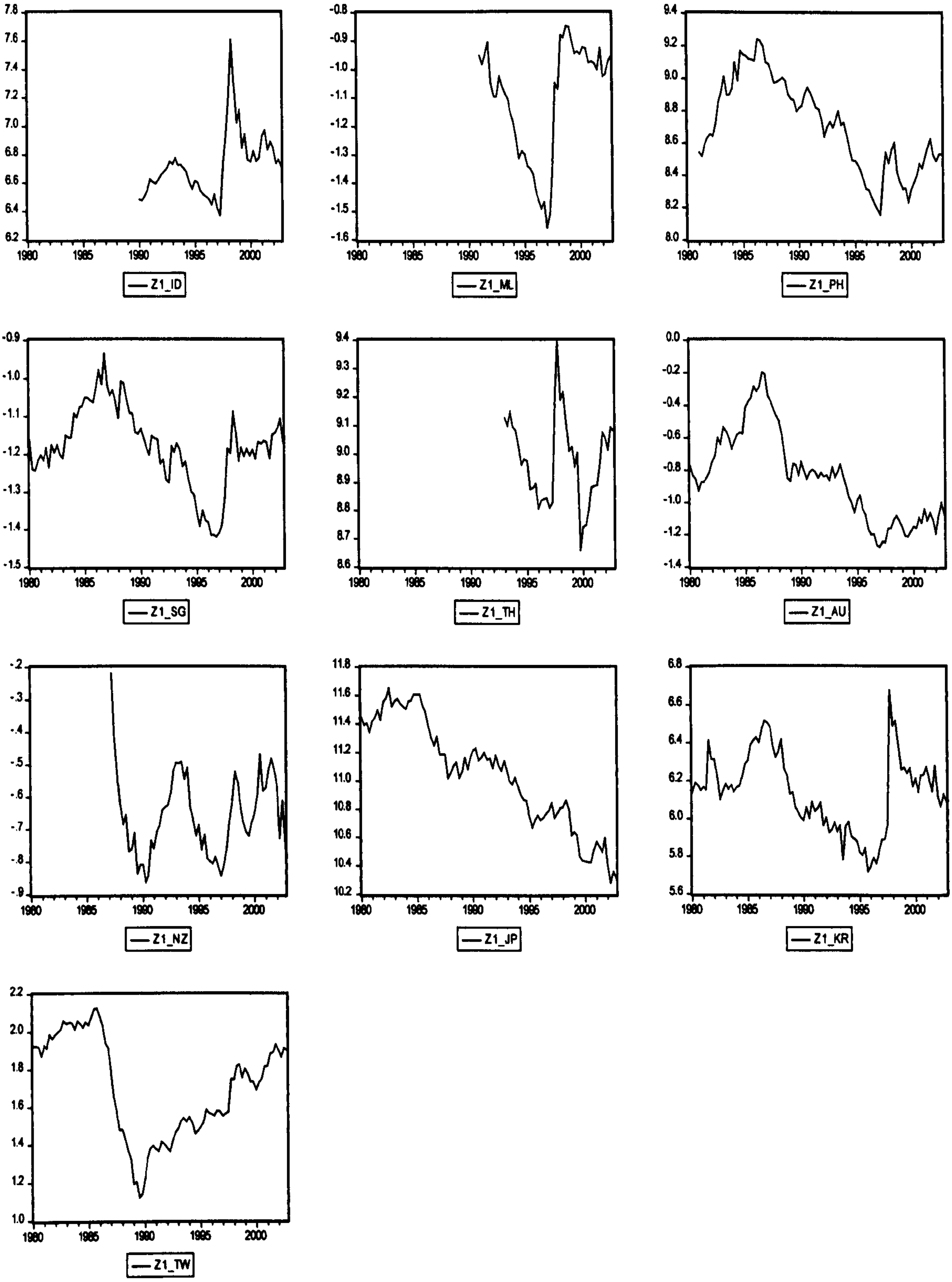
Next, we consider the graphical representation of real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$), presented in Figures 5.5 and 5.6, respectively. These figures provide evidence of weak mean reversion for each country in the panel, as all data look persistent. However, this may be the result of a slow speed of adjustment toward a long-run equilibrium. In addition, Figures 5.5 and 5.6 provide two interesting observations. First, the presence of time trends is evident in some countries, with the real exchange rates of Malaysia, the Philippines and Thailand depreciating over time. Second, the presence of a level shift is also observed in the period of the East Asian currency crisis (1997, quarter 2 to 3), which indicates the presence of a structural shift. Moreover, the data also indicate the possible structural breaks in the other periods of time in some of the countries.

Figure 5.5 Real exchange rates ($q_{i,t}$) in Asia Pacific countries



Note: See notes to Figure 5.1

Figure 5.6 Deviations from monetary fundamental ($z_{i,t}$) in Asia Pacific countries



Note: See notes to Figure 5.1

These observations raise two econometric issues in testing for the fundamental exchange rate theories. First, the traditional PPP hypothesis states that real exchange rate should reverse to its mean. The presence of time trends in the series implies depreciation (upward trend) or appreciation (downward trend) of real exchange rate over the period of time. Cuddingtion and Liang (2000) point out that either unit roots or deterministic time trends imply rejection of the PPP hypothesis. However, Lothian and Taylor (2000) argue that testing for PPP, allowing for the presence of time trends, provides strong evidence in support of PPP. The presence of time trends is possible in real exchange rates using the long-horizon data. The real effects, such as the Harrod-Balassa-Samuelson effect, would have made real exchange rates fall over the sample. In addition, Boyd and Smith (1998) point out that the measurement error problem in the data from developing countries may result in the presence of time trends in a long-run relationship. Second, the possibility of permanent shifts in real exchange rate can affect the performance of the tests for the existence of PPP and the monetary model. In Chapter 4, we noted that the presence of a structural change made it difficult to differentiate between a unit root process and a stationary process with level shifts. The effect of structural breaks in the empirical studies on PPP and the monetary model will be examined in Section 5.5.

Next, we apply the econometric method to investigate a long-run relationship between exchange rates and fundamentals. In the first step, the ADF tests are conducted to define the integration order of the variables ($s_{i,t}$, $p_{i,t}$, $m'_{i,t}$, $y'_{i,t}$). The ADF tests with constant are applied, to test for unit roots in nominal exchange rates ($s_{i,t}$), in which the presence of time trends is not obvious (see Figure 5.1). The price level ($p_{i,t}$), relative money supply ($m'_{i,t}$) and relative real income ($y'_{i,t}$) series are tested by the ADF tests with constant and trend. However, real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$) are tested, using both the ADF tests with only constant, as well as with constant and trend. The number of lagged difference terms is determined by the criterion suggested by Ng and Perron (1995) to ensure that there is no problem of serial correlation. This procedure is similar to that used to determine the number of augmented terms in the LM regression, discussed in Section 4.3. The maximum number of lag terms (k_{max}) is set equal to 9, which is approximately equal to $T^{1/2}$. The results of the ADF tests for the level and first difference of the s_t, $p_{i,t}$, $m'_{i,t}$ and $y'_{i,t}$ series are reported in Tables 5.2 and 5.3. The results show that all series are non-stationary in levels at the 5% significant level. When the series are first differenced, the null hypothesis of unit roots can be rejected for all series. Therefore, we conclude that all series ($s_{i,t}$, $p_{i,t}$, $m'_{i,t}$, $y'_{i,t}$) can be characterised as $I(1)$ process.

Table 5.2 The ADF results for the level of series

Countries	s^c	$p^{c,t}$	$m^{c,t}$	$y^{c,t}$
<u>AS5</u>				
Indonesia	-0.452 (4)	-1.226 (3)	-1.813 (2)	-1.985 (1)
Malaysia	-0.650 (0)	-2.450 (4)	-1.792 (4)	-2.397 (1)
Philippines	-1.554 (2)	-2.233 (2)	-2.846 (6)	-1.403 (1)
Singapore	-1.108 (3)	-2.377 (1)	-2.502 (6)	-1.753 (6)
Thailand	-0.855 (0)	-2.398 (4)	-2.595 (8)	-2.976 (5)
<u>AP5</u>				
Australia	-1.500 (0)	-2.070 (2)	-3.006 (3)	-2.722 (1)
New Zealand	-2.825 (6)	-2.024 (2)	-1.843 (2)	-2.832 (6)
Japan	-1.230 (3)	-0.422 (7)	-1.004 (0)	-1.480 (2)
Korea	-1.465 (3)	-2.528 (4)	-2.282 (5)	-1.447 (1)
Taiwan	-1.951 (9)	-2.323 (8)	-1.632 (5)	-2.576 (0)
United States		-1.644 (3)		

Note: The results are the ADF statistics of the level of variables. The figures in the parenthesis are the number of augmented term (p) in the ADF regressions. The 5% critical values of the ADF tests with constant (c) and with constant and trend (c,t) are equal to –2.894 and –3.461, respectively.

Table 5.3 The ADF results for the first difference of series

Countries	s^c	$p^{c,t}$	$m^{c,t}$	$y^{c,t}$
<u>AS5</u>				
Indonesia	-5.840 (3)	-4.690 (4)	-5.413 (0)	-4.695 (0)
Malaysia	-8.507 (0)	-4.265 (1)	-4.004 (3)	-4.702 (0)
Philippines	-5.034 (1)	-4.850 (5)	-9.216 (0)	-4.934 (7)
Singapore	-4.478 (2)	-4.377 (6)	-9.144 (0)	-4.441 (1)
Thailand	-5.545 (3)	-4.459 (4)	-4.486 (7)	-3.588 (0)
<u>AP5</u>				
Australia	-10.083 (0)	-3.935 (1)	-9.110 (0)	-6.198 (6)
New Zealand	-4.405 (3)	-3.799 (1)	-4.961 (1)	-7.768 (0)
Japan	-4.686 (2)	-5.689 (1)	-10.745 (0)	-4.022 (3)
Korea	-11.436 (0)	-4.252 (3)	-5.577 (3)	-7.712 (0)
Taiwan	-4.597 (3)	-6.669 (3)	-4.727 (1)	-8.519 (0)
United States		-4.004 (2)		

Note: See note to Table 2. The results are the ADF statistics of the first differences of variables.

Table 5.4 Empirical results of the ADF tests for the level of real exchange rates ($q_{i,t}$)

and deviations from monetary fundamental ($z_{i,t}$)

Countries	q^c	$q^{c,t}$	z^c	$z^{c,t}$
<u>AS5</u>				
Indonesia	-1.488 (4)	-2.556 (4)	-2.111 (4)	-3.609* (3)
Malaysia	-0.841 (0)	-2.191 (0)	-1.207 (0)	-1.570 (0)
Philippines	-2.177 (2)	-2.302 (2)	-1.053 (5)	-3.001 (5)
Singapore	-1.917 (5)	-1.883 (5)	-2.025 (3)	-2.299 (3)
Thailand	-1.240 (0)	-2.264 (0)	-3.813* (7)	-3.731* (7)
<u>AP5</u>				
Australia	-1.474 (0)	-1.959 (0)	-0.883 (0)	-2.979 (3)
New Zealand	-3.008* (8)	-3.030 (8)	-3.348* (8)	-3.292 (8)
Japan	-2.062 (4)	-2.374 (4)	-0.141 (4)	-4.225* (4)
Korea	-2.096 (3)	-2.212 (3)	-1.859 (8)	-1.849 (8)
Taiwan	-1.530 (9)	-1.172 (9)	-2.183 (9)	-1.850 (9)

Note: See note to Table 5.2. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

Table 5.5 Empirical results of the ADF tests for the first difference of real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	q^{c}	$q^{c,t}$	z^{c}	$z^{c,t}$
<u>AS5</u>				
Indonesia	-6.489* (3)	-6.477* (3)	-4.318* (3)	-4.280* (3)
Malaysia	-8.712* (0)	-8.660* (0)	-5.453* (0)	-5.494* (0)
Philippines	-5.751* (1)	-5.720* (1)	-4.500* (4)	-4.541* (4)
Singapore	-4.541* (2)	-4.511* (2)	-11.118* (0)	-11.040* (0)
Thailand	-9.559* (0)	-9.508* (0)	-6.124* (0)	-6.104* (0)
<u>AP5</u>				
Australia	-9.941* (0)	-9.876* (0)	-8.635* (0)	-8.624* (0)
New Zealand	-4.058* (3)	-4.022* (3)	-4.938* (1)	-4.778* (1)
Japan	-4.755* (2)	-4.717* (2)	-4.181* (5)	-4.241* (5)
Korea	-12.213* (0)	-12.143* (0)	-4.036* (7)	-4.025* (7)
Taiwan	-5.071* (3)	-5.142* (3)	-3.940* (1)	-4.029* (1)

Note: See note to Table 5.2. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

Table 5.6 Empirical results of the LM tests for the level of real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	$q^{c,t}$	$z^{c,t}$
<u>AS5</u>		
Indonesia	-2.575 (4)	-3.619* (3)
Malaysia	-2.207 (0)	-1.200 (0)
Philippines	-1.136 (6)	-1.066 (5)
Singapore	-1.781 (5)	-2.042 (3)
Thailand	-2.240 (0)	-3.851* (7)
<u>AP5</u>		
Australia	-1.974 (0)	-1.308 (0)
New Zealand	-2.866 (8)	-1.792 (8)
Japan	-2.313 (4)	-3.669* (4)
Korea	-2.189 (3)	-1.881 (8)
Taiwan	-1.734 (9)	-2.175 (9)

Note: The results are the LM statistics of the level of variables. The figures in the parenthesis are the number of augmented term (p) in the LM regressions. The 5% critical value of the LM test (with constant and trend (c,t)) is equal to –3.07. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

Next, we examine the mean reversion property of real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$). The results of the ADF tests are presented in Tables 5.4 and 5.5 for the level and first difference of series, respectively. In addition, we apply the LM unit root test of Schmidt and Phillips (1992) (SP), as discussed in Chapter 4. SP suggest that this LM unit root test is more powerful than the ADF test when the null hypothesis is close to being true. The results of the LM unit root test are reported in Table 5.6.

We first consider stationarity of real exchange rate ($q_{i,t}$), which implies the existence of long-run PPP. The results from Tables 5.4 to 5.6 appear to confirm that real exchange rates are non-stationary ($I(1)$) for most countries. Using both the ADF and LM tests, we cannot reject the null hypothesis of unit roots in real exchange rates for all countries, with the exception of New Zealand, where we reject the null hypothesis at the 0.05 significant level, using the ADF tests with constant.

Turning to deviations from monetary fundamental ($z_{i,t}$), stationarity of $z_{i,t}$ can be taken as evidence of mean reversion according to the monetary model. The results from Tables 5.4 and 5.6 show that there is more evidence in favour of the monetary model than in support of the PPP hypothesis. Using the ADF and LM tests, we reject unit roots in $z_{i,t}$ for four countries (Indonesia, Thailand, New Zealand and Japan) at the 0.05 significant level. Comparing the results from the ADF and LM tests (with constant and trend), both the ADF and LM tests provide similar results in unit root testing for both $q_{i,t}$ and $z_{i,t}$. Using the ADF test, we find marginally more evidence in support of stationarity in both $q_{i,t}$ and $z_{i,t}$ than for the LM test. Overall, the results from the ADF and LM tests on the level and first difference of $q_{i,t}$ and $z_{i,t}$ show that the $q_{i,t}$ and $z_{i,t}$ series ($s_{i,t}$, $p_{i,t}$, $m'_{i,t}$, $y'_{i,t}$) are still characterised as $I(1)$ process for most countries.

Next, we consider the presence of cross-sectional dependence in the errors. In Chapter 2, we discussed the effect of cross-correlations in testing for unit roots and cointegration in panel data. To illustrate this correlation, we present estimates of the cross-correlation matrices between the residuals of the ADF tests for $q_{i,t}$ and $z_{i,t}$ in Tables 5.7 and 5.8, respectively.

Table 5.7 The estimated cross-correlation matrix of the residuals from the ADF regressions for real exchange rates ($q_{i,t}$)

		AS5					AP5				
		ID	ML	PH	SG	TH	AU	NZ	JP	KR	TW
AS5	ID	1.000	0.421	0.212	0.422	0.333	0.192	0.257	0.232	0.313	0.222
	ML		1.000	0.436	0.606	0.762	0.249	0.210	0.249	0.448	0.353
	PH			1.000	0.231	0.486	0.235	0.137	-0.026	0.368	0.233
	SG				1.000	0.561	0.353	0.340	0.613	0.399	0.348
	TH					1.000	0.282	0.253	0.295	0.530	0.366
AP5	AU						1.000	0.481	0.204	0.273	0.086
	NZ							1.000	0.322	0.138	0.071
	JP								1.000	0.142	0.242
	KR									1.000	0.543
	TW										1.000

Note: The figures are the values of the cross-correlations in the residuals of the ADF regressions for real exchange rates in the panel.

Table 5.8 The estimated cross-correlation matrix of the residuals from the ADF regressions for deviations from monetary fundamental ($z_{i,t}$)

		AS5					AP5				
		ID	ML	PH	SG	TH	AU	NZ	JP	KR	TW
AS5	ID	1.000	0.453	0.224	0.300	0.438	0.057	0.206	0.127	0.182	0.058
	ML		1.000	0.309	0.307	0.471	0.183	0.299	0.056	0.401	0.287
	PH			1.000	0.343	0.282	0.223	0.244	0.130	0.263	0.402
	SG				1.000	0.212	0.225	0.175	0.333	0.381	0.296
	TH					1.000	0.099	0.241	0.050	0.273	0.132
AP5	AU						1.000	0.172	0.095	0.192	0.288
	NZ							1.000	0.130	0.206	0.174
	JP								1.000	0.142	0.218
	KR									1.000	0.310
	TW										1.000

Note: The figures are the values of the cross-correlations in the residuals of the ADF regressions for deviations from monetary fundamental in the panel.

The results from Tables 5.7 and 5.8 show that there are strong correlations among several countries, in particular, those of South-East Asian countries. For example, the correlations between the Malaysian ringgit/Thai baht and Malaysian ringgit/Singaporean dollar real exchange rates are equal to 0.762 and 0.606, respectively. The other countries in the AP5 panel have somewhat lower correlations. Australia and New Zealand are lowly correlated with the other countries. The presence of cross-correlation in testing for unit roots of $q_{i,t}$ and $z_{i,t}$ can arise from several factors. For example, the effect of common shocks (e.g. currency crisis) and strong trading ties between countries in our panel are possible causes of these cross-correlations. In addition, the use of same base currency (US dollar) in the construction of $q_{i,t}$ and $z_{i,t}$ leads to the cross-correlations because of the inclusion of common components (p_t^*, m_t^*, y_t^*) across countries. Therefore, the method that takes account of these cross-correlations is then applied to unit root testing for $q_{i,t}$ and $z_{i,t}$. In Chapter 2, the simulation results show that in the presence of cross-sectional dependence in the errors, the SURADF test can improve the power to reject the unit root null hypothesis over the standard ADF test. In addition, the SURADF test estimated in the small panel (N=5) has better size and power properties than that of the larger panel (N=10). Therefore, we estimate separately the SURADF tests for the countries in the AS5 and AP5 panels. The empirical results of the SURADF tests for $q_{i,t}$ and $z_{i,t}$ are reported in Table 5.9. Using the SURADF test, stationarity of $q_{i,t}$ ($z_{i,t}$) can be found only for New Zealand (Thailand, New Zealand and Japan). These results are similar to those of the standard ADF test. There is no improvement in the empirical results by applying the SURADF test over the standard ADF test.

Table 5.9 Empirical results of the SURADF tests for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	q^c	$q^{c,t}$	z^c	$z^{c,t}$
AS5				
Indonesia	-1.940	-3.203	-2.349	-3.765
	(-3.251)	(-4.029)	(-3218)	(-3.950)
Malaysia	-2.329	-4.115	-1.317	-2.212
	(-3.597)	(-4.642)	(-3.037)	(-3.818)
Philippines	-2.553	-3.280	-1.137	-3.503
	(-3.266)	(-4.076)	(-3.147)	(-3.927)
Singapore	-1.611	-3.249	-2.116	-2.852
	(-3.295)	(-4.143)	(-3.245)	(-3.990)
Thailand	-2.751	-4.899*	-4.864*	-5.462*
	(-3.594)	(-4.576)	(3.182)	(3.763)
AP5				
Australia	-1.620	-2.339	-0.745	-2.847
	(-3.263)	(-4.018)	(-2.803)	(-3.763)
New Zealand	-3.693*	-4.209*	-3.475*	-3.360
	(-3.207)	(-3.946)	(-3.077)	(3.745)
Japan	-2.888	-3.714	-0.085	-4.538*
	(-3.162)	(-3.892)	(-2.784)	(-3.851)
Korea	-2.410	-2.506	-2.167	-2.038
	(-3.294)	(-4.012)	(-3.048)	(-3.984)
Taiwan	-2.097	-1.845	-3.056	-2.403
	(-3.230)	(-3.943)	(-3.128)	(-4.071)

Note: The results are the ADF statistics of the level of variables estimated by the SUR method. The figures in the parenthesis are the 5% critical values of the tests, which are calculated from the bootstrap method. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

The weak evidence of PPP (monetary model) stationarity in Tables 5.4, 5.6 and 5.9, may be due to cointegration between exchange rate and price levels (monetary fundamental) with coefficients other than unity. In the light of this, the multivariate cointegration tests are applied to investigate the existence of long-run PPP and the monetary model without the imposition of unitary parameter restrictions. The findings of the bivariate and trivariate Engle and Granger (1987) two-step cointegration test (E-G) and the Johansen Likelihood Ratio test (*JLR*) are presented in Tables 5.10 and 5.11 for the PPP hypothesis and the monetary model, respectively. For the E-G test, the number of augmented terms in the ADF regression testing for unit roots on the residuals for each country is chosen by the Ng and Perron

(1995) method, similar to that of the univariate ADF test. For the *JLR* test, the number of lags in VARs is determined by the LR test for the significance of last augmented lags in all regressions. In addition, we perform the diagnostic tests for the residual in VARs. The results are presented in Appendix A. We do not reject the null hypothesis of no serial-correlation in the errors for all countries and of homoskedasticity for majority of countries in our panel. However, the null hypothesis of normality in the residuals is rejected for most countries. This may be due to the effect of outliers. Even though the Johansen method uses a Gaussian likelihood, the asymptotic properties of cointegration only depend on the errors being identical independent distributed. Therefore, normality failures do not have serious consequence for the cointegration properties of the data (see Johansen (1995))

We first consider the PPP hypothesis. Using the E-G test, there is no evidence of long-run PPP for any countries in our panel at the 5% significant level. However, the results from the *JLR* test indicate that for all countries in our panel, excluding the Philippines and Malaysia, the null hypothesis of no cointegration ($r = 0$) is rejected at the 5% significant level, based on the trace statistics using the trivariate system. In addition, the null hypothesis of one cointegrating vector ($r \leq 1$) is also rejected for Australia, New Zealand, Korea and Taiwan, which implies that there may exist more than one cointegrating vector in these cases. However, the results from the bivariate system provide weaker evidence, as the null hypothesis can be rejected only for Singapore and Taiwan.

Table 5.10 The Engle-Granger (E-G) and Johansen likelihood-ratio (*JLR*) test results for the PPP hypothesis

Countries	Bivariate system (k=2)			Trivariate system (k=3)			
	E-G	*JLR*		*E-G*	*JLR*		
		$r=0$	$r\leq 1$		$r=0$	$r\leq 1$	$r\leq 2$
AS5							
Indonesia	-2.382 (4)	15.200	5.475 (5)	-2.621 (4)	36.967*	17.948	5.192 (5)
Malaysia	-0.601 (0)	9.245	1.588 (5)	-2.148 (0)	34.044	17.931	7.058 (5)
Philippines	-2.292 (2)	11.704	2.471 (3)	-2.299 (2)	29.934	13.431	3.669 (4)
Singapore	-1.937 (5)	39.235*	2.422 (4)	-1.938 (3)	55.231*	14.906	2.923 (4)
Thailand	-1.916 (7)	10.438	1.444 (3)	-2.163 (0)	37.657*	19.689	8.332 (4)
AP5							
Australia	-1.496 (0)	17.488	3.565 (4)	-2.364 (2)	45.283*	22.558*	8.445 (6)
New Zealand	-3.308 (8)	19.941	4.729 (3)	-3.099 (8)	36.605*	21.183*	9.850* (4)
Japan	-2.501 (4)	16.178	3.043 (5)	-3.784 (4)	39.633*	19.302	3.188 (5)
Korea	-2.191 (3)	15.002	2.314 (2)	-2.208 (3)	41.699*	22.903*	7.535 (4)
Taiwan	-1.481 (9)	32.104*	1.156 (5)	-1.481 (9)	59.910*	33.865*	11.234* (5)

Note: The results are the ADF statistics of the E-G test and the trace statistics of the *JLR* test. The figures in the parenthesis are the number of augmented term (p) in the ADF regressions for the E-G test and the number of lags in the VAR for the *JLR* test. The 5% critical values of the E-G tests are equal to –3.403 and –3.835 for the bivariate and trivariate systems, respectively. The 5% critical values of the *JLR* test for r=0 are equal to 19.96 and 34.91 in the bivariate and trivariate systems, respectively. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

Table 5.11 The Engle-Granger (E-G) and Johansen likelihood-ratio (*JLR*) test results for the monetary model

Countries	Bivariate system (k=2)			Trivariate system (k=3)			
	E-G	*JLR*		*E-G*	*JLR*		
		$r=0$	$r\leq 1$		$r=0$	$r\leq 1$	$r\leq 2$
AS5							
Indonesia	-3.825* (3)	17.241	5.056 (4)	-3.754 (4)	43.774*	20.761*	8.199 (4)
Malaysia	-0.857 (0)	15.174	7.211 (1)	-1.568 (0)	41.314*	15.119	7.357 (2)
Philippines	-2.542 (0)	26.876*	5.903 (1)	-2.460 (2)	57.458*	19.908	2.289 (6)
Singapore	-2.023 (0)	20.319*	7.322 (6)	-1.610 (3)	39.725*	17.855	4.612 (6)
Thailand	-3.018 (7)	10.395	4.800 (1)	-2.818 (0)	33.240	9.711	3.731 (1)
AP5							
Australia	-1.989 (0)	12.149	2.614 (1)	-1.982 (0)	29.615	6.913	1.981 (1)
New Zealand	-2.765 (8)	23.666*	4.852 (3)	-2.302 (8)	36.084*	15.743	4.198 (3)
Japan	-1.258 (3)	21.816*	2.630 (2)	-1.351 (4)	33.047	10.872	1.963 (2)
Korea	-2.470 (0)	12.677	3.103 (2)	-2.324 (0)	40.076*	12.424	3.908 (2)
Taiwan	-1.811 (3)	14.206	2.253 (6)	-2.167 (3)	56.012*	15.944	2.787 (6)

Note: See note to Table 5.10.

With regard to the monetary model, we find similar evidence to that of the PPP hypothesis. The results from the residual-based E-G test provide evidence of long-run PPP only for Indonesia. However, applying the *JLR* test, the null hypothesis of no cointegration can be rejected for all countries in our panel, with the exception of Thailand, Australia and Japan. In addition, the results from the bivariate system provide weaker evidence of a cointegration relationship.

Subsequently, we are able to compare empirical results in this section with the simulation results from Chapters 2 and 3. In those chapters, we demonstrate that the ADF and E-G tests have low power to reject the null hypothesis in moderate sample sizes (*T*) when the series are persistent and the adjustment process towards a long-run equilibrium is slow. The power of the E-G test is extremely low, especially in the trivariate system. The *JLR* test is significantly more powerful than the univariate ADF test and the E-G test in these conditions (see Tables 2.2 and 3.2). Our empirical results in this section provide the results, which are not inconsistent with those simulation results. The empirical evidence of long-run relationships for both PPP and the monetary model is strongest when we apply the *JLR* test. The results from the univariate ADF test and the multivariate E-G test provide only little supporting evidence for PPP and the monetary model.

5.4.3 Empirical results of the panel data tests

In this section, we use the panel data methodology to investigate empirical evidence for PPP and the monetary model in the AS5 and AP5 panels. In addition, we consider a panel that combines both the AS5 and AP5 panels, denoted as ALL. We first consider the panel IPS, MW, SURIPS, CIPS and LM tests, to check for unit roots in $q_{i,t}$ and $z_{i,t}$, the results of which are presented in Tables 5.12 and 5.13. We also report the bootstrap critical values, which are calculated from the method discussed in Section 2.3, using 10,000 replications.

Table 5.12 Panel unit root test results for real exchange rates ($q_{i,t}$)

Panel	IPS		MW		SURIPS		CIPS		LM
	c	c,t	c	c,t	c	c,t	c	c,t	c,t
AS5	-0.066	-0.249	7.995	8.152	-1.867	-4.505*	-1.912	-2.582	-0.194
	(-1.884)	(-2.096)	(19.278)	(20.194)	(-3.153)	(-4.281)	(-2.543)	(-3.044)	
AP5	-1.394	-0.089	15.118	9.145	-2.659*	-2.211	-2.195	-2.697	-1.132
	(-1.720)	(-1.772)	(18.486)	(18.308)	(-2.481)	(-3.083)	(-2.543)	(-3.044)	
ALL	-1.045	-0.238	23.113	17.298	-4.220	-6.411*	-2.152	-2.662	-0.940
	(-1.993)	(-2.257)	(33.299)	(35.076)	(-4.249)	(-5.867)	(-2.329)	(-2.836)	

Note: The results are the panel statistics for the AS5, AP5 and ALL panels. The figures in the parenthesis are the 5% critical values of the tests, which are calculated from the bootstrap method. The 5% asymptotic critical value of the panel IPS and LM tests is equal to −1.645, while those of the MW test are equal to −18.307 and 31.410 for N = 5 (AS5, AP5) and 10 (ALL), respectively. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

Table 5.13 Panel unit root test results for deviations from monetary fundamental($z_{i,t}$)

Panel	IPS		MW		SURIPS		CIPS		LM
	c	c,t	c	c,t	c	c,t	c	c,t	c,t
AS5	-1.422	-1.947*	17.685	20.241*	-2.182	-3.817*	-1.634	-2.351	-1.634
	(-1.692)	(-1.705)	(18.470)	(18.422)	(-2.374)	(-2.885)	(-2.554)	(-3.075)	
AP5	-0.567	-2.023*	14.810	22.018*	-1.113	-2.550	-1.330	-2.673	-1.035
	(-1.079)	(-1.702)	(-15.753)	(18.479)	(-1.541)	(-2.787)	(-2.540)	(-3.055)	
ALL	-1.409	-2.806*	32.495	42.259*	-2.692	-4.740	-1.742	-2.503	-1.893*
	(-1.990)	(-2.376)	(-33.219)	(-35.684)	(4.270)	(-5.887)	(-2.333)	(-2.838)	

Note: See Note to Table 5.12

We first consider evidence for the PPP hypothesis. The results from Table 5.12 show that the panel IPS, MW, CIPS and LM statistics cannot reject the unit root null hypothesis at the 5% significant level for the AS5, AP5 and ALL panels, using both the asymptotic and bootstrap critical values. However, using the SUR method, we reject the null hypothesis of unit roots in heterogeneous panels for any panel. These results are sensitive to the presence of the time trend. For the AS5 and ALL panels, we can reject the unit root null hypothesis when the specification of the ADF tests with constant and trend is used across countries. For the AP5 panel, the empirical evidence of PPP is found when we consider the SURIPS test estimated without trend.

Turning to the monetary model, the results are also sensitive to the choice of panel test statistic and the time trend specification. For every panel (AS5, AP5 and ALL), the results from the IPS and MW tests (with intercept and trend) reject the unit root null hypothesis for deviations from monetary fundamental ($z_{i,t}$), using both the asymptotic and bootstrap critical values. However, there is no evidence of stationarity of $z_{i,t}$ when we consider the tests without trend. Evidence of the monetary model is weakened when we apply the SUR method, as the unit root null hypothesis can be rejected only in the AS5 panel. Using the CIPS test, we cannot reject the null hypothesis in any panels, a result similar to that of the PPP hypothesis. Finally, the results from the panel LM test show that we reject the null hypothesis for the ALL panel alone, and only marginally cannot reject the null hypothesis in the AS5 panel at the 0.05 significant level.

With regard to the panel cointegration tests, the empirical results of the residual-based tests of IPS, MW, CIPS and the likelihood-based test of LLL are presented in Tables 5.15 and 5.16 for the PPP hypothesis and the monetary model, respectively. At the 5% significant level, evidence of a long-run relationship is found only from the results of the panel LLL rank test for both the PPP hypothesis and the monetary model. We cannot find any significant evidence to validate the PPP hypothesis and monetary model, using the residual-based tests of IPS, MW and CIPS. These panel data results are similar to those of the time-series data reported in Tables 5.10 and 5.11. The empirical evidence of cointegration relationships is still evident only for the likelihood-based cointegration test.

Table 5.14 Panel cointegration test results for the PPP hypothesis

Panel	Bivariate system (k=2)				Trivariate system (k=3)			
	IPS	MW	LLL	CIPS	IPS	MW	LLL	CIPS
AS5	0.368	7.511	2.700	-1.365	0.499	5.660	5.416*	-2.187
	(-1.715)	(18.844)	(2.892)	(-2.885)	(-2.253)	(21.388)	(4.144)	(-3.254)
AP5	-0.737	13.294	4.324*	-2.726	-0.940	14.734	7.542*	-2.596
	(-1.520)	(17.546)	(2.916)	(-2.885)	(-2.422)	(22.697)	(4.416)	(-3.254)
ALL	-0.257	20.805	4.900*	-2.113	-0.292	20.394	9.163*	-2.595
	(-1.750)	(32.110)	(3.271)	(-2.662)	(-2.673)	(38.596)	(5.078)	(-3.042)

Note: See note to Table 5.12. The 5% asymptotic critical value of the IPS and LLL tests is equal to –1.645 and 1.645, respectively. The 5% asymptotic critical values of the MW tests is equal to –18.307 and 31.410 for N = 5 (AS5, AP5) and 10 (ALL), respectively. The asterisk (*) denotes the rejection of the null hypothesis at the 5% significant level.

Table 5.15 Panel cointegration test results for the monetary model

Panel	Bivariate system (k=2)				Trivariate system (k=3)			
	IPS	MW	LLL	CIPS	IPS	MW	LLL	CIPS
AS5	-1.366	19.624	3.130	-1.754	-0.345	13.407	6.989*	-1.602
	(-2.094)	(20.980)	(3.250)	(-2.621)	(-1.872)	(19.805)	(4.494)	(-2.916)
AP5	-0.264	10.134	1.538	-2.209	0.872	5.145	5.488*	-2.111
	(-1.873)	(19.186)	(3.260)	(-2.680)	(-1.637)	(18.317)	(3.230)	(-2.980)
ALL	-1.156	29.758	3.300	-1.666	0.366	18.552	8.823*	-1.535
	(-2.185)	(34.994)	(3.514)	(-2.307)	(-1.854)	(32.773)	(5.208)	(-2.571)

Note: See note to Table 5.14.

It is noted that the results of the monetary model from the AP5 panel might be weakened as Australia and New Zealand are different as in both countries the monetary policies have been based on the inflation targeting since the early 1990s. Therefore, exchange rate fluctuations might be related more to an interest rate reaction function and less to the monetary model (see Bernanke *et al.* (1999) for the details of the inflation targeting).

Next, we compare the empirical results in this section with the simulation results discussed in Chapters 2 and 3. In Chapter 2, we found that the SURIPS test was more powerful than the bootstrap IPS and MW tests and the CIPS test in the small panel (N=5) when the values of the cross-correlations are high. However, when the degree of cross-correlation is moderate, the bootstrap IPS and MW tests are more powerful than the SURIPS and CIPS tests. The estimated cross-correlation matrices from Tables 5.7 and 5.8 show that the values of the cross-correlations for the residuals of the ADF tests on $q_{i,t}$ are higher than those of $z_{i,t}$. Therefore, the SURIPS test provides supporting evidence of mean reversion only for the panel of real exchange rates ($q_{i,t}$). For the $z_{i,t}$ series, where the degree of cross-correlation is not high, the bootstrap IPS and MW tests perform better than the SURIPS and CIPS tests in terms of rejecting the null hypothesis.

In Chapter 3, the simulated power of the likelihood-based panel cointegration test of LLL was significantly higher than that of the residual-based panel cointegration tests of IPS, MW and CIPS. The empirical evidence in this section does not contradict the simulation results reported in Chapter 3. The panel LLL rank test provides more significant evidence for the fundamental determination based on the PPP hypothesis and the monetary model than that of the residual-based panel cointegration tests.

5.5 The impact of the 1997 East Asian currency crisis

In this section, we test for the presence of a long-run relationship between exchange rates and fundamentals in the presence of structural changes. In Chapter 4, we discussed the fact that the presence of structural breaks makes it difficult to distinguish between a unit root process and a stationary process with regime shifts. In Section 5.4.2, we noted that a sudden change in real exchange rates and deviations from monetary fundamental was observed in most countries in our panel during the periods 1997:2 and 1997:3.

The 1997 East Asian currency crisis started in June of 1997, when a number of currencies in South-East Asian countries were subjected to speculative attack. Consequently, the Bank of Thailand abandoned its fixed exchange rate regime on 2nd July 1997. The impact of the crisis quickly spread throughout the region, with most currencies in the Asia Pacific region suffering massive devaluation. For this reason, the period 1997:2 is used as a threshold point of structural change, triggered by the currency crisis in the AS5 panel, and 1997:3 is used as a threshold point of shift, prompted by the currency crisis in the AP5 panel.

5.5.1 Empirical results using the sample period before the 1997 crisis

To cope with a possible change as an effect of the 1997 crisis, we first exclude the post-crisis observations from our sample period, and repeat the time-series and panel data methods, to test for the validity of PPP and the monetary model. With regard to the monetary model, we do not consider the countries in which the span of data is incomplete, namely, Indonesia, Malaysia, The Philippines, Thailand and New Zealand. Therefore, in this section, we exclude these countries from our panel in testing for the monetary model. The results of the ADF, LM and SURADF tests for unit roots of real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$) are presented in Tables 5.16, 5.17 and 5.18, respectively.

Using the ADF and SURADF tests, we reject non-stationarity in real exchange rates ($q_{i,t}$) only for New Zealand and Japan. For deviations from monetary fundamental ($z_{i,t}$), the null hypothesis of unit roots is rejected only for Japan. Using the LM test, there is no evidence of long-run PPP and the monetary model for any country at the 5% significant level. These results are similar to those reported for the whole sample period (see Table 5.5 to 5.7). There is additional evidence in support of PPP only for Japan.

Table 5.16 Empirical results of the ADF tests for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$) in the pre-crisis period

Countries	q^c	$q^{c,t}$	z^c	$z^{c,t}$
<u>AS5</u>				
Indonesia	-1.652 (0)	-1.273 (0)	-	-
Malaysia	-1.718 (0)	-0.916 (0)	-	-
Philippines	-2.067 (2)	-1.963 (2)	-	-
Singapore	-0.629 (5)	-1.934 (4)	-0.251 (1)	-1.502 (1)
Thailand	-2.030 (0)	-1.525 (0)	-	-
<u>AP5</u>				
Australia	-2.611 (3)	-2.545 (3)	-0.055 (0)	-2.485 (3)
New Zealand	-2.082 (6)	-3.494* (8)	-	-
Japan	-1.586 (4)	-3.523* (4)	-0.776 (4)	-3.611* (4)
Korea	-2.411 (3)	-2.928 (3)	-1.101 (8)	-1.827 (8)
Taiwan	-1.234 (1)	-2.426 (5)	-1.682 (5)	-2.077 (9)

Note: See note to Table 5.4.

Table 5.17 Empirical results of the LM tests for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$) in the pre-crisis period

Countries	$q^{c,t}$	$z^{c,t}$
AS5		
Indonesia	-1.538 (0)	-
Malaysia	-1.302 (0)	-
Philippines	-1.786 (2)	-
Singapore	-1.575 (4)	-1.603 (0)
Thailand	-1.434 (0)	-
AP5		
Australia	-2.289 (0)	-2.326 (3)
New Zealand	-3.050 (8)	-
Japan	-2.589 (4)	-2.896 (4)
Korea	-2.297 (3)	-1.469 (8)
Taiwan	-1.530 (9)	-1.991 (5)

Note: See note to Table 5.6.

Next, we consider the results from the multivariate methods, reported in Tables 5.19 and 5.20. Again, we find results similar to those in Section 5.4.2 (see Tables 5.10 and 5.11). Using the E-G test, there is evidence supporting long-run PPP for Japan only, and support for the monetary model only for Australia. Considering the results from the *JLR* test, exclusion of the post-crisis period data leads only to marginally more support of long-run PPP for Malaysia.

Table 5.18 Empirical results of the SURADF tests for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$) in the pre-crisis period

Countries	q^{c}	$q^{c,t}$	z^{c}	$z^{c,t}$
AS5				
Indonesia	-1.423	-1.307	-	-
	(-3.296)	(-3.896)		
Malaysia	-1.178	-1.230	-	-
	(-3.110)	(-4.118)		
Philippines	-1.823	-1.755	-	-
	(-3.158)	(-3.916)		
Singapore	-1.957	-3.235	-0.196	-0.997
	(-3.129)	(-4.118)	(-3.245)	(-3.859)
Thailand	-2.204	-1.967	-	-
	(-3.243)	(-3.957)		
AP5				
Australia	-2.295	-2.343	-0.413	-1.595
	(-3.317)	(-4.064)	(-3.132)	(-3.957)
New Zealand	-2.888	-4.360*	-	-
	(-3.289)	(-4.064)		
Japan	-3.301	-5.075*	-1.576	-4.563*
	(-3.320)	(-3.988)	(-3.143)	(-4.033)
Korea	-2.330	-2.884	-1.252	-1.681
	(-3.242)	(-4.010)	(-3.272)	(-3.984)
Taiwan	-1.576	-2.759	-2.948	-2.581
	(-3.163)	(-3.992)	(-3.305)	(-4.046)

Note: See note to Table 5.7.

Table 5.19 The Engle-Granger (E-G) and Johansen likelihood-ratio (*JLR*) test results for the PPP hypothesis in the pre-crisis period

Countries	Bivariate system (k=2)			Trivariate system (k=3)			
	E-G	*JLR*		*E-G*	*JLR*		
		$r=0$	$r\le 1$		$r=0$	$r\le 1$	$r\le 2$
<u>AS5</u>							
Indonesia	-1.385 (0)	30.309*	4.344 (3)	-1.717 (0)	46.965*	22.979*	7.887 (4)
Malaysia	-2.592 (3)	29.202*	5.104 (2)	-3.016 (3)	64.671*	17.451	2.577 (2)
Philippines	-1.853 (2)	11.259	3.266 (3)	-2.183 (2)	32.980	17.845	6.234 (4)
Singapore	-1.301 (4)	40.012*	5.959(5)	-2.401 (0)	65.189*	33.075*	10.434* (4)
Thailand	-2.666 (0)	19.836	2.251 (4)	-2.189 (0)	34.998*	18.134	8.117 (4)
<u>AP5</u>							
Australia	-2.502 (3)	18.946	5.331 (2)	-2.491 (3)	38.609*	16.328	5.528 (5)
New Zealand	-3.007 (8)	27.036*	5.370 (5)	-3.270 (8)	66.277*	32.811*	8.982 (6)
Japan	-3.753* (4)	32.449*	10.385* (5)	-3.772 (4)	50.896*	27.350*	11.656* (5)
Korea	-2.728 (3)	24.693*	2.111 (4)	-2.539 (3)	53.987*	25.222*	9.670 (5)
Taiwan	-1.993 (1)	24.988*	8.375 (4)	-2.540 (5)	54.812*	28.423*	8.596 (5)

Note: See note to Table 5.10. The 5% critical values of the E-G test are equal to –3.425 and –3.865 for the bivariate and trivariate systems, respectively. The 5% critical values of the *JLR* test for r=0 are equal to 19.96 and 34.91 in the bivariate and trivariate systems, respectively.

Table 5.20 The Engle-Granger (E-G) and Johansen likelihood-ratio (*JLR*) test results for the monetary model in the pre-crisis period

Countries	Bivariate system (k=2)			Trivariate system (k=3)			
	E-G	*JLR*		*E-G*	*JLR*		
		$r=0$	$r\le 1$		$r=0$	$r\le 1$	$r\le 2$
Australia	-2.744 (3)	13.156	3.919 (1)	-4.337* (3)	61.131*	15.093	6.662 (6)
Singapore	-0.497 (1)	23.093*	10.030 (6)	-3.809 (4)	35.168*	17.047	7.169 (3)
Japan	-1.393 (5)	22.762*	7.140 (2)	-1.393 (5)	41.315*	12.910	2.692 (3)
Korea	-2.765 (3)	18.956	2.544 (5)	-2.393 (4)	39.007*	15.151	4.298 (5)
Taiwan	-1.228 (5)	12.150	1.934 (6)	-2.623 (0)	54.620*	17.398	3.883 (6)

Note: See note to Table 5.19.

Next, we consider the results from the panel data tests. The empirical results from the panel unit root and cointegration tests in the pre-crisis period are presented in Tables 5.21 and 5.22, respectively. The panel unit root test results from the pre-crisis sample provide mixed evidence for PPP. For the AS5 panel, we cannot reject the unit root null hypothesis, using the IPS, MW, SURIPS, CIPS and LM tests. However, for the AP5 panel, we find stationarity for real exchange rates ($q_{i,t}$) from the results of the IPS, MW and SURIPS tests. For the ALL panel, the results are similar to those reported in Table 5.12, where the unit root null hypothesis can be rejected, using the SURIPS test. For the monetary model, the results from Table 5.21 suggest that there is no evidence in support of stationarity for $z_{i,t}$.

The panel cointegration test results provide more supportive evidence than those of the panel unit root tests. We can reject the null hypothesis of no cointegration, using both the residual-based test of CIPS and the likelihood-based test of LLL test for all panels. In addition, we can reject the null hypothesis of no cointegration, using the residual-based tests of IPS and MW for the AP5 panel. For the monetary model, we still only find supporting evidence from the results of the panel LLL rank test.

Overall, using the pre-crisis sample, we find only marginally more evidential support for PPP than for the full sample. The supportive evidence for PPP is significant only for the AP5 panel. However, for the monetary model, the evidence is weaker than for the full sample. In addition, even though application of the sub-sample period can be used to cope with the impact of a structural shift, there are some disadvantages, as it reduces the sample size, causing the power of the tests to decrease. Therefore, in the next section, we will employ the panel LM unit root test with level shifts to investigate the existence of PPP and the monetary model in the presence of structural shifts.

Table 5.21 Panel unit root test results for the PPP hypothesis and the monetary model in the pre-crisis period

Panels	IPS		MW		SURIPS		CIPS		LM
	c	*c,t*	*c*	*c,t*	*c*	*c,t*	*c*	*c,t*	*c,t*
PPP									
AS5	-0.290	1.784	9.803	2.781	-0.541	0.720	-1.879	-2.987	1.613
	(-1.466)	(-1.776)	(17.301)	(18.220)	(-2.034)	(-2.972)	(-2.544)	(-3.054)	
AP5	-1.349	-2.363*	14.089	21.560*	-2.459	-3.713*	-1.898	-2.928	-1.620
	(-1.713)	(-1.727)	(18.281)	(18.796)	(-2.540)	(-3.033)	(-2.544)	(-3.054)	
ALL	-1.165	-0.476	23.171	24.341	-3.269*	-1.805	-2.119	-2.365	-0.017
	(-1.608)	(-1.770)	(30.280)	(31.385)	(-3.189)	(4.697)	(-2.325)	(-2.834)	
MM									
ASP5	1.743	0.152	3.262	9.853	0.485	-0.449	-1.533	-2.002	-0.549
	(-1.427)	(-1.529)	(17.127)	(17.546)	(-2.270)	(-2.850)	(-2.544)	(-3.054)	

Note: See note to Table 5.12. PPP and MM denote the PPP hypothesis and the monetary model, respectively.

Table 5.22 Panel cointegration test results for the PPP hypothesis and the monetary model in the pre-crisis period

Panels	Bivariate system (k=2)				Trivariate system (k=3)			
	IPS	MW	LLL	CIPS	IPS	MW	LLL	CIPS
PPP								
AS5	0.090	8.147	7.303*	-2.901*	0.105	7.860	9.121*	-2.954
	(-1.731)	(18.600)	(3.009)	(-2.882)	(-2.688)	(24.282)	(3.843)	(-3.260)
AP5	-2.401*	22.877*	7.045*	-3.014*	-1.989	19.409	10.550*	-2.891
	(-2.016)	(20.322)	(3.330)	(-2.882)	(-2.320)	(22.265)	(5.245)	(-3.260)
ALL	-1.613	31.024	10.146*	-2.774*	-1.292	27.269	13.910*	-2.898
	(-2.023)	(33.810)	(4.026)	(-2.667)	(-2.921)	(40.791)	(5.998)	(-3.043)
MM								
ASP5	0.607	8.501	3.660*	-1.750	-1.698*	21.142*	8.660*	-2.035
	(-1.393)	(17.195)	(3.161)	(-2.882)	(-1.562)	(19.223)	(4.498)	(-3.260)

Note: See note to Table 5.14. PPP and MM denote the PPP hypothesis and the monetary model, respectively.

5.5.2 Empirical results from the exogenous-break LM unit root tests

In this section, we employ the LM unit root test to control for the effect of the structural breaks in testing for stationarity of real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$). We first consider the exogenous break LM unit root test of Amsler and Lee (1995) (AL). To cope with the impact of the 1997 crisis, we specify the break point at 1997:2 and 1997:3 for countries in the AS5 and AP5 panels, respectively. In addition, the panel LM test of Im, Lee and Tieslau (2002) (ILT) is also calculated on the AS5, AP5 and ALL panels. The results of the exogenous break LM unit root test are given in Table 5.23.

The results from the individual time-series exogenous break LM unit root test provide additional evidence supporting PPP ($q_{i,t}$) for Indonesia and New Zealand over the results from the LM test without shifts, using both the null and pre-crisis sample. For the monetary model ($z_{i,t}$), evidence of mean reversion is found only for Japan, using the time-series test, but is not observed for any panel.

From the panel LM test, we reject the unit root null hypothesis of $q_{i,t}$ for the ALL panel, and only marginally accept the null hypothesis for the AP5 panel at the 5% significant level. However, for the monetary model, evidence of mean reversion of $z_{i,t}$ cannot be found for any panel. These results are similar to those from panel unit root testing using the pre-crisis sample, where we find evidence of fundamentals determinants only for the PPP in the AP5 and ALL panels.

These empirical results would imply that, for the AS5 panel, the PPP hypothesis and the monetary model do not hold before the 1997 crisis. Exchange rate movements after the crisis, when the flexible exchange rate regimes are applied in most countries would then provide evidence of the adjustment toward the fundamental equilibrium. For the AP5 panel, where the majority of countries have the flexible exchange rate regimes, the PPP hypothesis holds throughout the sample period and the structural shifts should be occurred for several countries in the panel due to the effect of the currency crisis.

Table 5.23 Empirical results of the exogenous break LM unit root test for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	$q^{c,t}$	$z^{c,t}$
AS5		
Indonesia	-3.494* (1)	-2.919 (0)
Malaysia	-2.407 (2)	-1.760 (2)
Philippines	-1.554 (2)	0.728 (5)
Singapore	-1.713 (5)	-1.612 (1)
Thailand	-1.727 (1)	-2.376 (5)
AP5		
Australia	-2.154 (0)	-1.280 (0)
New Zealand	-3.124* (8)	-1.999 (8)
Japan	-2.531 (4)	-3.522* (4)
Korea	-2.044 (6)	-1.295 (8)
Taiwan	-1.846 (9)	-2.240 (9)
Panel statistics		
AS5	-0.863	-0.246
AP5	-1.642	-0.673
All	-1.772*	-0.647

Note: See note to Table 5.6. The 5% critical value of the panel LM statistic is equal to –1.645.

5.5.3 Empirical results from the endogenous-break LM unit root tests

In this section, we apply several endogenous break selection procedures to determine the number and location of structural breaks from the data. We first consider the two-break test. If any dummy variable that indicates a break point is insignificant, the one-break test is then applied. If it is still possible to indicate insignificance of the one break dummy, the LM test without shifts is used. The min-t_{β}, max-$|t_{\delta}|$ and min-*SBC* procedures are applied to estimate the break points (see Section 4.3.3 for the details of the endogenous break selection procedures). The results of the individual time-series and panel LM statistics and the estimates of the break points for the min-t_{β}, max-$|t_{\delta}|$ and min-*SBC* tests are reported in Tables 5.24 to 5.26, respectively.

The results from Tables 5.24 to 5.26 show that the presence of at least one structural break is found for most countries in the panel, using the endogenous break selection procedures. The break dates corresponding to the 1997 crisis (1997:2, 1997:3) are selected for many countries when the max-$|t_{\delta}|$ and min-*SBC* procedures are applied. These results confirm that the impact of the 1997 crisis results in a level shift in $q_{i,t}$ and $z_{i,t}$ for most of the countries in the panel. However, there is no evidence of structural shift in the period of the crisis (1997:2, 1997:3) for Japan and New Zealand, using either the max-$|t_{\delta}|$ or the min-*SBC* tests for either $q_{i,t}$ or $z_{i,t}$. These results indicate that the 1997 currency crisis may not have had any significant impact for Japan and New Zealand. The estimated break points from the min-t_{β} test are often different from those of the max-$|t_{\delta}|$ and min-*SBC* tests. In addition, the break dates selected by the min-t_{β} test are insignificant for the New Zealand and

Korea real exchange rate ($q_{i,t}$) series and the Thailand and Korea deviation from monetary fundamental ($z_{i,t}$) series. These differences in the selection of the break dates may result from the difference in the accuracy of break point estimation across the tests, highlighted in Section 4.6.2 (see Tables 4.41 and 4.42).

Table 5.24 Empirical results of the endogenous break min-t_β LM unit root test for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	$q^{c,t}$		$z^{c,t}$	
	LM statistics	Break date	LM statistics	Break date
AS5				
Indonesia	-4.052* (3)	86:2, 99:2	-4.855* (3)	98:4, 99:2
Malaysia	-3.046 (3)	91:4, 98:1	-2.144 (2)	97:4
Philippines	-2.421 (2)	93:3, 97:4	-1.638 (1)	84:1, 84:3
Singapore	-2.207 (5)	85:3, 98:2	-2.269 (3)	91:3
Thailand	-2.487 (0)	99:2	-3.851* (7)	-
AP5				
Australia	-3.241 (3)	88:1, 88:3	-2.143 (3)	82:4, 88:3
New Zealand	-2.866 (8)	-	-1.999 (8)	97:3
Japan	-2.818 (4)	87:3	-4.432* (4)	82:2, 97:1
Korea	-2.189 (3)	-	-1.881 (8)	-
Taiwan	-1.351 (1)	97:3	-2.370 (9)	93:3
Panel statistics				
AS5	-1.139		-2.068*	
AP5	-1.129		-1.051	
All	-1.601		-2.148*	

Note: See note to Table 5.6. The 5% critical values of the one- and two-break min-t_β LM tests are equal to –3.500 and –3.918, respectively, when T=92. The 5% critical value of the LM test (without shifts) is equal to –3.07. The 5% critical value of the panel LM statistics is equal to –1.645.

Table 5.25 Empirical results of the endogenous break max-$|t_{\delta}|$ LM unit root test for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	$q^{c,t}$		$z^{c,t}$	
	LM statistics	Break date	LM statistics	Break date
AS5				
Indonesia	-2.348 (2)	86:2, 98:2	-2.302 (0)	97:2, 98:1
Malaysia	-1.426 (3)	97:2, 98:1	-1.848 (0)	97:3, 98:1
Philippines	-1.242 (6)	97:2, 98:3	-1.129 (0)	84:4, 97:2
Singapore	-1.087 (2)	97:3, 98:1	-1.654 (1)	97:3, 98:1
Thailand	-1.783 (1)	84:3, 97:2	-1.596 (0)	97:2, 99:3
AP5				
Australia	-1.982 (0)	84:4, 97:3	-1.485 (0)	84:4, 88:3
New Zealand	-2.450 (8)	84:2, 85:2	-1.519 (7)	89:2, 94:1
Japan	-2.290 (4)	87:3, 98:3	-3.830* (4)	87:3, 98:3
Korea	-1.525 (8)	97:3, 99:3	-1.295 (8)	97:3
Taiwan	-1.013 (4)	97:3, 98:3	-2.126 (9)	85:4, 97:3
Panel statistics				
AS5	1.423		1.246	
AP5	0.397		-0.276	
All	1.289		0.699	

Note: See note to Table 5.6. The 5% critical values of the one- and two-break max-$|t_{\delta}|$ LM tests are equal to –3.097 and –3.240, respectively, when T=92. The 5% critical value of the LM test (without shifts) is equal to –3.07. The 5% critical value of the panel LM statistics is equal to –1.645.

Table 5.26 Empirical results of the endogenous break min-*SBC* LM unit root test for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$)

Countries	$q^{c,t}$		$z^{c,t}$	
	LM statistics	Break date	LM statistics	Break date
AS5				
Indonesia	-2.348 (2)	86:2, 98:2	-3.784* (1)	98:1, 99:1
Malaysia	-2.638 (2)	97:2, 98:1	-1.848 (0)	97:3, 98:1
Philippines	-1.620 (2)	82:4, 97:2	-1.129 (0)	84:4, 97:2
Singapore	-1.087 (2)	97:3, 98:1	-1.654 (1)	97:3, 98:1
Thailand	-1.783 (1)	84:3, 97:2	-1.596 (0)	97:2, 99:3
AP5				
Australia	-1.982 (0)	84:4, 97:3	-1.485 (0)	84:4, 88:3
New Zealand	-2.450 (8)	84:2, 85:2	-1.906 (2)	89:2, 00:3
Japan	-2.421 (3)	82:3, 87:3	-3.830* (4)	87:3, 98:3
Korea	-2.169 (3)	97:3, 00:3	-1.137 (8)	93:3, 97:3
Taiwan	-1.013 (4)	97:3, 98:3	-2.175 (9)	-
Panel statistics				
AS5	1.031		0.799	
AP5	0.424		-0.071	
All	1.032		0.539	

See note to Table 5.6. The 5% critical values of the one- and two-break min-*SBC* LM tests are equal to –3.388 and –3.696, respectively, when T=92. The 5% critical value of the LM test (without shifts) is equal to –3.07. The 5% critical value of the panel LM statistics is equal to –1.645.

Structural changes are also found at other times in most countries. In Section 5.4.2, we noted the presence of sudden changes in nominal exchange rates, which is regularly observed in some countries, e.g. Indonesia, the Philippines and Thailand, due to the devaluation or changes in the exchange rate regimes. For example, the max-$|t_\delta|$ and min-*SBC* procedures indicate the break points during 1984:3 and 1997:2 for Thailand, which represents the devaluation of the Thai baht in 1984 and the 1997 crisis, respectively.

Next, we consider the results of the LM statistics testing for the unit root null hypothesis in the presence of structural breaks. For the PPP hypothesis, we reject the unit root null hypothesis only for Indonesia, using the min-t_β test. The results of the max-$|t_\delta|$ and min-*SBC* tests cannot reject the null hypothesis in any country. In addition, the results from the panel statistics also fail to reject the unit root null hypothesis of $q_{i,t}$ for all panels. These results are similar to those of the individual and panel LM test (without shifts), reported in Tables 5.6 and 5.12, suggesting that there is no evidence of stationarity for $q_{i,t}$. For the monetary model, using the min-t_β test, we find similar results to those for the LM test without shifts (see Table 5.6), in which the unit root null hypothesis is rejected for Indonesia, Thailand and Japan. Using the max-$|t_\delta|$ (min-*SBC*) test, we reject the null hypothesis only for Japan (Indonesia and Japan). The results from the panel min-t_β test provide results similar to those for the panel LM test (without shifts), as shown in Table 5.6. The panel min-t_β test rejects the null hypothesis for the AS5 and ALL panels. However, the panel max-$|t_\delta|$ and min-*SBC* tests still do not provide evidence of mean reversion in $z_{i,t}$.

Overall, the break point selection procedures usually indicate the presence of structural breaks in the data. However, the selected break dates and the panel LM test results both vary according to the method used to estimate the break points. Even though the max-$|t_\delta|$ and min-*SBC* tests select the break dates corresponding to the 1997 crisis, they provide weaker evidence supporting the monetary model than the min-t_β test for either time-series or panel data. For the PPP hypothesis, the time-series (panel data) endogenous-break LM tests fail to reject the unit root null hypothesis in all countries (panels), with the exception of Indonesia, using the min-t_β test. These results suggest that the exogenous break LM unit root test is more useful than the endogenous break counterparts when we examine the effect of structural break due to the specific events such as the currency crisis. In Chapter 4, we show that the endogenous break LM tests are sensitive to the magnitude of breaks and the gap between location of breaks. The estimated break points may not represent the interesting events and be imprecise if the size of the breaks is not significantly large.

5.6 Conclusion

In this chapter, the empirical evidence of long-run PPP and the monetary model was studied for a panel of Asia Pacific countries. We first employed standard time-series methods to test for the existence of PPP and the monetary model. The unit root tests suggested that the null hypothesis could not be rejected for real exchange rates ($q_{i,t}$) and deviations from monetary fundamental ($z_{i,t}$), with the exception of New Zealand for $q_{i,t}$ and Indonesia, Thailand, New Zealand and Japan for $z_{i,t}$. The results from the two-step cointegration test provided evidence supporting the monetary model for Indonesia, but did not find evidence for PPP in any country. However, there was substantial evidence of cointegration relationships, using the *JLR* test for most countries.

Next, we applied the panel data techniques to improve the power of the tests over the standard time-series tests. The bootstrap IPS and MW, SURIPS and CIPS panel unit root tests were used to take account of the presence of cross-sectional dependence in the errors. The results from the panel unit root tests rejected the unit root null hypothesis of $q_{i,t}$ and $z_{i,t}$ for all panels. However, these results were sensitive to the choice of panel statistics. For the PPP hypothesis, supporting evidence was found only from the results of the SURIPS test; however, only the results from the bootstrap IPS and MW tests rejected the unit root null hypothesis for the monetary model in any panel. For the panel cointegration tests, we found evidence to support PPP and the monetary model from only the panel LLL rank test. The results from the panel cointegration tests of IPS, MW and CIPS failed to find cointegration relationships for either PPP or the monetary model in any panel. These results showed that the panel unit root and cointegration tests were more powerful than the individual time-series counterparts. The performance of the panel unit root tests was affected by the presence of cross-sectional dependence. Therefore, the bootstrap IPS and SURIPS tests were recommended to control the effect of cross-sectional dependence over the CIPS test when the panel size was not large enough ($N \leq 10$). We also recommended applying the SUR method in the highly cross-correlated panels. For the panel cointegration test, the panel rank test was still recommended over the residual-based panel cointegration test. However, the bootstrap method was required, as the empirical critical values of the panel rank test were significantly different from those of the asymptotic values. Overall, the empirical results were not inconsistent with those of the simulation results discussed in Chapters 2 and 3.

We considered the impact of the 1997 East Asian currency crisis. The standard time-series and panel data methods were again applied, to test for a long-run relationship for the pre-crisis sample to eliminate the effect of the crisis. The results from the standard time-series tests were similar to those results for the full sample, where the unit root tests and the residual-based tests provided supporting evidence for the fundamental exchange models for few countries, and the *JLR* test rejected the null hypothesis of no cointegration in most countries. However, the similarity between the results of the pre-crisis sample and those of the full sample might be due to co-breaking where the effect of the breaks present in the different series cancel each other out (see Hendry and Mizon (1998)).

However, the results from the panel unit root tests generated some interesting results. Evidence supporting PPP was strong for the AP5 panel, as we rejected the unit root (no cointegration) null hypothesis, using the bootstrap IPS and MW tests and the SURIPS test (the LLL rank test and the residual-based test of IPS, MW and CIPS). However, the results indicated that there was no significant evidence of stationarity either for the monetary model or for PPP in the case of the AS5 panel.

Next, we used the exogenous break LM unit root tests to control the effect of structural breaks. The results from the exogenous break LM test were similar to those obtained from the tests in the pre-crisis sample, in that we found stationarity for real exchange rates ($q_{i,t}$) only in the AP5 and ALL panel. However, it was impossible to find stationarity either for $q_{i,t}$ in the AS5 panel or for $z_{i,t}$ in any panel.

Next, we considered the LM unit root test with the endogenous break selection procedure, to determine the number and location of structural changes from the data. The results from the endogenous break selection procedures indicated the presence of level shifts for the majority of countries. The break points were often selected at 1997:2 and 1997:3, representing the impact of the 1997 currency crisis. However, using the panel max-$|t_\delta|$ and min-*SBC* tests, there was no significant evidence in favour of long-run PPP or the monetary model. The results from the panel min-t_β test rejected the non-stationrity for $z_{i,t}$ alone in the AS5 and ALL panels. However, the min-t_β test barely selected the realistic break points corresponding to the currency crisis. Therefore, the exogenous break panel LM unit root test was recommended over the endogenous break tests when we investigated the effect of some specific events. The empirical results from the endogenous break tests were sensitive to the choice of the endogenous break selection procedures, especially when the size of breaks was not significantly large.

Appendix A

Table A.5.1 The Diagnostic tests for VAR models of the PPP hypothesis

Countries	Bivariate system (k=2)				Trivariate system (k=3)			
	Serial-correlation	Hetero-Skedasticity	Normality		Serial-correlation	Hetero-skedasticity	Normality	
AS5								
Indonesia	0.875	0.004	0.000	(5)	0.720	0.052	0.030	(5)
Malaysia	0.961	0.752	0.000	(5)	0.315	0.313	0.000	(5)
Philippines	0.452	0.000	0.000	(3)	0.381	0.155	0.000	(4)
Singapore	0.808	0.000	0.556	(4)	0.998	0.045	0.147	(4)
Thailand	0.233	0.010	0.000	(3)	0.296	0.108	0.000	(4)
AP5								
Australia	0.630	0.278	0.044	(4)	0.659	0.406	0.022	(6)
New Zealand	0.059	0.215	0.000	(3)	0.302	0.026	0.000	(4)
Japan	0.128	0.799	0.036	(5)	0.243	0.355	0.002	(5)
Korea	0.293	0.078	0.000	(2)	0.341	0.045	0.000	(4)
Taiwan	0.541	0.675	0.000	(5)	0.713	0.351	0.007	(5)

Note: The figures are the p-values of the diagnostic tests for the residuals of VAR models. The figures in the parenthesis are the number of lags in VAR models for the *JLR* test.

Table A.5.2 The Diagnostic tests for VAR models of the monetary model

Countries	Bivariate system (k=2)				Trivariate system (k=3)			
	Serial-correlation	Hetero-Skedasticity	Normality		Serial-correlation	Hetero-skedasticity	Normality	
AS5								
Indonesia	0.374	0.206	0.303	(4)	0.949	0.086	0.021	(4)
Malaysia	0.873	0.000	0.000	(1)	0.800	0.013	0.001	(2)
Philippines	0.111	0.153	0.000	(1)	0.189	0.021	0.067	(6)
Singapore	0.410	0.385	0.072	(6)	0.520	0.687	0.003	(6)
Thailand	0.868	0.298	0.000	(1)	0.697	0.068	0.000	(1)
AP5								
Australia	0.643	0.368	0.000	(1)	0.928	0.133	0.000	(1)
New Zealand	0.802	0.215	0.545	(3)	0.688	0.177	0.315	(3)
Japan	0.355	0.189	0.000	(2)	0.312	0.133	0.000	(2)
Korea	0.340	0.958	0.000	(2)	0.450	0.077	0.000	(2)
Taiwan	0.152	0.183	0.000	(6)	0.646	0.412	0.000	(6)

Note: See notes to Table A.5.1.

Table A.5.3 The Diagnostic tests for VAR models for the PPP hypothesis in the pre-crisis sample period

Countries	Bivariate system (k=2)				Trivariate system (k=3)			
	Serial-correlation	Hetero-Skedasticity	Normality		Serial-correlation	Hetero-skedasticity	Normality	
AS5								
Indonesia	0.445	0.030	0.000	(3)	0.428	0.185	0.000	(4)
Malaysia	0.223	0.497	0.712	(2)	0.075	0.582	0.887	(2)
Philippines	0.889	0.002	0.001	(3)	0.792	0.203	0.065	(4)
Singapore	0.624	0.043	0.262	(5)	0.327	0.278	0.078	(4)
Thailand	0.909	0.053	0.000	(4)	0.865	0.365	0.000	(4)
AP5								
Australia	0.336	0.414	0.127	(2)	0.303	0.287	0.028	(5)
New Zealand	0.367	0.152	0.000	(5)	0.697	0.007	0.002	(6)
Japan	0.550	0.456	0.078	(5)	0.524	0.300	0.118	(5)
Korea	0.421	0.231	0.144	(4)	0.128	0.041	0.103	(5)
Taiwan	0.904	0.116	0.010	(4)	0.843	0.152	0.036	(5)

Note: See notes to Table A.5.1

Table A.5.3 The Diagnostic tests for VAR models for the monetary model in the pre-crisis sample period

Countries	Bivariate system (k=2)				Trivariate system (k=3)			
	Serial-correlation	Hetero-Skedasticity	Normality		Serial-correlation	Hetero-skedasticity	Normality	
Australia	0.680	0.721	0.007	(1)	0.571	0.337	0.001	(6)
Singapore	0.396	0.544	0.067	(6)	0.378	0.539	0.229	(3)
Japan	0.693	0.193	0.751	(2)	0.264	0.200	0.502	(3)
Korea	0.666	0.058	0.331	(5)	0.588	0.520	0.465	(5)
Taiwan	0.370	0.418	0.124	(6)	0.464	0.329	0.001	(6)

Note: See notes to Table A.5.1

Chapter 6

Conclusion and Directions for Future Research

6.1 Concluding remarks

In this thesis, we investigated the finite sample performance of several heterogeneous panel unit root and cointegration tests, based on a number of different experiments. We then applied these tests in an empirical study of fundamental exchange rate modelling in Asia Pacific countries. The main findings from the thesis can be summarised as follows.

Chapter 2 undertook an investigation of the finite sample performance of the panel unit root tests of Im, Perasan and Shin (2003) (IPS), and Maddala and Wu (1999) (MW). Monte Carlo simulations were conducted, based on different assumptions of the correlation structure in the error terms and the number of stationary series in the panel. We considered the case of a moderate sample size (T) corresponding to quarterly data for the post-Bretton Woods period, and for a slow speed of mean reversion.

The simulation results indicated that the panel IPS and MW unit root tests increased the power over the standard ADF test, with the IPS test proving slightly more powerful than the MW test. The simulation results showed the need for caution in applying heterogeneous panel unit roots. First, the inclusion of non-stationary series in the panel considerably weakened the performance of both tests, with the power of the tests depending on the proportion of stationary series in the panel. Therefore, the power of rejecting the unit root null hypothesis would be sensitive to the inclusion of some cross-section units in the panel. These results suggested that the exclusion of some cross-section units, which are likely to be non-stationary, could improve the power performance of panel unit root tests in the empirical study. Second, the IPS and MW tests were both over-sized in the presence of cross-correlated error terms. This size distortion problem was particularly serious when the values of the cross-correlations were high and the panel size (N) was large. The MW test was slightly less size-distorted than the IPS test in cross-correlated panels.

Next, we compared the performance of three alternative methods of controlling for the effect of cross-correlation in the errors: the bootstrap method, the Seeming Unrelated Regression method (SUR) and the Cross-sectionally augmented IPS test (CIPS). The bootstrap method was used to calculate the empirical critical values of the standard IPS and MW tests, while the IPS-type t-bar statistic estimated by the SUR method (SURIPS) was applied to extract additional information from cross-correlations in the errors. The CIPS test of Pesaran (2003) augmented the standard ADF regression with the cross-section average of lagged levels and first difference of the individual series to control for cross-correlations. Comparing these three methods, with regard to the small panel (N=5), the SURIPS test was the most powerful in highly cross-correlated panels. However, the bootstrap IPS test provided the best power performance when the degree of cross-correlation (ϖ) was moderate. Therefore, the SUR method was recommended in the presence of strong cross-correlation in the errors, and the bootstrap IPS and MW tests were recommended when the degree of cross-correlation was not high. However, the bootstrap method was sensitive to the specification of deterministic terms (intercept, trend) and the exclusion of lag terms. In addition, it was also computationally expensive. In the larger panel, ($N \geq 10$), the CIPS test had the best power performance and therefore, was recommended over the bootstrap test and the SUR method.

In Chapter 3, we extended a simulation study of the panel IPS and MW unit root tests, to test for the existence of cointegration relationships, using the residual-based cointegration approach of Engle and Granger (1987). We also considered the panel rank test of Larsson, Lyhagen and Lothgren (2001) (LLL), based on the Johansen (1988) cointegration approach. The simulation results showed that the panel LLL rank test outperformed the residual-based panel tests of IPS and MW in terms of higher power, even when there was a mixture of cointegrated and non-cointegrated relationships in the panel. However, the panel LLL test was slightly over-sized for moderate sample sizes. The effect of cross-sectional dependence rendered all of the three panel cointegration tests over-sized. The bootstrap method and the CIPS test were then applied to correct the size distortions. The bootstrap panel LLL test produced size reasonably close to the nominal level, and remained more powerful than the bootstrap residual-based tests. The residual-based panel cointegration test of CIPS produced the correct size in cross-correlated panels. The empirical power of the CIPS test was slightly higher than that of the bootstrap panel test of IPS in the bivariate system with highly cross-correlated errors. In light of this, the panel LLL rank test was recommended over the residual-based tests of IPS, MW and CIPS. However, the bootstrap method was acknowledged as necessary to correct for size distortions, occurring in moderate sample sizes and in the presence of cross-sectional dependence.

In Chapter 4, we turned our attention to the matter of structural breaks in panel data unit root testing. The panel LM unit root test of Im, Lee and Tieslau (2002) (ILT) was applied, to control for the effect of structural breaks. Monte Carlo experiments were used, to evaluate the finite sample properties of this panel LM test with and without shifts. The panel LM test without breaks was markedly more powerful than the individual time-series tests. The power of the panel LM test was similar to that of the IPS and MW tests. However, the invariance property, where the asymptotic properties were unaffected by the presence of breaks in any location, was practically useful in constructing the panel statistic. The panel LM test with level shifts could be standardised, using the same adjustment parameters (mean, variance) as those of the panel LM test without any shifts. The simulation results showed that when the break points were correctly specified, the size and power performance of the exogenous break panel LM test was similar to that of the test without shifts. However, incorrectly specifying the number and/or location of breaks resulted in size distortions.

In view of this, several endogenous break selection procedures were applied to estimate the break dates from the data. We first investigated the finite sample properties of the endogenous break LM test. The results showed that finite sample means and variances of the endogenous break LM test varied according to the methods used to estimate the break points. In addition, the magnitude of breaks under the DGP affected the properties of the tests. These differences in the finite sample properties of the endogenous break tests depended on the accuracy with which the true break points were estimated. Comparing the tests across several break selection methods, the max-$|t_\delta|$ test, which selects the break points by maximising the statistic testing for the significance of the break dummy variables, had the best performance in terms of the power and accuracy of true break point selection. However, the min-*SBC* test, which estimates the break dates by minimising the *SBC* information criterion, also performed well, differing only marginally from the max-$|t_\delta|$ test. The min-t_β test, which minimises the LM statistic for testing the unit root null hypothesis, had significantly lower power, and estimated the true break dates less accurately than both the max-$|t_\delta|$ and min-*SBC* tests. In addition, the simulation results suggested that the endogenous break panel LM unit root test was sensitive to incorrect utilisation of the adjustment parameters (means, variance). The means and variances of the endogenous break test varied according to the break point estimation methods and the magnitude of breaks under the DGP. The max-$|t_\delta|$ test was less sensitive to the choice of incorrect adjustment parameters than the other tests. For this reason, the max-$|t_\delta|$ test was recommended in preference to the other tests.

Finally, in Chapter 5, we performed an empirical analysis of fundamental exchange rate modelling, to implement the simulation results and evaluate the performance of the tests in the actual data. Purchasing power parity (PPP) and the monetary model were used as the fundamental determinants of exchange rate movements. We considered a panel of five Pacific rim countries (AP5): Australia, New Zealand, Japan, Korea and Taiwan, and a panel of five South-East Asian countries (AS5): Indonesia, Malaysia, the Philippines, Singapore and Thailand. In light of this, the chapter focused only on the case of the small panel (N = 5, 10). The results from the standard time-series unit root and cointegration tests provided evidence in support of long-run PPP and the monetary model for only a few of the countries when the standard time-series unit root tests and the two-step Engle and Granger (1987) cointegration test (E-G) were used. Evidence in support of PPP was found for New Zealand, and the monetary model was observed for Indonesia, Thailand, New Zealand and Japan. However, the results from the Johansen Likelihood Ratio test (*JLR*) established support for both PPP and the monetary model for most countries. Next, the panel data methodology was applied in an attempt to improve the power over the standard time-series tests. For the panel unit root tests, results supporting the existence of PPP and the monetary model were obtained. However, the empirical results differed across the panel unit root tests applied (IPS, MW, SURIPS, LM and CIPS). The results from the bootstrap IPS and MW tests provided evidence of mean-reversion for the panels using the monetary model ($z_{i,t}$) where there were moderate cross-correlations. The SURIPS test rejected non-stationarity for the panels of real exchange rates ($q_{i,t}$) where the errors were highly cross-correlated. For the cointegration tests, we rejected the null hypothesis of no cointegration for PPP and the monetary model only when the bootstrap panel LLL rank test, which was more powerful than the residual-based tests of IPS, MW and CIPS, was used. These results were not inconsistent with the simulation results reported in Chapters 2 and 3.

We considered the impact of the 1997 East Asian currency crisis. The pre-crisis sample period was used to test for long-run relationships, to eliminate the effect of the crisis. For the individual time-series results, we found more evidential support for the PPP hypothesis for Japan, using the ADF test. In addition, the results from the panel tests provided strong evidence in support of PPP for the AP5 panel, but for the AS5 panel, the results failed to reject non-stationarity for $q_{i,t}$. Next, the exogenous and endogenous break LM unit root tests were employed, to account for level shifts. Using the exogenous break LM test, the results were similar to those panel results from the pre-crisis sample, in that significant support for the fundamental exchange rate models was found only for PPP in the AP5 and ALL panels. The implication of these results would imply that, before the 1997 crisis, the PPP hypothesis and the monetary model did not hold in the AS5 panel where most countries in the panel fixed their nominal exchange rates with either the US dollar or the basket of currency. Exchange rate movements after the crisis, when the flexible exchange rate regimes were implemented in most countries, would then accommodate some adjustments towards long-run relationships. For the AP5 panel, which had flexible exchange rate regimes in the majority of countries, the PPP hypothesis held throughout the sample period, and the effect of the crisis resulted in level shifts for several countries in the panel.

Using the endogenous break selecting procedures, most countries in the panel were found to have at least one structural shift. The break dates corresponding to the 1997 crisis (1997:2, 1997:3) were usually selected by the max-$|t_\delta|$ and min-*SBC* procedures. However, we could not find evidence for mean reversion in either $q_{i,t}$ or $z_{i,t}$, when the panel max-$|t_\delta|$ and min-*SBC* tests were used, although there was some evidence to support the monetary model for the AS5 panel, using the panel min-t_β test.

6.2 Directions for future research

The analysis of panel unit root and cointegration tests in the recent literature suggests the following possibilities for future research in this area.

Given the discussion in the recent literature on cross-sectional dependence (see Section 2.3), a number of useful tests promoting the defactoring of the data to eliminate the effect of cross-correlations before applying panel unit root tests, are proposed (see, for example, Phillips and Sul (2003), and Moon and Perron (2004)). It would be interesting to compare the performance of these methods with those considered in our thesis (the bootstrap method, the SUR method and the CIPS test), which are based on tests with ADF parameterisation.

For panel cointegration tests, the effect of cross-sectional dependence is also an important concern. In the thesis, we simply considered the bootstrap method and the CIPS test, to account for the cross-correlations. Recently, Groen and Kleibergen (1999) have proposed an alternative method that relaxes the assumption of diagonal block of the cross-correlation matrix in the errors. These methods use the generalised method of moments framework to construct maximum likelihood estimators of the cointegrating vectors. Groen and Kleibergen (1999) note that their proposed method can be interpreted as the SUR estimation in the vector error-correction model. The simulation results from panel unit root tests showed that application of the SUR method increases the power of the tests when the degree of cross-correlation is high. Therefore, it would also be useful to investigate the way in which additional information from the SUR-type procedure can improve the power of the test over the standard estimators in the panel cointegration tests.

With regard to the panel unit root test with structural breaks, in the thesis, we studied the finite sample properties of the endogenous break LM unit root test, using Monte Carlo simulations. Im, Lee and Tieslau (2002) derive the asymptotic distribution of the exogenous break panel LM unit root test, which is shown to be the same as that of the test without level shifts. However, the asymptotic distribution of the endogenous break panel LM unit root test has not been fully investigated. In view of this, it would be interesting in a future study, to extend the results of the asymptotic distribution from the panel LM unit root test of Im, Lee and Tieslau (2002) to the tests with endogenous break selection procedures.

Another possible direction of future research is the consideration of the effect of cross-sectional dependence. In this thesis, the presence of cross-sectional dependence as the effect of common shocks was implicitly assumed. Under such an assumption, the effect of cross-correlations can be controlled when we allow for structural shifts in panel data unit root testing. However, the presence of cross-sectional dependence can be the result of other effects, such as the use of common base currency or model mis-specification (see Section 2.3). In light of the fact that the exogenous break panel LM test was size-distorted when the number or location of breaks were mis-specified, as highlighted in Chapter 4, and that size distortion in panel data root testing could be corrected by means of the bootstrap method, as discussed in Chapter 2, it would be interesting for future research to consider the way in which the bootstrap method could be applied to correct the size distortion problem in the panel unit root test with structural breaks.

Next, we should also consider the recent applications in the time-series unit root and cointegration testing, which can be applied to the panel data framework. An interesting subject in this respect is the nonlinear unit root test. The importance of the nonlinear unit root test, which takes account of the asymmetries in the adjustment process toward a long-run equilibrium, has been widely acknowledged in the recent literature (see van Dijk, Teräsvirta and Frances (2002)). Fok, van Dijk and Frances (2004) introduce a Smooth-Transition Regression (STR) for a panel time-series, to examine the potential presence of common nonlinear features in US industrial production modelling. This proposed test provides an interesting framework which can be used to test for unit roots in panel data, allowing for nonlinearity based on the STR model. In addition, it would be interesting to examine asymmetries and structural breaks jointly by the STR model where a deterministic trend plays the role of the transition variable.

Another research direction is exploration of the possibility that the order of integration of the series is fractional, $I(d)$, rather than integer $I(1)$ versus $I(0)$. The long-memory economic variables, such as real exchange rate, may be characterised as the fractional integration processes (see Diebold, Husted and Rush (1991)). Testing for fractional integration is well documented in the literature (see Tanaka (1999)). These tests allow the integration order of the series to adopt any value on the real line. The applications of the fractional integration to panel data testing could also be an interesting area of future research.

References

Amsler, C., and Lee, J. (1995) "An LM test for a Unit Root in the Presence of a Structural Change", *Econometric Theory*, 11, 359-368.

Andrews, D.W. (1991) "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation ", *Econometrica,* 59, 817-858.

Baharumshah, A.Z. and Ariff, M. (1997) "Purchasing Power Parity in South East Asian Countries Economies: A Cointegration Approach", *Asian Economic Journal*, 11, 141-153.

Bahmani-Oskooee, M. (1993) "Purchasing Power Parity Based on Effective Exchange Rate and Cointegration: 25 LDCs' Experience with its Absolute Formation", *World Development*, 21, 1023-1031.

Bai, J. and Ng, S. (2004) "A PANIC Attack on Unit Roots and Cointegration", *Econometrica*, 72, 1127-1177.

Banerjee, A. (1999) "Panel Data Unit Roots and Cointegration: An Overview", *Oxford Bulletin of economics and Statistics*, 61, 607-629.

Banerjee, A., Lumsdaine, R.L. and Stock, J.H. (1992) "Recursive and Sequential Tests of Unit-root and Trend-break Hypothesis: Theory and International Evidence", *Journal of Business and Economic Statistics*, 10, 271-287.

Basher, S.A. and Mohsin, M. (2002) "PPP Tests in Cointegrated Panels: Evidence from Asian Developing Countries", mimeo, University of York.

Baum, C.F, Barkoulas, J.T. and Caglayan, M. (1999) "Long Memory or Structural Breaks: Can Either Explain Nonstationary Real Exchange Rates under the Current Float?", *Journal of International Financial Markets, Institutions, and Money*, 9, 359-376.

Bernanke, B.S., Laubach, T., Mishkin, F.S. and Posen, A.S. (1999) *Inflation Targeting: Lessons from the International Experience*, Princeton, New Jersey: Princeton University Press.

Bornhorst, F. (2003) "On the Use of Panel Unit Root tests on Cross-Sectionally dependent Data: an Applicatin to PPP", Working Paper, European University Institute.

Boyd, D. and Smith, R. (1998) "Testing for Purchasing Power Parity: Econometric Issues and an Application to Developing Countries", Working Paper, University of East London.

Breitung, J. and Candelon, B. (2003) "Purchasing Power Parity during Currency Crisis: A Panel Unit Root Test under Structural Breaks", Working Paper, University of Bonn.

Breitung, J. and Meyer, W. (1994) "Testing for Unit Roots in Panel Data: Are wages on different bargaining levels cointegrated?", *Applied Economics*, 26, 353-361.

Breuer J.B., McNown, R.F. and Wallace, M.S. (2001) "Misleading Inferences from Panel Unit-Root Tests with an Illustration from Purchasing Power Parity", *Review of International Economics*, 9, 482-493.

Cerrato, M. (2001) "The Cross-sectional Dependence Puzzle", Working Paper, London Guildhall University.

Chang, Y. (2004) "Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency", *Journal of Econometrics*, 120, 263-293.

Chin, M.D. (1999) "On the Won and Other East Asian Currencies", *International Journal of Finance and Economics*, 4, 113-127.

Chin, M.D. (2000) "Before the Fall: Were East Asian Currencies Overvalued?", *Emerging Markets Review*, 1, 101-126.

Choi, I. (2001) "Unit Root Tests for Panel Data", *Journal of International Money and Finance*, 20, 249-272.

Choi, I. (2002) "Combination Unit Root Tests for Cross-Sectionally Correlated Panels", mimeo, Hong Kong University of Science and Technology.

Choong, C-K. *et al.* (2000) "The Validity of PPP Theory in ASEAN-Five: Another Look on Cointegration and Panel Data Analysis", Working Paper, Universiti Tunku Abdul Rahman.

Clemente, J., Montanes, A. and Reyes, M. (1998) "Testing for a Unit Root in Variables with a Double Change in the Mean", *Economics Letters*, 59,175-182.

Coakley, J. and Fuertes, A. (1997) "New Panel Unit Root Tests of PPP", *Economics Letter*, 57, 17-22.

Coakley, J. and Fuertes, A. (2000) "Is There a Base Currency Effect in Long-Run PPP?", *International Journal of Finance and Economics,* 5, 253-263.

Cuddington, J.T. and Liang, H. (2000) "Purchasing Power Parity over Two Centuries?", *Journal of International Money and Finance*, 19, 753-757.

Dekle, R. and Pradhan, M. (1999) "Financial Liberalization and Money Demand in the ASEAN Countries", *International Journal of Finance and Economics*, 4, 205-215.

Diebold, F.X., Husted, S. and Rush, M. (1991) "Real Exchange Rates under the Gold Standard", *Journal of Political Economy*, 99, 1252-1271.

Elliott, G., Rothemberg, T.J., Stock, J.H. (1996) "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813-836.

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

Esaka, T. (2003) "Was It Really a Dollar Peg?: The Exchange Rate Policies of East Asian Countries, 1980-1997", *Journal of Asian Economics*, 13, 787-809.

Evans, P. and Karras, G. (1996) "Do Economies Converge? Evidence from a Panel of U.S. States", *Review of Economics and Statistics*, 78, 384-388.

Fok, D., van Dijk, D. and Franses, P.H. (2004) "A Multi-Level Panel STAR Model for US Manufacturing Sectors", Working Paper, Erasmus University Rotterdam.

Frankel, J.A. and A.K. Rose (1996) "A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries", *Journal of International Economics*, 40, 209-224.

Fujii, E. (2002) "Exchange Rate and Price Adjustments in the Aftermath of the Asian Crisis", *International Journal of Finance and Economics*, 7, 1-14.

Groen, J.J.J (2000) "The Monetary Exchange Rate Model as a Long-Run Phenomenon", *Journal of International Economics*, 52, 299-319.

Groen, J.J.J., and Kleibergen, F.R. (1999) "Likelihood-based Cointegration Analysis in panels of Vector Error Correction Model", Discussion Paper, Tinbergen Institute.

Hadri, K. (2000) "Testing for Stationarity in Heterogeneous Panel Data", *Econometrics Journal,* 3, 148–161.

Hall, A. (1994) "Testing for a Unit Root in Time Series with Pretest Data-based Model Selection", *Journal of Business and Economic Statistics*, 12, 461-70.

Henry, D.F and Mizon, G.E. (1998) "Exogeneity, Causality and Co-breaking in Economic Policy Analysis of a Small Econometric Model of Money in the UK", *Empirical Economics*, 23, 267-294.

Husted, S. and MacDonald, R. (1998) "Monetary-based Models of the Exchange Rate: A Panel Perspective", *Journal of International Financial Markets, Institutions and Money*, 8, 1-19.

Husted, S. and MacDonald, R. (1999) "The Asian Currency Crash: Were Badly Driven Fundamental to Blame?", *Journal of Asian Economies*, 10, 537-550.

Im, K.S., Lee, J. and Teislau, M. (2002), "Panel LM Unit Root Tests with Level Shifts", Working Paper, University of Central Florida.

Im, K.S., Pesaran, M.H., and Shin, Y. (2003) "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, 115, 53-74.

Isard, P. (1995) *Exchange Rate Economics*, Cambridge: Cambridge University Press.

Johansen, S. (1988) "Statistical Analysis of Cointegrating Vectors", *Journal of Economic Dynamics Control*, 12, 231-254.

Johansen, S. (1995) *Likelihood-based Inference in Cointegrated VAR Models*, Oxford: Oxford University Press.

Kao, C. (1999) "Spurious Regression and Residual-based tests for Cointegration in Panel Data", *Journal of Econometrics*, 90, 1-44.

Kao, C. and Chiang, M-H. (2000) "On the Estimation and Inference of a Cointegrated Regression in Panel Data", *Advances in Econometrics*, 15, 179-222.

Kao, C., Chiang, M-H. and Chen, B. (1999) "International R&D Spillovers: An Application of Estimation and Inference in Panel Cointegration", *Oxford Bulletin of Economics and Statistics*, 61, 691-709.

Kilian, L. (1999) "Exchange Rates and Monetary Fundamentals: What do We Learn from Long-Horizon Regressions?", *Journal of Applied Econometrics*, 14, 491-510.

Kim, J.O. and Anders, W. (1991) "Real and Monetary Causes of Real Exchange Rate Movements in the Pacific Rim", *Southern Economic Journal*, 57, 1061–1070.

Kim, Y. (1990) "Purchasing Power Parity in the Long Run: A Cointegration Approach", *Journal of Money, Credit and Banking*, 22, pp.491-503.

Krugman, P. (1994) "The Myth of Asia's Miracle", *Foreign Affairs*, 73, 62-77.

Kwiatkowski, D., Phillips, P.C.B, Schmidt, P. and Shin, Y. (1992) "Testing the null hypothesis of stationarity against the alternative of a unit root", *Journal of Econometrics*, 54, 91–115.

Larsson, R., Lyhagen, J. and Löthgren, M. (2001) "Likelihood-based Cointegration Tests in Heterogeneous Panels", *Econometrics Journal*, 4, 109-142.

Lee, D.Y. (1999) "Purchasing Power Parity and Dynamic Error Correction Evidence from Asia Pacific Economies", *International Review of Economics and Finance*, 8, 199-212

Lee, J., and Strazicich, M.C. (2001) "Break Point Estimation and Spurious Rejections with Endogenous Unit Root Tests", *Oxford Bulletin of economics and Statistics*, 63, 535-558.

Lee, J., and Strazicich, M.C. (2003) "Minimum Largrange Multiplier Unit Root Test with Two Structural Breaks", *Review of Economics and Statistics*, 85, 1082-1089.

Lee, K., Pesaran, M.H. and Smith, R. (1997) "Growth and Convergence in a Multi-country Empirical Stochastic Solow Model", *Journal of Applied Econometrics*, 12, 357-392.

Levin, A. and Lin, C.F. (1992) "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", Discussion Paper, University of California at San Diego.

Levin, A. and Lin, C.F. and Chu, C.S. (2002) "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", *Journal of Econometrics*, 108, 1-24.

Li, H., and Maddala, G.S. (1996) "Bootstrapping Time Series Models", *Econometric Review*, 15, 115-158.

Lothian, J.R., and Taylor, M.P. (1996) "Real Exchange Rate Behavior: The Recent Float from the Perpective of the Past Two Centuries", *Journal of Political Economy*, 104, 488-509.

Lothian, J.R., and Taylor, M.P. (1997) "Real Exchange Rate Behaviour", *Journal of International Money and Finance*, 16, 945-954.

Lothian, J.R., and Taylor, M.P. (2000) "Purchasing Power Parity over Two Centuries: Strengthening the Case of Real Exchange Rate Stability a Reply to Cuddington and Liang", *Journal of International Money and Finance*, 19, 753-757.

Lumsdaine, R.L. and Papell, D.H. (1997) "Multiple Trend Breaks and the Unit Root Hypothesis", *Review of Economics and Statistics*, 79, 212-218.

MacDonald, R. and Taylor, M.P. (1993) "The Monetary Model to Exchange Rate: Rational Expectation, Long-Run Equilibrium and Forecasting", *IMF Staff Papers*, 40, 89-107.

MacDonald, R. and Taylor, M.P. (1994) "The Monetary Model of the Exchange Rate: Long-run Relationships, Short-run Dynamics and How to Beat a Random Walk", *Journal of International Money and Finance*, 13, 276-290.

Maddala, G.S., and Kim, I. (1999), *Unit Roots, Cointegration and Structural Change*, Cambridge: Cambridge University Press.

Maddala, G.S. and Wu, S. (1999) "A Comparative Study of Unit Root Tests with Panel Data and A New Simple Test", *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

Mark, N.C. (1990) "Real and Nominal Exchange Rates in the Long Run: An Empirical Investigation", *Journal of International Economics*, 28, 115-136.

Mark, N.C. (1995) "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability", *American Economic Review*, 85, 201-218.

McCoskey, S. and Kao, C. (1998) "A Residual-Based Test of the Null of Cointegration in Panel Data", *Econometric Reviews*, 17, 57-84.

McCoskey, S. and Kao, C. (1999) "The Monte Carlo Comparison of Tests for Cointegration in Panel Data", Working Paper, Syracuse University.

Meese, R.A. and Rogoff, K. (1983) "Empirical Exchange Rate Models of the Seventies: Do They Fit out of Sample?", *Journal of International Economics*, 14, 3-24.

Moon, H.R., and Perron, B. (2004) "Testing for a Unit Root in Panels with Dynamic Factors", *Journal of Econometrics*, 122, 81-126.

Moon, H.R., Perron, B. and Phillips, P.C.B. (2003) "Incidental Trends and the Power of Panel Unit Root Tests", mimeo, University of Southern California.

Ng, S. and Perron, P. (1995) "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag", *Journal of American Statistical Association*, 90, 268-281.

Nunes, L.C., Newbold, P. and Kuan, C-M. (1997) "Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered", *Oxford Bulletin of economics and Statistics*, 59, 435-448.

O'Connell, P.G. (1998) "The Overvaluation of Purchasing Power Parity", *Journal of International Economics*, 44, 1-19.

Oh, K-Y. (1996) "Purchasing Power Parity and Unit Root Test using Panel Data", *Journal of International Money and Finance*, 15, 405-418.

Oh, K-Y. (1999) "Are Exchange Rates Cointegrated with Monetary Model in Panel Data", *International Journal of Finance and Economics*, 4, 147-154.

Osterwald-Lenum, M. (1992) "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics", *Oxford Bulletin of economics and Statistics*, 54, 461-472.

Papell, D.H. (1997) "Searching for Stationarity: Purchasing Power Parity under the Current Float", *Journal of International Economics*, 43, 313-332.

Papell, D.H. and Theodoridis, H. (2001) "The Choice of Numeraire Currency in Panel Tests of Purchasing Power Parity", *Journal of Money, Credit, and Banking*, 33, 790-803.

Pedroni, P. (1999) "Critical Values for Cointegration tests in Heterogeneous Panels With Multiple Regressors", *Oxford Bulletin of economics and Statistics*, 61, 653-678.

Pedroni, P. (2001) "Purchasing Power Parity in Cointegrated Panels", *Review of Economic and Statistics*, 83, 727-731.

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 and Economic Statistics*, 8, 153-162.

Perron, P. (1997) "Further Evidence on Breaking Trend Functions in Macroeconomic Variables", *Journal of Econometric*, 10, 251-270.

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

Pesaran, M.H. (2003) "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence", Woking paper, University of Cambridge.

Phillips, P.C.B. and Sul, D. (2003) "Dynamic Panel Estimation and Homogeneity Testing under Cross-Section Dependence", *Econometrics Journal*, 6, 217-259.

Phylaktis K. and Kassimatis, Y. (1994) "Does the Real Exchange Rate follow a Random Walk? The Pacific Basin Perspective", *Journal of International Money and Finance*, 13, 476-495.

Quah, D. (1994) "Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data", *Economic Letters*, 77, 9-19.

Radelet, S. and Sachs, J. (1998) "The Onset of the East Asian Financial Crisis", Manuscript, Harvard Institute for International Development.

Razzaghipour, A., Fleming, G. and Heaney, R. (2001) "Deviations and Mean Reversion to Purchasing Power Parity in the Asian Currency Crisis of 1997", *Applied Economics*, 33, 1093-1100.

Rogoff, K. (1996) "The Purchasing Power Parity Puzzle", *Journal of Economic Literature*, 34, 647-668.

Rossi, B. (2004) "Confidence Intervals for Half-life Deviations from Purchasing Power Parity", Working Paper, Duke University.

Sarantis, N. (1994) "The Monetary Exchange Rate Model in the Long Run: An Empirical Investigation", *Weltwirtschaftliches Archiv*, 130, 678-711.

Sarno, L. and Taylor, M.P. (1998) "Real Exchange Rates under the Recent Float: Unequivocal Evidence of Mean Reversion", *Economic Letters*, 60, 131-137.

Schmidt, P. and 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.

Strazicich, M.C., Tieslau, M. and Lee, J. (2001) "Hysteresis in Unemployment? Evidence from Panel Unit Root Tests with Structural Change", Working Paper, University of North Texas.

Tanaka, K. (1999) "The Nonstationary Fractional Unit Root", *Econometric Theory*, 15, 549-582.

Taylor, M.P. (1995) "The Economics of Exchange Rates", *Journal of Economic Literature*, 33, 13-47.

Taylor, M.P. and Peel, D.A. (2000) "Nonlinear Adjustment, Long Run Equilibrium and Exchange Rate Fundamental", *Journal of International Money and Finance*, 19, 33-53.

Taylor, M.P. and Sarno, L. (1998) "The Behaviour of Real Exchange Rates During the Post-Bretton Woods Period", *Journal of International Economics*, 46, 281-312.

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

Vogelsang, T.J. and Perron, P. (1998) "Additional Tests for a Unit Root Allowing for a Break in the Trend Function at Unknown Time", *International Economic Review*, 39, 1073-1100.

Wang, P. (2000) "Testing PPP for Asian Economies during the Recent Floating Period", *Applied Economics Letter*, 7, 545-548.

Wu, J-L. and Chen, S-L. (1999) "Are Real Exchange Rates Stationary Based on Panel Unit Root Tests? Evidence from Pacific Basin Countries", *International Journal of Finance and Economics*, 4, 243-252.

Wu, J-L. and Wu, S. (2001) "Is Purchasing Power Parity Overvalued?", *Journal of Money, Credit, and Banking*, 33, 804-812.

Wu, Y. (1996) "Are Real Exchange Rates Nonstationary? Evidence from a Panel-Data Test", *Journal of Money Credit and Banking*, 28, 34-61.

Zivot, E. and Andrews, D. (1992) "Further Evidence on the Great Crash, the Oil-price Shock and Unit Root Hypothesis", *Journal of Business and Economic Statistics*, 10, 251-270.