Finite Sample Distributions and Non-normality in Second Generation Panel Unit Root Tests

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Qi Sun

Department of Economics

University of Leicester

December 2010

Finite Sample Distributions and Non-normality in Second Generation Panel Unit Root Tests

by

Qi Sun

Abstract

As a remarkable advantage, panel unit root testing statistics present Gaussian distribution in the limit rather than the complicated functionals of Wiener processes compared with traditional single time series tests. Therefore, the asymptotic critical values are directly used and the finite sample performance is not given proper attention. In addition, the unit root test literature heavily relies on the normality assumption, when this condition fails, the asymptotic results are no longer valid.

This thesis analyzes and finds serious finite sample bias in panel unit root tests and the systematic impact of non-normality on the tests. Using Monte Carlo simulations, in particular, the application of response surface analysis with newly designed functional forms of response surface regressions, the thesis demonstrates the trend patterns of finite sample bias and test bias vary closely in relation to the variation in sample size and the degree of non-normality, respectively. Finite sample critical values are then proposed, more importantly, the finite sample critical values are augmented by the David-Johnson estimate of percentile standard deviation to account for the randomness incurred by stochastic simulations. Non-normality is modeled by the Lévy-Paretian stable distribution. Certain degree of non-normality is found which causes so severe test distortion that the finite sample critical values computed under normality are no longer valid. It provides important indications to the reliability of panel unit root test results when empirical data exhibit non-normality.

Finally, a panel of OECD country inflation rates is examined for stationarity considering its feature of structural breaks. Instead of constructing structural breaks in panel unit root tests, an alternative and new approach is proposed by treating the breaks as a type of non-normality. With the help of earlier results in the thesis, the study supports the presence of unit root in inflation rates.

Acknowledgements

I am deeply grateful and indebted to my supervisor, Professor Wojciech Charemza, for enlightening me and guiding me throughout my PhD research, for his continuous encouragement and invaluable advice on various aspects of my study life in Leicester.

I would like to express my gratitude to Doctor Emi Mise, my PhD advisor, for her immense support and advice on my study and part time teaching at the Department of Economics. Many thanks are due to Professor Stephen Hall, Professor Peter Burridge, Professor Badi Baltagi, Professor Yoosoon Chang and Doctor Richard Handle for their invaluable suggestions and help on my research. My special thanks are to Professor Michael Tretyakov from Mathematics department, who kindly invited me to attend his Financial Mathematics lectures.

I am also indebted to my friend Doctor Ayman Gendy, who had patiently and unconditionally taught me the essential programming skills required for my research at the very beginning. Many thanks are due to my dear friends Ya'ning Raymond Wang, for his sincere support on pursuing the PhD study, and Oliver Garn, for his continuous encouragement and kind advice. My warm thanks are due to Glynn and Sue Richerby, my host family, for their kindness and love to me as of my own family.

My deep gratitude to my dear mother and father is beyond the expression with words, for their entirely unreserved love, support and encouragement throughout all my studies and life.

Finally, I would like to thank the Department of Economics, University of Leicester for offering me the Graduate Teaching Assistant position to fund my PhD study.

Contents

Approximate number of words: 43 000

Introduction

Panel data are datasets that involve observations on individual economic agents (in the cross section dimension N) over time (in the time dimension T). Compared with time series and cross section studies, panel data have been in growing favour among researchers as they gain advantages by obtaining more information and combining the information from the two dimensions in panel. For example, the changes of cross sections over time can be studied with the help of panel data; panel data can control for individual heterogeneity in the presence of certain individual invariant or time invariant variables, whereas time series or cross section studies not controlling this may lead to biased results.

The early panel data analysis lies in the field of microeconomics with the start of the construction of household panel surveys initially launched in the Unite States in the 1960s. Micro panels usually consist of large N and small T. Since the beginning of the 1990's, studies using macroeconomic panels have become increasingly popular, such as analyses involving inflation rates, interest rates, exchange rate, etc. Compared with micro panels, macro panels usually bear the feature of large N and large T. The long time series dimension in macro panels raises the concern of non-stationary panel and thus propels the development of panel unit root tests.

Introduction

Stationarity has long been one of the main subjects of time series studies. A non-stationary series (or a series containing a unit root) suggests once receiving a shock, the series will eventually drift away from its mean. This feature poses serious considerations for empirical applications, e.g. the inflation targeting, or can provide a foundation for further analysis, e.g. the cointegration analysis. Testing for unit root in a series requires unit root tests. Single time series tests have started to develop since the pioneer work of Dickey and Fuller in 1979, the Dickey-Fuller test. Afterwards, there have been considerable amount of efforts to either improve the statistical performance of the test or develop new tests to account for problems that are relevant in empirical studies, e.g. the presence of structural breaks (the abrupt, relatively large and long-lasting changes in time series).

Since attention has been raised to non-stationary panel, the study of testing for unit root in panel data has become popular, with Levin and Lin (1992) being the initial seminal contribution in the field of panel unit root tests. The main advantage of applying panel unit root tests over traditional single time series tests is the gain in statistical performance due to the increased sample size in relation to the additional cross section dimension N. In addition, panel unit root test statistics have the Gaussian distribution in the limit rather than the complicated functionals of the Wiener processes in the case of time series.

The early panel unit root tests use pooled regression models that restrict the autoregressive parameter to be homogenous across the cross section dimension N, *i.e.* assume the panel individuals to have the same behaviour. These are called the homogenous tests in the literature. However, panel data generally introduce a substantial amount of unobserved heterogeneity, which renders the parameters of the model individually specific across the cross section dimension. Subsequently the heterogeneous tests are proposed in the literature to allow the individual autoregressive parameters across panel to differ from each other. Nevertheless, these tests share a crucial and restrictive assumption that the panel individuals act independently from one another. In many empirical applications, especially in macroeconomic and finance studies such as research on inflation rates, exchange rates, etc., the imposition of cross section independence assumption is inappropriate. To overcome this difficulty, a new generation of panel unit root tests have been developed to allow for different forms of cross section dependence/correlation (by which panel individuals are related with one another) and solve the problem through various approaches.

Among the new generation of panel unit root tests the Chang and Song (2005, 2009) test (CS hereafter) is proposed as a most general test which is robust to all the different forms of cross section dependence. The test adopts an approach based on instrument variable regression, using a nonlinear transformation of the variable of interest as instrument. It is a development of the unique approach set up in Chang (2002) test (CH hereafter), a test that can only deal with the weak form of dependence and is criticized for failure in the presence of the strong form of dependence. The remarkable conclusion of the two studies is that after applying the nonlinear instrument variable technique, the *t*-statistic of the autoregressive parameter of each panel individual has standard normal limiting distribution under the unit root null hypothesis and is independent from each other. The panel unit root testing statistic, the average of the individual *t*-statistics, obviously has standard normal distribution in the limit.

Since panel unit root testing statistics generally have the Gaussian limiting distribution, one of the gaps in the literature is that their finite sample performance is not yet given proper attention. Empirical applications simply apply the asymptotical critical values of the tests. However, the well known knowledge is that finite sample performance can be substantially different from the asymptotics, in particular, when the convergence rate of the test statistic is slow. In addition, in the case of panel data the situation becomes much more complicated due to the additional cross section dimension *N* (details will be discussed in Chapter 1). Furthermore, regarding the asymptotic properties of the CH test, Im and Pesaran (2003) criticizes that for its asymptotic properties to hold, a much more restrictive condition is required, *i.e.*, $\frac{N \ln(T)}{\sqrt{T}} \to 0$ as $N, T \to \infty$. This suggests that to ensure the performance of the test, *N* needs to be very small relative to *T*, which is particularly restrictive in practice. Considering that the CS test adopts the nonlinear instrument variable approach in the same concept, the similar problem is also conjectured to the CS test by the thesis.

In econometrics the highly likely difference between the performance of finite sample and large sample is hard to evaluate analytically but can be conveniently examined through Monte Carlo experiments. Monte Carlo method is a computational approach that provides approximate solutions by performing repeated statistical sampling experiments. In this thesis, using the outputs of a series of Monte Carlo experiments, the finite sample performance of panel unit root tests (the CH and CS tests) is analyzed through the response surface method, which helps observe the pattern of the changes in finite sample distributions as sample size varies. MacKinnon (1994, 1996 and 2000) applied the response surface method to a number of univariate unit root tests and cointegration tests. The method can help provide test critical values for any given sample size; compute numerical approximations to finite sample distribution functions and thus the *p*-value for any given percentile can also be computed. In Chapter 2, the thesis firstly examines the finite sample bias of the two tests by response surface regressions through newly designed regression specification, involving the relevant factors that affect bias as explanatory variables, such as sample sizes and the condition claimed to be essential to ensure the asymptotic properties by Im and Pesaran (2003). On observing the bias, the chapter then computes the finite sample critical values. Instead of providing the traditional point critical values, the finite sample critical values are augmented by the David-Johnson estimate of percentile standard deviation to account for the randomness caused by stochastic simulations.

Nevertheless, the above unit root test literature heavily relies on the normality assumption that the error terms in the model follow Gaussian process. When this condition fails the asymptotic results are no longer valid (Hamilton, 1994). However, there is abundant evidence showing that many economics variables are non-normally distributed with heavy/fat tails and a number of studies suggest use the Lévy-Paretian stable distribution (also called the α stable distribution) to model these economic data. A few contributions had been made to the asymptotic property of single time series unit root tests under the assumption of α stable distribution in the error terms, whereas in the case of panel data almost no attention is given. Following Chapter 2 the next task in this thesis is to analyze the magnitude of bias in panel unit root tests caused by non-normality which is represented by the α stable distribution of the error terms. Chapter 3 uses a fresh design of response surface regression to discover the sensitivity of panel unit root tests to non-normality as the degree of non-normality increases. Meanwhile it aims to find out the degree of non-normality which causes so severe size distortion that the finite sample critical values computed in Chapter 2 under normality can no longer be used. The results can provide important implication and caution on the reliability of panel unit root test conclusion when empirical data present non-normality feature.

Another considerable benefit could be provided by applying the α stable distribution is that the α stable random variables can be used to simulate processes with structural breaks/changes (as mentioned earlier, the sudden and large changes in a process). Regarding the impact of structural breaks on unit root test, Perron (1989) raised the caution that structural breaks in time series tend to bias unit root tests toward not rejecting the unit root hypothesis. Research indicates that inflation rates over long time span are likely to experience structural breaks due to certain economic shocks or monetary policy changes. Various traditional single time series tests cannot reach a consensus if there is a unit root in inflation. More recent literature applies panel unit root tests to take the advantage of panel data, whereas the results are still mixed. Given the contradictory conclusions in the literature, Chapter 4 re-examines the stationarity property of inflation rate by the most robust panel unit root test, taking into account the structural break character. The chapter adopts a brand new approach to deal with structural changes: instead of modelling and estimating the breaks in the test which usually incurs complications and limitations, the appearance of breaks is treated as a form of non-normality represented by the α stable distribution.

The rest of the thesis is organized as follows: Chapter 1 reviews panel data and non-stationary panels. A number of contemporary popular panel unit root tests are then introduced in two categories divided by whether cross section dependence is considered. Chapter 2 applies the response surface method based on Monte Carlo simulation results and analyzes the finite sample bias of the CH and CS tests. Finite sample critical values of the tests are computed and augmented by the David-Johnson estimate of percentile standard deviation. Chapter 3 raises the question of the impact of non-normality on panel unit root tests. The normality assumption is relaxed for panel unit root tests and error terms are assumed to follow the α stable distribution. The bias incurred by non-normality is analyzed and illustrated through Monte Carlo simulations. The chapter continues to find out the degrees of non-normality which so severely distort the tests that the finite sample critical values computed in Chapter 2 are invalid. Chapter 4 re-examines the stationarity property of inflation rates using an inflation rate panel of 15 OECD countries. The problem of structural breaks in a process is dealt with by an alternative approach which uses the representation of non-normality to capture the abrupt changes in the process. Finally the thesis finishes with the conclusions and some suggestions for further research in this area.

Chapter 1 Panel Data and Panel Unit Root Tests in Retrospect

---

1. Panel Data and Some Examples

Panel data, or longitudinal data, refer to datasets that contain observations of cross-section individuals, such as countries, firms, households, etc. over some time periods. Let y_{it} denotes a panel data variable, where $i = 1,2,\ldots,N$; $t = 1,2,\ldots,T$; N is the number of individuals in the panel and T is time length. Contemporary economic data are often provided in the form of panels so that empirical studies are able to simultaneously analyze time series data set of states, industries or countries, e.g. Baltagi and Griffin (1983) studied gasoline demand in the OECD using 18 OECD countries over 18 years; Munnell (1990) investigated productivity growth by applying a panel of 48 contiguous states over the period 1970-86; Banerjee et al. (2005) surveyed the empirical literature of purchasing power parity (PPP) and found that PPP holds when tested in panel data, which is in contrast to those tested under time series. This may be due to the higher power of panel unit root tests compared with time series tests (more discussion is given in Section 5).

Panel data can be classified into micro panels and macro panels. Micro panels include individuals such as households, communities, firms and so on; macro panels usually contain individuals of countries or industries. Early panel data development is more focused on micro panels, e.g. studies of households, whereas in recent years growing interest has been shown in macro panels such as inflation rates, exchange rates, etc. There is increasing amount of practice in pooling times series from a number of countries or industries to analyze them simultaneously. Some examples of both micro and macro panels are illustrated below.

1.1 Micro Panels

Household panels are typical examples of micro panels. Household panel survey has become a major resource for understanding the behavior of households in each country. The variables usually cover the issues of income, labour market behaviour, social and political values, health, education, housing and household organization. The first household panel survey was launched in the United States in 1968, the Panel Survey of Income Dynamics (PIDS). By 2003, the PIDS had collected information on 65,000 individuals over 36 years (Baltagi, 2008).

The first European household panel survey started in (West) Germany, the German Socio-Economic Panel (GSOEP) in 1984. Then surveys from other countries have followed such as in the Netherlands, the Socio-Economic Panel (SEP), in Belgium, the Panel Study on Belgian Households (PSBH), in Luxembourg, the Panel Study of Economic Life in Luxembourg (PSELL) and in Britain, the British Household Panel Survey (BHPS). Each of these surveys was designed independently to meet the data needs perceived in each host country.

The BHPS (British Household Panel Survey) is carried out by ISER (Institute for Social & Economic Research) at the University of Essex. The main objective of the survey is to further understand the social and economic change at individual and household level in Britain. It intends to identify, model and forecast these changes, causes and consequences in relation to a range of social economic variables including organization, employment, accommodation, tenancy, income and wealth, housing, health, socio-economic values, residential mobility, marital and relationship history, social support, and individual and household demographics. It started in 1991 with 5,500 households and keeps the interview each year since. Recently large new samples from Scotland, Wales and Northern Ireland were added to the main sample.

The BHPS provided the UK component of the European Community Household Panel (ECHP). The ECHP was launched in 1994 by the Statistical Office of the European Communities (EuroStat) for cross-European comparability. It started survey for parallel comparable micro level data with 60,000 samples among the 12 member countries in the European Union (EU) and follows the same samples (persons and households) each year since then. It contains data on income, work, housing, health and many other indicators indicating living conditions. Three more members were added in when they joined the EU, namely, Austria, Finland and Sweden, which brought in 13,000 more households. The series continues through from 1994 to 2001 and had ended now.

1.2 Macro Panels

As we can see one important character of micro panels is that they usually have large N and small T, whereas macro panels are usually presented with large N and large T.

There are many organizations (such as the IMF, World Bank, OECD, etc.) that provide long time series macro data sets for countries all over the world or countries within a certain area. The IMF was established to promote international monetary cooperation and assist countries in balance of payment difficulties, and it has more than 180 member countries. Some major databanks provided by the IMF are the Direction of Trade Statistics (DOTS), International Financial Statistics (IFS), Government Finance Statistics (GFS) and Balance of Payments Statistics (BOPS). For example, the IFS data set covers

200 countries worldwide and contains approximately 32,000 time series on economic topics including balance of payments, banking and financial systems, labour, exchange rates, fund position, government finance, interest rates, international liquidity, national accounts, population, prices, production and trade. Most annual data began in 1948; quarterly and monthly data generally started in 1957. Large time dimension can be obtained through using monthly data. Numerous PPP studies have taken exchange rate data from the IFS.

The World Bank Group was founded in 1944. Its aim is to reduce poverty and improve the living standards of people in the developing world. Two important World Bank databases are World Development Indicators and Global Development Finance. For example, the World Bank Global Development Finance database provides annual data of more than 200 debt and financial flows indicators for 136 countries from 1970 to 2013 (where available).

The UN was established in 1945 by countries that committed on preserving peace through international cooperation and collective security. The UN produces a very large database, the UN common Database (UNCDB), which covers economic, social, finance and development issues for 280 countries. Data usually are available from 1970-1980, published every half year.

The OECD is an international alliance of national governments that acts as a forum for member countries to develop economic and social policies. It currently has 30 member countries. The OECD produces many databases, the OECD Education Statistics, the OECD Main Economic Indicators (MEI), the OECD International Development, etc. For example, the MEI contains annual, quarterly and monthly data for around 3,800 economic indicators for the OECD member countries and 6 non-member countries. Monthly data are available from January 1960 to the current month.

Some other international organizations established regarding specific topics also provide databases. For example, the International Energy Agency (IEA) was established in 1974 in response to the oil crisis. It has databases such as the Coal information, Electricity information, World Energy Statistics and Balances and so on. The IEA World Energy Statistics and Balances data set consists of four separate databases, containing the annual energy balance data from 1960 for 30 OECD countries and over 100 non-OECD countries.

2. Advantages of Panel Data

Baltagi (2008) summarizes the various advantages of using panel data. The advantages are:

1. As panel data present information for a number of individuals over time, there is bound to be heterogeneity across the individuals. Ignoring heterogeneity time series and cross section studies may be exposed to the risk of biased results (c.f. Moulton, 1986), while panel data can take into account heterogeneity. This will be discussed in more detail in the next section.

2. By combining cross section units, panel data can give more informative data, more variability, more degrees of freedom and more efficiency. For example, when time series studies are plagued with multicollinearity among independent variables, it is less likely with a panel since the cross section dimension in panel data adds a lot of variability and more informative data on the variables.

3. Compared with cross sectional data that are only observed on one point of time, panel data are better suited to study the dynamic adjustment with its time dimension.

4. Panel data are better in identifying and measuring effects that are simply not detectable in pure time series or cross section data. Suppose that a cross-section data set of

students with 80% exam passing rate. The fact might be (a) in any year there are 20% of the students fail or (b) there are 20% of the students fail all through the years. While cross section data cannot distinguish the cases, panel data could solve the problem (c.f. Baltagi, 2008).

5. Panel data models allow us to study more complicated behavioral models. For example, technical efficiency is better modeled with panel (c.f. Schmidt and Sickles, 1984; Baltagi and Griffin 1988; Koop and Steel, 2001).

6. With the additional cross section dimension in panel data, panel unit root tests have standard asymptotic distributions rather than the nonstandard distributions of time series unit root tests. This will be discussed in greater detail in the following literature of panel unit root tests.

3. Static Linear Panel Models

A typical linear panel regression model takes the following form

$$y_{it} = \alpha + X'_{it}\beta + \varepsilon_{it} \qquad i = 1,\ldots,N;\ t = 1,\ldots,T_i \tag{1.1}$$

where i denotes panel individuals, households, firms, countries, etc. in the cross section dimension and t denotes time series dimension; α is scalar; β is $k \times 1$ vector of coefficients; $X_{it} = (x_{1t},..., x_{kt})$ is $k \times 1$ vector of k explanatory variables and x_{it} is the it^{th} observation; ε_{it} are error terms and have independent identical distributions with mean 0 and variance $\sigma^2_{i\varepsilon}$, $\varepsilon_{it} \sim IID(0, \sigma^2_{i\varepsilon})$.

When the intercept term α and the coefficients of independent variables β remain constant across panel individuals, the model is homogeneous. Pooled OLS (ordinary least squares) regression (*i.e.* pool individuals across the cross section dimension) provides consistent and efficient estimations of α and β. However, heterogeneity appears when the intercept terms or the coefficients of explainable variables differ across individuals, then pooling individuals across panel is inappropriate. Heterogeneous panels refer to the cases when the parameters estimated from each unit model are different across panel individuals. Since early panel data studies mainly apply micro panels which contain large amount of individuals and short time span, they are more oriented on cross section analysis. Therefore, heterogeneity is often the central topic of panel study.[1]

3.1 Heterogeneity from the Constant Terms

The *fixed effect* and *random effect* panel data models are developed to account for the heterogeneity arising from constant term[2]. If heterogeneity is unobserved and correlated with independent variables X_{it}, the least squares estimation of β is biased and inconsistent. In this case, the *fixed effect* model can be considered

$$y_{it} = \alpha_i + X'_{it}\beta + \varepsilon_{it} \tag{1.2}$$

where α_i is taken as the group specific constant term for each individual.

If the unobserved heterogeneity is assumed to be uncorrelated with the independent variables, then the *random effect* model can be formulated as

$$y_{it} = c + X'_{it}\beta + \alpha_i + \varepsilon_{it} \tag{1.3}$$

where c is the constant term for the panel regression model, α_i denotes the unobservable individual specific effect and are drawn from identical independent distributions with mean 0 and variance σ_α^2 in each time period, $\alpha_i \sim IID(0, \sigma_\alpha^2)$. α_i and ε_{it} are independent.

[1] C.f. Green (2007).
[2] Since static linear models are not the main issue of this thesis, they are not elaborated here. For more readings, refer to Gujarati (2009), Green (2007) and Baltagi (2008).

Therefore, the essential difference between the two models is whether the elements that cause the unobserved individual effects (*i.e.* the heterogeneity) are correlated with the independent variables or not.

3.2 Heterogeneity from the Slope Coefficients

Both *fixed effect* and *random effect* models assume that the slope coefficients β are homogeneous, whereas this assumption is testable and is quite often rejected (Baltagi, 2008). Roberson and Symos (1992) studied the properties of some heterogeneous estimators which are assumed to be homogeneous. They found severe biases in dynamic estimation for both stationary and nonstationary regressors. Pesaran and Smith (1995) estimated a dynamic heterogeneous model with four procedures, and showed that when both N and T are large, cross section regression (*i.e.* average the estimators of each unit) will yield consistent estimate of the mean value of the estimators from each individual. However, Maddala et al. (1997) on studying the elasticities of residential demand for electricity found that heterogeneous panel estimation is inaccurate and even can yield wrong signs for the coefficients, and panel data estimation is invalid when homogeneity is rejected. Nevertheless, in the reconsideration using the same data sets by Baltagi et al (2002), evidence is added that in out-of-sample forecasts simple homogeneous panel data estimation outperforms individual estimates and the *shrinkage* estimators suggested by Maddala et al (1997).

4. Nonstationary Panels

In recent years with the growing interests in topics such as PPP, interest rate, inflation rate, growth convergence and so on, the focus of panel data study has shifted toward macro panels. Since macro panels often contain relatively large N and large T, more attention is therefore devoted to the stationarity property of macro panels on considering their long time period.

A time series y_{it} is stationary if the means and variances of the process are constant over time, and the covariance between two periods depends only on the gap between the periods, *i.e.*

$$E(y_{it}) = \mu; \qquad Var(y_{it}) = \sigma^2; \qquad Cov(y_{it}, y_{i,t+j}) = \gamma_j$$

If one or more of the conditions are not fulfilled, the process is nonstationary. Assume the time series $\{y_{i0}, ..., y_{iT}\}$ is generated by an *AR*(1) process

$$y_{it} = \mu + \rho_i y_{i,t-1} + \varepsilon_{it} \tag{1.4}$$

where $\varepsilon_{it} \sim IID(0, \sigma_{i\varepsilon}^2)$. When $N = 1$, (1.4) reduces to an *AR*(1) of single time series

$$y_t = \mu + \rho\, y_{t-1} + \varepsilon_t \tag{1.5}$$

It can be easily proved that as $|\rho| < 1$, the series is stationary; when $\rho = 1$ the series is explosive with increasing variance as time proceeds, so y_t is nonstationary (c.f. Enders 2004). The nonstationary process is often called a *unit root process* or a *process with a unit root*.

When $N > 1$, the single time series is extended to panel data as in (1.4). If $\rho_i < 1$ for all $i = 1,..., N$, the panel is stationary; if $\rho_i = 1$ for all $i = 1,..., N$, y_{it} is a nonstationary panel. However, due to the possible heterogeneity in ρ_i across the cross section dimension, so far there is not a clear definition regarding the stationarity of a panel for the situation when ρ_i 's behave differently.

As the use of panel data has become popular, numerous studies have been extended from single time series to panel data. To examine the stationarity property of panel data, the field of panel unit root test is progressing productively and is applied on various topics (such as those stated at the start of this section).

5. Unit Root Tests

5.1 Testing for Unit Root in Single Time Series

Refer to (1.5) the *AR*(1) process, the null hypothesis of a unit root test is $\rho = 1$ against the alternative $|\rho| < 1$. The OLS estimator $\hat{\rho}$ of ρ is consistent, however, Dickey and Fuller (1979) demonstrates that under the null $\rho = 1$, the *t*-statistic of $\hat{\rho}$ presents a distribution skewed to the right (with a long tail on the right-hand side) rather than the standard asymptotic distribution due to the nonstationarity of unit root process

$$t_{\hat{\rho}} \Rightarrow \frac{\int_0^1 W(r)dW(r)}{\sqrt{\int_0^1 W(r)^2 dr}} = \frac{\frac{1}{2}\left(W(1)^2 - 1\right)}{\sqrt{\int_0^1 W(r)^2 dr}} \tag{1.6}$$

where $W(\cdot)$ is standard Brownian motion/Wiener process; $r \in [0,1]$. Consequently, the *t*-distribution table or standard normal distribution table can not be used. Dickey and Fuller calculated the critical values of different rejection percentiles for the appropriate distribution. Their test is referred to as the *Dickey-Fuller* test and has been widely used to testing for unit root of time series ever since. Thereafter, a wave of unit root tests were developed, such as Sargan and Bhargava (1983), Phillips (1987), Cochrane (1988), Elliott et al. (1996), Ng and Perron (2001) and Perron (1997), etc. The later tests are developed either to improve the statistical performance of some existing tests or to consider some special situations that frequently appear in practice (e.g. structural breaks).

5.2 Testing for Unit Root in Panel Data

Although testing for unit root in single time series had become popular at early stage, testing for unit root in panel data is only recent since the seminal work of Levin and Lin (1992, 1993). The additional cross section dimension in panel data provides advantage and meanwhile poses problems and difficulties compared with time series.

5.2.1 Advantages of Panel Unit Root Test

Since adding cross section dimension has greatly increased the number of observations, one of the main advantages of applying panel unit root test is the gain in statistical power compare with the poor power performance of univariate series due to their relatively limited sample size. For example, a number of studies on PPP using panel data find that the real exchange rate is stationary (c.f. Flood and Talor; 1996, Choi, 2001), whereas this is normally rejected by the traditional ADF (Augmented Dicky-Fuller) test. This may result from the higher power of panel unit root tests than that of time series tests. As noted by Baltagi and Kao (2000), nonstationary panel data econometrics aims at combining *the best of the two worlds: the method of dealing with non-stationary data from the time series and the increased data and power from the cross-section*. In addition, a distinctive feature of panel unit tests is that the test statistics have normal limiting distributions rather than the complicated functional forms in the case of time series. Under certain assumption the cross section dimension can act as repeated draws from the same distribution, thus the distribution of the test statistics converge to standard normal distribution as the number of individuals increases. Panel data can also help avoid the problem of spurious regression while produce a consistent estimates of the true values as both N and T go to ∞ (Baltagi 2008).

5.2.2 Problems in Panel Unit Root Test

Although panel unit root tests exhibit significant advantages over time series tests, there are also more complicated problems. For example, early panel unit root tests impose the restrictive assumption of homogeneity. The following wave of the tests is criticized by ignoring cross section dependence (which will be introduced below). Due to the critics, two generations of tests are divided in the literature by the consideration of cross section correlation. In addition, panel unit root tests involve more complicated asymptotic properties caused by the two dimensions (N and T).

5.2.2.1 Heterogeneity and Cross Section Dependence

As discussed above, panel data generally introduce significant amount of heterogeneity. Ignoring this problem by pooling data across individuals can cause serious inconsistency (c.f. Pesaran and Smith, 1995). In addition, to assume the individuals in the cross section dimension independent from each other is inappropriate in many empirical applications such as exchange rate, inflation rate, etc. If this assumption is violated, it can lead to size distortions and low power (c.f. Banerjee et al, 2004). Therefore, the recent wave of panel unit root tests tries to include cross section dependence into the model through various approaches. Cross section dependence (or correlation) refers to the situation when panel individuals are related with one another. It suggests that while certain individuals receive shocks in the panel, others will be affected as well. The dependence can be represented by the variance covariance matrix of the error terms or common factor(s) that drive the common trend(s) among individuals (this will be presented in detail in section 6.2). In the literature the tests considering cross section dependence are categorized as the second generation panel unit root tests. Thus the first generation tests refer to those that assume cross section independence. In the application to the stationarity property of real interest rate, in contrast to the earlier results from the first generation tests, the second generation tests tend to reject unit root (c.f. Pesaran, 2007).

5.2.2.2 Asymptotic Properties

Regarding the asymptotic properties, there are more serious complications in panel models due to the additional cross section dimension *N*. Conventional limit theory cannot be directly applied because it has only one index tending to infinity. Phillips and Moon (2000) overview three asymptotic approaches used on panel limit theory according to the way that *N* and *T* pass to infinity. (a) Sequential Limits, which fixes one index (say *N*) and allows the other passes to infinity (say *T*) to obtain an intermediate limit. Then let *N* go to infinity subsequently to obtain a sequential limit theory. This approach is used by, for example, Im, Pesaran and Shin (1997) test which is reviewed in the next section. (b) Diagonal Path Limits, which allow both *N* and *T* to pass to infinity along a specific diagonal path in the two dimension array. This path can be determined by a monotonically increasing functional relation of the type $T = T(N)$ which applies as the index $N \rightarrow \infty$. Levin and Lin (1992, 1993) test uses this approach in finding the limits of panel unit root test statistics. (c) Joint Limits, which allow both *N* and *T* to pass to infinity simultaneously without specific diagonal path restrictions on the divergence, although some control over the relative rate of expansion of the two indexes may be necessary to obtain definitive results. The joint limits approach suggests $T = N$ as $N \rightarrow \infty$, so it is a special case of the diagonal path limits approach.

Although asymptotic properties guarantee the convergence of testing statistics, they are based on sample size tending to infinity. This poses a serious problem for empirical applications, since infinity is not achieved in reality and very often even the available sample observations can be very limited. Consequently critical values given by asymptotic distributions may distort empirical results and therefore finite sample performance of the tests is demanded.

6. Review of Panel Unit Root Tests

As was discussed above, cross section dependence acts as a dividing line to separate the first and second generation panel unit root tests. Meanwhile among the first generation tests, the problem of heterogeneity in panels was concerned as the literature develops. Therefore, after the pioneer seminal work of Levin and Lin (1992, 1993), a homogeneous test that is based on pooled estimator of the autoregressive parameter, a number of tests considering heterogeneous models were developed, such as Im, Pesaran and Shin (1997, 2003), Maddala and Wu (1999), Choi (2001) and Hadri (2000). The second generation tests include Harvey and Bates (2003), Jönsson (2005a), Breitung and Das (2005), Chang (2002), Bai and Ng (2004), Moon and Perron (2004), Phillips and Sul (2003), Pesaran (2007), Chang and Song (2005, 2009), etc. So far the first generation tests have been programmed in commercial packages, e.g. Eviews, which can be conveniently used for empirical applications. The second generation tests are not yet available in the commercial packages. Nevertheless, Gauss or Matlab programmes for most of the tests are available from the authors.

6.1 First Generation Panel Unit Root Tests

6.1.1 Levin and Lin (1992) (LL) and Levin, Lin and Chu (2002) (LLC) Tests

The basic models of the LL tests are

$$y_{it} = \rho y_{i,t-1} + \varepsilon_{it} \tag{1.7}$$

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \varepsilon_{it} \tag{1.8}$$

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \gamma_i t + \varepsilon_{it} \tag{1.9}$$

where $\varepsilon_{it} \sim IID(0,\sigma^2)$ for all i = 1,2,…,N and all t = 1,2,…,T. These models can be presented equivalently as those in LLC test

$$\Delta y_{it} = \delta y_{i,t-1} + \zeta_{it} \tag{1.10}$$

$$\Delta y_{it} = \alpha_i + \delta y_{i,t-1} + \zeta_{it} \tag{1.11}$$

$$\Delta y_{it} = \alpha_i + \delta y_{i,t-1} + \gamma_i t + \zeta_{it} \tag{1.12}$$

where $\varsigma_{it} \sim IID(0,\sigma^2)$. The models therefore allow for fixed effects (in α_i) and individual-specific time trend (in γ_i) to capture some heterogeneity; while ρ or δ, the coefficient of $y_{i,t-1}$, is restricted to be homogeneous.

The null hypothesis is $H_0 : \rho = 1$ against the alternative $H_0 : \rho < 1$ in (1.7)-(1.9), or $H_0 : \delta = 0$ against $H_0 : \delta < 0$ in (1.10)-(1.12). The asymptotic properties were investigated under the assumptions whether the model has fixed effects and individual-specific time trend. In the simplest case of the LL test (1.7), under the null hypothesis $\rho = 1$, as $N\to\infty$ and $T\to\infty$, the asymptotic distribution of the t-statistic of pooled OLS estimator $\hat{\rho}$ is given by

$$t_{\hat{\rho}} \Rightarrow N(0,1)$$

Levin and Lin found that in the cases of individual-specific fixed effects as (1.8) and serial correlation in the disturbances, the t-statistic of $\hat{\rho}$ diverges. However, with a transformation of the t-statistic, it does converge to N(0,1). As $N\to\infty$, $T\to\infty$ and $\frac{\sqrt{N}}{T} \to 0$, the asymptotic distributions of $t_{\hat{\rho}}$ is given by

$$\sqrt{1.25}t_{\hat{\rho}} + \sqrt{1.875N} \Rightarrow N(0,1)$$

Model (1.9) allows for different time trend across units. Under the null hypothesis $\rho = 1$ and $\gamma_i = 0$ as $N \to \infty$, $T \to \infty$ and $\frac{\sqrt{N}}{T} \to 0$, $t_{\hat{\rho}}$ has the asymptotic distributions as

$$\sqrt{\frac{448}{277}}\left\{t_{\hat{\rho}} + \sqrt{3.75N}\right\} \Rightarrow N(0,1)$$

The scaling such as $\sqrt{1.25}$ and $\sqrt{\frac{448}{277}}$ is derived from the moments of the terms involving Brownian Motions; the expressions such as $\sqrt{1.875N}$ and $\sqrt{3.75N}$ are centering corrections needed in order for the statistics to have mean zero asymptotically. In the case with serial correlation, Levin and Lin derived the asymptotics through estimating the average variance of the dependent variable and of the error terms as in Philips (1987).

The significance of LL test is, as Banerjee (1999) noted, that it is the first formal demonstration of asymptotic normality of panel unit root testing statistics subject to suitable scaling and corrections; and that it is the first piece of work that focuses on the rates at which T and N tend to infinity and develops a *joint limit* theory of panel unit tests. Later the LL test was extended to the LLC test to consider more general serial correlation and heteroscedasticity in the error terms. Levin et al. (2002) proposed a three-step estimation procedure; under the null hypothesis $\delta = 0$ in (1.10), $t_{\hat{\delta}}$ has standard normal limiting distribution as $N \to \infty$ and $T \to \infty$. Since the presence of specific-individual fixed effects, expressed through α_i as in (1.11) and (1.12), causes $t_{\hat{\delta}}$ to diverges to negative infinity, an adjusted t-statistic $t^*_{\hat{\delta}}$ was suggested. It is shown that $t^*_{\hat{\delta}}$ also converges to $N(0,1)$ as $N \to \infty$ and $T \to \infty$. Detailed estimation procedures and results are not introduced in the thesis. They can be found in Levin et al. (2002).

The LL test assumes the autoregressive parameter to be homogenous across panel individuals. This means that either all the individuals collectively have a unit root or they do not. It is very restrictive for the empirical analysis, e.g. in growth convergence study, it does not make sense to assume that all countries will converge at the same rate if they do converge (Maddala and Wu, 1999). Later work that relaxes this point includes Im et al. (1997, 2003), Maddala and Wu (1999) and Choi (2001).

6.1.2 Im, Pesaran and Shin (1997, 2003) (IPS) Test

In a similar form to (1.11) in LLC test, the basic model of the IPS test is

$$\Delta y_{it} = \alpha_i + \delta_i y_{i,t-1} + \varepsilon_{it} \tag{1.13}$$

where $\varepsilon_{it} \sim N(0, \sigma_i^2)$ for i = 1,…,N; t = 1,…,T; so heterogeneity is further presented in σ_i across panel individuals. The null hypothesis is $H_0 : \delta_i = 0$ for all i against the alternative $H_0 : \delta_i < 0$ for i = 1,2,…N_1 and $\delta_i = 0$ for $i = N_1+1, N_1+2,\ldots,N$, $0<N_1<N$. The alternative hypothesis therefore allows for heterogeneous δ_i. The test is based on the average of individual Dickey-Fuller (DF) test statistics. Let t_{iT} be the standard DF statistic for the i^{th} individual, so t_{iT} converges to DF distribution (1.6). The average statistic which is referred to as *t-bar* statistic is

$$t\_bar_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_{iT}$$

Due to the cross section independence assumption, the individual DF *t*-statistics, t_{iT}, i =1,…,N, are identically and independently distributed with mean $E(t_T)$ and finite variance $Var(t_T)$. Under the null hypothesis $\delta_i = 0$ when T is fixed (T>5) and as $N\to\infty$, the limiting distribution of the standardized *t-bar* statistic

$$Z_{tbar} = \frac{\sqrt{N}\left\{t\_bar_{NT} - E(t_T)\right\}}{Var(t_T)}$$

is standard normal, *i.e.*

$$Z_{tbar} \overset{N}{\Rightarrow} N(0,1)$$

However, with fixed *N* and *T*, the sample distribution of *t-bar* statistic under the null is nonstandard. So Im et al. computed the sample critical value through simulation experiments.

In the case with serial correlated errors, the model is generalized as the Augmented Dickey Fuller (ADF) regression with lag orders of q_i

$$\Delta y_{it} = \alpha_i + \delta_i y_{i,t-1} + \sum_{j=1}^{q_i} \beta_{ij} \Delta y_{i,t-1} + \varepsilon_{it} \tag{1.14}$$

The t-bar statistic, $\tilde{t}\_bar_{NT}$, is calculated as the average of individual ADF *t*-statistics $t_{iT}(q_i, \beta_i)$

$$\tilde{t}\_bar_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_{iT}(q_i, \beta_i)$$

When *T* is fixed, t_bar_{NT} will depend on the nuisance parameters β_i. This will invalid the standardization of using $E[t_{iT}(q_i, \beta_i)]$ and $Var[t_{iT}(q_i, \beta_i)]$. Since when *T* and *N* are sufficiently large the t-bar type test is free from the nuisance parameters, a *sequential limit* theory is developed as $T \to \infty$ followed by $N \to \infty$.

As $T \to \infty$ the individual ADF statistics converge to the DF distribution, say η_i, denote

$$x_{iT}(q_i, \beta_i) = \frac{t_{iT}(q_{i,} \beta_i) - E(\eta_i)}{\sqrt{Var(\eta_i)}}$$

for fixed *N*. Let the t-bar statistic, $\tilde{t}\_bar_{NT}$, distribute with mean $E(\eta)$ and variance $Var(\eta)$. In Im et al. (2003, p.64) it shows that the standardized *t-bar* statistic

$$Z_{\tilde{t}bar}(\mathrm{q},\beta)=\frac{\sqrt{N}\left\{\tilde{t}\_bar_{NT}-E(\eta)\right\}}{\sqrt{Var(\eta)}}=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}x_{iT}(q_i,\beta_{\mathrm{i}})$$

converges to a certain distribution

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}x_{iT}(q_i,\beta_{\mathrm{i}})\overset{T}{\Rightarrow}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}x_i$$

where $x_i=\dfrac{\eta_i-E(\eta_i)}{Var(\eta_i)}$, the limiting distribution of x_{iT}. Since $x_i\sim IID(0,1)$, then let $N\to\infty$, the standard normal distribution is obtained

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}x_i\overset{N}{\Rightarrow}N(0,1)$$

Therefore the sequential normal limiting distribution of the standardized t-bar statistic as $T\to\infty$ followed by $N\to\infty$ can be denoted as

$$Z_{\tilde{t}bar}(\mathrm{q},\beta)\overset{T,N}{\Rightarrow}N(0,1)$$

Im et al. (2003) also suggested an alternative sequential limit theory which calculates the standardized t-bar statistic using the means and variances of $t_{iT}(p_i,0)$ evaluated under $\beta_i=0$, namely $E[t_{iT}(q_i,0)|\beta_i=0]$ and $Var[t_{iT}(q_i,0)|\beta_i=0]$. The values of $E[t_{iT}(q_i,0)|\beta_i=0]$ and $Var[t_{iT}(q_i,0)|\beta_i=0]$ for different q and T are computed via Monte Carlo simulations. A diagonal path convergence result is conjectured by the authors as N and $T\to\infty$ simultaneously while $\dfrac{N}{T}\to k$, k being a finite non-negative constant. Monte Carlo results indicate that the small sample performance of t-bar test is generally better than that of LL test.

6.1.3 Maddala and Wu (1999) (MW) and Choi (2001) Fisher Type Tests

Maddala and Wu reviewed the LL and IPS tests and pointed out some limitations of the two tests.[3] They then proposed a Fisher type test, combining the *p*-values (or significant levels) of the unit root test statistics from each panel individual. Due to Fisher (1932), the MW panel unit root test statistic is given by

$$P = -2\sum_{i=1}^{N} \ln(p_i)$$

where p_i is the *p*-value of unit root test statistic from the i^{th} panel individual; N is the number of individuals. P is then distributed as χ^2 with $2N$ degrees of freedom with the assumption of cross section independence. The test is a non-parametric and exact test.

Since the alternative hypothesis of MW test is the same as that of IPS test, *i.e.* more general than LL test, they are directly comparable. Considering that IPS test is restricted in only using ADF statistics as a base, Banerjee (1999) notes that the simplicity of MW test and its robustness to individual unit root test statistic choice, different lag length and sample size of each individual makes it extremely attractive. Through their Monte Carlo simulations, Maddala and Wu showed that in the case that the panel is a mixture of stationary and non-stationary series as an alternative hypothesis, MW test has the highest power in distinguishing the null and the alternative compared with the IPS and LL tests.

Choi (2001) suggested a similar combining *p*-value Fisher type test. In addition to MW test with finite N as $T\rightarrow\infty$, he also derived the sequential limit results of the modified test statistic in the case of infinite N, *i.e.* $N\rightarrow\infty$. Under the null hypothesis as $T\rightarrow\infty$ followed by $N\rightarrow\infty$, the modified statistic has standard normal distribution.

[3] See Maddala and Wu (1999) for the limitations.

$$P_m = \frac{1}{2\sqrt{N}} \sum_{i=1}^{N} \left(-2\ln(p_i) - 2 \right) = -\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\ln(p_i) + 1 \right)^4$$

$$P_m \Rightarrow N(0,1)$$

Although the Fisher type tests are simple and robust, there is a disadvantage that the *p*-values have to be derived by Monte Carlo simulations (Baltagi and Kao, 1999). In addition, they still impose the restrictive assumption of cross section independence which is often violated in practice.

6.1.4 Residual Based Stationarity Test: Hadri (2000) Test

All the tests introduced above have unit root as their null hypothesis. Hadri (2000) notes that unless there is strong evidence of stationarity, classical unit root tests have low power. Hadri reckons that it would be useful to perform a stationarity test, *i.e.* set the null hypothesis to stationarity. By extending the work of Kwiatkowski et al. (1992), he proposed a residual based Lagrange multiplier (LM) test for panel data. He claims that in contrast to the previous single time series stationarity tests whose either critical values or asymptotic distribution moments are calculated by simulations, the moments of the asymptotic distribution in this paper are calculated exactly.

The basic model is

$$y_{it} = r_{it} + \beta_i t + \varepsilon_{it} \tag{1.15}$$

and r_{it} is a random walk

$$r_{it} = r_{i,t-1} + u_{it} \tag{1.16}$$

---

4 Let $P_i = -2\ln(p_i)$, then $E(P_i) = 2$ and $Var(P_i) = 4$. $P_m = \frac{1}{\sqrt{N}} \left\{ \frac{\sum_{i=1}^{N} (P_i - E(P_i))}{\sqrt{Var(P_i)}} \right\}$

where $\varepsilon_{it} \sim IIDN\left(0,\sigma_\varepsilon^2\right)$ and $u_{it} \sim IIDN\left(0,\sigma_u^2\right)$ for i = 1,…,N; t = 1,…,T. After back substitution, (1.15) can be written as

$$\begin{aligned} y_{it} &= r_{i0} + \beta_i t + \sum_{i=1}^{N} u_{it} + \varepsilon_{it} \\ &= r_{i0} + \beta_i t + e_{it} \end{aligned} \tag{1.17}$$

where $e_{it} = \sum_{i=1}^{N} u_{it} + \varepsilon_{it}$. If $\sigma_u^2 = 0$, then r_{it} reduces to a constant, and e_{it} is identical to ε_{it} and is therefore stationary. On the other hand, if $\sigma_u^2 \neq 0$, then r_{it} is a random walk, and e_{it} is nonstationary. So the null hypothesis is simply $\sigma_u^2 = 0$. More specifically, $H_0 : \lambda = 0$ against the alternative $H_0 : \lambda > 0$, where $\lambda = \dfrac{\sigma_u^2}{\sigma_\varepsilon^2}$. The LM statistic is then given by

$$LM = \frac{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \dfrac{1}{T^2} \sum_{t=1}^{T} S_{it}^2}{\hat{\sigma}_\varepsilon^2}$$

where S_{it} is the partial sum of the residuals, and $\hat{\sigma}_\varepsilon^2$ is a consistent estimator of σ_ε^2 under the null hypothesis

$$S_{it} = \sum_{j=1}^{t} \hat{\varepsilon}_{ij} \quad \text{and} \quad \hat{\sigma}_\varepsilon^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\varepsilon}_{it}^2$$

Under the null hypothesis as $T\rightarrow\infty$ followed by $N\rightarrow\infty$, the standardized LM test statistic has standard normal limiting distribution.

Hadri then generalized the model by allowing for serial correlation and heteroscedasticity in the error term across panel individuals. However, the cross section independence assumption still applies.

6.2 Second Generation Panel Unit Root Tests

As the contemporaneous correlation among panel individuals is often found present in empirical analysis[5], the second generation panel unit root tests aim to relax the independence assumption and model the cross section dependence through various approaches.

A though panel unit root tests review is provided by Breitung and Pesaran (2008). Breitung and Pesaran distinguish two cases of cross section dependence, the weak and strong forms of dependence. The general form of dependence can be presented as

$$y_{it} = \alpha_i + \rho_i y_{i,t-1} + u_{it} \tag{1.18}$$

or equivalently as

$$\Delta y_{it} = \alpha_i + \delta_i y_{i,t-1} + u_{it} \tag{1.19}$$

$$u_{it} = \boldsymbol{\gamma}_i \mathbf{F_t} + \xi_{it} \tag{1.20}$$

or

$$\mathbf{u_t} = \boldsymbol{\Gamma}\mathbf{F_t} + \boldsymbol{\xi}_\mathbf{t} \tag{1.21}$$

where $\mathbf{F_t}$ is an $m \times 1$ vector of serially uncorrelated unobserved common factors with covariance matrix $\mathbf{I_m}$; $\boldsymbol{\Gamma}$ is an $N \times m$ matrix of factor loadings defined as $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, ..., \boldsymbol{\gamma}_N)'$; $\boldsymbol{\xi}_\mathbf{t} = (\xi_{1t}, ..., \xi_{Nt})'$ is an $N \times 1$ vector of serially uncorrelated errors with zero mean and positive definite covariance matrix $\boldsymbol{\Omega}_\xi$. $\mathbf{F_t}$ and ξ_{it} are assumed to be independently distributed[6]. The covariance matrix of $\mathbf{u_t}$ is given by $\boldsymbol{\Omega} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}' + \boldsymbol{\Omega}_\xi$.

Breitung and Pesaran (2008) specify the two cases of cross section dependence: (*i*) *Weak dependence*. Under this assumption the unobserved common factors are excluded. The dependence is only generated by the covariance matrix of $\boldsymbol{\xi}_\mathbf{t}$, $\boldsymbol{\Omega}_\xi$. (*ii*) *Strong de-*

[5] C.f. O'Connell (1998) for PPP and Phillips and Sul (2003) for output convergence.
[6] If $\gamma_1 = \gamma_2 = \ldots = \gamma_N$ and $\boldsymbol{\Omega}_\xi$ is diagonal (no correlation in ξ_{it}), then $\theta_t = \gamma'\mathbf{f}_t$ (a time effect) can be removed by demeaning across panel units, and thus eliminate dependence such as in IPS test. However, if the assumptions are violated, size distortion still remains (c.f. Strauss and Yigit, 2003).

pendence. In this case the unobserved common factors exist. The dependence could arise from both $\mathbf{F}_t$ and $\boldsymbol{\xi}_t$ and is represented by $\boldsymbol{\Omega} = \boldsymbol{\Gamma\Gamma}' + \boldsymbol{\Omega}_\xi$. However, when $\boldsymbol{\Omega}_\xi$ is a diagonal matrix, the dependence only arises from the unobserved factor(s).

6.2.1 Tests with Weak Dependence

Refer to (1.18)-(1.21), in the weak form of cross section dependence, dependence is only generated from $\boldsymbol{\xi}_t$ represented by its non-diagonal variance covariance matrix

$$\boldsymbol{\Omega}_\xi = E\left(\mathbf{u}_t \mathbf{u}_t{}'\right)$$

As (1.18) or (1.19) can also be seen as a seemingly unrelated regression (SUR) system, the GLS and OLS estimation is applicable (O'Connell, 1998).

6.2.1.1 Harvey and Bates (2003) GLS Based Test

Harvey and Bates (2003) consider the homogeneous multivariate model in vector form

$$\mathbf{y}_t = \boldsymbol{\Phi}\mathbf{y}_{t-1} + \mathbf{u}_t, \qquad t = 1, \ldots, T \tag{1.22}$$

where $\mathbf{y}_t$ is $N \times 1$ vector and y_0 is fixed but unknown; $\boldsymbol{\Phi} = \phi\, \mathbf{I}_N$ is an $N \times N$ matrix of autoregressive parameters and ϕ a is scalar, so the model is homogeneous and the test statistic is invariant to pre-multiplication of $\mathbf{y}_t$ by a nonsingular $N \times N$ matrix; $\mathbf{u}_t$ is a Gaussian $N \times 1$ disturbance vector with positive definite covariance matrix $\boldsymbol{\Omega}_u$. The intention is to generalize ADF test, based on the *t*-statistic of the feasible GLS estimator of $\pi = \phi - 1$. Thus the test hypothesis is $H_0 : \pi = 0$ (or equivalently $\phi = 1$), against the alternative $H_0 : \pi < 0$.

The feasible GLS estimator $\hat{\pi}$ obtained through maximum likelihood (ML) estimation is

$$\hat{\pi} = \frac{\sum_{t=2}^{T} \mathbf{y}'_{t-1} \hat{\mathbf{\Omega}}_{u}^{-1} \Delta \mathbf{y}_{t}}{\sum_{t=2}^{T} \mathbf{y}'_{t-1} \hat{\mathbf{\Omega}}_{u}^{-1} \mathbf{y}_{t-1}}$$

where $\hat{\mathbf{\Omega}}_{u} = T^{-1} \sum_{t=2}^{T} \left(\Delta \mathbf{y}_{t} - \hat{\pi}\ \mathbf{y}_{t-1}\right)\left(\Delta \mathbf{y}_{t} - \hat{\pi}\ \mathbf{y}_{t-1}\right)'$. The t-statistic (referred to as *multivariate homogeneous Dickey-Fuller* statistic, MHDF) is then given as

$$t_{gls}(N) = \frac{\sum_{t=2}^{T} \mathbf{y}'_{t-1} \hat{\mathbf{\Omega}}_{u}^{-1} \Delta \mathbf{y}_{t}}{\sqrt{\sum_{t=2}^{T} \mathbf{y}'_{t-1} \hat{\mathbf{\Omega}}_{u}^{-1} \mathbf{y}_{t-1}}}$$

If $\mathbf{y}_{t}$ is pre-multiplied by $\hat{\mathbf{\Omega}}_{u}^{-1/2}$ in the first place, then OLS estimation can also be applied to the pooled observations.

Harvey and Bates derived the sequential limit of $t_{gls}(N)$. Under the null hypothesis for fixed N as $T \rightarrow \infty$, $t_{gls}(N)$ converges to an intermediate limit associated with Wiener process; as $N \rightarrow \infty$, it converges to standard normal distribution.

However, Chang (2004) demonstrates that the limiting distributions of the test statistics obtained from either GLS or OLS estimation are non-standard and heavily depend on nuisance parameters in the error covariance matrix that defines cross section dependence and heterogeneous serial correlation (Chang, 2004, p.273). She then applied bootstrap methodology and computed the critical values by simulations.

6.2.1.2 Breitung and Das (2005) and Jönsson (2005) OLS Based Tests

Breitung and Das (2005) point out that the GLS estimator can only be used when $N \le T$, since otherwise $\hat{\mathbf{\Omega}}_{\mathbf{u}}$ is singular. In addition, the performance of GLS is very poor unless *T* is substantially larger than *N*.[7] Jönsson (2005a) also shows through Monte Carlo simulations that as the degree of cross section correlation increases, the size distortion in MHDF test becomes larger within finite sample size. So tests based on *panel corrected standard errors* (PCSE) were developed by the authors to handle different degrees of cross section correlation and different sizes of *N* and *T* (*i.e.* to enable the test to be applicable when *N*>*T*).

Consider the equivalent homogeneous model to (1.22)

$$\mathbf{\Delta y_t} = \mathbf{\delta y_{t-1}} + \mathbf{u_t} \tag{1.23}$$

where all terms are $N \times 1$ vectors and $\mathbf{u_t}$ has positive definite covariance matrix $\mathbf{\Omega_u}$. The test statistic is based on pooled OLS regression. The estimator for the variance of the OLS estimator is given as

$$\hat{v}_{\hat{\delta}} = \frac{\sum_{t=1}^{T} \mathbf{y}'_{\mathbf{t-1}} \hat{\mathbf{\Omega}}_{\mathbf{u}} \mathbf{y}_{\mathbf{t-1}}}{\left(\sum_{t=1}^{T} \mathbf{y}'_{\mathbf{t-1}} \mathbf{y}_{\mathbf{t-1}} \right)^2}$$

where $\hat{\mathbf{\Omega}}_{\mathbf{u}} = \frac{1}{T} \sum_{t=1}^{T} \left(\mathbf{\Delta y_t} - \hat{\mathbf{\delta}}\ \mathbf{y_{t-1}} \right) \left(\mathbf{\Delta y_t} - \hat{\mathbf{\delta}}\ \mathbf{y_{t-1}} \right)'$, $\hat{\mathbf{\delta}}$ is the OLS estimator of $\mathbf{\delta}$ in (1.23).

The corresponding testing statistic is

$$t_{ols} = \frac{\hat{\delta}}{\sqrt{\hat{v}_{\hat{\delta}}}} = \frac{\sum_{t=1}^{T} \mathbf{y}'_{\mathbf{t-1}} \mathbf{\Delta y_t}}{\sqrt{\sum_{t=1}^{T} \mathbf{y}'_{\mathbf{t-1}} \hat{\Omega}_u \mathbf{y}_{\mathbf{t-1}}}}$$

[7] See the empirical sizes in Breitung and Das (2005).

Breitung and Das proved that as $T \to \infty$ followed by $N \to \infty$ under the null hypothesis $\boldsymbol{\delta} = 0$ the test statistic t_{ols} has standard normal limiting distribution.

6.2.1.3 Chang (2002) Nonlinear IV Test

As an alternative approach, Chang (2002) proposed a panel unit root test based on non-linear instrument variable (IV) estimation of the ADF regression of each individual. It uses the nonlinear transformations of the lagged levels of the series as instruments. In contrast to the non-standard DF distribution, Chang showed that the nonlinear IV *t*-statistic for each cross section individual has standard normal limiting distribution under the unit root hypothesis, even if cross section correlation exists. Therefore the test is free from the need of modeling dependence.

Consider the model

$$y_{it} = \rho_i y_{i,t-1} + \sum_{k=1}^{q_i} \beta_{i,k} \Delta y_{i,t-k} + \varepsilon_{it}, \qquad i = 1,\ldots,N;\ t = 1,\ldots,T_i \tag{1.24}$$

the number of observations for each individual, T_i, may differ across panel. The order of serial correlation is q_i. Cross section dependence is represented through ε_{it}. ε_{it}'s are assumed to be identical independent across time period, but cross sectionally dependent. The null hypothesis is $H_0 : \rho_i = 1$ for all *i* against the alternative $H_1 : \rho_i < 1$ for some *i*.

The IV estimation is performed on (1.24). The cross section dependence is handled through using the instrument generated by a nonlinear integrable *instrument generating function* (IGF) $F(y_{i,t-1})$. The *t*-statistic of the IV estimator $\hat{\rho}_i$ for the i^{th} individual is

$$Z_i = \frac{\hat{\rho}_i - 1}{s(\hat{\rho}_i)}$$

where $s(\hat{\rho}_i)$ is the standard error of $\hat{\rho}_i$. Chang (2002, p.270) Theorem 3.3 shows that under the null hypothesis as $T_i \to \infty$, the limiting distribution of Z_i is standard normal if a regularly integrable function is used as an IGF. The IV t-ratios Z_i's are asymptotically independent across the dependent panel individuals. The independence of Z_i's follows from the asymptotic orthogonality for the nonlinear transformations of integrated process by an integrable function, which is established in Chang et al. (2001). The panel unit root testing statistic is given as the average of Z_i

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i$$

Since T_i's are allowed to differ across panel, Chang proved that under the null as $T_{\min} \to \infty$ and $\frac{T_{\max}^{1/4} \log T_{\max}}{T_{\min}^{3/4}} \to 0$

$$S_N \Rightarrow N(0,1)$$

It deserves to note that due to the independent standard normal distribution of Z_i, only the T-asymptotics is used in deriving the limit theory of panel unit root test statistic and N may take any value, whereas usual panel unit root tests apply sequential limit theory[8]. Meanwhile, the problems of cross section dependence and unbalanced panel are also properly handled.

However, applying the nonlinear IV estimation is not without critics. Regarding the nonlinear IV panel unit root test statistic Im and Pesaran (2003) proved that the test needs a much more restrictive condition for its asymptotic property to hold, *i.e.*

$$\frac{N \ln T}{\sqrt{T}} \to 0 \text{, as } N, T \to \infty$$

So for the best performance of the test, N needs to be very small relative to T. Im and Pesaran further criticized the low degree of cross section dependence designed in

[8] Refer to the review of limit theories in section 5.2.2.2.

Chang's simulation experiments and showed severe size distortion of the test if strong dependence is structured.

Moreover, Breitung and Das (2005) point out that the integrability of the IGF can causes a tradeoff between the size and power of the test, so the power of the test may suffer from the nonlinear transformations if the choice of IGF is to produce good size. Furthermore, the optimal choice of transformation is unclear.

6.2.2 Tests with Strong Dependence

As discussed in section 6.2, in the case of strong dependence, the unobserved common factor (s) exist. Tests that can only handle the weak form of dependence experience size distortion in the presence of strong dependence. So some recent tests aim to include the common factor structure and solve the problem it causes.

6.2.2.1 Bai and Ng (2004) (BN) Test

Bai and Ng (2004) suggest test for unit root separately in the common factors and the error terms of a panel, since when a series is the sum of two components with different dynamic properties, e.g. a weak nonstationary and a strong stationary components, testing results will be affected. They define that a series with a factor structure is nonstationary if one or more of the common factors are nonstationary, or the error terms are nonstationary, or both. They consider the following factor model

$$y_{it} = c_i + \gamma_i t + \lambda_i' F_t + e_{it} \tag{1.26}$$

$$F_{mt} = \tau_m F_{m,t-1} + \zeta_{mt} \tag{1.27}$$

$$e_{it} = \rho_i e_{i,t-1} + \varepsilon_{it} \tag{1.28}$$

where $i = 1,\ldots, N$, $t = 1,\ldots, T$, $m = 1,\ldots, r$; F_t is an $r \times 1$ vector of common factors; λ_i is a vector of factor loadings; e_{it}, ζ_{mt} and ε_{it} are idiosyncratic error terms. Serial correlation is allowed in both (1.27) and (1.28). The strong dependence is represented by F_t, while e_{it} is assumed to be weakly correlated across individuals with non-diagonal variance covariance matrix. There are r_0 stationary factor(s) and r_1 common trend(s) (*i.e.* nonstationary factors), $r = r_0 + r_1$. y_{it} is nonstationary if F_t contain unit root(s), or e_{it} contain a unit root, or both.

Due to the problem that F_t and e_{it} are both unobserved, a robust procedure is developed to consistently estimate the factors F_t through the method of principle component. The consistency holds even without imposing stationarity on the errors.

In the intercept only case, the model is given as

$$y_{it} = c_i + \lambda_i' F_t + e_{it} \tag{1.29}$$

first difference (1.29) generating

$$\Delta y_{it} = c_i + \lambda_i' \Delta F_t + \Delta e_{it} \tag{1.30}$$

The estimated factors $\Delta\hat{F}_t$ and factor loadings $\hat{\lambda}_i$ are obtained by the method of principle component.[9] The estimated residuals are $\Delta\hat{e}_{it} = \Delta y_{it} - \hat{\lambda}_i' \Delta\hat{F}_t$.

Regarding testing for the dynamic properties of the factors F_t, Bai and Ng (2004) consider two cases:

(1) When $r = 1$, *i.e.* there is only one common factor. The ADF regression is conducted on the following model with an intercept

[9] Refer to Bai and Ng (2004) for detailed estimation procedure.

$$\Delta\hat{F}_t = c + \delta_0 \hat{F}_{t-1} + \sum_{j=1}^{q} \delta_j \Delta\hat{F}_{t-j} + error \tag{1.31}$$

where q is the lag order of the series. Let $ADF_{\hat{F}}^{c}$ be the t-statistic of for testing $\delta_0 = 0$. As $T \rightarrow \infty$, the $ADF_{\hat{F}}^{c}$ has the same limiting distribution as the DF test for the constant only case.

(2) When $r > 1$, *i.e.* there are more than one factors, two statistics are considered. The two statistics MQ_c^c and MQ_f^c computed from demeaned $\hat{F}_t$ are modified versions of Stock and Watson (1988)'s Q_c^c and Q_f^c statistics. The tests start with $m = r$. If $H_0 : m = r$ is rejected, set $m = m - 1$ and redo the test, until it fails to reject $m = r_1$ and stop. Thus the number of stochastic trends in F_t is $\hat{r}_1$. Bai and Ng provided the critical values of MQ_c^c and MQ_F^c by simulations.

Regarding testing for stationarity of e_{it}, firstly the ADF regression is conducted on the following model without deterministic terms

$$\Delta\hat{e}_{it} = d_{i0}\hat{e}_{i,t-1} + \sum_{j=1}^{q} d_{ij} \Delta\hat{e}_{i,t-j} + error \tag{1.32}$$

where q is the lag order in the series. Let $ADF_{\hat{e}}^{c}(i)$ be the t-statistic for testing $d_{i0} = 0$. $T \rightarrow \infty$, the $ADF_{\hat{e}}^{c}(i)$ coincides with the limiting distribution of the DF test with no constant. The pooled test for the panel adopts the Fisher type combining p-value test. The null ohypothesis is $H_0 : \rho_i = 1$ for all i (in (1.28)), against the alternative $H_1 : \rho_i < 1$ for some i. Let $p_{\hat{e}}^{c}(i)$ be the p-values corresponding to $ADF_{\hat{e}}^{c}(i)$. Under the null hypothesis as $T \rightarrow \infty$ followed by $N \rightarrow \infty$, the Fisher type test statistic converges to standard normal distribution, *i.e.*

$$p_{\hat{e}}^{c} = \frac{-2\sum_{i=1}^{N}\log p_{\hat{e}}^{c}(i) - 2N}{\sqrt{4N}} \Rightarrow N(0,1)$$

6.2.2.2 Moon and Perron (2004) (MP) and Phillips and Sul (2003) (PS) Tests

Moon and Perron (2004) and Phillips and Sul (2003) propose similar tests that are based on de-factored data. Instead of separating the common factors and error terms as in the Bai and Ng (2004) test, they use orthogonalization procedure to de-factor the data and obtain the panel that is cross sectinally independent. Usual independent panel unit root test can then be applied on the de-factored series after orthogonalization transformation. Phillips and Sul considered only one factor structure, whereas Moon and Perron (2004) construct it more generally by including K factors, so here the Moon and Perron (2004) test is introduced.

The model with fixed effects is given as

$$y_{it} = \alpha_i + y_{it}^{0} \tag{1.33}$$

$$y_{it}^{0} = \rho_i y_{i,t-1}^{0} + u_{it} \tag{1.34}$$

$$u_{it} = \lambda_i' F_t + e_{it} \tag{1.35}$$

where F_t is a $K \times 1$ vector of unobservable random factors representing cross section dependence, and λ_i is factor loadings determining the extent of correlation. e_{it} are idiosyncratic shocks. The number of factors K is unknown. The null hypothesis is $H_0 : \rho_i = 1$ for all *i*, against the alternative $H_1 : \rho_i < 1$ for some *i*.

Moon and Perron (2004) also intend to study the test performance under local alterative hypothesis, so they combine the hypotheses by the near unit root model

$$\rho_i = 1 - \frac{\theta_i}{\sqrt{NT}}$$

where random variable θ_i is nonnegative and is *iid* with mean μ_θ. Therefore the hypotheses are equivalent to $H_0 : \mu_\theta = 0$ or $\theta_i = 0$ for all i against the local alternative $H_1 : \mu_\theta > 0$ for some *i*.

Write (1.33), (1.34) and (1.35) in matrix form

$$\mathbf{Y} = \mathbf{l_T}\boldsymbol{\alpha} + \mathbf{Y^0} \tag{1.36}$$

$$\mathbf{Y^0} = \mathbf{Y^0_{-1}} + \mathbf{F}\,\boldsymbol{\beta}' + \mathbf{e} \tag{1.37}$$

where $\mathbf{Y}$, $\mathbf{Y^0}$, $\mathbf{Y_{-1}}$, $\mathbf{F}$ and $\mathbf{e}$ denote the corresponding matrix for y_{it}, y^0_{it}, $y_{i,t-1}$, F_t and e_{it}, respectively; $\mathbf{l_T}$ is a $T \times 1$ vector of ones; $\boldsymbol{\beta}$ is $N \times K$ matrix of factor loadings. Denote the long-run variance of e_{it} as $\omega^2_{e,i}$. As $N\to\infty$, let $\omega^2_e = \lim_n \frac{1}{N}\sum_{i=1}^{N}\omega^2_{e,i}$ and $\phi^4_e = \lim_n \frac{1}{N}\sum_{i=1}^{N}\omega^4_{e,i}$.

The pooled estimator of ρ is

$$\hat{\rho}_{pool} = \frac{tr\left(\mathbf{Y}'_{-1}\mathbf{Y}\right)}{tr\left(\mathbf{Y}'_{-1}\mathbf{Y}_{-1}\right)}$$ [10]

Moon and Perron (2004) show that the limit distribution of $T\left(\hat{\rho}_{pool} - 1\right)$ depends on the common factor(s), and therefore they suggest multiply (1.37) by the projection matrix Q_β [11] to eliminate the factor. Thus under the null hypothesis, (1.37) becomes

$$\mathbf{Y}^0 Q_\beta = \mathbf{Y}^0_{-1} Q_\beta + \mathbf{e} Q_\beta \tag{1.38}$$

The modified pooled OLS estimator of ρ is given as

$$\hat{\rho}^+_{pool} = \frac{tr\left(\mathbf{Y}_{-1} Q_\beta \mathbf{Y}'\right) - NT\lambda^n_e}{tr\left(\mathbf{Y}_{-1} Q_\beta \mathbf{Y}'_{-1}\right)}$$

[10] *tr* denotes *trace*.

[11] $Q_\beta = I - P_\beta$, where $P_\beta = \beta(\beta'\beta)^{-1}\beta'$

where the modification $\lambda_e^n = \frac{1}{N}\sum_{i=1}^{N}\lambda_{e.i}$ is to deal with serial correlation in $\mathbf{e}Q_\beta$ and $\lambda_{e,i}$ is the one-sided long-run variance of e_{it}. It is proved that as N, $T\to\infty$ with $N/T\to 0$,

$$\sqrt{N}T\left(\hat{\rho}_{pool}^{+} - 1\right) \Rightarrow N\left(-\mu_\theta, \frac{2\phi_e^4}{\omega_e^4}\right)$$

The distribution shows that under both the null and the local alternative hypotheses, $\hat{\rho}_{pool}^{+}$ is $\sqrt{N}T$ -consistent and asymptotically normal, and $\sqrt{N}T\left(\hat{\rho}_{pool}^{+} - 1\right)$ is unbiased under the null.

Two feasible panel unit root test statistics are then proposed. The number of factors is estimated by the criterion function in Bai and Ng (2002); the factor loadings $\boldsymbol{\beta}$ are estimated using principle component method; consistent kernel estimators are applied to the long-run variances $\lambda_{e,i}$ and $\omega_{e,i}^2$. The two statistics are,

$$t_a^* = \frac{\sqrt{N}T\left(\hat{\rho}_{pool}^{*} - 1\right)}{\sqrt{2\hat{\phi}_e^4 / \hat{\omega}_e^4}}$$

$$t_b^* = \sqrt{N}T\left(\hat{\rho}_{pool}^{*} - 1\right)\sqrt{\frac{1}{NT^2}tr\left(\mathbf{Y}_{-1}Q_{\hat{\beta}_k}\mathbf{Y}'_{-1}\right)}\left(\frac{\hat{\omega}_e}{\hat{\phi}_e^2}\right)$$

where

$$\hat{\rho}_{pool}^{*} = \frac{tr\left(\mathbf{Y}_{-1}Q_{\hat{\beta}_K}\mathbf{Y}'\right) - NT\hat{\lambda}_e^n}{tr\left(\mathbf{Y}_{-1}Q_{\hat{\beta}_K}\mathbf{Y}'_{-1}\right)}$$

Under the null hypothesis and various assumptions, as N, $T\to\infty$ with $N/T\to 0$,

$$t_a^*, t_b^* \Rightarrow N\left(-\mu_\theta\sqrt{\frac{\omega_e^4}{2\phi_e^4}}, 1\right)$$

In the later Monte Carlo simulations, it is shown that when there is no incidental trend, the tests have substantial power. However, when incidental trends are added to the model, the tests have no power.

Phillip and Sul (2003) propose a panel unit root test with cross section dependence in the same spirit, whereas they only consider the case of a single factor, or the time effect. Instead of using principal component method, they adopt a moment based procedure to estimate the factor loading(s) and variance of the idiosyncratic shocks. Similarly, the original panel is multiplied by the estimated projection matrix and thus cross section dependence is removed. Since the de-factored panel is cross sectionlly independent, they use the Fisher type combining *p*-values test or the IPS test to compute panel unit root test statistics.

6.2.2.3 Pesaran (2007) Test

On considering the single factor model, instead of de-factoring the data through orthogonalization method, Pesaran (2007) proposes a simple procedure that augments the ADF regressions. The factor term in the model is replaced by the averages of individuals' lagged levels and first differences as a proxy. The individual *cross section augmented ADF* (CADF) statistics are shown to be independent of the factor loadings, so a modified IPS panel unit root test is then developed using the CADF statistics.

The heterogeneous model with single factor is given as

$$y_{it} = \alpha_i + \rho_i y_{i,t-1} + u_{it} \tag{1.39}$$

$$u_{it} = \lambda_i F_t + \varepsilon_{it} \tag{1.40}$$

where F_t is the unobserved common factor; ε_{it} are independently distributed across both dimensions of the panel with mean zero and variance σ_i. (1.39) and (1.40) can be written as

$$\Delta y_{it} = \alpha_i + \delta_i y_{i,t-1} + \lambda_i F_t + \varepsilon_{it} \tag{1.41}$$

where $\delta_i = (1-\rho_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$. The unit root null hypothesis based on (1.41) is $H_0 : \delta_i = 0$ for all *i* against the alternative $H_1 : \delta_i < 0$ for $i = 1,\ldots,N_1$; it is assumed that the fraction N_1 / N is non-zero and tends to the fixed value θ such that $0 < \theta < 1$ as $N \to \infty$.

To handle the common factor, Pesaran uses the cross sectional mean of y_{it}, $\bar{y}_{it} = N^{-1}\sum_{i=1}^{N} y_{it}$ and its lagged values as proxy to substitute F_t, so (1.41) becomes

$$\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_i \Delta\bar{y}_t + e_{it}\text{[12]} \tag{1.42}$$

The *t*-statistic of the augmented DF regression (CADF) (1.42) is

$$t_i(N,T) = \frac{\mathbf{\Delta y}_i' \overline{\mathbf{M}}_w \mathbf{y}_{i,-1}}{\hat{\sigma}_i \left(\mathbf{y}_{i,-1}' \overline{\mathbf{M}}_w \mathbf{y}_{i,-1}\right)^{1/2}}$$

where $\mathbf{y}_i$, $\mathbf{y}_{i,-1}$ and $\mathbf{\Delta y}_i$ are vector forms of y_{it}, $y_{i,t-1}$ and Δy_{it}; also let $\bar{\mathbf{y}}_{-1}$, $\mathbf{\Delta y}$ denote the matrix form of $y_{i,t-1}$ and Δy_{it},

$$\overline{\mathbf{M}}_w = \mathbf{I}_T - \overline{\mathbf{W}}\left(\overline{\mathbf{W}}'\overline{\mathbf{W}}\right)^{-1}\overline{\mathbf{W}}' , \; \overline{\mathbf{W}} = (\boldsymbol{\kappa}, \mathbf{\Delta}\bar{\mathbf{y}}, \bar{\mathbf{y}}_{-1}), \; \boldsymbol{\kappa} = (1,\ldots,1)',$$

$$\hat{\sigma}_i^2 = \frac{\mathbf{\Delta y}_i' \mathbf{M}_{i,w} \mathbf{\Delta y}_i}{T-4}$$

$$\mathbf{M}_{i,w} = \mathbf{I}_T - \mathbf{G}_i(\mathbf{G}_i'\mathbf{G}_i)^{-1}\mathbf{G}_i', \; \mathbf{G}_i = \left(\overline{\mathbf{W}}, \mathbf{y}_{i,-1}\right)$$

[12] In the serially uncorrelated case $\bar{y}_t$ and $\bar{y}_{t-1}$ (or equivalently $\bar{y}_{t-1}$ and $\Delta\bar{y}_t$) are asymptotically sufficient for remove the effects of unobserved common factor.

It is proved that as $N,T\to\infty$ $t_i(N,T)$ has both sequential and joint limit distributions, CADF$_i$. The joint limit requires $N/T\to k$ (0<k<∞). CADF$_i$ are free from nuisance parameters, but are correlated due to the presence of common factor. The distribution of CADF$_i$ is more skewed to the left than DF distribution, presenting substantially negative mean and less than unity variance. However, the standardized version of CADF$_i$ is remarkably close to standard normal distribution, although statistically rejected.

Since the CADF$_i$ statistics are asymptotically independent of nuisance parameters, a cross sectionally augmented version of the IPS *t*-bar test is adopted for panel unit root test

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

Considering the mean deviations,

$$D(N,T)=t\_bar=N^{-1}\sum_{i=1}^{N}\left[t_i(N,T)-CADF_i\right]$$

where CADF$_i$ is the stochastic limit of $t_i(N,T)$ as $N,T\to\infty$ such that $N/T\to k$. Due to the technical difficulty of establishing the moment conditions of $D(N,T)$, Pesaran bases the t-bar test on a truncated version of the CADF$_i$ statistics, $t_i^*(N,T)$, as the standardized CADF$_i$ statistics $t_i(N,T)$ are very close to standard normal distribution. $t_i^*(N,T)$ is decided as

$$t_i^*(N,T)=t_i(N,T),\ \text{if}\quad -K_1<t_i(N,T)<K_2$$

$$t_i^*(N,T)=-K_1,\ \text{if}\quad t_i(N,T)\le -K_1$$

$$t_i^*(N,T)=K_2,\ \text{if}\quad t_i(N,T)\ge K_2$$

where K_1 and K_2 are positive constants sufficiently large to let $\Pr\left[-K_1<t_i(N,T)<K_2\right]$, so that it exceeds 0.9999. The values of K_1 and K_2 are provided for different cases, *i.e.*

models with or without constant and linear trend. The truncated panel unit root test statistic is then given as,

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

The distributions of both CIPS and CIPS* are non-standard due to the dependence of $CADF_i$, so critical values are computed through simulation experiments. In addition, the finite sample distributions of CIPS and CIPS* are found very similar and indistinguishable for T>20. The Monte Carlo simulations indicate that the tests have satisfactory size and power even for relatively small N and T.

6.2.2.4. Chang and Song (2005, 2009) (CS) Nonlinear IV Test

6.2.2.4.1 The Problem of Cross Unit Cointegration

Banerjee et al. (2004, 2005) raise the caution of ignoring *cross unit cointegration* in constructing both panel unit root tests and panel cointegration tests. Cross unit cointegration refers to the case that two or more panel individuals share at least one common stochastic trend and therefore the individual series in the panel cointegrate. It is also called the *long-run* dependence (a special case of the strong form of dependence) and is usually represented by common factors in panel models. Economic theory and certain empirical data (e.g. exchange rates) tend to strongly suggest the presence of cross unit cointegration. Therefore, if panel individuals are nonstationary only due to nonstationary common factor(s), they share common stochastic trend(s) and are cointegrated. This situation invalids some of the second generation panel unit root tests. For example, the Moon and Perron (2004) and Phillips and Sul (2003) tests propose to remove the common stochastic trends and conduct testing on the de-factored data. However, if nonstationarity is caused by the common factors, the tests are unable to detect unit root. Pesaran (2007) may also lead to misleading results if the panel presents cross unit cointegration (c.f. Breitung and Pesaran, 2008). Bai and Ng (2004) can properly deal with the situation since it separately tests for unit root in the common factor(s) and the error terms.

Using simulations Banerjee et al. found sever size distortion of some first generation tests when the assumption is relaxed. The unit root hypothesis is even rejected when the true process is nonstationary. Their empirical section of purchasing power parity (PPP) study shows opposite results when unit root is tested with and without cross unit cointegration restriction.

6.2.2.4.2 Chang and Song (2005, 2009) (CS) Test

As an alternative approach that enables the test to cope with both short-run and long-run dependence, Chang and Song (2005, 2009) propose an improved version of the Chang (2002) nonlinear IV panel unit root test[13]. The Chang (2002) test which is based on only one *instrument generating function* (IGF) for all panel individuals is invalid in the presence of strong dependence. Therefore, Chang and Song (2009) suggest use a set of orthogonal functions as IGFs to handle all forms of dependence. The *t*-statistics of the IV estimators for panel individuals are shown to be asymptotically independent of one another and normally distributed.

The basic model is,

$$y_{it} = \rho_i y_{i,t-1} + u_{it} \tag{1.43}$$

where u_{it} are specified later. Three sets of hypotheses are proposed,

(A) $H_0 : \rho_i = 1$ for all i against $H_1 : \rho_i < 1$ for all i ;

(B) $H_0 : \rho_i = 1$ for all i against $H_1 : \rho_i < 1$ for some i ;

[13] The Chang (2002) test is reviewed in section 6.2.1.3.

(C) $H_0 : \rho_i = 1$ for some i against $H_1 : \rho_i < 1$ for all i .

Hypotheses (A) is as constructed as in the homogeneous tests and (B) is as in the heterogeneous test. Hypotheses (C) is new. Rejection of the null indicates that all panel individuals are stationary.

It is assumed that under the null hypothesis there are (*N-M*) cointegrating relationships in the unit root process y_{it}, represented by cointegrating vectors c_j, j=1,…, *N-M*. The short-run dynamics of y_{it} can be presented by error correction representation, let $\mathbf{y}_t = (y_{1t},..., y_{Nt})'$

$$\Delta y_{it} = \sum_{j=1}^{N}\sum_{k=1}^{P_i} a_{ij}\Delta y_{j,t-k} + \sum_{j=1}^{N-M} b_{ij}c_j'\mathbf{y}_{t-1} + \varepsilon_{it} \tag{1.44}$$

where ε_{it} are white noise. Under the null hypothesis of unit root $u_{it} = \Delta y_{it}$, combining (1.43) and (1.44) obtains

$$y_{it} = \rho_i y_{i,t-1} + \sum_{k=1}^{P_i} \alpha_{i,k}\Delta y_{i,t-k} + \sum_{k=1}^{Q_i} \beta_{i,k}'\omega_{i,t-k} + \varepsilon_{it} \tag{1.45}$$

where ω_{it} are explained as covariates added to the ADF regression for the i^{th} individual. (1.43) and (1.44) are rewritten as (1.45) with several lagged differences of other cross section individuals and linear combinations of the lagged levels of all cross sections as covariate. Chang and Song (2005, 2009) notes that both Hasen (1995) and Chang et al. (2001) show using covariates offers a great potential in power gain for a unit root test. The following testing statistics will be based on (1.45).

The choice of *instrument generating functions* (IGF) is a set of orthogonal *Hermite functions* $G_k(x)$ of odd orders $k = 2i - 1$, $i = 1,..., N$. The Hermite function $G_k(x)$ of order k ($k = 0,1,2,...$) is defined as

$$G_k(x) = \left(2^k k!\sqrt{\pi}\right)^{-1/2} H_k(x)e^{-x^2/2}$$

where H_k is the Hermite polynomial of order *k* given by

$$H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$$

Let F_i be the IGF for the i^{th} individual. Then the IGF's (F_i 's) are defined as

$$F_i = G_{2i-1}$$

for $i = 1,..., N$. So the instrument is $F_i(y_{i,t-1}) = G_{2i-1}(y_{i,t-1})$. The IV *t*-statistic computed from (1.45) for testing unit root in (1.43) or (1.45) is,

$$\tau_i = \frac{\hat{\rho}_i - 1}{s(\hat{\rho}_i)}$$

where $\hat{\rho}_i$ is the nonlinear IV estimator of ρ_i; $s(\hat{\rho}_i)$ is the standard error of IV estimator $\hat{\rho}_i$. Chang and Song (2009, p.917) Lemma 1 shows that under the null hypothesis as $T_i \rightarrow \infty$, τ_i 's have standard normal limiting distribution and are asymptotically independent of one another across the cross section dimension. They note that the asymptotic independence of the individual IV t-ratio τ_i 's follows from the orthogonality of the IGF's even in the presence of cross unit cointegration. The normality and independence properties of τ_i 's are discussed in detail in Section 2.3 of Chang and Song (2009).

The panel unit root test statistics for hypotheses (A) – (C) are respectively,

$$S = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \tau_i ; \qquad S_{\min} = \min_{1 \le i \le N} \tau_i ; \qquad S_{\max} = \max_{1 \le i \le N} \tau_i .$$

Chang and Song (2009, p.917) Theorem 1 implies that the test using statistics S and $S_{\min}$ with critical values $c(\lambda)$ and $c_{\min}(\lambda)$, respectively, have the exact size λ asymptotically under the null hypotheses in Hypotheses (A) and (B). However, the rejection probabilities of the test relying on $S_{\max}$ with critical values $c(\lambda)$ may not be exactly λ even asymptotically under the null hypothesis in Hypothesis (C).

The later section of Monte Carlo simulations confirmed the fact that tests such as Moon and Perron (2004) have severe size distortion and poor power in the presence of cross unit cointegration. The Chang and Song test can cope with the situation with reasonable size and power. However, the performance is generally not sufficiently satisfactory for application. The empirical sizes tend to distort downward as N becomes large, and so the nominal critical values from standard normal distribution are not appropriate to apply. In Chapter 3 the finite sample bias is analyzed and empirical critical values for various sample sizes are computed through numerical method.

6.3 Summary

Section 6 has surveyed the popular contemporary panel unit root tests. The tests are divided into two generations by the consideration of cross section dependence/correlation. The first generation tests that ignore cross section dependence started to develop from homogeneous tests to heterogeneous tests, due to the fact that panel data generally introduce substantial amount of heterogeneity. Since some recent studies found that the assumption of cross section independence is inappropriate in many fields, a second generation panel unit root tests are developed to solve the dependence. Earlier second generation tests focus on the weak form of dependence generated by the covariance matrix of error terms. Nevertheless, in the presence of the strong form of dependence (i.e. when common factor(s) exist(s) across panel individuals), tests that can only handle weak dependence experience size distortions. Therefore, the later second generation tests are constructed to cope with the strong form of dependence. To sum up, Table 1.1 in the next page provides an overlook of the tests surveyed in this chapter.

Table 1.1 Panel unit root tests surveyed in section 6

	Test(s)	Contribution	Weakness
First Generation Tests	Levin and Lin (1992) (LL) Levin, Lin and Chu (2002) (LLC)	• LL test -- the first formal demonstration of asymptotic normality of panel unit root testing statistics; • LLC test – extension to the LL test to consider more general serial correlation and heteroscedasticity in the error terms.	• Homogeneous tests; • Ignore cross section dependence
	Im, Pesaran and Shin (1997, 2003)	• Early heterogeneous test	• Ignore cross section dependence
	Maddala and Wu (1999) Choi (2001)	• Non-parametric and exact test; • Heterogeneous test	• The *p*-values required in the tests have to be derived by Monte Carlo simulations; • Ignore cross section dependence.
	Hadri (2000)	• Stationarity test ; • Heterogeneous test	• Ignore cross section dependence
Second Generation Tests	Harvey and Bates (2003)	• GLS based test to solve weak dependence	• Homogeneous test; • Unable to handle strong dependence
	Breitung and Das (2005) Jönsson (2005)	• OLS based test to solve weak dependence	• Homogeneous test; • Unable to handle strong dependence

Table 1.1 Panel unit root tests surveyed in section 6 (Cont'd)

	Test(s)	Contribution	Weakness
Second Generation Tests	Chang (2002)	• Solve weak dependence; • Non-linear IV technique applied, free from modeling dependence.	• Unable to handle strong dependence; • More condition required for the test asymptotic properties to hold.
	Bai and Ng (2004)	• Solve strong dependence; • Separately test for unit root in the common factor(s) and the error terms. • Able to cope with cross unit cointegration	• Need to estimate common factor(s)
	Moon and Perron (2004) Phillips and Sul (2003)	• Solve strong dependence	• Through de-factoring the data to eliminate common factor(s); if unit root exists in the common factor(s), the test fails to detect the unit root. • Only one common factor is considered in Phillips and Sul (2003).
	Pesaran (2007)	• Solve strong dependence; • Able to cope with cross unit cointegration.	• Only one common factor is considered.
	Chang and Song (2005, 2009)	• Solve strong dependence; • Developed non-linear IV technique applied, free from modeling dependence; • Able to cope with cross unit cointegration.	• Poor finite sample performance; • Conjectured by the thesis that more condition is required for the asymptotic properties to hold, similar to the Chang (2002) test.

Chapter 2 Finite Sample Distributions of Nonlinear IV Panel Unit Root Tests in the Presence of Cross Section Dependence

1. Introduction

As reviewed in Chapter 1 the first generation panel unit root tests ignore cross section dependence. The dependence frequently exists in empirical data and can lead to size distortions and low power of the tests. The second generation tests propose a number of approaches to either model or eliminate the effect of dependence. Among these, the Chang and Song (2005, 2009) (CS hereafter) test is one of the most general tests that are able to handle all forms of dependence. In particular, it can cope with the problem of *cross unit cointegration* within the panel, a special case of strong dependence when panel units share common stochastic trend(s), which leads to cointegration of the panel individuals and forms the long run dependence. Based on the early version Chang (2002) (CH hereafter) test, CS test uses the developed nonlinear IV technique and obtains the cross sectionally independent individual *t*-statistics of the IV estimators which also have asymptotic standard normal distribution. The panel unit root statistic, the standardized sum of the individual *t*-statistics, is thus asymptotically normally distributed.[14] This approach overcomes the weakness of some other methods that solve dependence by

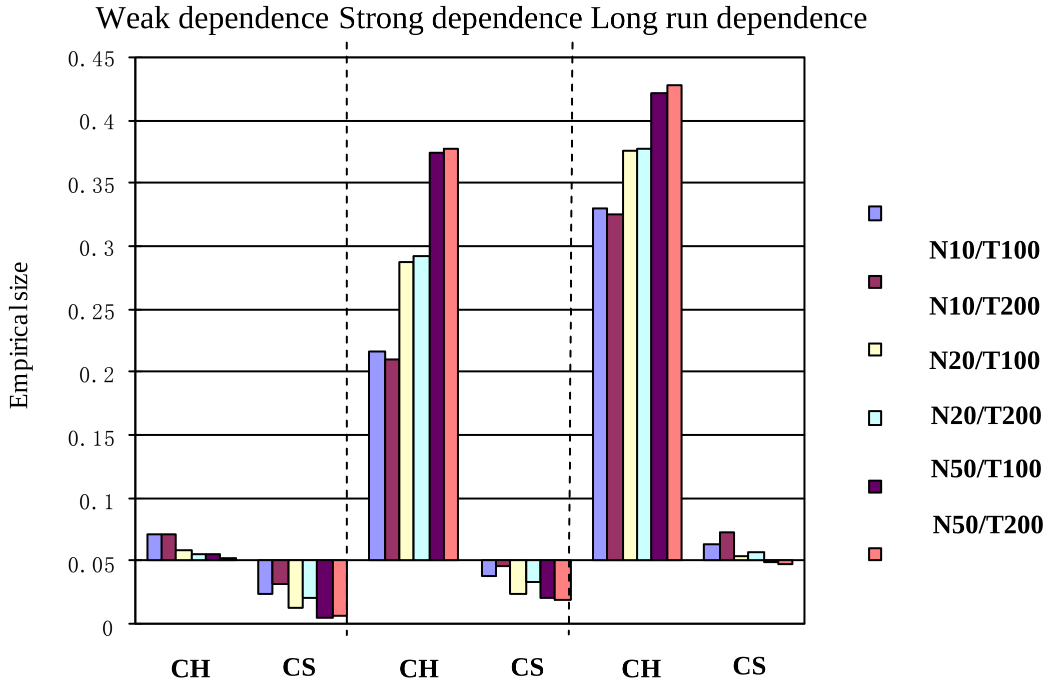
[14] The CS test proposes three hypotheses and thus there are three testing statistics. However, for simplicity and comparison with CH test, only the average statistic, *i.e.* the standardized sum of individual IV *t*-statistics, is considered.

de-factoring the model and thus may eliminate unit root if unit root exists in the common factors (see the review in Chapter 1). The CS test also gains improvement over the CH test which can only deal with the weak form of dependence.

As panel unit root test statistics usually exhibit standard normal distribution asymptotically, issues of finite sample distribution or finite sample critical values are not given much attention in the literature. Empirical research generally directly applies the nominal critical values of panel unit root tests. However, it is well known that finite sample performance can substantially differ from the corresponding asymptotic properties, in particular, when the rate of convergence is slow. Moreover, the finite sample distribution of panel unit root testing statistic can vary across the two dimensions, the cross section and time dimensions. An effective approach to observe the variation in small sample distributions is the response surface method, using points simulated by a number of Monte Carlo experiments under a variety of sample sizes. The response surface method explores the relationships between several explanatory variables and one or more response variables. The early development and applications of response surface method are by MacKinnon (1991, 1994) where the percentiles of the statistic distributions in several tests under a range of finite sample sizes are examined as well as the finite sample (or numerical) distribution functions of the statistics. The method later gained favour in investigating the finite sample performance of statistical tests and more applications include Cheung and Lai (1995a,b), Sephton (1995), Carrion et al. (1999), MacKinnon et al. (1999), Ericsson and MacKinnon (2002), Presno and López (2003), etc.

In the field of panel unit root test, Jönsson (2005a) tabulated the critical values for a panel corrected standard error (PCSE) based Levin and Lin (1992) (LL) homogeneous test by the response surface method. The PCSE correction is introduced to deal with the cross section dependence in the weak form. Jönsson (2005b) continues to augment the test with serial correlation and estimates response surfaces for critical values. However, the LL test assumes homogeneity which hardly holds in practice and the weak form of cross section dependence is often not sufficient to represent the magnitude of dependence in empirical data.

Figure 2.1 5% finite sample test sizes of the CH and CS tests



Source: Chang and Song (2005)
Note: *N* and *T* denote cross section and time dimension, respectively; CH and CS denote the Chang (2002) and Chang and Song (2005, 2009) test, respectively.

Due to the robustness to heterogeneity and all forms of cross section dependence, the CS and CH tests are favourable for applications. However, the finite sample performance of the two tests does not present satisfactory results. Figure 2.1 plots the 5% finite sample sizes of the tests based on the Monte Carlo simulation results in Chang and Song (2005). The CS test has reasonable sizes in the presence of long run dependence, but in other cases the sizes deteriorate, in particular, as sample size increases. This seems to contradict the theory that as sample size grows large, the test performance is expected to behave close to the asymptotics. In addition, since the CS test is a developed version based on the CH test using the same non-linear IV technique with only different IGFs, it is conjectured that the same condition to sustain the asymptotic properties of CH test, as pointed out by Im and Pesaran (2003), is also needed for CS test[15]. The conjecture will be dealt with by numerical method in this chapter. Figure 2.1 also confirms the argument that CH test is not able to cope with strong dependence, whereas when only weak dependence exists CH significantly outperforms CS. Therefore, CH is also analyzed and is recommended for empirical work where dependence among panel individuals is weak.

Given the poor finite sample performance this chapter applies response surface method to analyze the finite sample bias of the two tests and provide finite sample critical values. A series of Monte Carlo experiments are conducted on the CH and CS tests under a variety of sample sizes. The relationship between finite sample bias (as the response variable) and several functional forms of sample size (as the explanatory variables) is investigated through a new design of regression specification. Instead of calculating the traditional point critical values, the chapter provides the indecisive range around point critical values caused by the uncertainty that companies Monte Carlo simulations. Formulas to compute the upper and lower limits of a critical value inverval are provided using the David-Johnson estimate of percentile standard deviation. The numerical distribution functions of the testing statistics are presented to calculate the *p*-value of any given percentile. Finally, graphs of the numerical distributions under various sample sizes are given to highlight finite sample properties of the statistics.

[15] Refer to the literature review in Chapter 1, section 6.2.1.3.

The rest of this chapter is organized as follows. Section 2 specifies the different forms of cross section dependence. Section 3 briefly reviews the CH and CS tests. Section 4 presents the design of Monte Carlo experiments. Response surface estimation is provided in section 5 and section 6 discusses the results. Section 7 concludes.

2. Specification of Cross Section Dependence

Without considering residual serial correlation a general form of cross section dependence can be given as

$$y_{it} = \rho_i y_{i,t-1} + u_{it} \qquad i = 1,\ldots,N;\ t = 1,\ldots,T_i \tag{2.1}$$

$$u_{it} = \upsilon_i' F_t + \eta_{it} \tag{2.2}$$

where $F_t = (\mathrm{f}_{1t}, \mathrm{f}_{2t}, \ldots, \mathrm{f}_{mt})'$ is an $m \times 1$ vector of serially uncorrelated unobserved common factor(s); $\upsilon_i = (\upsilon_{i1}, \upsilon_{i2}, \ldots, \upsilon_{im})'$ is an $m \times 1$ vector of factor loadings; η_{it} is serially uncorrelated process with mean zero and positive definite covariance matrix V_η. For generality set the covariance matrix of F_t as identity matrix I_m. It is assumed that F_t and η_{it} are independently distributed.

Initially two cases of cross section dependence can be distinguished. (ⅰ) Weak dependence. This assumption rules out the presence of unobserved common factor(s) F_t ($\upsilon_i = 0$). Dependence only arises from spatial correlation among cross section individuals which is represented by the covariance matrix of η_{it}, V_η. (ⅱ) Strong dependence. In this case unobserved common factor(s) exist and dependence is generated from two sources, F_t and η_{it}. The covariance matrix of error u_{it} is thus given by $V = \upsilon_i' \upsilon_i + V_\eta$.

Given the model (2.1), under the null hypothesis of a unit root, *i.e.* $\rho_i = 1$ for all *i* = 1,…, *N*, in the case of strong dependence y_{it} contains nonstationary *cumulated* common factor(s) F_t and nonstationary *cumulated* errors η_{it}. If the process built upon η_{it} is stationary, *e.g.* substitute η_{it} by $\Delta\eta_{it}$ and write u_{it} in (2.2) as

$$u_{it} = \upsilon_i' F_t + \Delta\eta_{it} \tag{2.3}$$

then the nonstationarity of y_{it} is only driven by the nonstationary *cumulated* common factor(s) F_t and therefore F_t serve as common stochastic trend(s). This leads to the cointegration relationship between any pair of y_{it} and y_{jt}, with $(N-m)$ linearly independent cointegrating relations among *N* cross section individuals. The case is the *cross unit cointegration* and it drives long run dependence.

3. Brief Review of the CH and CS tests

Details of the two tests have been reviewed in Chapter 1, so they are only briefly introduced in this section. Using nonlinear IV technique the CH and CS tests are developed to deal with cross section dependence. The basic model considering autocorrelation is

$$y_{it} = \rho_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k} \Delta y_{i,t-k} + \varepsilon_{it}, \qquad i = 1,\dots,N;\ t = 1,\dots,T_i \tag{2.4}$$

where ε_{it} are white noise. The instrument is generated by a nonlinear integrable *instrument generating function* (IGF) $F(y_{i,t-1})$ in CH test or a set of orthogonal IGFs in CS test. For each *i* = 1,…, *N* under the unit root hypothesis $H_0 : \rho_i = 1$, the *t*-statistic of the nonlinear IV estimator $\hat{\rho}_i$ is constructed as

$$\tau_i = \frac{\hat{\rho}_i - 1}{se(\hat{\rho}_i)}$$

where $se(\hat{\rho}_i)$ is the standard error of $\hat{\rho}_i$. It is shown that τ_i's have standard normal limiting distribution and are cross sectionally independent of one another. Therefore the test is free from the need of modeling dependence. The panel unit root test statistic is the average of the individual *t*-statistics τ_i's

$$S = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \tau_i$$

As $T_i \to \infty$, the statistic S has standard normal limiting distribution.

Given the poor finite sample performance of the two tests illustrated in the introduction in Figure 2.1 and the critique by Im and Pesaran (2003), the finite sample bias is analyzed in the following sections through numerical methods.

4 The Simulation Experiments

4.1. The DGPs

For simulation experiments the DGPs in Chang and Song (2005, 2009) are used to formulate non-stationary panels with cross section weak dependence, strong dependence and cross unit cointegration, respectively. They are referred to DGP1, DGP2 and DGP3 hereafter. The basic model considered is

$$y_{it} = \rho_i y_{i,t-1} + u_{it} \quad i = 1,\ldots, N;\ t = 1,\ldots,T$$

Balanced panels are used in the simulation experiments, so each panel cross section unit has the same time length *T*. Under the null hypothesis of a unit root $\Delta y_{it} = u_{it}$. The innovations u_{it} are given as following to generate different forms of dependence

DGP1: $u_{it} = \beta_i u_{i,t-1} + \eta_{it}$ (2.5)

DGP2: $u_{it} = \beta_i u_{i,t-1} + \upsilon_i \xi_t + \eta_{it}$ (2.6)

DGP3: $u_{it} = \beta_i u_{i,t-1} + \upsilon_i \xi_t + \Delta\eta_{it}$ (2.7)

where β_i is the AR coefficient; ξ_t is common factor and υ_i are factor loadings ; η_{it} are innovations with symmetric and nonsingular covariance matrix V which is specified soon after.

Note that in DGP1 only weak dependence exists, given by the covariance matrix of η_{it}. Besides η_{it}, DGP2 has stronger level of dependence generated by a common stochastic trend built upon ξ_t. Under the unit root hypothesis y_{it} contains both nonstationary *cumulated* ξ_t and nonstationary *cumulated* η_{it}, *i.e.* $\sum \xi_t$ and $\sum \eta_{it}$, so the cross section individuals in the panel are not cointgrated. However, in DGP3 cross unit cointegration is present. Since $\sum \Delta\eta_{it}$ is stationary, the nonstationarity of y_{it} arises only from the nonstationary *cumulated* common stochastic trend, $\sum \xi_t$. Hence, there is cointegrating relationship between any pair of the N panel individuals with $(N-1)$ linearly independent cointegrating relations.

Recall the performance of CH and CS tests under different forms of dependence as shown in Figure 3.1. The CH test outperforms CS test on weak dependence but is unable to cope with strong dependence, not mention the long run dependence; while the CS test has reasonable sizes under the strong form and long run dependence. Therefore, the CH test is applied on DGP1 and CS test is applied on DGP2 and 3. Since in the Monte Carlo simulations the DGPs are known in advance and the appropriate tests are directly applied on the corresponding simulated data. In practice when the characteristic of empirical data is unknown, some examination measures can be applied on the data (e.g.

compute the covariance matrix of panel individuals) to detect of the degree of dependence across panel, which can provide an indicator for the choice of test.

The parameters in the DGPs are generated as follows. The AR coefficient of u_{it}, β_i, is randomly drawn from uniform distribution [0.2, 0.4]. The factor loadings, υ_i, are randomly drawn from uniform [0.5, 3]. The processes ξ_t and η_{it} are independent and drawn from *iid* **N**(0,1) and *iid* **N**(0,*V*), respectively. The $N \times N$ covariance matrix *V* of the innovations η_{it} is symmetric positive definite. To ensure this property, *V* is generated according to the steps in Chang (2002):

(1) Generate an $N \times N$ matrix Ψ from uniform distribution [0,1];

(2) Construct from Ψ an orthogonal matrix $\mathrm{H} = \Psi(\Psi'\Psi)^{-1/2}$;[16]

(3) Generate a set of *N* eigenvalues, $\lambda_1,...,\lambda_N$. Let $\lambda_1 = r > 0$, $\lambda_N = 1$ and draw $\lambda_2,...,\lambda_{N-1}$ from uniform distribution [*r*,1];

(4) Form a diagonal matrix Λ with $(\lambda_1,...,\lambda_N)$ on the diagonal;

(5) Construct the covariance matrix *V* using the spectral representation $V = H\Lambda H'$.

4.2. The Monte Carlo Experiments

Simulations are conducted under various combinations of *N* and *T* on the CH and CS tests. Certain percentiles of the empirical distributions of the statistics are calculated for further response surface analysis. The choice of sample sizes is to facilitate running response surface regressions. The sample sizes are

[16] $(\Psi'\Psi)^{-1/2}$ is obtained by Cholesky decomposition of the inverse of $(\Psi'\Psi)$.

$$T \in \{50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500\}$$
$$\text{and} \quad N \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$$

So there are altogether $19 \times 10 = 190$ experiments for each DGP. Each experiment consists of $np = 10000$ replications. For each experiment the 0.01, 0.05 and 0.10 percentiles are recorded. The Monte Carlo simulations are programmed in Gauss 7.0.

5. Response Surface Estimation

The general procedure of response surface method is firstly to identify the factors that affect the response variable; once the important factors have been identified, the next step is to determine the settings or functional forms in which these factors result in the optimum value of the response variable by the outputs of regressions. On the studies of exploring the finite sample performance of statistical tests, MacKinnon (1996, 2000) point out that response surface regression coefficients can help researchers estimate the percentile for any given sample size and furthermore derive the numerical (or empirical or finite sample) distribution functions, so the empirical *p*-value for any given percentile can also be calculated.

5.1. Representation of Response Surface Regression

In this study, the dependent variable (or response variable) is the finite sample bias, calculated as the difference between a percentile in the numerical distribution (obtained by Monte Carlo experiments) and its corresponding percentile in the asymptotic distribution (standard normal distribution). The response surface regression takes the form

$$\Delta q_i^\alpha = \theta_1^\alpha \frac{1}{T} + \theta_2^\alpha \frac{1}{N} + \theta_3^\alpha \frac{1}{N^2} + \theta_4^\alpha \frac{N \ln(T)}{\sqrt{T}} + \varepsilon_i^\alpha \tag{2.8}$$

$$\Delta q_i^\alpha = \tilde{q}_i^\alpha - q^\alpha \tag{2.9}$$

where ε_i^α are error terms; $\tilde{q}_i^\alpha$ denotes the α percentile of the numerical distribution under sample size *N* and *T* which is obtained from the i^{th} experiment, so *i* is the note for sample size; q^α is the α percentile of standard normal distribution, so the dependent variable Δq_i^α represents finite sample bias.

Generally there is no specific rule on the design of the regression representation, rather it is a empirical and trial procedure. The functional form is basically determined by the goodness-of-fit of the regression and the significance of the coefficients of independent variables after a number of trials. Ericsson (1986) notes that the functional form of response surface regression can be justified on the grounds of the significant coefficients obtained and the generally high R^2 values. A general form to start with can be:

$$\Delta q_i^\alpha = \theta_1^\alpha \frac{1}{T} + \theta_2^\alpha \frac{1}{T^2} + \theta_3^\alpha \frac{1}{N} + \theta_4^\alpha \frac{1}{N^2} + \theta_5^\alpha \frac{1}{NT} + \theta_6^\alpha \frac{N\ln(T)}{\sqrt{T}} + e_i^\alpha \qquad (2.10)$$

For generalization a uniform specification for the 1%, 5% and 10% significant levels is opted for all the three DGPs rather than optimize the functional form for each one. The variable $\frac{N\ln(T)}{\sqrt{T}}$ is chosen according to Im and Pesaran (2003) criticism. They argue that the nonlinear IV panel unit root test statistic of the CH test needs a much more restrictive condition for its asymptotic property to hold, *i.e.* it requires $\frac{N\ln T}{\sqrt{T}} \to 0$, as *N*, $T \to \infty$. Since the only technical difference between CS and CH tests is the choice of IGF, *i.e.* a single IGF vs. a set of orthogonal IGFs, a similar problem is also conjectured to the CS test. Therefore this term is included in the regression to observe its influence on finite sample bias. Since it is panel data concerned, $\frac{1}{NT}$ is introduced as an interaction term according to Jönsson (2005), whereas the coefficient of this variable is insignificant for DGP2, 3 for all the 3 percentiles (1%, 5% and 10%), and the coefficient of $\frac{1}{T^2}$ is insignificant for DGP1 for all the 3 percentiles. With the presence of $\frac{1}{NT}$ and/or $\frac{1}{T^2}$, the values of R^2 are not essentially improved. So the two terms are excluded from the regression. Later it is shown that the chosen functional form has reasonably R^2 and good performance in terms of significance of the coefficients for all response surface regressions.

5.2. Estimation of Response Surface Regression

The response surface regression (2.8) is estimated by OLS. However, the errors in (2.8) are heteroskedastic due to the deterministic increasing values of *N* and *T*. Therefore the variance of the errors depends systematically on the sample sizes (*N* and *T*) and heteroskedasticity exists in the regression. To account for the heteroskedasticity, the covariance estimator developed by MacKinnon and White (1985) is applied.

Let $\hat{\theta}$ denote the vector of estimators and *X* be the matrix of regressors in (2.8). The covariance estimator of $\hat{\theta}$ is given as

$$\hat{V}(\hat{\theta}) = n^{-1}(n-1)(X'X)^{-1}\left(X'\hat{\Omega}X - n^{-1}X'\hat{u}\hat{u}'X\right)(X'X)^{-1} \tag{2.11}$$

where n is the number of observations in (2.8), *i.e.* $19 \times 10 = 190$; $\hat{\Omega}$ is an $(n \times n)$ diagonal matrix with diagonal elements $\hat{u}_j^2$; $\hat{u}_j = (1-k_{jj})^{-1}\hat{e}_j$ with k_{jj} as the j'th diagonal element of $X(X'X)^{-1}X'$; e_j are the residuals of (2.8) and $\hat{u}$ is the vector of $\hat{u}_j$.

Since the computation of percentiles (from simulation experiments) involves the two dimensions (N and T) of panel data, one may consider using panel data regression. The suitable panel data estimation for (2.8) is the fixed effect. Nevertheless, if the fixed effect is used, the variables $\frac{1}{N}$ and $\frac{1}{N^2}$ will be dropped during estimation process, which causes loss of information and results in inefficiency. This also can be evidenced by the goodness-of-fit of the regression, the value of R^2 from OLS estimation being higher than that from the fixed effect panel data estimation.

5.3. Augmentation on the Point Critical Values

After estimate the response surface regression (2.8), the finite sample bias is smoothed as

$$\Delta\hat{q}_i^\alpha = \hat{\theta}_1^\alpha \frac{\mathbf{1}}{T} + \hat{\theta}_2^\alpha \frac{\mathbf{1}}{N} + \hat{\theta}_3^\alpha \frac{\mathbf{1}}{N^2} + \hat{\theta}_4^\alpha \frac{N\ln(T)}{\sqrt{T}} \tag{2.12}$$

where $\Delta\hat{q}_i^\alpha$ is the fitted finite sample bias under any given sample size N and T. According to (2.9) the corresponding finite sample critical value is calculated as:

$$\hat{q}_i^\alpha = q^\alpha + \Delta\hat{q}_i^\alpha \tag{2.13}$$

However, there exist two types of uncertainty caused by Monte Carlo simulations and response surface regression. Firstly, sample percentiles exhibit certain distribution in the same spirit as the distribution of sample mean. To capture this randomness resulting from Monte Carlo experiments, the 95% confidence intervals are provided for the 1%, 5% and 10% percentiles using David-Johnson (1954) estimate of percentile standard deviation. David and Johnson (1954) derived the formula to calculate the standard deviation of percentile distribution on the basis of its sample size. The sample size in the study here is the number of replications in a simulation experiment. Secondly, since the bias is smoothed by response surface regressions, the residuals unavoidably become another source of uncertainty. This uncertainty can be represented by the standard error of regression residuals.

The 95% confidence interval of a sample percentile is given as

$$ul\hat{q}_i^{\alpha} = \hat{q}_i^{\alpha} \pm \mathbf{1.96}\left(se(\alpha) + se\left(\varepsilon^{\alpha}\right)\right) \tag{2.14}$$

where $\hat{q}_i^{\alpha}$ is the critical value of α percentile calculated from (2.12) and (2.13) with sample size *N* and *T*; $se(\alpha)$ is David-Johnson (1954) estimate of the standard deviation of α percentile distribution; $se\left(\varepsilon^{\alpha}\right)$ is the standard error of the response surface regression residuals in (2.8). Thus, the upper and lower limits of a critical value interval are respectively given by

$$u\hat{q}_i^{\alpha} = \hat{q}_i^{\alpha} + \mathbf{1.96}\left(se(\alpha) + se\left(\varepsilon^{\alpha}\right)\right)$$

and $$l\hat{q}_i^{\alpha} = \hat{q}_i^{\alpha} - \mathbf{1.96}\left(se(\alpha) + se\left(\varepsilon^{\alpha}\right)\right)$$

This means if an empirical testing statistic falls between lq_i^{α} and uq_i^{α}, it is indecisive if the null hypothesis is to be rejected or not.

5.4. *P*-values of the Numerical Distributions of CH and CS Testing Statistics

The response surface coefficients computed from (2.8) can also be used to help estimate *p*-values of finite sample distributions. The approximation procedure is according to MacKinnon (1996, 2000). MacKinnon (1996, 2000) approximate a finite sample distribution under sample size *N* and *T* by 221 fitted values of percentiles $\hat{q}_i^{\alpha}$ from (2.13). The chosen 221 α 's are 0.0001, 0.0002, 0.0005, 0.001, 0.002, …, 0.01, 0.015, …, 0.99, 0.991, …, 0.999, 0.9995, 0.9998, 0.9999. However, the result of augmented critical values indicates that due to the relatively large standard error of the response surface regression residuals, the percentiles smoothed by response surface regressions are not appropriate to use. This is shown is section 6. So the 221 point estimates are taken directly from Monte Carlo experiment results. For sample sizes that are not considered in the experiments, the nearest in their neighborhood can be used as approximation.

To interpolate between the 221 points in order to calculate *p*-value for any given test statistic, the procedure involves an estimation

$$\Phi^{-1}(\alpha) = \gamma_0 + \gamma_1 \tilde{q}_i^{\alpha} + \gamma_2 \left(\tilde{q}_i^{\alpha}\right)^2 + \gamma_3 \left(\tilde{q}_i^{\alpha}\right)^3 + \upsilon^{\alpha} \tag{2.15}$$

where $\Phi^{-1}(\alpha)$ is the inverse of cumulative standard normal distribution at α, since the asymptotic distributions of CH and CS tests are standard normal; $\tilde{q}_i^{\alpha}$ is the point percentile estimate from Monte Carlo experiment under sample size *N* and *T*.

The idea of regressing (2.15) is to use a small number of points in the neighborhood of the test statistic to estimate the relationship between the empirical distribution and standard normal distribution. For example, suppose a numerical distribution based on DGP3 under sample size $N = 10, T = 100$ is the interest, and the empirical testing statistic is $\hat{q}_{NT} = \text{-}2.7112$. Among the estimated 221 percentiles from the corresponding distribution computed by Monte Carlo experiment, the closest one to this statistic is $\tilde{q}_{NT}(\alpha = \mathbf{0.006}) = \text{-}2.73216$. If 9 points are used, (2.15) is to be estimated with the percentiles $\tilde{q}_i^{\alpha}$ for

$$\alpha = \{0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.010\}^{17}.$$

Using the estimators in (2.15), the *p*-value of an empirical test statistic, $\hat{q}_{NT}$, can be computed from

$$p = \Phi\left(\hat{\gamma}_0 + \hat{\gamma}_1 \hat{q}_{NT} + \hat{\gamma}_2 \left(\hat{q}_{NT}\right)^2 + \hat{\gamma}_3 \left(\hat{q}_{NT}\right)^3\right) \tag{2.16}$$

[17] MacKinnon (2001) suggests that shown by experiments, 9, 11 and 13 points are reasonable numbers to use and 9 points is a good choice. Hence, 9 points is used here.

MacKinnon (1996) confirms the reliable accuracy of the approximation of *p*-values with experiments.

Since $\Phi\left(\hat{\gamma}_0 + \hat{\gamma}_1 \hat{q}_{NT} + \hat{\gamma}_2 (\hat{q}_{NT})^2 + \hat{\gamma}_3 (\hat{q}_{NT})^3\right)$ approximates the cumulative distribution function of statistic $\hat{q}_{NT}$, the approximate density function can also derived by taking the first derivative of (2.16),

$$f(\hat{\tau}) \approx \phi\left(\hat{\gamma}_0 + \hat{\gamma}_1 \hat{q}_{NT} + \hat{\gamma}_2 (\hat{q}_{NT})^2 + \hat{\gamma}_3 (\hat{q}_{NT})^3\right)\left(\hat{\gamma}_1 + \hat{\gamma}_2 \hat{q}_{NT} + \hat{\gamma}_3 (\hat{q}_{NT})^2\right) \tag{2.17}$$

where $\phi(.)$ denotes the standard normal probability density function.

6. Results

Following section 4.1, for simplicity, in the result section ‘DGP1’ refers to the CH test applied on DGP1; ‘DGP2’ and ‘DGP3’ refer to the CS test applied on DGP2 and 3, respectively.

6.1. Estimation Results of Response Surface Regressions

The results of response surface estimations are presented in Table 2.1. The table contains the coefficients of variables in the regression and the value of R^2 for each estimation. Only the standard errors of $\hat{\theta}_4^{\alpha}$ (the coefficient of $\frac{N\ln(T)}{\sqrt{T}}$) is provided, since $\frac{N\ln(T)}{\sqrt{T}}$ is the condition pointed out by Im and Pesaran (2003) for the asymptotic properties of CH test to hold and the similar situation is also conjectured to CS test as discussed in the introduction.

Almost all the coefficients in each regression are highly significant, in particular, $\hat{\theta}_4^{\alpha}$, with practically zero *p*-values (so they are not printed in the table). The standard errors of $\hat{\theta}_4^{\alpha}$ in all the three DGPs (both CH and CS tests) are substantially small. This suggests the estimation accuracy of $\hat{\theta}_4^{\alpha}$ and implies the important role of $\frac{N\ln(T)}{\sqrt{T}}$ in the performance of both tests. In addition, in the trials, if constant term is added to (2.8) and exclude $\frac{N\ln(T)}{\sqrt{T}}$, the constant term is significant. This obviously contradicts the asymptotic theory, since by the asymptotic properties as $N, T \to \infty$, the magnitude of bias computed through (2.8) should pass to zero. However, if $\frac{N\ln(T)}{\sqrt{T}}$ is included in (2.8), the constant term becomes insignificant. This suggests that $\frac{N\ln(T)}{\sqrt{T}} \to 0$ is required for the bias to diminish as $N, T \to \infty$, which is consistent with Im and Pesaran (2003)'s critique.

Table 2.1 Response surface regression estimates

	α	$\hat{\theta}_1^{\alpha}$	$\hat{\theta}_2^{\alpha}$	$\hat{\theta}_3^{\alpha}$	$\hat{\theta}_4^{\alpha}$	$se(\hat{\theta}_4^{\alpha})$	R^2
DGP1	0.01	-5.8576	-4.4796	32.4385	0.0031	(0.0003)	0.6410
	0.05	-6.2600	-1.8075	10.4013	0.0046	(0.0002)	0.7979
	0.10	-6.8405	-0.2919	-0.9532	0.0057	(0.0002)	0.8363
DGP2	0.01	27.2984	-13.2460	42.8065	0.0118	(0.0085)	0.8455
	0.05	20.2953	-7.8281	24.0752	0.0082	(0.0006)	0.8347
	0.10	17.2754	-5.8285	20.4715	0.0066	(0.00047)	0.8213
DGP3	0.01	152.5797	-112.8447	772.0970	-0.0302	(0.0022)	0.8309
	0.05	72.3938	-46.0001	266.5182	-0.0137	(0.0012)	0.8419
	0.10	44.3161	-23.1107	117.0348	-0.0081	(0.00084)	0.8040

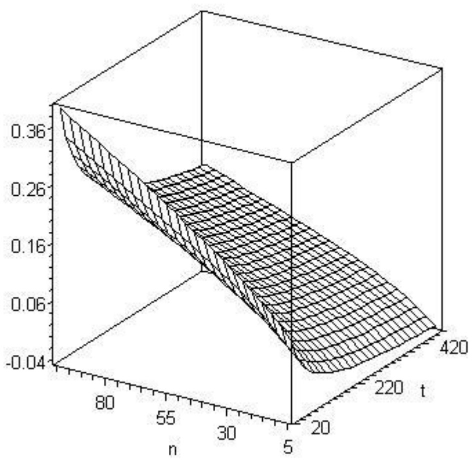
Note: $se(\hat{\theta}_4^{\alpha})$ denotes the standard error of $\hat{\theta}_4^{\alpha}$

The goodness-of-fit of response surface regressions is reasonably high. Regressions for the CS test have better performance than that of the CH test in terms of R^2. The values of R^2 of regressions for all the significant levels for DGP2 and DGP3 are higher than 0.8.

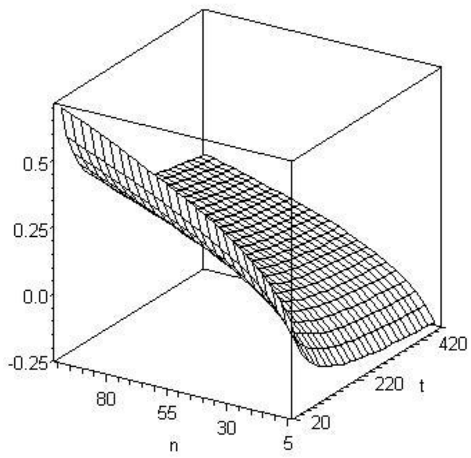
6.2 Finite Sample Bias

Figure 2.2 (a)-(i) plot the trends of bias of the 1%, 5% and 10% percentiles for the three DGPs as sample size N and T increase (for illustration the estimates of 5% percentile, (b), (e) and (h), are presented in the text; the rest are attached in the Appendix B). The values used to plot Figure 2.2, $\Delta\hat{q}_i^{\alpha}$, are computed through response surface regression estimates in (2.12) using N = {5,15,25,…,105} and T = {25,50,75,…,500}.

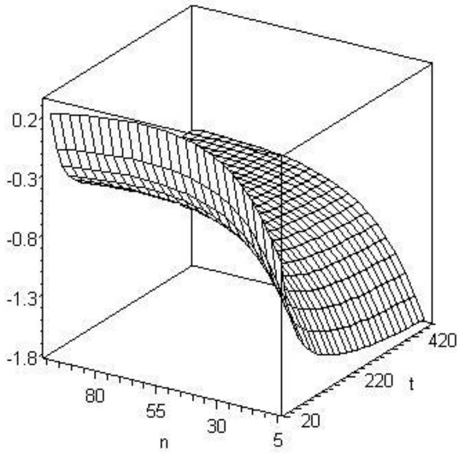
Figure 2.2 Estimates of finite sample bias of the 5% percentiles for the three DGPs as sample sizes N and T vary (different scales for N and T)



(b) DGP1-5%



(e) DGP2-5%



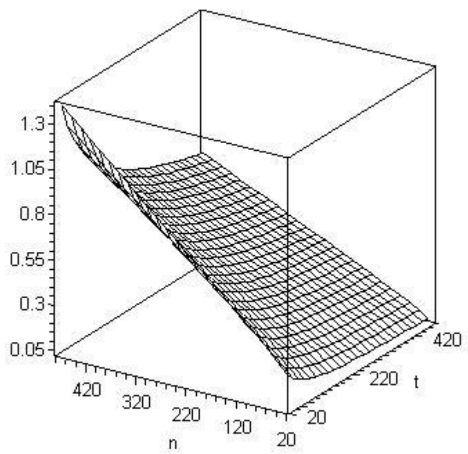
(h) DGP3-5%

A few properties can be observed from the figures. For all the DGPs as *T* increases it tends to drag down the magnitude of bias, whereas as *N* increases the magnitude of bias tends to rise. The effect caused by *T* is similar to all DGPs. When the size of *T* is small, an increase in *T* causes large downward movement of the value of bias; as *T* becomes large, the effect mitigates. When *N* increases, the response of bias value shows the same constant growing pattern in DGP1 and 2; whereas its growing rate in DGP3 slows down after certain *N*. Moreover, all graphs illustrate significant magnitude of bias with small *T* and large *N*, which suggest that as both *N* and *T* go large, the bias tend to be eliminated.

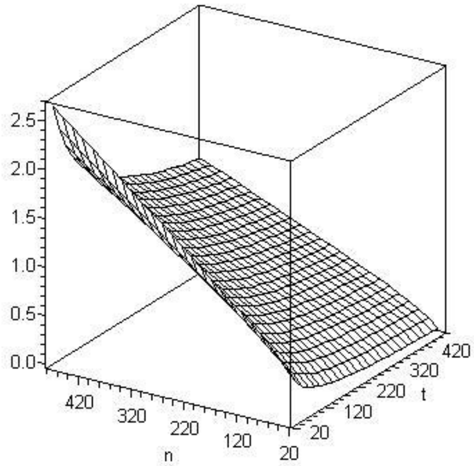
However, Figure 2.2 is plotted on different scales of *N* and *T* with *T* growing faster than *N* due to the choice of sample size in Monte Carlo simulations. Figure 2.3 unifies the scales of *N* and *T* and shows some different feature regarding the trend of bias. An additional view for DGP3 from a different angle is provided to give better visibility. As *N* and *T* both increase at the same rate, the size of bias keeps growing constantly. The bias in DGP3 goes in the opposite direction against that in DGP1 and 2 as shown in

Figure 2.3. The position where bias tends to disappear is where T is large and N is small, which is consistent with the point of Im and Pesaran (2003).

Figure 2.3 Estimates of finite sample bias of the 5% percentile for the three DGPs as sample sizes N and T vary (same scale for N and T)

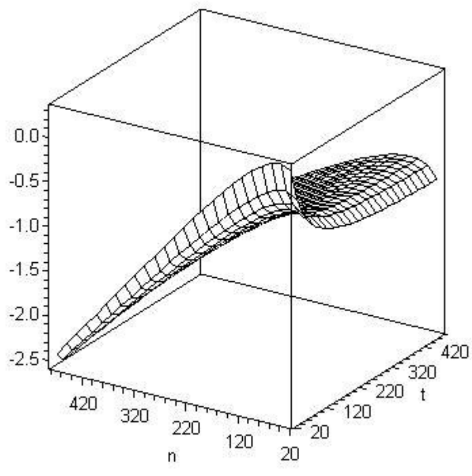


(a) DGP1-5%

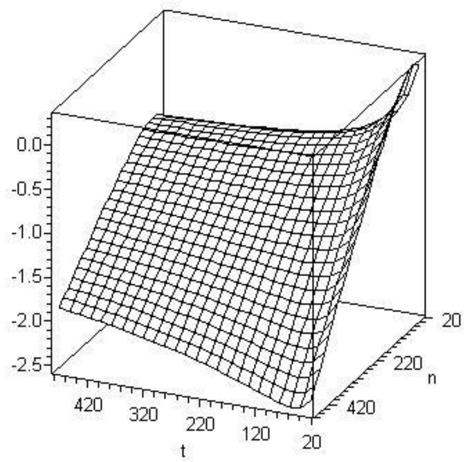


(b) DGP2-5%

Figure 2.3 Estimates of finite sample bias of the 5% percentile for the three DGPs as sample sizes N and T vary (same scale for N and T) **(Cont'd)**



(c) DGP3-5%



(d) DGP3-5% (different angle)

6.3 The Augmented Critical Values

According to the discussion in section 5.3, to account for the uncertainties resulting from Monte Carlo simulations and regression estimation, confidence intervals for percentiles are computed. To see the magnitude of the uncertainties, Table 2.2 presents the gap between the upper limit and lower limit of the percentile confidence intervals, *i.e.* $\left(uq^{\alpha} - lq^{\alpha}\right)$.

Table 2.2 Gap between the upper limit and lower limit of percentile confidence interval $\left(uq^{\alpha} - lq^{\alpha}\right)$

α	DGP1	DGP2	DGP3
0.01	0.3437	0.7865	1.5630
0.05	0.2182	0.5196	0.8259
0.10	0.1865	0.4170	0.6130

As expected, the gap increases with decreasing percentiles, since empirical percentiles tend to be more volatile in the extremes. However, it is found that due to the relatively large standard error of response surface regression residuals, the gap is so wide in some intervals that they overlap with their neighborhood percentile interval. For example, the upper limit of a 1% percentile interval is higher than the lower limit of that of 5%. This makes it difficult to distinguish between 1% and 5% significant level. The problem is particularly serious for DGP3. Therefore, the response surface estimates are not applicable in computing finite sample critical value intervals. As a result the point estimates from Monte Carlo simulations are directly used, then only the randomness from simulation experiments needs to be considered, *i.e.*

$$\tilde{q}_{ul}^{\alpha} = \tilde{q}_{i}^{\alpha} \pm \mathbf{1.96} \times se(\alpha)$$

where $\tilde{q}_i^{\alpha}$ denotes the α percentile obtained from the i^{th} Monte Carlo experiment with sample size *N* and *T*; $se(\alpha)$ is David-Johnson (1954) estimate of the standard deviation of α percentile distribution. The finite sample critical value intervals are provided in Table 2.3 in Appendix B.

6.4 The Numerical Distributions

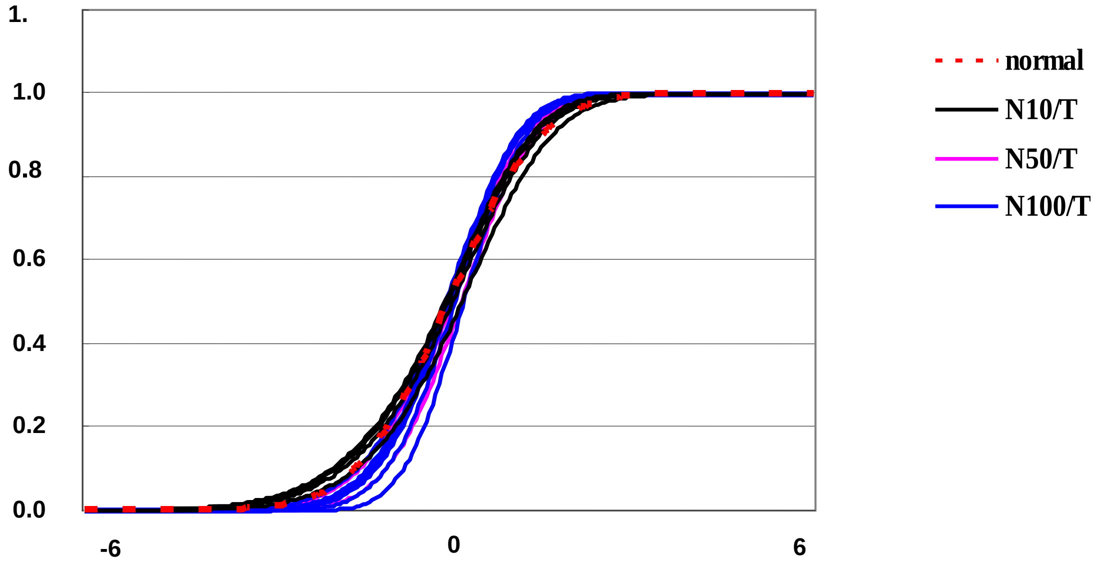
Figure 2.4-2.6 are the plots of finite sample cumulative density functions (CDF) and probability density functions (PDF) for the three DGPs along with the plot of those of standard normal distribution for comparison (for illustration Figure 2.5 for DGP2 is presented in the text; Figure 2.4 and 2.6 for DGP 1 and 3 are attached in Appendix B). All figures are plotted with 221 points. The calculation of empirical *p*-values for the CDFs is discussed in section 5.4. For simplicity, the values for the PDFs are taken from the Monte Carlo simulation estimates. Plots with N = {10,50,100} and T = {50,100,200,300,500} are chosen for illustration as shown in Figure 2.4-2.6. To clearly observe the trend as *T* increases, plots with N = {50} and T = {50,100,200,300,500} are extracted and provided in Figure 2.7-2.9 (for the same reason, Figure 2.8 for DGP2 is presented in the text; Figure 2.7 and 2.9 for DGP 1 and 3 are attached in Appendix B).

The finite sample distributions of all the three DGPs suggest substantial difference from standard normal distribution and each DGP has its own feature. Keep *T* constant and let *N* increase, the numerical CDF and PDF for DGP1 move to the right (Figure 2.4); the CDF for DGP2 tends to move anticlockwise and the PDF grows taller and thinner (Figure 2.5); the left tail of the CDF for DGP3 in Figure 2.6 is particularly heavy and move toward that of standard normal distribution.

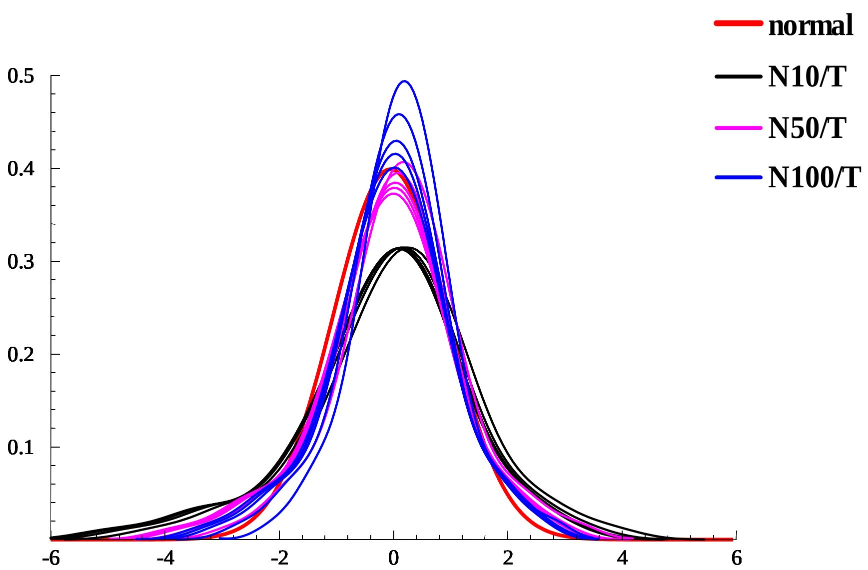
Let T grow given constant N, the numerical CDF and PDF for DGP1 both shift to the right slowly first and then back toward the left as shown in the enlarged views in Figure 2.7. The numerical CDF for DGP2 tends to move to the left and the PDF is pressed flatter, just in the opposite direction to that as N increases (Figure 2.8). The similar patterns of the numerical distributions to DGP2 are also found in DGP3 (Figure 2.9).

In general, increase in either N or T drives the finite sample distributions away from that of standard normal distribution; whereas growth in the other dimension tends to offset the previous effect and brings the numerical distributions back to normal distribution. However, it appears that the speed of convergence is seriously slow. The graphs in this section provide additional evidence that simply applying critical values from standard normal distribution for the CH and CS tests is highly unreliable.

Figure 2.5 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP2

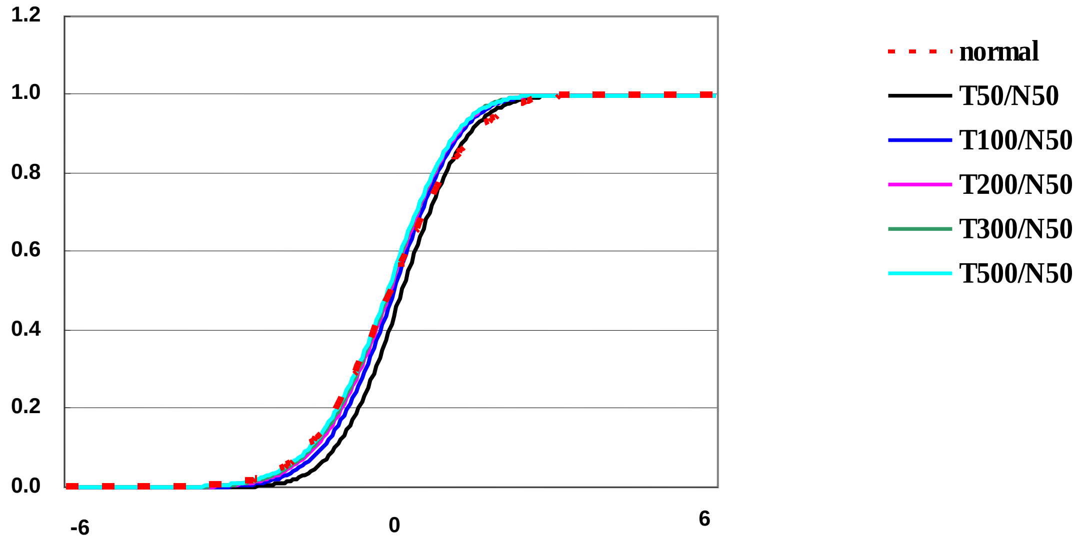


(a) DGP2-CDF

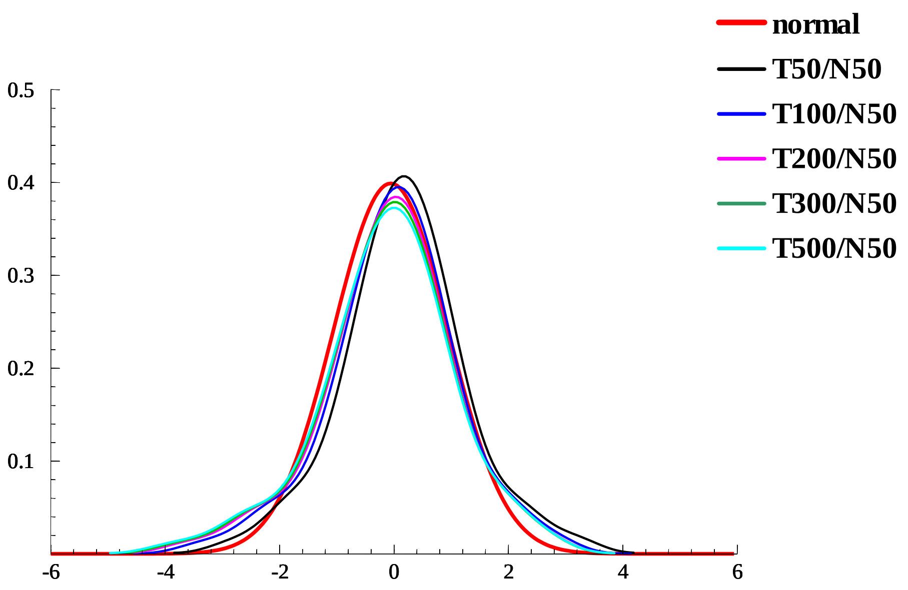


(b) DGP2-PDF

Figure 2.8 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP2, $N = 50$



(a) DGP2-CDF



(b) DGP2-PDF

7. Conclusion

This chapter has assessed the finite sample performance of the recently developed CH and CS panel unit root tests using numerical methods. Although compared with traditional time series unit root tests, panel unit root testing statistics exhibit standard normal distribution asymptotically, it is found in this chapter that the finite sample performance of panel unit root tests substantially differ from their asymptotic properties. Simply applying the critical values given by asymptotic distribution can seriously mislead the results.

A number of Monte Carlo experiments are conducted on the CH and CS tests and provide a base for carrying out the response surface analysis. The response surface regression examines the finite sample bias of the two tests in relation to sample size and the condition concerning the asymptotic properties of the tests. The finite sample bias and the numerical distributions under selected sample sizes are plotted and reveal that the finite sample performance of the tests systematically depends on sample size. These results clearly highlight the substantial differences between finite sample and large sample properties and raise caution for empirical studies that ignore the problem.

The 95% confidence intervals of critical values are provided by augmenting the Monte Carlo point estimates with David-Johnson estimate of percentile standard deviation to take into account the randomness incurred by simulation experiments. In addition, the formula to compute finite sample *p*-value of test statistic is also presented. Due to the slow convergence rate of the test statistics (as shown in the numerical distributions), the chapter recommends that even with larger number of observations ($N > 100$ and/or $T > 500$), the critical values computed under the largest finite sample size ($N = 100$ and/or $T = 500$) are preferred for applications.

The results are consistent with the Im and Pesaran (2003) critique and the conjecture in this chapter that $\frac{N\ln(T)}{\sqrt{T}} \to 0$ is an important condition for the asymptotic properties of the CH and CS tests to hold, since the coefficients of $\frac{N\ln(T)}{\sqrt{T}}$ in the response surface regressions are highly significant and precisely estimated; the presence of the variable also provides support to the high value of R^2. Nevertheless, by applying the finite sample critical values provided in this chapter, this problem should not pose a danger to applications.

However, the study experiences some limitations due to the goodness-of-fit of response surface regressions. Although the goodness-of-fit of the regressions have reached reasonably high values (over 0.8) after experimenting with a number of potential independent variables (factors that affect the response variable), they are still not sufficiently high for smoothing finite sample critical values. During augmenting the smoothed critical values (by response surface regressions), the gaps between the upper and lower limits of some critical value intervals are excessively enlarged due to the relatively large residual standard errors in response surface regressions, so some neighbour intervals overlap. Therefore the point estimates from Monte Carlo simulations are used instead. This problem suggests that unless the response surface regression has near perfect goodness-of-fit, smoothing critical values by the regression adds excessive uncertainty and it is not recommended. The problem also reflexes the complexity of the finite sample performance of panel unit root tests. In order to increase to goodness-of-fit of response surface regressions to the desired level so that the finite sample critical value of panel unit root tests can be conveniently smoothed, more work is needed in future research on the model selection issue.

A general notable limitation for unit root tests is that the idiosyncratic terms in the model are assumed to be normally distributed, whereas this assumption is very often hard to maintain in empirical data. If normality is relaxed, the corresponding asymptotic distribution does not hold. An extended analysis of relaxing the normality assumption in panel unit root tests is the interest in the next study.

Chapter 3 Panel Unit Root Tests with Infinite Variance Errors

1. Introduction

In the literature of panel unit root tests as well as most time series tests, the crucial assumption of Gaussian error terms is imposed. The limiting distribution of single time series test statistic is expressed in terms of Brownian motion/Wiener process (see Chapter 1 section 5.1). The Gaussian law plays a crucial role in developing the asymptotic properties that are associated with Wiener process -- on each time step the value change in the process is assumed to be normally distributed (c.f. Hamilton, 1994, Chap 17). Recall in Chapter1 in the procedures of developing the asymptotics of panel unit root tests, normality plays important role. For example, the heterogeneous tests usually adopt the sequential approach which allows T to pass to infinity first to obtain an intermediate limit for each individual time series in the panel, a similar procedure to that for time series test. So if the Gaussian assumption cannot be fulfilled, the asymptotic properties will be affected. The difficulty is that in practice, normality is hardly found in most empirical data and the common phenomenon of non-normality has to be taken into account.

One popular candidate used to model non-normality is the Lévy-Paretian stable distribution, also called the α stable distribution. It is well known that financial variables are usually distributed with heavy tails. Mandelbrot (1963, 1967) postulate infinite variance and suggest considering stable laws for certain commodity prices. Fama (1965) notes that the empirical distribution of stock price changes are leptokurtic and are governed by stable laws. More recently McCulloch (1997) provide some empirical support for stable modeling on financial data. Other studies in favor of the stable process for certain economic variables include Mandelbrot (1961), DuMouchel (1983), Cambanis and Fotopoulos (1995), Fotopoulos (1998), etc.

Due to the importance of examining the stationarity property of time series in economics, single time series unit root test has been analysed in a few studies on the basis of non-normality. Chan and Tran (1989) develop the asymptotic theory for time series regression with a unit root under the assumption that the distribution of errors belong to the domain of attraction of a stable law. Phillips (1990) extends the results of Chan and Tran (1989) to weak dependence and heterogeneity in the errors. The limiting distribution formulas in the two studies are presented in terms of functionals of a stable process instead of Brownian motion without any other modification compared with the finite variance case. Callegari et al. (2003) study the case of random walk with drift under non-normality. Ahn et al. (2001) analyze a few unit root tests associated with infinite variance and assess the small sample performance of the test statistics. However, in the field of panel unit root test, no attention has yet been given to the problem caused by non-normality.

The aim of this chapter is to study how sensitive panel unit root tests are in response to non-normality. The α stable distribution is adopted to capture the non-normality. More specifically, the error terms in panel unit root test models are drawn from α stable distribution. Using Monte Carlo simulations and regarding the parameter that controls the leptokurtosis of α stable distribution as an indicator of the degree of non-normality, the chapter intends to find the certain degree which causes so severe size distortion to the test that the critical value intervals computed in the previous chapter under normality can no longer be used.

For application purpose the framework panel unit root tests used are the Chang (2002) (CH hereafter) and Chang and Song (2005, 2009) (CS hereafter) tests due to their advantage in robustness to all the different forms of cross section dependence (*i.e.* the weak form, strong form and long run dependence). Reviews of the tests and the problem of cross section dependence are provided in Chapter 1. Monte Carlo experiments are conducted with a selection of sample sizes and degrees of non-normality. The results suggest that the tests in different situations in terms of dependence respond to non-normality very differently and each has its own benchmark tolerance to non-normality. In addition, the results are seriously influenced by the randomness caused by different sets of random numbers. Each test under different circumstance also has its distinctive response to this impact. This problem will be discussed in great detail in the results section. To illustrate the trend of test bias incurred by non-normality as the degree of non-normality varies, response surface regression with a special functional form design is applied using the simulation results. Due to the fact that the computations of Monte Carlo experiments are remarkably time consuming, the response surface analysis of bias is only performed on the CS test for strong dependence as an example.

The remainder of the chapter is organized as follows. Section 2 introduces the Lévy-Paretian stable distribution. Section 3 presents the Monte Carlo experiments. Section 4 and

5 set out the estimation of response surface regression and the decision rule to find the benchmark α, respectively. Results are analyzed in section 6. Section 7 concludes.

2. Lévy-Paretian Stable Distribution

2.1 Definitions of Stable Distribution

There are a few equivalent definitions of a stable distribution based on either its stability property or characteristic function. Details can be found in several textbooks and monographs such as Samorodnitsky and Taqqu (1994), Sato (1999), Feller (1971), Zolotarev (1986). Here two definitions are introduced on the stability property and characteristic function, respectively.

Definition 1.1 A random variable X is *stable* or *stable in the broad sense* if for any positive numbers *a* and *b*, there is a positive number *c* and a real number *d* such that

$$aX_1 + bX_2 \overset{d}{=} cX + d \tag{3.1}$$

where X_1 and X_2 are independent copies of X; '$\overset{d}{=}$' denotes equality in distribution, which suggests the expressions on both sides have the same probability law. The random variable is *strictly stable* or *stable in the narrow sense* if (3.1) holds with $d = 0$. A random variable is *symmetric stable* if it is stable and symmetrically distributed around 0, e.g. X and $-X$ have the same distribution. A symmetric stable random variable is obviously strictly stable.

It is familiar to us that (3.1) is an important property of normal or Gaussian random variables which is supported by the Central Limit Theorem (CLT). Later it is shown that normal distribution is a special case of stable distribution and stable distribution generalizes the CLT. The word '*stable*' used is due to the important property that the shape of X is preserved or unchanged (up to scale and shift) under addition as it is shown in (3.1).

The probability density of stable distribution is not available explicitly except for three special cases[18]. Stable distributions are usually described by their characteristic function.

Definition 1.2 A random variable X is *stable* if there are parameters $0<\alpha\le 2$, $\gamma>0$, $-1<\beta\le 1$ and μ real such that its characteristic function φ has the following form

$$\varphi(t)=E\exp(i\theta X)=\begin{cases}\exp\left\{-\gamma^{\alpha}|\theta|^{\alpha}\left[1-i\beta(\mathbf{sign}\theta)\tan\frac{\pi\alpha}{2}\right]+i\mu\theta\right\} & \alpha\neq\mathbf{1}\\ \exp\left\{-\gamma|\theta|\left[1+i\beta\frac{2}{\pi}(\mathbf{sign}\theta)\ln|\theta|\right]+i\mu\theta\right\} & \alpha=1\end{cases} \tag{3.2}$$

where $i^2=-1$ and $\mathbf{sign}\theta$ represents the sign function such that

$$\mathrm{sign}\theta=\begin{cases}1 & \text{if } \theta>0\\ 0 & \text{if } \theta=0\\ \text{-}1 & \text{if } \theta<0\end{cases}$$

There are four parameters in the characteristic function. The characteristic exponent or the *index of stability* α is the measure of leptokurtosis of the distribution; the *skewness* parameter β determines the skewness of the distribution; the *scale* or *dispersion* parameter γ measures the width of the distribution; and the *location* parameter μ controls the shift of the distribution. When the distribution is standardized, the scale $\gamma=1$ and the location $\mu=0$; by definition if the distribution is symmetric, the skewness $\beta=0$ and the location

[18] The three cases are the Gaussian distribution ($\alpha=2$), the Cauchy distribution ($\alpha=1$) and the Lévy distribution ($\alpha=1/2$)

$\mu = 0$. Since (3.2) is characterized by the four parameters, we denote a stable distribution by $S_\alpha(\gamma, \beta, \mu)$ and write

$$X \sim S_\alpha(\gamma, \beta, \mu)$$

to indicate that *X* has stable distribution $S_\alpha(\gamma, \beta, \mu)$. When *X* is symmetric α -stable, *i.e.* when $\beta = \mu = 0$, it is written as $X \sim S\alpha S$ and obviously $X \sim S_\alpha(\gamma, 0, 0)$. A random variable *X* is *standard* $S\alpha S$ if $\gamma = 1$. It is observed from (3.2) that if *X* is $S\alpha S$, then its characteristic function takes a particularly simple form

$$\varphi(t) = E\exp(i\theta X) = \exp\left(-\sigma^\alpha |\theta|^\alpha\right) \tag{3.3}$$

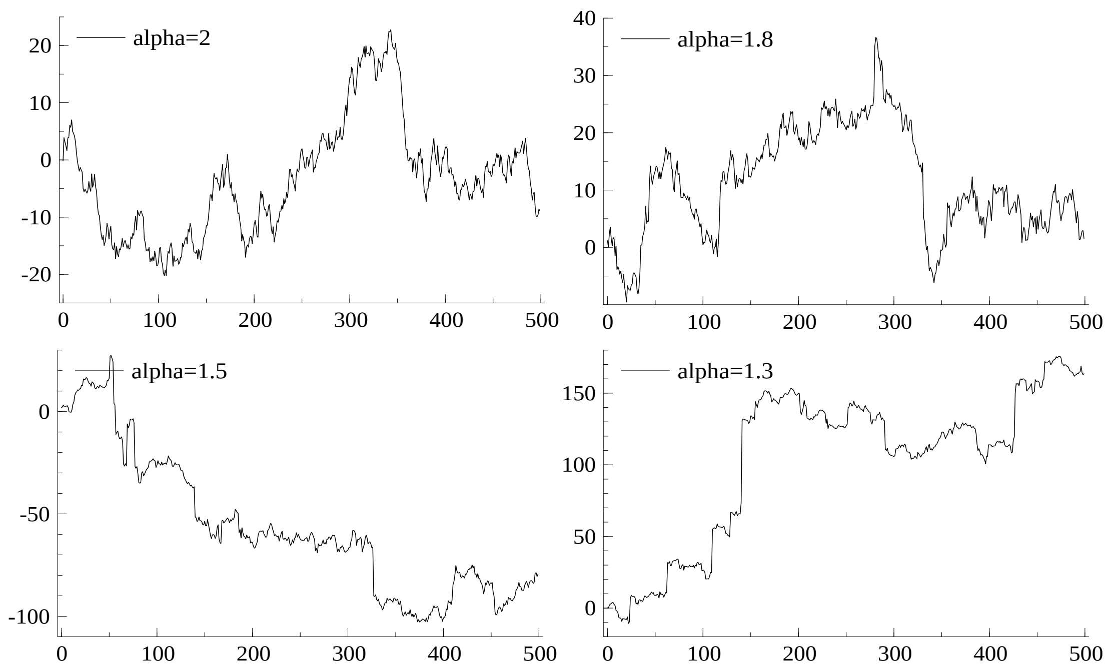
In this study the *standard* symmetric stable distribution is considered, $X \sim S\alpha S$ or $X \sim S_\alpha(1,0,0)$. This directs the focus on the most important and interesting parameter α. The index of stability α measures the tail thickness and peakedness at the origin. As α becomes smaller, the shape of the distribution becomes higher, more peaked and shows fatter tails. For $\alpha = 2$, the stable distribution reduces to normal (Gaussian) distribution. When $\alpha < 2$ leptokurtosis appears, and the tails become so heavy that the variance is infinite; for $\alpha < 1$ even the first moment does not exist. (Rachev et al, 2007) notes that in empirical finance α usually takes the values in the interval $(1, 2)$, which suggests some financial data modeled with stable laws exhibit finite means but infinite variance.

More recently some studies find the tails of α stable distribution too heavy for empirical study, while the tempered stable distribution can be opted as a compromise between normal distribution and α stable distribution (c.f. Kim et al, 2008). The tempered stable distribution (or truncated Lévy distribution) has tails heavier than normal distribution and thinner than α stable distribution; moreover, it has finite moments for all orders, compared with α stable distribution which has infinite second moment and infinite first moment when $\alpha < 1$ due to the excessive heaviness in its tails. Mantegna and Stanley (1997) showed that the probability density function (PDF) of the Standard & Poor's 500 index can be described by a tempered stable distribution. Some recommended readings on this subject can be found in Koponen (1995), Boyarchenko and Levendorskii (2000) and Carr et al. (2002).

2.2 Advantages of Applying Stable Distribution

The infinite variance of stable distribution suggests that the tails of the distributions may be too heavy for empirical data, whereas Rachev and Mittnik (2000) account for comprehensive applications of stable distribution in finance. One significant advantage of the stable laws is that the sums of α stable random variables are still α stable (as in definition 1.1). In financial modeling this property implies that the cumulative *IID* (identically independent distributed) daily returns over a period have the same distributional shape as the individual daily returns. Furthermore, according to the Central Limit Theorem (CLT), the normalized sum of *IID* random variables with finite variance converges to normal distribution. The Generalized Central Limit Theorem (GCLT) relaxes the finite variance assumption and states that the *only* possible resulting limit distribution is stable, *i.e.* the sums of *IID* random variables with infinite variance converge only to stable distribution. In addition and more importantly, the stable random variables are useful in simulating structural breaks. Figure 3.1 presents several plots of stable processes with different values of α. It shows that as α becomes smaller (*i.e.* the distribution of the variable becomes more leptokurtotic), the number of breaks increases and the magnitude of the breaks appear to be enlarged.

Figure 3.1 Plots of stable processes with different values of index of stability α



3. Monte Carlo Analysis on the Impact of Non-normality on Panel Unit Root Tests

As most of the distributions of empirical data do not exhibit normal distribution, e.g. the distributions of inflation rate, exchange rate, asset prices, etc., non-normality must be considered while conducting econometric testing. In the field of unit root test, the derivation of asymptotic properties usually experiences complicated procedures and yet is still not able to provide analytical solutions. In particular, when non-normality is involved the analysis becomes even more complex. In terms of panel unit root tests, even though under normality the testing statistics have normal distribution in the limit, the additional cross section dimension *N* adds extra complications to the finite sample performance (as found in Chapter 2) as well as the analytics (refer to the reviews in Chapter 1), not mention it under non-normality. Given that empirical work almost always experiences the constraint of limited number of data observations and Monte Carlo experiments are convenient to apply, in particular with the advantage of modern advanced digital computers, the Monte Carlo method is obviously a good choice to provide approximate solutions to empirical problems.

This section uses Monte Carlo simulation analysis to detect the influence of non-normality on the CH and CS panel unit root tests. The error terms in the unit root models are assumed to follow symmetric α stable distribution instead of normal distribution. The aim is to investigate the response of the tests as the value of the index of stability α decreases, *i.e.* as the distribution differs further and further from normal distribution and non-normality appears to be more severe.

3.1 The DGPs

3.1.1 The CH and CS Tests under Non-normality

The basic model of panel unit root test is

$$y_{it} = \rho_i y_{i,t-1} + u_{it} \qquad i = 1,\ldots,N;\ t = 1,\ldots,T_i \tag{3.4}$$

Following the CH and CS tests, under the null hypothesis of a unit root (then $\Delta y_{it} = u_{it}$), the innovations u_{it} are generated by the three DGPs below to incorporate the weak form, strong form and long run cross section dependence (in the same spirit as the DGPs in Chapter 2 under normality):

$$\text{DGP1:}\quad u_{it} = \beta_i u_{i,t-1} + \eta_{it} \tag{3.5}$$

$$\text{DGP2:}\quad u_{it} = \beta_i u_{i,t-1} + \upsilon_i \xi_t + \eta_{it} \tag{3.6}$$

$$\text{DGP3:}\quad u_{it} = \beta_i u_{i,t-1} + \upsilon_i \xi_t + \Delta\eta_{it} \tag{3.7}$$

where β_i is the AR coefficient; ξ_t is common factor and υ_i are factor loadings; η_{it} are innovations following symmetric α stable distribution with symmetric and nonsingular covariance matrix, say V. DGP1 embodies weak dependence given by the covariance matrix of η_{it}. DGP2 presents a mixture of the strong and weak forms of dependence through time effect ξ_t and innovations η_{it}. DGP3 exhibits the long run dependence. More explanations of the dependence in the DGPs and the choice of parameters as well as the generation of V are in Chapter 2, section 4.1. For the same reason as stated in Chapter 2, the CH test is applied on DGP1 and CS test is applied on DGP2 and 3. Under the assumption that η_{it} follow Gaussian process, the CH and CS tests both conclude with the standard normal asymptotic distributions of the testing statistics. Nevertheless, the η_{it} in this chapter are generated from the symmetric stable distribution.

3.1.2 Generation of α Stable Random Variables

It is assumed that $\eta_{it} \sim S\alpha S$ or $\eta_{it} \sim S_\alpha(1,0,0)$. Since the stable distributions do not have densities in closed form, sampling from stable distributions is not a straightforward affair. A common procedure of generating stable random variables is based on the following representation: if Z is uniformly distributed over $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and W is exponentially distributed with mean 1, assume Z and W are independent, then

$$X = \frac{\sin(\alpha Z)}{(\cos(Z))^{1/\alpha}} \left(\frac{\cos((1-\alpha)Z)}{W} \right)^{(1-\alpha)/\alpha} \tag{3.8}$$

is $S_\alpha(1,0,0)$. See p.42 of Samorodnitsky and Taqqu (1994) for a proof. In the Cauchy case when $\alpha = 1$, (3.8) reduces to $X = \tan(Z)$; in the Gaussian case when $\alpha = 2$, (3.8)reduces to $X = W^{1/2}\dfrac{\sin(2Z)}{\cos(Z)} = 2W^{1/2}\sin(Z)$, which is the Box-Muller method of generating a $N(0, 2)$ random variable (Box and Muller, 1958). In this study the standard symmetric stable random variables are considered, so the random variables are sampled using formula (3.8).

A $S_\alpha(\gamma, \beta, \mu)$ random variable with arbitrary shift and scale parameters can be obtained from a variable $X \sim S_\alpha(1, \beta, 0)$ through the following propertys

$$\gamma X + \mu \sim S_\alpha(\gamma, \beta, \mu), \quad \text{if } \alpha \neq 1$$

$$\gamma X + \frac{2}{\pi}\beta\gamma\ln(\gamma) + \mu \sim S_\alpha(\gamma, \beta, \mu), \quad \text{if } \alpha = 1$$

Chambers, Mallows and Stuck (1976) describe a method of generating α stable random variables for any $0 < \alpha \leq 2$ and $-1 \leq \beta \leq 1$ based on the formulas of type (3.8) and provide a Fortran programme. The program applied in this chapter is encoded in GAUSS by McCulloch (1996)[19].

3.2 Monte Carlo Experiments

Since the computations conducted in the previous chapter are enormously time consuming and the aim of this chapter is to find out the α which starts seriously affecting the results computed in previous chapter under normality, only certain representative sample sizes are

[19] Stable Random Number Generators are encoded in GAUSS by J. H. McCulloch (1996), Ohio State University, Economics department, based on the method of Chambers, Mallows and Stuck (1976).

considered rather than the broad and small-stepped selection as those for calculating critical values. The sample sizes are

$$N \in \{10, 30, 50, 70, 100\} \text{ and } T \in \{50, 100, 200, 300, 400, 500\}.$$

Hence, there are altogether $5 \times 5 = 25$ experiments for each α on each DGP. Each experiment consists of $n = 10000$ replications. For each experiment the 0.01, 0.05 and 0.10 percentiles are recorded.

The choices of α 's, however, are different for each DGP. This is because the DGPs respond to the change of α very differently according to some preliminary trial experiments. For example, when $\alpha = 1.7$, it has almost no effect on DGP1, whereas this is disastrous for DGP 2 and 3. Due to the enormous amount of computation time, DGP2 is chosen as an example for the response surface regression analysis on test bias with respect to α . Therefore, DGP2 is computed with a wider range of α 's. The α 's for each DGP are

DGP1: $\alpha \in \{2, 1.9, 1.8, 1.7, 1.6, 1.5, 1.4, 1.3, 1.2, 1\}$ [20]

DGP2: $\alpha \in \{2, 1.95, 1.9, 1.85, 1.8, 1.7, 1.5, 1.2, 1, 0.7, 0.5\}$

DGP3: $\alpha \in \{2, 1.95, 1.9, 1.85, 1.8\}$

The Monte Carlo simulations were programmed in Gauss 7.0.

4. Response Surface Estimation

To illustrate the relationship between the magnitude of test bias caused by non-normality and α (which controls the degree of non-normality), response surface regressions are esti-

[20] The reason for keeping $\alpha = 2$, *i.e.* the distribution is normal, is to observe the difference of normal random number generators between Gauss and McCulloch (1996) (when $\alpha = 2$). This will be discussed in great detail in the following section.

mated. It helps observe the trend of test bias as the distribution of error terms in the model deviates further and further away from normal distribution.

The response surface regression takes a novel form as following

$$\begin{aligned}\Delta q_i^\alpha &= \delta_1(2-\alpha)+\delta_2(2-\alpha)^2+\delta_3\frac{(2-\alpha)}{T}+\delta_4\frac{(2-\alpha)^2}{T^2}+\delta_5\frac{(2-\alpha)}{N}\\ &+\delta_6\frac{(2-\alpha)^2}{N^2}+\delta_7(2-\alpha)\frac{N\ln(T)}{\sqrt{T}}+\varepsilon_i\end{aligned} \tag{3.9}$$

$$\Delta q_i^\alpha = \tilde{q}_i^\alpha - \tilde{q}_i \tag{3.10}$$

where $\tilde{q}_i^\alpha$ denotes the percentile obtained from the i^{th} experiment with sample size N, T (so i is the note for sample size) and the index of stability α ; $\tilde{q}_i$ is the corresponding percentile of the numerical distribution under normality (computed in Chapter 2); so the dependent variable or response variable Δq_i^α represents the test bias incurred by non-normality. The expression $(2-\alpha)$ represents the degree of non-normality, by how much the distribution is away from Gaussian distribution. When $\alpha = 2$, the distribution is normal and test bias does not exist. As the value of α decreases from 2, the degree of non-normality starts growing and is embodied in the increasing value of $(2-\alpha)$. The expression $\frac{N\ln(T)}{\sqrt{T}}$ involved in certain variable is according to the Im and Pesaran (2003) critique on the asymptotic distribution of the CH and CS tests. Detailed discussion of this issue is provided in Chapter 1. Moreover, the results from numerical examination in Chapter 2 are also consistent with the critique.

The functional form is determined by the goodness-of-fit of the regression and the significance of variable coefficients after a number of experimentations. Ericsson (1986) notes that the functional form of response surface regression can be justified on the grounds of the significant coefficients and the generally high R^2 value. A general form to start with can be

$$\Delta q_i^{\alpha} = \delta_1(2-\alpha) + \delta_2(2-\alpha)^2 + \delta_3(2-\alpha)^3 + \delta_4\frac{(2-\alpha)}{T} + \delta_5\frac{(2-\alpha)^2}{T^2} + \delta_6\frac{(2-\alpha)}{N}$$
$$+ \delta_7\frac{(2-\alpha)^2}{N^2} + \delta_8\frac{(2-\alpha)}{NT} + \delta_9\frac{(2-\alpha)^2}{NT} + \delta_{10}(2-\alpha)\frac{N\ln(T)}{\sqrt{T}} + \delta_{11}(2-\alpha)^2\frac{N\ln(T)}{\sqrt{T}} + \varepsilon_i \tag{3.11}$$

For generalization a uniform specification is opted for the 1%, 5% and 10% significant levels rather than optimize the functional form for each one. The trial estimations show that the coefficients of $(2-\alpha)^2\frac{N\ln(T)}{\sqrt{T}}$, $\delta_3(2-\alpha)^3$ and variables involving $\frac{1}{NT}$ are generally insignificant and do not contribute to value of R^2, so they are excluded from the regression.

The response surface regression (3.9) is estimated by ordinary OLS. However, due to the deterministic nature of the values of N, T and α, the errors in (3.9) are heteroskedastic. Therefore the variance of the errors depends systematically on sample size (N and T) and α. To deal with the heteroskedasticity, the covariance estimator of the coefficients' estimators in a regression developed by MacKinnon and White (1985) is applied. The procedure is described in Chapter 2, section 5.2.

5. The Decision Rule of Locating the Benchmark α and the Associated Problems

Recall the critical value intervals computed in Chapter 2. The 95% confidence interval of a critical value is given as

$$\tilde{q}_{ul}^{\theta} = \tilde{q}_i^{\theta} \pm 1.96 \times se(\theta) \tag{3.12}$$

where $\tilde{q}_i^{\theta}$ is the point critical value of θ percentile computed from the i^{th} Monte Carlo simulation with sample size N and T (so i is the note for sample size); $se(\theta)$ is David-Johnson (1954) estimate of the standard deviation of the θ percentile distribution. The 1%, 5% and 10% percentiles computed from the Monte Carlo simulations in this chapter are checked with the corresponding critical value intervals. As α becomes smaller, the aim is to find the α which starts making the percentile fall outside of the corresponding critical value interval computed under normality, *i.e.* the percentile is either smaller than the lower limit lq^{α} or larger than the upper limit uq^{α} of the interval. So the critical value interval can no longer tolerate the percentile due to the non-normality influence on the tests.

However, some preliminary results suggest there is a serious problem associated with sampling distribution. It shows that even when α is set as large as 2 (*i.e.* the errors terms in the DGPs drawn from McCulloch (1996) are Gaussian), there are still a remarkable amount of percentiles falling out of the critical value intervals. This reveals that the impression incurred by sampling distribution has not been sufficiently accounted for by David-Johnson percentile standard deviation estimates. For the statistics analyzed in this study, even different sets of random numbers generated by the same random number generator can still cause serious difference.

To illustrate this, some trials are conducted on DGP1, 2 and 3 for sample size $N = 10$, $T = 200$ using the Gauss normal random number generator.10,000 replications are used in each Monte Carlo experiment and 100 experiments are conducted on each DGP. The 1%, 5% and 10% percentiles are computed, so there are overall $100 \times 3 = 300$ percentiles for each DGP. The percentiles are checked with the critical values intervals whose mid-point estimates were computed under the same assumptions and procedures (by the Monte Carlo simulations in Chapter 2). Results show that the violation of critical value intervals is so severe that all the percentiles from the experiments based on DGP2 and 3 fall out; those for DGP1 experience about 200 violations. This also indicates that the David-Johnson estimate of percentile standard deviation has not captured enough randomness/precision caused by Monte Carlo simulations. Meanwhile, the problem is also case sensitive.

Therefore the critical value intervals are required to take into account more randomness. A proxy of the extra randomness is considered, the empirical difference in the standard deviations of percentile distributions (denoted as $se(d_\theta)$). Since it is suggested by the trail experiments that the problem is also case sensitive, proxies are computed for each DGP. 100 experiments are carried out on each DGP with error terms generated by Gauss normal random number generator and another 100 experiments with the random numbers generated by McCulloch (1996) stable random number generator setting $\alpha = 2$. Again due to the lengthy computation time and the fact that it is the relative difference that is of interest, the sample size is fixed at $N = 10$, $T = 200$. The three percentiles of 1%, 5% and 10% are taken, so there are 100 observations for each percentile. Standard deviations are calculated for each set of the 100 observations. Denote the standard deviations obtained based on Gauss random number generator as $se(G_\theta)$ and those based on McCulloch (1996) as $se(M_\theta)$, the proxy $se(d_\theta)$ is thus calculated as

$$se(d_\theta) = \left| se(G_\theta) - se(M_\theta) \right| \tag{3.13}$$

So the modified critical value interval is given as

$$Mul\tilde{q}_i^\theta = \tilde{q}_i^\theta \pm 1.96 \times \left[se(\theta) + se(d_\theta) \right] \tag{3.14}$$

Again the Monte Carlo simulation outputs from sections 3.2 (the 1%, 5% and 10% percentiles) are checked with the modified critical value interval $Mul\tilde{q}_i^{\alpha}$.

6. Results

Following section 3.1.1, for simplicity, in the result section 'DGP1' refers to the CH test applied on DGP1; 'DGP2' and 'DGP3' refer to the CS test applied on DGP2 and 3, respectively.

6.1 Estimation Results of Response Surface Regressions (DGP2 chosen as example)

Since the pattern of response surface estimation results of the three DGPs are similar, only the results from DGP2 are reported. Table 3.1 shows the regression outcomes from DGP2. The table contains the coefficients of variables in the regression and the value of R^2 for each estimation. Due to the important role of the condition $\frac{N\ln(T)}{\sqrt{T}}$ in the asymptotic properties of the test, only the standard errors of $\hat{\delta}_7$ (coefficient of $(2-\alpha)\frac{N\ln(T)}{\sqrt{T}}$) is provided.

Although the values of R^2 for all the three percentiles are not relatively very high, around 0.5 (which indicates that about half of the variation in the test bias can be explained by the variables involving α), the majority of the coefficients are significant at 1% significance level. So as α decreases, it does systematically cause test bias. The standard errors of $\hat{\delta}_7$ for all the three percentiles are remarkably small, which suggests the estimation accuracy of $\hat{\delta}_7$ and again implies the importance of $\frac{N\ln(T)}{\sqrt{T}}$ in the test performance in addition to the findings in Chapter 2.

Table 3.1 Results of response surface regression estimation of DGP2

Percentile	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{\delta}_4$	R2
0.01	3.598**	2.819**	2.819**	67.820**	0.474
0.05	2.107**	1.633**	1.633**	41.531**	0.503
0.10	1.434**	1.090**	1.090**	24.572*	0.491

	$\hat{\delta}_5$	$\hat{\delta}_6$	$\hat{\delta}_7$	$se(\hat{\delta}_7)$
0.01	-12.283**	119.523**	-0.075**	(0.0150)
0.05	-6.578**	65.636**	-0.043**	(0.0086)
0.10	-4.573*	45.193**	-0.026**	(0.0067)

Note: 1. '**' denotes significance at 1% level; '*' denotes significance at 5% level;
2. $se(\hat{\delta}_7)$ denotes the standard error of $\hat{\delta}_7$

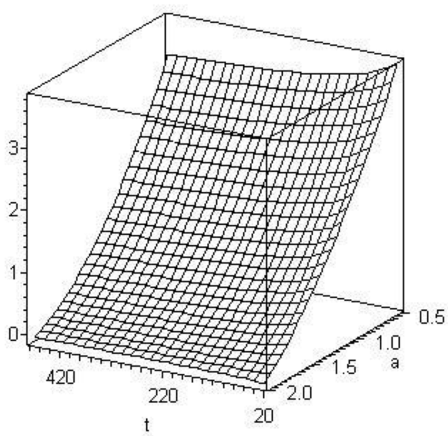
6.2 Test Bias Incurred by Non-normality (DGP2 chosen as example)

Since the pattern of test bias for the three percentiles, 1%, 5% and 10% are similar, Figure 3.2 (a)-(i) only plot the estimates of bias for 5% percentile as an example. The bias is smoothed by response surface regressions and plotted with axes of α and T at different values of N. Figure 3.3 (a)-(d) are the smoothed estimates of bias for 5% percentile with axes of α and N at different values of T. The values used to plot Figure 3.2 and 3.3, $\Delta\hat{q}_i^{\theta}$, are computed through response surface regression estimates obtained from (3.9) and (3.10).

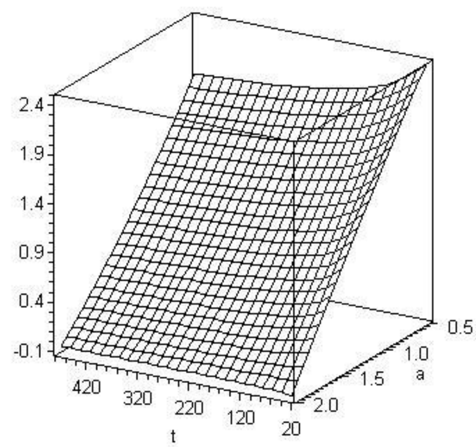
Figure 3.2 clearly shows that when N is small, the decrease in α leads to different patterns of bias at each N. Generally, when $N = \{5,6,7\}$, the bias increases straight away from zero as α reduces, while starting form $N = 8$, the bias seems to reach a maximum where $\alpha \in (1,\ 1.5)$ and then drops. The magnitude of bias later grows again at very small α with negative sign. It is also observed that in the areas where T is small, as sample size T rises, bias is noticeably brought down.

Figure 3.3 presents a similar pattern of bias to those in Figure 3.2 as α varies, *i.e.* a maximum magnitude of bias is reached where $\alpha \in (1,\ 1.5)$, whereas the growth in N does not essentially make any influence. The increase in T reduces the bias on the positive side but extends it on the negative side as shown in plots (a)-(d).

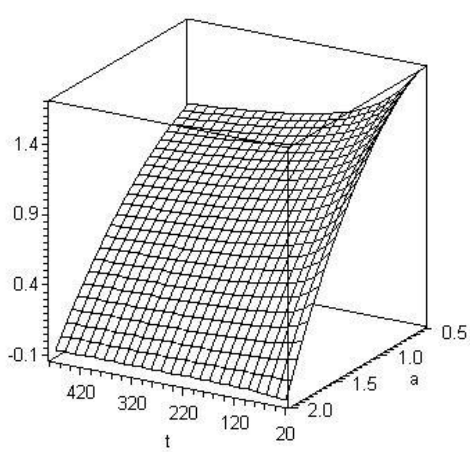
Figure 3.2 Estimated test bias for 5% percentile with α and T at different N, **DGP2**



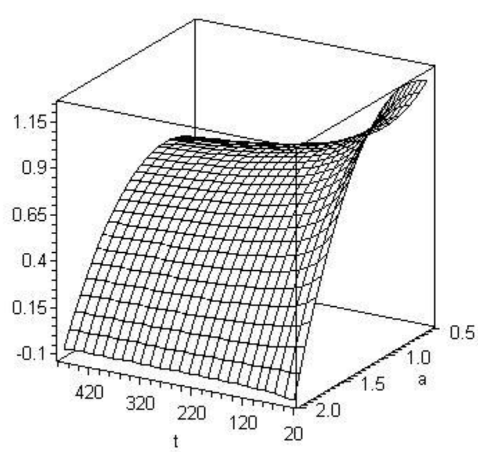
(a) *N* = 5



(b) *N* = 6

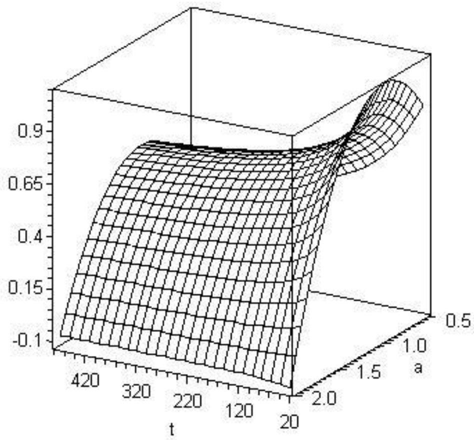


(c) *N* = 7

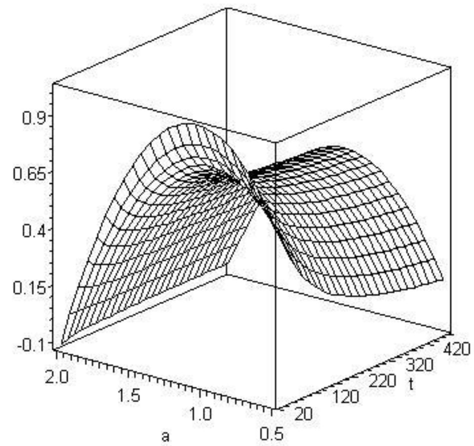


(d) *N* = 8

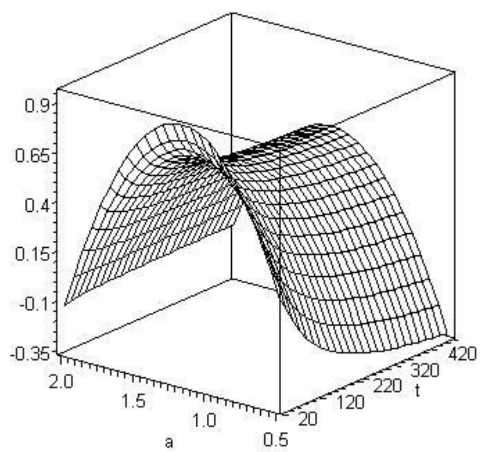
Figure 3.2 Estimated test bias for 5% percentile with α and T at different N, **DGP2 (Cont'd)**



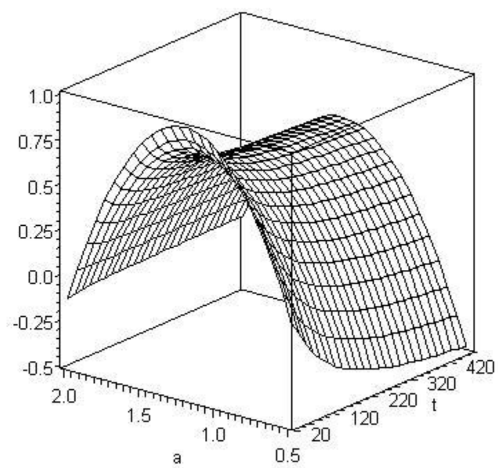
(e) N **= 9**



(f) N **= 10**

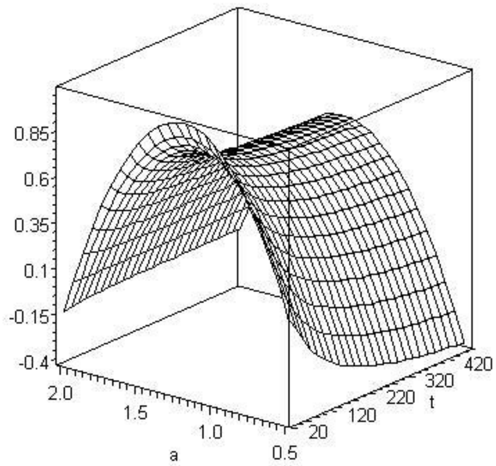


(g) N **= 15**



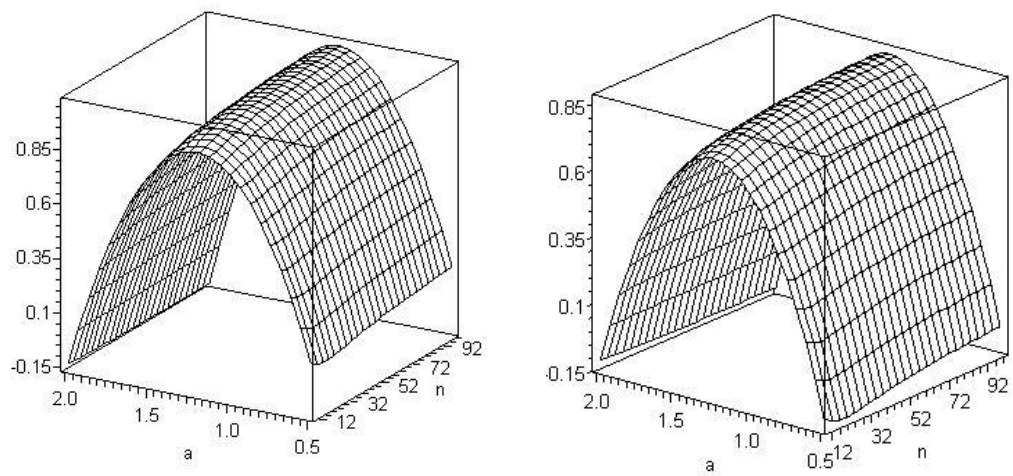
(h) N **= 30**

Figure 3.2 Estimated test bias for 5% percentile with α and T at different N, **DGP2 (Cont'd)**



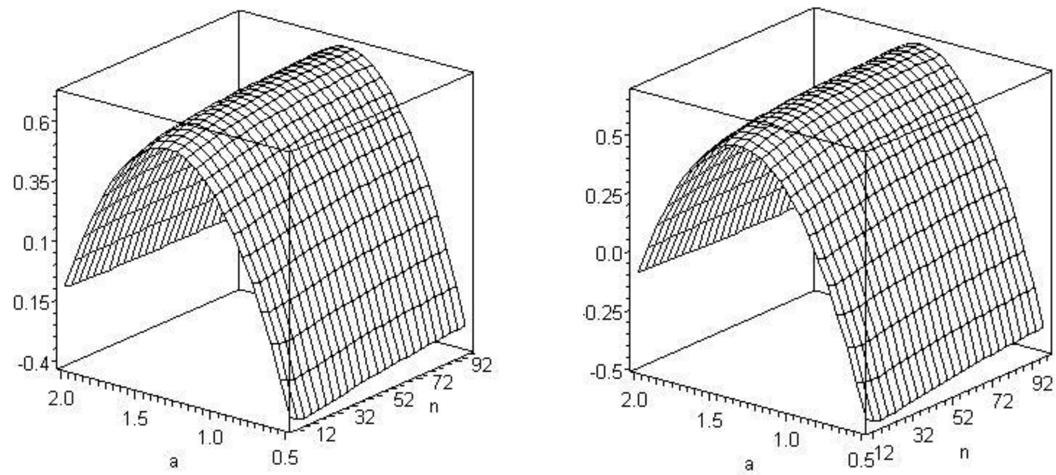
(i) N = 100

Figure 3.3 Estimated test bias for 5% percentile with α and N at different T, **DGP2**



(a) T = 50 **(b) T = 100**

Figure 3.3 Estimated test bias for 5% percentile with α and N at different T, **DGP2**
(Cont'd)



(c) *T* = 300 **(d) *T* = 500**

6.3 The Benchmark α

Recall the Monte Carlo experiments in section 3.2 and refer to the results in Table 3.2, 3.3 and 3.4. The results reveal distinctive features for each DGP. There were 25 experiment for each DGP under a particular α, so there are 25 observations for each percentile and 3 percentile are considered in each experiments. Then under each α, $25 \times 3 = 75$ observations or points appear in the tables (based on one DGP). When one observation falls outside of the corresponding critical value interval computed under normality, it is called an ***outlier*** and will appear in the table as "**1**"; otherwise, it is an ***insider*** and is shown as "0". Clearly the randomness/imprecision caused by sampling distribution is still not completely removed. This can be observed from the columns under $\alpha = 2$, where the three DGPs still have certain numbers of outliers. DGP1 performs the best among the three with only a few statistics land outside of critical value intervals when $\alpha = 2$; whereas DGP2 has a similar performance only for panels with larger number of individuals ($N \geq 50$); in the case of

DGP3 the numbers of outliers and insiders compete with each other under all N. However, with regards to *T*, not a particular characteristic is captured from the tables under $\alpha = 2$.

The problem of randomness unavoidably affects the observation on the trend in the results as α decreases. One straightforward approximate solution can be: calculate the number of outliers in the column under $\alpha = 2$; subtract this number from the number of outliers in other columns under $\alpha < 2$. The results are printed in the last row of the tables. This gives a clear view of the increase in the number of outliers as α becomes smaller. DGP1, which represents the CH test under weak dependence, seems to be most tolerable with stable error terms. It almost does not affect DGP1 even when α is as small as 1.7. More points of violations start appearing as α decreases below 1.7, while they grow in a very moderate rate. DGP2, which denotes the CS test under strong dependence, is far less generous to stable innovations than DGP1. The results remain unaffected only when $\alpha = 1.95$, and then are followed by sharp growth of outliers. DGP3, which represents the CS test under long run dependence, is seriously influenced by randomness and is not tolerable to non-normality at all. Even when $\alpha = 1.95$, there are still 10 outliers. Whereas comparing the growth rate in the violation points, DGP3 seems to be milder than DGP2.

The ultimate aim of this study is to locate the α which invalidates the results computed under normality assumption, *i.e*. the appearance of outliers in the above discussion. In the strict sense, $\alpha = 1.7$, $\alpha = 1.95$ are the benchmarks for DGP 1 and 2, respectively, since outliers start emerging afterwards. Whereas for DGP3 the previous results would be invalid once α is below 2. Nevertheless, if some flexibility is allowed, a few outliers can be permitted. For example, if 10 outliers are permitted, then $\alpha = 1.4$, $\alpha = 1.9$ and $\alpha = 1.95$ are the benchmarks for DGP 1, 2 and 3, respectively. However, according to the tables, permitted violations with more than 10 points are not recommended.

(Table 3.2-3.4 are printed from the next page)

Table 3.2 Results of percentiles computed under stable disturbances checked with the corresponding critical value intervals computed under normality, **DGP1**

N	**T**	**alpha=2**			**alpha-1.9**			**alpah=1.8**			**alpha=1.7**			**alpha=1.6**		
		1%	**5%**	**10%**	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**
10	**50**	0	0	0	0	0	0	0	0	0	0	**1**	**1**	**1**	**1**	**1**
	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	200	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	**50**	**1**	0	0	**1**	0	0	**1**	0	0	**1**	0	0	**1**	0	0
	100	**1**	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0
	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	500	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0
50	**50**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	200	0	0	**1**	0	0	**1**	0	0	**1**	0	0	**1**	**1**	0	**1**
	300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
70	**50**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0	**1**	0	0	**1**	0	0	**1**	0
	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	**50**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0
	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
no. of outliers				**5**			**2**			**4**			**5**			**9**
no. of outliers after deduct the no. of outliers under alpha=2							**-3**			**-1**			**0**			**4**

Note: '**1**' denotes the percentile computed under stable disturbances falling out of the corresponding critical value interval computed under normality; '0' denotes the opposite.

Table 3.2 Results of percentiles computed under stable disturbances checked with the corresponding critical value intervals computed under normality, **DGP1 (Cont'd)**

N	T	alpah=1.5			alpha=1.4			alpha=1.3			alpha=1.2			alpha=1		
		1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
10	50	0	1	1	0	1	1	0	1	1	0	1	0	0	0	0
	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	300	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	50	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	200	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
	300	0	1	0	1	1	0	1	1	1	1	1	1	1	1	0
	500	0	0	1	1	0	1	1	0	1	1	0	1	1	1	0
50	50	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	100	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	200	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
	300	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
70	50	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	100	0	1	0	0	1	0	1	1	0	1	1	0	1	1	0
	200	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	300	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
	500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	50	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1
	100	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1
	200	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0
	300	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
	500	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0
no. of outliers				**13**			**16**			**22**			**25**			**31**
no. of outliers after deduct the no. of outliers under alpha=2				**4**			**8**			**11**			**17**			**20**

Note: '**1**' denotes the percentile computed under stable disturbances falling out of the corresponding critical value interval computed under normality; '0' denotes the opposite.

Table 3.3 Results of percentiles computed under stable disturbances checked with the corresponding critical value intervals computed under normality, **DGP2**

N	T	alpha=2			alpha=1.95			alpha=1.9			alpha=1.85			alpha=1.8		
		1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
10	50	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	100	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	200	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1
	300	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
	500	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1
30	50	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
	100	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0
	200	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	300	1	0	0	0	0	0	0	0	0	1	0	0	1	1	1
	500	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
50	50	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1
	100	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1
	200	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1
	300	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	500	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
70	50	0	0	0	0	1	0	0	1	1	0	1	1	0	1	1
	100	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	200	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	300	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0
	500	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
100	50	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	100	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
	200	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1
	300	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	500	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
no. of outliers				24			19			34			49			60
no. of outliers after deduct the no. of outlier under alpha=2							-5			10			25			36

Note: '**1**' denotes the percentile computed under stable disturbances falling out of the corresponding critical value interval computed under normality; '0' denotes the opposite.

Table 3.4 Results of percentiles computed under stable disturbances checked with the corresponding critical value intervals computed under normality, **DGP3**

N	T	alpha=2			alpha=1.95			alpha=1.9			alpha=1.85			alpha=1.8		
		1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
10	50	1	0	0	0	0	1	0	1	1	1	1	1	1	1	1
	100	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
	200	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1
	300	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1
	500	1	1	1	1	1	1	0	0	1	1	0	1	1	1	0
30	50	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1
	100	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
	200	1	1	1	1	1	1	0	1	1	0	0	1	1	0	0
	300	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1
	500	1	0	0	0	1	1	0	0	1	1	1	1	1	1	1
50	50	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0
	100	0	0	0	0	1	0	1	1	0	1	1	1	1	1	1
	200	1	1	1	0	0	1	0	0	1	1	0	0	1	1	0
	300	1	1	0	1	0	0	1	0	0	1	1	1	1	1	1
	500	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1
70	50	0	0	0	0	1	1	0	1	1	0	1	0	0	1	0
	100	0	0	0	0	1	0	0	1	0	1	1	0	1	1	0
	200	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1
	300	1	0	0	0	0	1	0	1	1	1	1	1	1	1	1
	500	1	1	1	0	1	1	0	1	1	1	0	1	1	0	0
100	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	100	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	200	1	1	1	1	1	1	0	1	1	0	0	1	0	0	1
	300	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1
	500	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1
no. of outliers				**34**			**44**			**51**			**56**			**58**
no. of outliers after deduct the no. of outliers when alpha=2							**10**			**17**			**22**			**24**

Note: '**1**' denotes the percentile computed under stable disturbances falling out of the corresponding critical value interval computed under normality; '0' denotes the opposite.

7. Conclusion

This chapter analyzed the impact of non-normality to panel unit tests through Monte Carlo simulation analysis. Empirical data frequently bear the non-normal distribution characteristic, whereas unit root tests are usually developed on the normality assumption. In the field of panel unit tests this problem has not yet been considered. Empirical test results can be seriously misleading if non-normality is simply ignored.

The chapter uses the Lévy-Paretian Stable distribution to model leptokurtosis or heavy tails that are often found in empirical data. The aim is to examine the sensitivity of panel unit root tests to non-normality modeled by stable distribution. The errors in the two panel unit tests, the CH and CS tests, are assumed to follow the α stable distribution. Using the outputs of a series of Monte Carlo experiments, response surface regressions with a novel design of functional form are estimated and help reveal the relationship between test bias and the degree of non-normality. Results suggest that non-normality indeed causes serious test bias and attacks the tests differently in different situations. As the index of stability α decreases and drives the distribution further and further away from Gaussian distribution, the magnitude of test bias becomes more and more severe; whereas as sample size N and T increases respectively, the trend of bias shows varying patterns.

Let DGP1, DGP2 and DGP3 denote the CH test applied on panel data with weak dependence, the CS test applied on data with strong and long run dependence, respectively. In terms of the decision as to which α starts to severely influence the test and invalids the critical value interval computed under normality, results reveal that each DGP has different reaction. If one chooses to be restrictive, the benchmarks are determined as $\alpha = 1.7$, $\alpha = 1.95$ for DGP 1 and 2, respectively; whereas DGP3 does not tolerate with any α less than 2. A more generous choice can be set as $\alpha = 1.4$, $\alpha = 1.9$ and $\alpha = 1.95$ for the three DGPs, respectively, which allows for 10 outliers out of 75 points (*i.e.* for each DGP, among the 75 percentile point estimates computed under certain degree of non-normality, 10 points fall out of the corresponding critical value interval computed under normality). However, permission of allowing for more than 10 outliers is not recommended.

The results provide important indications for applications that employ non-normally distributed data. When the degree of non-normality in data is higher than the benchmarks, the reliability of testing results are questioned. A limitation in this study is that due to the remarkably lengthy computation time, the chapter is unable to provide critical values for the tests under non-normality with a comprehensive range of α's and sample sizes.

Although much has been done in section 5 to eliminate the problem of randomness/imprecision resulting from sampling distribution, it still partially remains in the results, which indicates that the David-Johnson estimate of percentile standard deviation has not captured sufficient amount of randomness. The problem is also case sensitive in this study. DGP1 is affected at the least level; DGP2 copes with it well only with larger *N*; and DGP3 is seriously influenced with all sample sizes. Further analysis on this issue can be carried out in a general setting and search for more solutions.

Chapter 4 Is Inflation Stationary? Evidence from Panel Unit Root Tests with Structural Breaks

1. Introduction

In the area of macroeconomic research, the dynamic properties of macroeconomic series have been one of the central issues for decades. Being regarded as a key macroeconomic variable, inflation is one of the most analysed subjects. Since the seminar work of Nelson and Plosser (1982) the stationarity characteristic of inflation has received substantial amount of attention. The intuitive explanation of a stationary series (denoted as $I(0)$) is that it shows temporary memory and the impact of a shock will eventually disappear. On the other hand, a non-stationary series (containing a unit root, denoted as $I(1)$[21]) has permanent memory, so a shock has permanent effect on it and drives the series away from its mean.

Whether inflation is stationary or not leads to important implications in terms of economic theory and policy. For example, according to Fisher (1930) inflation and nominal interest rate are linked through real interest rate, and real interest rate plays a central role in saving and investment decisions in the economy. Given the fact that nominal interest rate is usually treated as non-stationary, inflation rate also needs to be non-stationary and thus cointegrates with nominal interest rate in order for real interest

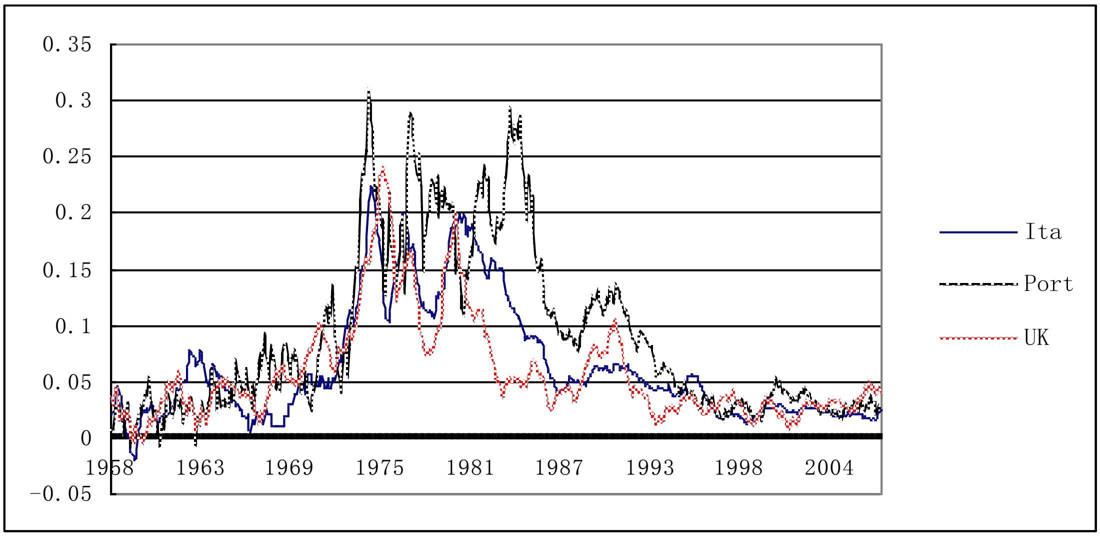
[21] For simplification it is assumed that nonstationary series are integrated of order 1, *i.e.* they can be transformed stationary by differencing it once (or first differencing), so here uses $I(0)$ to denote stationary series and $I(1)$ to denote non-stationary series.

to be stationary. In addition, from the view of acceleration hypothesis, low unemployment rate below its natural rate would be at the high cost of an ever-increasing level of inflation (*i.e.* inflation contains a unit root) in the long run. Regarding monetary policies, some researchers advocate that unless inflation and real income growth are cointegrated, if inflation is non-stationary, so is the nominal income growth. Therefore, monetary policy should target either the level or growth rate of nominal income (c.f. Hall, 1984; Taylor, 1985). More recently, inflation targeting has grown popular in monetary policies. The dynamic nature of inflation is then the essential prime knowledge since serious potential danger will be posed if the target is set on an explosive inflation process.

A great amount of work has been done to examine the stationarity of country inflation series whereas results are very mixed. Earlier research applying traditional single time series tests on inflation rates cannot reach a consensus but with preference for unit root. Baillie (1989), Ball and Cecchetti (1990), Johansen (1992), Nelson and Schwert (1997), etc. fail to reject the unit root hypothesis, whereas Rose (1988), Baillie et al. (1996), Choi (1994), etc. find inflation stationary. Following the insights of Perron (1989), the caution has been raised that the presence of breaks in the trend function of time series (often called structural change or structural break, e.g. a shift in the mean of a process) tend to cause unit root tests to bias towards not rejecting the unit root null hypothesis. This means that unit root test may well suggest a process is $I(1)$, whereas in fact it is $I(0)$ subject to structural changes. In addition, research indicates that inflation rates over long time span are prone to structural breaks due to monetary policy changes and other macroeconomic shocks like the oil crises (c.f. Garcia and Perron (1996); Rapach and Wohar (2005); Romero-Ávila and Usabiaga, 2008). Figure 4.1 below plots the monthly inflation rate series of Italy, Portugal and UK from 1998 to 2007. It illustrates the processes experience sharp increases in mid 70s and fluctuate at high level until mid

80s and then slump. Culver and Papell (1997) apply the sequential break unit root test of Perron and Vogelsang (1992) to monthly inflation series of 13 OECD countries and provide evidence of stationarity in inflation for only 4 countries.

Figure 4.1 Plot of inflation rates series of Italy, Portugal and UK



Due to the criticism of low power of single time series unit root test, the recent literature sees the application of panel data. Panel unit root tests can gain more statistical power essentially by increasing sample size through pooling time series in the cross section dimension. Many studies pool certain number of country inflation series as panel data and apply panel unit root tests. However, these still cannot conclude with the same results. Culver and Papell (1997) also apply the homogenous panel unit root test by Levin and Lin (1992) and reject unit root in inflation. While the rather strong assumptions of homogeneity and cross section independence of the test are questioned[22], Lee and Wu (2001) and Otero (2008) apply heterogeneous tests considering cross section dependence and also find inflation stationary. However, Ho (2008) using the Chang (2002) test discovers that there appears to be a unit root in inflation. Given the evidence

[22] Refer to Chapter 1 for the review of cross section dependence/correlation in panel data.

of structural breaks in inflation rates provided by some studies, Costantini and Lupi (2007), Lee and Chang (2007) and Romero-Ávila and Usabiaga (2008) apply panel unit root tests considering structural breaks and conclude that inflation is $I(0)$.

In the literature, to what extent a jump (the abrupt and substantial change) in a process is considered as a break is not clearly defined. Therefore the estimation of breaks seems rather arbitrary. In addition, modelling and estimating the breaks often involves intricate and tedious procedures, and the number of breaks a test can deal with is usually very limited. Rather than estimate the breaks, in this chapter the breaks are captured by an alternative and new approach of implementing the Lévy-Paretian Stable distribution[23] (also called the α stable distribution) in the panel unit root test models. Refer to Figure 3.1 in Chapter 3 section 2.2, when the index of stability α (the parameter that measures the tail fatness and peakness of the distribution) becomes smaller, the α stable process (*i.e.* the process built on cumulated α stable random variables) shows more and sharper jumps. This feature can be well adopted to simulate structural breaks. Charemza et al. (2005) concern the non-normal empirical distribution of inflation rates and study their dynamic behaviour by the augmented Dickey-Fuller (ADF) test under stable innovations. The results are in favour of non-stationarity. Since the inflation rates of most OECD countries exhibit very similar patterns (which will be illustrated in section 4), it is appealing to pool the inflation series into a panel to take the advantages of panel data and re-examine the issue.

This chapter assumes the error terms in panel unit root test models to follow α stable distribution rather than the restrictive normal distribution. On estimating the value of α from the data, refer to the results in Chapter 3 and conduct some further Monte Carlo simulations for critical values when necessary, the decision can be conveniently

[23] The introduction of α stable distribution can be found in Chapter 3.

made on the dynamic property of the series. The procedure will be elaborated in the following sections.

The framework tests used in this study are the Chang (2002) test (CH hereafter) and the Chang and Song (2005, 2009) test (CS hereafter) due to their robustness to all different forms of cross section dependence in panel data. For comparison, some other popular panel unit root tests are also applied as well as a number of single time series tests used on each individual in the panel. With still very mixed results, the conclusion gives preference to the presence of a unit root which is supported by the most general panel unit root tests and the majority of single time series tests.

The remainder of this chapter is organized as follows. Section 2 introduces the CH and CS tests incorporating structural breaks. Section 3 presents the estimation process of the index of stability α in order to detect the influence level of structural breaks on the tests. Section 4 illustrate the features of inflation rates among OECD countries and explains the poolability of the series. Section 5 presents the empirical results and section 6 concludes.

2. The CH and CS Tests under Non-normality to Account for Structural Breaks

2.1 The CH and CS Tests

Detailed reviews of the two tests are provided in Chapter 1. So here only the framework for application is presented. The CS test is a developed version of CH test with regard to the robustness to cross section dependence. To detect unit root in inflation, the following equation is estimated in CH and CS tests

$$Ifl_{it} = \rho_i Ifl_{i,t-1} + \sum_{k=1}^{p} \delta_{i,k} \Delta Ifl_{i,t-k} + \varepsilon_{it} \qquad i = 1,\ldots,N;\ t = 1,\ldots,T_i \qquad (4.1)$$

where Ifl_{it} denotes inflation panel data variable; i and t denote the countries in the panel and time period, respectively; the time length for each individual T_i are the same in this study, *i.e.* the panel is balanced; ε_{it} are white noise; Δ is difference operator. The nonlinear IV estimation is applied on (4.1). In CH test the instrument is generated by a single nonlinear integrable *instrument generating function* (IGF) $F(Ifl_{i,t-1})$ and in CS test it is a set of orthogonal IGFs $F_i(Ifl_{i,t-1})$. The panel unit root testing statistic is

$$S = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \tau_i$$ [24]

where τ_i is the nonlinear IV *t*-statistic of $\hat{\rho}_i$ for each individual in (4.1). S shows standard normal distribution asymptotically under the null hypothesis that all panel individuals have a unit root. Due to the choice of orthogonal IGFs, the CS test claims to be able to handle all the different forms of cross section dependence. Nevertheless, as shown in the introduction of Chapter 2, in terms of weak dependence, the CH test has better performance; however, it is unable to deal with strong dependence.

2.2. Models with α Stable Innovations to Incorporate Structural Breaks

On estimating (4.1) and if unit root in inflation is not rejected, the distribution of the first difference of inflation needs to be checked for α through the following models as an indicator for the presence of potential structural breaks. The chapter is not in favour of testing for structural breaks, as explained in the introduction of this chapter that so far there is not a clear definition of structural break and the estimation of breaks seems rather arbitrary. So the chapter incorporates the potential breaks as part of a process (the α stable process).

In the similar forms to the DGPs of Monte Carlo simulations on the CH and CS tests in

[24] As noted in Chapter 2, for simplification only the average S statistic is considered for CS test.

Chapter 2 and 3, the representations are

$$CH: \ \Delta Ifl_{it} = \sum_{k=1}^{p_i} \beta_{i,k} \Delta Ifl_{i.t-k} + \eta_{it} \tag{4.2}$$

$$CS1: \ \Delta Ifl_{it} = \sum_{k=1}^{p_i} \beta_{i,k} \Delta Ifl_{i.t-k} + \upsilon_i \xi_t + \eta_{it} \tag{4.3}$$

$$CS2: \ \Delta Ifl_{it} = \sum_{k=1}^{p_i} \beta_{i,k} \Delta Ifl_{i.t-k} + \upsilon_i \xi_t + \Delta \eta_{it} \tag{4.4}$$

where $\beta_{i,k}$ are the AR coefficients; ξ_t is the scalar common stochastic trend or common factor and υ_i are factor loadings; η_{it} are α stable innovations with symmetric and nonsingular covariance matrix. CH shows weak dependence generated from the covariance matrix of η_{it}; CS1 presents the strong and weak forms of dependence through time effect ξ_t and innovations η_{it}; CS2 exhibits long run dependence or cross unit cointegration, driven by the common stochastic trend $\sum \xi_t$. If the CH test is applied, α is estimated through (4.2) with weak dependence assumption; when the CS test is applied, (4.2) and (4.3) are used to estimate α, assuming strong dependence and long run dependence, respectively.

Under the assumption of α stable distribution, η_{it} belong to the domain of attraction of symmetric stable distribution with index of stability $\alpha \in (1, 2]$. When $\alpha = 2$ η_{it} are in the domain of attraction of normal distribution and the equations reduce to those analyzed in the original CH and CS tests. However, as α becomes smaller than 2, according to the findings in Chapter 3 the distribution of testing statistics can be effectively affected by a certain small α, so that the critical values computed under normality assumption cannot be used.

3. Estimation of α

α is estimated on η_{it} or $\Delta\eta_{it}$ in (4.2)-(4.4). According to the stability property of stable distribution introduced in Chapter 3, the α in η_{it} and $\Delta\eta_{it}$ are the same,. So the α estimated from (4.3) is for both CS1 and CS2. Estimate α_i from η_{it} (for $i = 1, 2, ..., N$) and the average of α_i's is taken as the α estimator for the panel[25]. α_i is estimated by the procedure developed in McCulloch (1996). η_{it} are obtained through OLS estimation of (4.2) and (4.3) for CH and CS tests, respectively.

In linear regression model the effect of infinite variance in regressors and error terms can be substantial. In the autoregression

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-1} + ... + \beta_k y_{t-k} + \varepsilon_t$$

the least square estimates are shown to be consistent as long as y_t is stationary (c.f. Burridge and Hristova, 2008; Knight, 1989). In the regression

$$y_i = \beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_n x_{n,i} + \varepsilon_i$$

as long as $\alpha > 1$, the OLS estimate is still consistent yet the convergence rate is slower than the case of error terms with finite variance (c.f. McCulloch, 1998). Knight (1993) notes that theoretical and empirical evidence suggests the robust M-estimates to linear regression with infinite variance error terms are more efficient than the OLS estimates (c.f. Knight (1993) for the M-estimate approach). Nevertheless, the below Monte Carlo experiments results suggest that with increasing sample size, the efficiency of OLS estimates is effectively improved. Given the relatively large sample size of the data used in this chapter (which will be introduced in section 4), the OLS estimates are used.

[25] In Chapter 3 where the influence of α on panel unit root tests is studied, homogeneous α is applied to all panel individuals. However, this is very restrictive in the real world. In addition the model is heterogeneous, so here the average of individual α's is computed rather than estimate α on the data pooled across panel.

The finite sample performance of OLS estimates with infinite variance error terms is examined by the following Monte Carlo experiments. Data are generated by

$$y_t = \beta\, y_{t-1} + \gamma\, \xi_t + \varepsilon_t \qquad t = 1,2,\ldots,T \tag{4.5}$$

where the coefficients β and γ are set as 0.8 and 1.5, respectively; ξ_t are drawn from *iid* ***N***(0,1) and ε_t are random draws from α stable distribution. The sample sizes and values of α considered are

$$T \in (50,100,200,300,500) \text{ and } \alpha \in \{2,1.9,1.8,1.7,1.5,1.4,1.2,1\}.$$

The number of replications is 5000. OLS estimation is applied. The results are displayed in Table 4.1. It shows that lager sample size can help increase the efficiency of estimators. The AR coefficient $\hat{\beta}$ is not obviously affected by the infinite variance in error terms; whereas the performance of $\hat{\gamma}$ deteriorates with growing standard errors as α decreases; in particular, when $\alpha = 1.2$ the estimators $\hat{\gamma}$ start to drift away from γ, and at $\alpha = 1$ the estimators are so biased that they cannot be used to serve application. The experiment results are consistent with the above theories. Some preliminary trial estimations indicate that the α estimated from inflation data is around 1.5, so the least square estimation should produce reasonable results.

Table 4.1 Simulation results of OLS regression with stable error terms

T	α	$\hat{\beta}$	se	$\hat{\gamma}$	se	α	$\hat{\beta}$	se	$\hat{\gamma}$	se
	2					**1.9**				
50		0.79	0.052	1.502	0.146		0.789	0.056	1.502	0.17
100		0.795	0.035	1.499	0.101		0.795	0.038	1.498	0.12
200		0.798	0.024	1.5	0.071		0.797	0.026	1.499	0.085
300		0.799	0.02	1.499	0.058		0.798	0.022	1.501	0.07
500		0.799	0.015	1.5	0.045		0.799	0.017	1.501	0.055
	1.8					**1.7**				
50		0.788	0.06	1.503	0.203		0.788	0.065	1.504	0.25
100		0.794	0.041	1.498	0.146		0.794	0.044	1.497	0.183
200		0.797	0.029	1.499	0.105		0.796	0.031	1.497	0.133
300		0.798	0.024	1.502	0.087		0.798	0.026	1.503	0.112
500		0.799	0.018	1.501	0.069		0.798	0.02	1.502	0.09
	1.5					**1.4**				
50		0.789	0.077	1.51	0.432		0.792	0.087	1.518	0.627
100		0.793	0.05	1.493	0.329		0.793	0.053	1.488	0.486
200		0.795	0.036	1.492	0.247		0.795	0.039	1.486	0.368
300		0.797	0.03	1.511	0.213		0.797	0.032	1.52	0.321
500		0.798	0.023	1.507	0.176		0.798	0.025	1.515	0.273
	1.2					**1**				
50		0.821	0.139	1.594	1.846		0.832	0.525	2.914	12.248
100		0.793	0.059	1.448	1.433		0.794	0.062	1.314	8.488
200		0.795	0.042	1.446	1.105		0.795	0.044	1.262	6.59
300		0.797	0.034	1.592	0.97		0.798	0.035	2.224	5.453
500		0.798	0.026	1.61	0.894		0.798	0.027	3.204	6.126

Note: ‘se’ denotes standard error

Refer to (4.3) and (4.4), the representations indicate that an appropriate proxy for the scalar common stochastic trend ξ_t is required for CS test on estimating α. Let y_{it} denote a panel data variable. Pesaran (2007) panel unit root test suggests capture the strong dependence that arises from common factor by lagged cross sectional mean ($\bar{y}_{t-1}$) and its first difference ($\Delta\bar{y}_t$).[26] In the case of serial correlation, the regression must be augmented by $\sum_{k=1}^{p}\delta_{i,k}\Delta y_{i,t-k}$ as usual and also by $\sum_{j=1}^{p}\gamma_j\Delta\bar{y}_{t-j}$, where p is the lag order. The procedure is applied here to construct a proxy to substitute ξ_t in (4.3) and (4.4).

[26] The Pesaran (2007) panel unit root test is reviewed in Chapter 1.

Since only (4.3) is used for estimation, it is rewritten with the proxy of ξ_t as

$$\Delta Ifl_{it} = \sum_{k=1}^{p} \delta_{i,k} \Delta Ifl_{i,t-k} + d_0 \overline{Ifl}_{t-1} + \sum_{j=0}^{p} \gamma_j \Delta \overline{Ifl}_{t-j} + e_{it} \tag{4.6}$$

Then α can be estimated from the residuals e_{it}.

4. Data

4.1 Data

The original collection of data is monthly Consumer Price Index (CPI) of 15 OECD countries from the International Monetary Fund's *International Financial Statistics Database* (IFS). Data range covers period from 1957m1 to 2007m12. The countries include Austria, Belgium, Canada, Finland, France, Italy, Japan, Luxembourg, Norway, Portugal, Spain, Sweden, Switzerland, UK and US.

The inflation in this study is annual inflations adjusted monthly, also called Y/Y. It is calculated as $(lp_t - lp_{t-12})$, where *lp* is the logarithm of CPI. Therefore, in each sample group the first 12 observations are dropped. The AR order is selected by both AIC and BIC criterion with a maximum of 18 lags. Although the CH and CS tests allow for different lag orders among panel individuals, due to the construction of the proxy for common factor in (4.6), the same lag order has to be applied across panel. The lag lengths of panel individuals selected by both criteria are particularly long, from 12 to 18. This is because there is overlapping of observations, $(lp_t - lp_{t-12})$, which generates excessive autocorrelation. Since the same lag length is required across the panel, to ensure the performance of the estimation, the longest lag order is chosen from the individual estimates and is applied on the panel.

4.2 The Similar Patterns of Inflation Rates across OECD Countries

The inflation rates of OECD countries have shown great similarities. A graphic overview of the country inflation series is provided in Figure 4.2. Figure 4.3 split the series into 5 groups to provide a clearer vision. It illustrates that the fluctuations of inflation rates have been strikingly similar across OECD countries. They have risen progressively in the 1960s and early 1970s and then experienced sharp increases. During part of the 1970s and part of the 1980s inflations were high and volatile and then declined during the 1990s and remained low and stable thereafter. More importantly, in terms of the appearance of breaks in inflation, Corvoisier and Mojon (2005) suggest that since 1960 inflation of every OECD country has two or three breaks in its mean. The remarkable feature is that the breaks largely coincide throughout the OECD countries around 1970, 1982 and, to a lesser extent, around 1992.

4.3 Some Interpretations of the Similarities of Inflation Rates of OECD Countries

The similarities are analysied in several recent empirical studies which have found that inflation, at least in industrialized countries, is largely a global phenomenon. Ciccarelli and Mojon (2005) show that inflations of 22 OECD countries have moved together over the last several decades and have a common factor that alone accounts for nearly 70% of the variability of country specific inflation; national inflation rates tend to be pulled back to Global Inflation by a robust “error correction mechanism”. The co-movement of inflation is particularly obvious during the last two decades, as can be clearly observed in Figure 4.2, series clustering together. This is largely due to the effects of international inflation transmission, globalization and similar monetary policy changes across countries.

Figure 4.2 Plot of original selection of inflation series

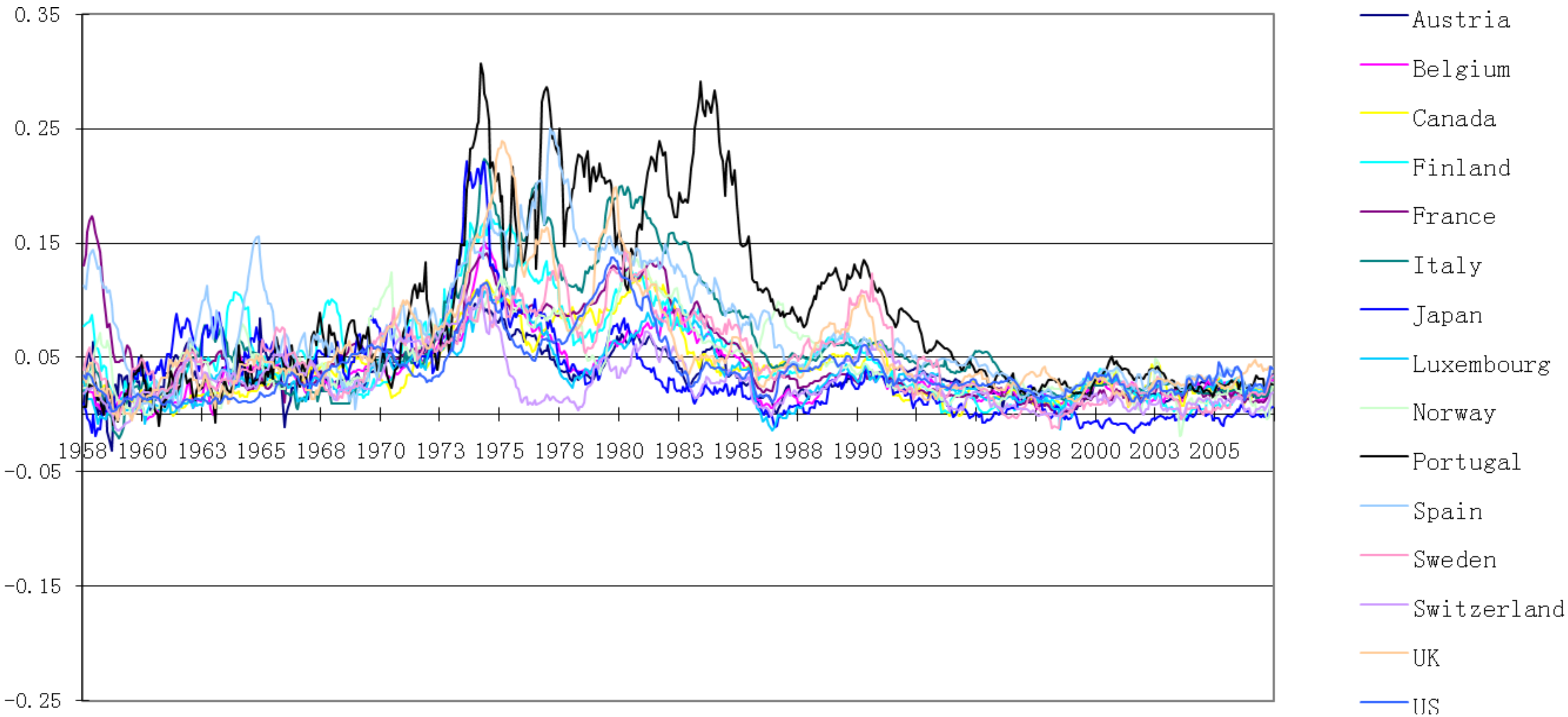
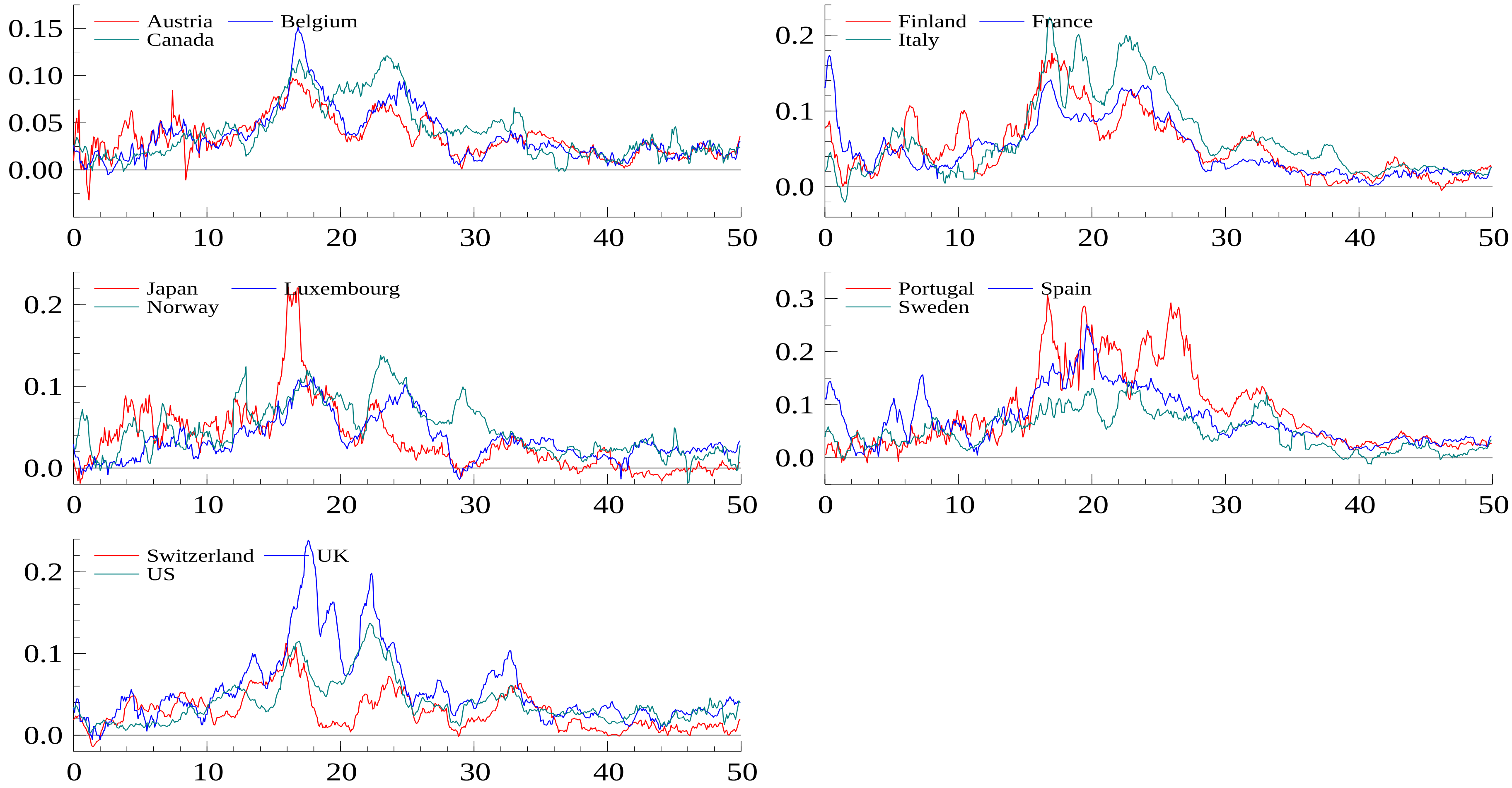


Figure 4.3 Plot of original selection of inflation series, split view



Although country specific inflation dynamics and persistence are basically determined by country specific price rigidities, these are also affected by the inflation dynamics and persistence in other countries as a result of international inflation transmission. The early literature of international inflation transmission focused on how the United States inflation was transmitted abroad under the Bretten Woods system of fixed exchange rates. Evidence had been found in Brunner and Meltzer (1977), Cassese and Lothian (1982) and Darby (1983). After the collapse of Bretten Woods system, however, the degree of exchange rates flexibility have been challenged (c.f. Reinhart and Rogoff, 2003; Calvo and Reinhart, 1998). Nevertheless, Boschen and Weise (2003) found empirical evidence that U.S. inflation policy continued to influence OECD inflations in the flexible exchange rate period.

During the last two decades, the world has experienced a remarkable process of disinflation, which is most observable in the industrialized countries. Globalization, the tremendous increase in economic integration, is argued to have made effective contributions since 1980s. Globalization can affect national inflation through various channels such as international trade, labour, capital market and exchange rate fluctuations. Inflation rate has become much less prone to domestic parameters, especially the domestic output gap; whereas global factors such as the output gap of main trading partners became more important in determining national inflation rate (Pehnelt, 2007). For example, some studies suggest that the Phillips curve in major developed countries has flattened in the last couple of decades due to the fact that the short-term relationship between domestic economics factors such as unemployment rate and domestic output gap has weakened (c.f. Debelle & Wilkinson 2002; Benati 2005), whereas the importance of foreign output gaps has increased (Borio & Filardo, 2006).

Apart from international inflation transmission and globalization, similar monetary policy is another driver that enables inflation rates to move together. Recall in Figure 4.2 the high and volatile inflation experienced by most countries during the 1970s and part of the 1980s is known as the 'Great Inflation'. This phenomenon was criticized by some studies as a result of over-emphasis of Keynesian theories in the 1960s and 1970s (e.g. Rogoff, 2003). Policy makers were driven by Phillips curve and stressed the trade-off between inflation and unemployment (c.f. Taylor, 1992 and Sargent, 1999). Facing the pressure of Great Depression, politicians of the U.S. in the 1960s and early 1970s placed high value on keeping unemployment very low, whereas no tightening monetary policy was applied to tackle the high inflation afterwards (De Long, 1997). The effect was also transmitted abroad as discussed above.

Since the mid 1980s the world has been experiencing the disinflation process along with increased emphasis of monetary policy objective on maintaining low and stable rate of inflation. This is more obvious in the countries that adopted the policy of inflation targeting. Woodford (2003) argues that it is because instability of the general level of prices causes substantial real distortions-- leading to inefficient variation both in aggregate employment and output and in the sectoral composition of economic activity -- that price stability is important. As central bank independence has been strengthened in many countries, they have become more committed to the primary goal of price stability and concerned with inflation risks. Besley (2008) illustrates the similar stance of policy across nine OECD countries which played a role in the moderation of inflation. The high and volatile inflation was negatively associated with real interest rate, the policy rate adjusted for inflation, which can be interpreted as a relaxed monetary policy; in contrast, the most recent period of low and stable inflation is characterised by positive and higher real rates of interest.

Table 4.2 Correlation coefficient matrix of inflation rates of 15 OECD countries

	Austria	Belgium	Canada	Finland	France	Italy	Japan	Luxembourg	Norway	Portugal	Spain	Sweden	Switzerland	UK	US
Austria															
Belgium	0.844														
Canada	0.677	0.801													
Finland	0.773	0.868	0.838												
France	0.677	0.796	0.882	0.783											
Italy	0.731	0.830	0.856	0.841	0.894										
Japan	0.769	0.741	0.597	0.710	0.516	0.632									
Luxembourg	0.808	0.935	0.762	0.829	0.780	0.815	0.660								
Norway	0.638	0.713	0.797	0.759	0.758	0.779	0.526	0.683							
Portugal	0.615	0.760	0.795	0.767	0.821	0.841	0.506	0.710	0.682						
Spain	0.624	0.740	0.789	0.785	0.802	0.863	0.517	0.710	0.730	0.805					
Sweden	0.675	0.719	0.804	0.767	0.756	0.828	0.560	0.722	0.767	0.792	0.770				
Switzerland	0.719	0.656	0.581	0.602	0.487	0.521	0.691	0.636	0.511	0.452	0.381	0.601			
UK	0.741	0.796	0.808	0.843	0.774	0.804	0.707	0.761	0.703	0.694	0.747	0.766	0.542		
US	0.644	0.696	0.872	0.721	0.848	0.787	0.616	0.653	0.662	0.691	0.677	0.730	0.542	0.825	

5. Empirical Results

5.1 Tests Applied

Although the variance covariance matrix does not exist in the α stable world, yet to give a clue of the degree of dependence among the panel individuals, the covariance matrix is computed through the data. Table 4.2 reports the covariance matrix of inflations rates in the 15 OECD countries and its grand average is 0.722 with standard deviation 0.108. So there is clear evidence of relatively strong dependence among the countries. In addition, the possibility of long run dependence (cross unit cointegration) cannot be simply excluded due to the similar behaviour of the series over time as discussed in section 4. Given the strong level of dependence, in addition to CH and CS tests, some other popular second generation tests, the Bai & Ng test (BN), Moon & Perron test (MP) and Pesaran test, are applied for comparison. Moreover, the results of some popular first generation tests, the Levin, Lin & Chu test (LLC), Im, Pesaran & Shin test (IPS), ADF Fisher test, PP Fisher test and Hardri stationarity test, are also presented. Besides, a number of single time series unit root tests are applied on each individual series in the panel. These tests include the augmented Dickey-Fuller test (ADF), Phillips-Perron test (PP), modified Phillip-Perron test (MPP), modified Sargan-Bhargava test (MSB), modified Elliott, Rothemberg and Stock feasible point optimal test (MERS), KPSS stationarity test, augmented Perron Additive Outlier (Crash) test (PAO) and augmented Perron Innovative Outlier test (PIO). The latter two tests consider structural breaks in the process.

5.2 Results under Normality Assumption

Panel unit root tests results are given in Table 4.3. As for CH and CS tests, the finite sample critical value intervals computed in Chapter 2 are applied. The CS test is confidently in favour of a unit root at 5% significant level with lag orders chosen either by AIC or BIC criteria. The CH test rejects unit root when the lag orders are chosen by AIC criteria and is in favour of unit root when BIC criteria is used to select lag orders. Among other second generation tests, the BN test separates the common factor(s) and error terms in the model and tests for unit root in the two processes respectively. It fails to reject unit root in either processes. The MP and Pesaran tests support stationarity in inflation. Nevertheless, recall that the technique used in MP test to deal with cross section dependence is to de-factor the panel. If unit root is only caused by common factor, *i.e.* the case of long run dependence or cross unit cointegration, the results of MP test can be misleading (refer to the review in Chapter 1, section 6.2.2.2).

Among first generation tests, the LLC and Hadri tests strongly advocate the unit root hypothesis; whereas there are more tests clearly in favour of stationarity. However, given the relatively high degree of dependence among the individual inflation series in the panel and the fact that the first generation tests are unable to cope with strong dependence, these outcomes might be unconvincing.

Finally refer to Table 4.4, the results of single time series unit root tests applied on each individual in the panel. There is sheer agreement on the presence of a unit root in the majority of the series. For each country, either structural breaks are considered or not, the conclusion remains the same. Austria, Belgium, Canada, Finland, France, Italy, Norway, Portugal, Spain, Sweden and the UK all give strong evidence of a unit root. Stationary inflation is suggested for Luxembourg, Switzerland and the US by more than half of the tests applied. However, take an overall point of view on the strong agreement

Table 4.3 Results of panel unit root tests

First Generation Tests

		***t*-stat**
	Levin, Lin and Chu (LLC)	2.1746
	Im, Pesaran and Shin (IPS)	-1.9608**
ADF Fisher	ADF - Fisher Chi-square	35.9667
ADF Fisher	ADF - Choi Z-stat	-1.7391**
PP Fisher	PP - Fisher Chi-square	61.1417**
PP Fisher	PP - Choi Z-stat	-4.0946**
Hardri	Hadri Z-stat	16.8795
Hardri	Heteroscedastic Consistent Z-stat	17.0575

Second Generation Tests

	***t*-stat (BIC)**	$\hat{\alpha}$	***t*-stat (AIC)**	$\hat{\alpha}$
Chang (CH)	-1.307	**1.531**	-2.647**	**1.550**
Chang and Song (CS)	-0.455	**1.546**	-0.401	**1.558**
Pesaran	**CADF** (max lag 12)	**CADF** (max lag 18)		
	-3.460**	-3.735**		
	***t*-stat (pool)**	**factor**	***t*-stat (factor)**	
Bai and Ng (BN)	85.902	1	-2.31	
	***t*-stat a**		***t*-stat b**	
Moon and Perron (MP)	-28.5558**		-8.1323**	

Note: '**' denotes significant at 5% significance level.

Table 4.4 Results of single time series tests

	Austria	Belgium	Canada	Finland	France	Italy	Japan	Luxembourg	Norway	Portugal	Spain
adf	-1.307	-1.8026	-1.3272	-1.3658	-0.0996	-0.9349	-1.911	-2.069**	-1.5593	-0.9738	-0.9576
pp	-3.3159	-6.2761	-3.6662	-3.7542	0.1227	-1.6001	-7.8393	-8.5425**	-4.4871	-1.9592	-2.4043
mpp	-3.1703	-6.252	-3.6505	-3.7361	0.1266	-1.5965	-7.7992	-8.4873**	-4.4457	-1.9448	-2.3923
msb	0.3795	0.2826	0.3699	0.3414	0.9384	0.5596	0.2531	0.2426	0.3353	0.506	0.4077
mers	7.6639	3.9221	6.7117	6.6133	51.6979	15.344	3.1435**	2.8915**	5.5113	12.5748	9.5749
kpss	14.2177	10.993	11.4556	18.3333	19.2016	11.4754	23.3773	8.3612	15.2105	11.826	15.6277
pao	-4.2844	-4.3256	-3.024	-4.4116	-3.4835	-3.8689	-5.7108**	-4.587	-3.432	-3.8897	-3.6277
pio	-3.1788	-2.6701	-1.6861	-1.9263	-1.6579	-1.8296	-1.9478	-2.6055	-1.8498	-2.0727	-1.4363

	Sweden	Switzerland	UK	US
adf	-1.854	-2.1444**	-1.4407	-2.194**
pp	-6.8193	-9.3734**	-4.2014	-9.4346**
mpp	-6.7463	-9.3305**	-4.1896	-9.4203**
msb	0.272	0.2315**	0.3453	0.2302**
mers	3.6386	2.6258**	5.8488	2.608**
kpss	15.3996	12.3485	10.4342	8.3244
pao	-3.6802	-4.9824**	-3.9272	-3.9723
pio	-2.6849	-3.3529	-2.3997	-2.7929

Notes: 1. **adf**- Augmented Dickey-Fuller test; **pp**- Phillips-Perron test; **mpp**- modified Phillip-Perron test; **msb**- modified Sargan-Bhargava test; **mers**- modified Elliott, Rothemberg and Stock feasible point optimal test; **kpss**- Kwiatkowski-Phillips-Schmidt-Shin test; **pao**- Augmented Perron Additive Outlier test by Vogelsang and Perron (1998) criteria; **pio**- Augmented Perron Innovative Outlier test by Vogelsang and Perron (1998) criteria. 2. '**' denotes significant at 5% significance level.

on unit root among the series, the stationary cases can be regarded as country specific characteristic.

5.3 Results under Non-normality Assumption to Incorporate Structural Breaks

In general, the CH and CS tests fail to reject unit root, the index of stability α are estimated through models (4.2) and (4.6). Residuals η_{it} are computed and α 's are estimated. The estimators are $\hat{\alpha}_{CH} = 1.585$ and $\hat{\alpha}_{CS} = 1.583$, which indicates that there are considerable structural breaks contained in the process. This provides additional support to the finding in some studies that structural breaks exist in inflation series as discussed in section 4. Refer to the results in Chapter 3, the α 's are small enough to affect CH test and severely impact CS test[30]. However, the earlier panel unit root test conclusion may not be overturned since the test statistics are relatively far away from the critical values computed under normality.

To confirm this point, Monte Carlo experiments are conducted again to compute the critical value intervals of the tests under $\alpha = 1.5$ and $\alpha = 1.6$. The procedure and programs follow from Chapter 3. Sample size is $N = 15$ and $T = 500$, since there are 15 countries and 600 observations in the time dimension; the number of replications is 10,000, the same as the computation for critical values. The 5% critical value intervals augmented by David-Johnson estimate of percentile standard deviation are presented in Table 4.5. As predicted, both tests still hold on to the same conclusion as that under normality.

Although the conclusion is not overturned, to give an idea about how likely structural breaks or non-normality can overturn the previous test results which are concluded in

[30] Recall in Chapter 3 that the conservative benchmarks of the α 's are $\alpha_{CH} = 1.7$, $\alpha_{CS1} = 1.95$ and $\alpha_{CS2} = 2$; a generous choices are $\alpha_{CH} = 1.4$, $\alpha_{CS1} = 1.9$ and $\alpha_{CS2} = 1.95$.

the absence of the problems, a measure can be proposed as following. For convenience the midpoints of the critical value intervals are used and are printed below the corresponding intervals in Table 4.5. Let q_{nor} and q_{st} denote the midpoints of the critical value intervals computed under normality and non-normality assumptions, respectively; t_{stat} denotes a testing statistic. Then a relative measure for test distortion can be formulated as

$$dm = \frac{|q_{st} - q_{nor}|}{t_{stat} - q_{nor}}\text{[31]}, \quad t_{stat} \neq q_{nor} \tag{4.7}$$

The measure (4.7) indicates the relative likelihood that how likely the test distortion caused by structural breaks or non-normality can overturn the previous results concluded under normality, given a testing statistic. The larger the distance between q_{nor} and q_{st} is, *i.e.* larger value of $|q_{st} - q_{nor}|$, the more severe the test distortion is; since for CH and CS tests $q_{st} > q_{nor}$, when $t_{stat} < q_{nor}$ (or $t_{stat} - q_{nor} < 0$), non-normality obviously cannot impact the previous conclusion; if $t_{stat} > q_{nor}$, the closer t_{stat} is to q_{nor}, *i.e.* a smaller value of $(t_{stat} - q_{nor})$, the more likely the previous result could be overturned. So overall, a negative value of *dm* means there is no danger that non-normality would affect the testing conclusion; a relatively high value of *dm* indicates that given a testing statistic t_{stat}, the result concluded under normality is more likely to be overturned by the impact of non-normality.

[31] Note that for CH and CS tests the critical value computed under non-normality is larger than the corresponding one computed under normality, *i.e.* $q_{st} > q_{nor}$, so in (4.7) the denominator is $t_{stat} - q_{nor}$; otherwise, the denominator is $q_{nor} - t_{stat}$.

Table 4.5 Finite sample ($N = 15, T = 500$) 5% critical value intervals and interval midpoints of CH and CS tests

Under normality assumption

CH (under weak dependence)			CS (under strong dependence)			CS (under long run dependence)		
cv interval	-1.725	-1.631	cv interval	-2.190	-2.096	cv interval	-3.412	-3.318
midpoint	-1.678		midpoint	-2.143		midpoint	-3.365	

Under non-normality assumption

	CH (under weak dependence)			CS (under strong dependence)			CS (under long run dependence)		
$\hat{\alpha} = 1.6$	cv interval	-1.515	-1.433	cv interval	-0.903	-0.821	cv interval	-1.362	-1.280
	midpoint	-1.474		midpoint	-0.862		midpoint	-1.321	
		BIC	AIC		BIC	AIC		BIC	AIC
	dm	0.550	-0.206	dm	0.759	0.735	dm	0.702	0.690
$\hat{\alpha} = 1.5$	cv interval	-1.496	-1.414	cv interval	-0.895	-0.813	cv interval	-1.366	-1.284
	midpoint	-1.455		midpoint	-0.854		midpoint	-1.325	
		BIC	AIC		BIC	AIC		BIC	AIC
	dm	0.601	-0.223	dm	0.764	0.740	dm	0.701	0.688

Note: BIC and AIC denote that calculation of the values below use the statistics computed with lag orders selected by the BIC and AIC criteria.

The measures *dm* calculated for the CH and CS tests and are provided in Table 4.5. It shows that the earlier testing conclusion of CH test with lag order selected by AIC criteria is not affected by structural breaks or non-normality at all, since the *dm* values are negative; on the other hand, the CS test results tend to be relatively highly likely influenced, although the conclusions still remain the same under non-normality assumption.

The CH test fails to reject unit root under one of the lag order selection criteria (BIC), however, considering the fact that the CH test is unable to handle strong dependence and the CS test is most robust to even the long run dependence, in addition, another general test, the BN test, also supports the unit root hypothesis, the chapter concludes that the inflation rates among OECD countries are non-stationary.

6. Conclusion

The dynamic properties of inflation rate play an important role in macroeconomics analysis. This chapter re-examines the stationarity of inflation rates by the recently developed panel unit root tests using a panel of 15 OECD countries. The presence of structural breaks is found in inflation and its impact on unit root tests are stressed by many studies. Instead of modelling structural breaks, an alternative and new approach is suggested, treating the presence of breaks in a process as a type of non-normality. The α stable distribution, which has been shown to be useful in simulating structural breaks, is implemented in the panel unit root test models. The new method is easy and convenient to apply and avoids the complications of estimating breaks.

The chapter uses the largest possible span of monthly inflation data during the period 1958-2007. Either considering structural breaks or not, the CS panel unit root test, which is most general to cross section dependence, supports the unit root hypothesis. Under the two different lag order selection criteria, the AIC and BIC criteria, the CH test results in opposite decisions; whereas either the presence of structural breaks is assumed or not, the decision also remains the same as by the CS test. Given the unchanged conclusion with and without structural breaks, a measure is constructed to show the relative likelihood that how likely the test bias caused by structural breaks or non-normality overturns the results obtained under normality. It suggests that the CS test is relatively more likely to be affected.

Applications of some other popular second generation and first generation tests produce very mixed outcomes. Among these, the BN test, which is another most robust test due to its unique treatment of cross section dependence (see the review in Chapter 1), is also in favour of unit root. The single time series tests applied on the individual series provide strong evidence on the presence of unit root even when structural breaks are considered. Overall, with preference on the most general panel unit root tests together with the substantial support from single time series tests, the chapter concludes that the inflation rates among OECD countries are non-stationary.

The testing results also seem to suggest that with sufficient number of observations, single time series test could outperform panel unit root test, since panel unit root tests experience extra difficulties such as cross section dependence and heterogeneity. Panel tests may also unnecessarily generalize the situation, since usually the null hypothesis in panel unit root tests is that all individuals have a unit root. So when only a small proportion of the individuals in the panel behave differently from other members, either a panel unit root test is able to reject the null or not (subject to its power), there is still open question left. For example, in this chapter 3 out of the 15 inflation series are found stationary by more than half of the single time series tests. If this is the true situation, panel tests can only conclude that either there is unit root in all individuals (generalization), or reject the unanimous presence of unit root with unknown numbers of stationary series in the panel. Under this circumtance, single time series tests might be a better choice, or at least can provide some assistance examination for the panel on each panel individual.

Conclusions

This thesis has used numerical methods to solve the previously ignored problems of finite sample bias and the non-applicable assumption of normality in panel unit root tests. Carrying the findings it continues to re-examine the long disputed stationarity property of inflation time series. Given the fact that structural breaks have been observed and analyzed in inflation rates, the thesis takes an alternative and brand new approach which regards the presence of breaks as a type of non-normality, instead of modeling breaks in unit root tests.

Using the outputs of Monte Carlo simulations, the response surface regressions take on newly designed representation forms to examine the impact of finite sample size and non-normality on panel unit root tests. Results find serious magnitude of bias incurred by the problems as well as illustrate the trends of bias. In particular, the finding is consistent with the criticism of Im and Pesaran (2003) and the conjecture in the thesis that more condition is required for the asymptotic properties of the CH and CS tests to hold. This clearly suggests that the asymptotic critical values of the tests are inappropriate for empirical applications.

The finite sample critical values are then computed for the CH and CS test. Rather than provide the conventional point estimates of critical values, the thesis improves the calculation by augmenting the point critical values with David-Johnson estimate of percentile standard deviation to account for the randomness incurred by simulation experiments, so that it effectively improves on testing precision. Due to the apparently slow convergence rate of the testing statistics, it is recommended that even when empirical data have larger number of observations ($N > 100$ and/or $T > 500$), the critical value computed under the largest finite sample size ($N = 100$ and/or $T = 500$) is used as approximation.

During the augmentation of the smoothed critical values (by response surface regressions), two sources of uncertainty have been included in the augmentation, the Monte Carlo simulation and the response surface regression. However, a problem of overlapping between neighbor critical value intervals occurs, which makes the two intervals indistinguishable from each other. This is due to the relatively large degree of uncertainty from regression, despite the values of R^2's are reasonably high. This suggests that, unless the response surface regressions have near perfect goodness-of-fit, smoothing critical values by response surface regression is inappropriate and point estimates from Monte Carlo simulations are preferred. Moreover, compared with single time series, it also reflects the complexity associated with the additional cross section dimension in panel data.

Non-normality is modeled by applying the α stable distribution. Following the detection by response surface regression that the magnitude of test distortion advances with increasing degree of non-normality, the thesis further finds the benchmark degrees of non-normality which distort the tests so severely that the critical value intervals computed under normality can no longer be applied. Empirically when unit root is suggested in a process by the CH and/or CS tests, the degree of non-normality in the error terms of the unit root test models needs to be evaluated and taken to check for the reliability of testing results (as the procedure applied on inflation rates in Chapter 4). If the degree of non-normality appears to be higher than the benchmark, more computations of the critical value under the corresponding degree of non-normality are needed.

Another fresh point made on the purposes which fat tailed non-normal distribution can serve is that the non-normal random variables are eligible to be used on simulating structural breaks in a process. Due to the caution that structural breaks in time series tend to bias unit root tests toward not rejecting the unit root hypothesis, traditionally structural breaks are modeled and estimated. This usually involves intricate procedures and the number of breaks it can reach is limited. As an alternative and brand new approach, the thesis treats the appearance of structural breaks in a process as a type of non-normality and implements it into the panel unit root test models. Instead of estimating the breaks, the degree of non-normality is estimated. The testing conclusion can be conveniently drawn upon the findings in Chapter 2 and 3 of this thesis.

The thesis finally turns to the ongoing debate of the stationarity property of inflation rate. Neither single time series tests nor panel unit root tests have been able to reach a consensus on the dynamic property of inflation rate. With the concern of structural changes in focus, a panel of 15 OECD countries is tested for unit root by the robust panel unit root tests, the CH and CS tests. The presence of structural breaks is treated as non-normality and represented by the α stable distribution as proposed in the methodology. The estimated relatively high degree of non-normality indicatively confirms the existence of breaks and more importantly, the critical values computed under normality are not appropriate to use. Further computations of critical values under the corresponding degree of non-normality provide support for the unit root hypothesis.

For comparison, a number of contemporary popular panel unit root tests and single time series tests are also applied. The panel tests produce very mixed results whereas the single times series tests provide sheer agreement on the presence a unit root. Considering that support for unit root hypothesis is advocated by the most robust panel unit root tests and by single time series tests on the majority of the countries, it concludes that the inflation rates of OECD countries are non-stationary.

To this end, a point might be raised regarding the choice between panel unit root tests and single time series tests for applications. Although panel unit root tests present statistical advantages over single time series tests (essentially due to the enlarged sample size in relation to the cross section dimension), they also experience more difficulties due to the additional dimension (as reviewed in Chapter1) as well as more complicated finite sample performance (as analyzed in Chapter 2), which in turn might produce misleading results. Moreover, usually the null hypothesis in panel unit root tests is that all individuals have a unit root; whereas when a very small proportion of the individuals are stationary, the power of the tests to reject the null poses another question. Even in such a situation a test with high power is able to reject the null, the number of stationary or non-stationary individuals in the panel is unknown. Therefore, when the number of observations in empirical data is reasonably large, single time series tests are still a good choice.

The non-stationary inflation rates conclusion also sheds some light on certain policy concerns. For example, on inflation targeting, if inflation rate is stationary, the Taylor rule or Svensson rule can be comfortably implemented. However, if inflation rate is found non-stationary, a moving target has to be constructed to form a stationary platform together with inflation rate (e.g. an error correction system derived from cointegration analysis); or the first difference of inflation process can be used instead.

Nevertheless, there are still some limitations in this thesis. The α stable distribution has been the traditional approach to model non-normality. However, more recent research finds that the tails of α stable distribution are often too heavy for empirical study, *i.e.* the distributions of certain economic variables have heavier tails relative to the normal distribution but thinner tails than the α stable distribution. In addition, it is assumed in the thesis that the distribution of inflation rates is symmetric, whereas more empirical studies suggest it tends to be skewed. So further research in this area might search for improved modeling frame to better resemble empirical data as well as relax the symmetric assumption. For example, the tempered stable distribution (or truncated Lévy distribution) is recently proposed in the forefront as a compromise between the normal distribution and the α stable distribution, whose tails are heavier than the normal distribution but thinner than the α stable distribution.

In addition, the imprecision caused by sampling distribution in Monte Carlo experiments is found particularly influential in the analysis in Chapter 3. It seems the David-John estimate of percentile standard deviation has not captured sufficient amount of randomness incurred by stochastic simulation. It is also case sensitive in terms of the study in the chapter. Although some remedial measures were taken to resolve the problem, the impact is still not completely removed. Some thorough examination on this issue may be carried out in a more general setting and search for more solutions.

Furthermore, due to the immense amount of computation time required, the thesis is not able to provide a comprehensive range of critical values for the tests under non-normality with different degrees of non-normality and various sample sizes. It is hoped that in the near future with the accessibility to super-power computers, this restraint can be easily resolved.

Appendix A An Introduction to Monte Carlo Method in Econometrics

Monte Carlo methods (or Monte Carlo experiments, Monte Carlo simulations) provide approximate solutions to a variety of mathematical problems by performing statistical sampling experiments on a computer. Through conducting Monte Carlo experiments the analytically intractable deterministic mathematical problems can be tackled by substituting an equivalent stochastic problem and solving the later. Essentially, Monte Carlo experiments sample randomly from a universe of possible outcomes and take the fraction of random draws that fall in a given set as an estimate of the set's volume. The law of large numbers ensures that this estimate converges to the correct value as the number of draws (say n) increases.

The method was initially employed by physicists who worked on nuclear weapon projects in the 1940s. As the problems of neutron diffusion could not be solved by analytical solutions, John von Neumann and Stanislaw Ulam suggested that these be solved by performing sampling experiments using random walk models on digital computer. Being secret, the work was given the code name "Monte Carlo". Using this proposal, instead of deriving the cumbersome analytical solutions, one conducts sampling experiments to obtain solutions to the integro-differential equations (c.f. Metropolis and Ulam, 1949; Hurd, 1985; Metropolis, 1987; Hoffman 1998).

The concept of employing sampling experiments on a computer later came to prevail in many other scientific disciplines such as mathematics, engineering, chemistry, econometrics, operational research, etc. More recently the method has also gained favour in the field of financial mathematics due to the convenient and easy calculations involved to produce solutions. Since the purpose of this thesis is to investigate certain problems in the field of econometrics and Monte Carlo method is a substantially broad subject, this introduction is devoted to the aspects of Monte Carlo approach that are relevant to the studies in the thesis. Extensive exploration of Monte Carlo simulations can be found in Fishman (1996), Rubinstein and Kroese (2007), Robert and Casella (2004), MacKeown (1997), Glasserman (2004), Hammersley and Handscomb (1975), Niederreiter (2004), etc.

Econometric studies in favour of Monte Carlo method are mainly due to two reasons. One is as mentioned above, when the analytical solution of a problem is hard to obtain and derivation of the theoretical analysis is too expensive (e.g. complicated estimation or testing procedures), results based on computer experiments are inexpensive and easy to produce (c.f. Mariano and Brown, 1993; Keane, 1993; Gouriéroux et al., 1993; Gouriéroux and Monfort, 1996, Gallant and Tauchen, 1996, etc.). In particular, at the stage of the rapid and advanced development of computers, computations have become even faster and more efficient. The other important role that Monte Carlo method plays in econometrics is to analyze the finite sample performance of statistical tests and compute finite sample critical values for empirical applications (e.g. the calculation of the critical values of Dickey-Fuller test). Even some tests have well developed asymptotic theories (e.g. panel unit root tests), the rate of convergence can be slow and thus their finite sample distributions experience apparent different behavior from large sample properties. While this usually poses difficulties to analytics, Monte Carlo experiments can effectively provide complementary numerical analysis to fill in the gap in the literature (c.f. Hylleberg et al., 1990; Beaulieu and Miron, 1993; Franses and Hobijn, 1997; Banerjee et al., 1998; Harbo et al., 1998; Nabeya and Tanaka, 1998; Pesaran et al., 2000; MacKinnon, 1996, 2002, etc.). This role is one of the subjects in this thesis. It is particularly practical for empirical study as very often the availability of empirical data is limited.

The Data Generation Process (DGP)

To conduct a Monte Carlo experiment, the *data generation process* (DGP) must be formulated using some predetermined parameter values and random numbers in the first place. The DGP is the random process decided by the setting of the problem and provides a framework for analysis. It is important to note that the DGP is fully known to researchers, both its functional form and parameters. This information can assist researchers in improving the efficiency of Monte Carlo experiments.

As mentioned above, the DGP embodies two sources of information, the particular functional form under econometric assumption and the parameters. The functional form is chosen for the desired analytical results and represents the setting of the problem in study. For example, to investigate non-stationary problems such as unit root tests and cointegration tests, the DGP in the simulation certainly needs to be set as a random walk process; for the study of nonlinear topics, the DGP must be a nonlinear form to set up the framework of the investigation.

The parameters of DGP have different natures. They are either the numerical values to be examined by simulation or randomly generated actual values to approximate a certain distribution. For example, in the study of estimation and hypotheses testing of OLS regression, the coefficients of independent variables are preassigned to act as the true parameters in order to observe the properties of the corresponding estimators in simulation; the error terms are assumed to be normally distributed in OLS regression, and so in the DGP the error terms are randomly drawn from normal distribution.

As a simple example, consider the problem of the consistency property of OLS estimators in classical linear case. Let there be two independent variables. To set up the framework for the analysis a DGP can be formulated as following

$$y = 2x_1 + 2x_2 + u \qquad u \sim NIID\left(0, \sigma_u^2\right) \tag{A.1}$$

where *NIID* denotes identical independent normal distribution; y, x_1, x_2 and u are $T \times 1$ vectors; the values of x_1, x_2 and σ_u^2 are priorly known. So u are randomly drawn from normal distribution as required by the assumptions of the OLS estimation.

Next consider a perfectly specified linear regression model in accordance with the DGP

$$y = \beta_1 x_1 + \beta_2 x_2 + u \qquad u \sim NIID\left(0, \sigma_u^2\right) \tag{A.2}$$

Let $X = (x_1, x_2)$ and $\beta = \left(\beta_1, \beta_2\right)'$. The research interest is in the OLS estimator $\hat{\beta}$. It is calculated as

$$\hat{\beta} = \left(X'X\right)^{-1} X'y \tag{A.3}$$

For illustration purpose, suppose that the consistency property of $\hat{\beta}$ is analytically unachievable, so that the analysis resorts to Monte Carlo experiment. The objective is to calculate $E\left(\hat{\beta}\right)$ and analyze the finite sample bias of $\hat{\beta}$.

Pseudorandom Numbers

One fundamental element of a Monte Carlo simulation is a sequence of apparently random numbers used to drive the simulation, as the u in the above example, which are randomly drawn from normal distribution. As noted in the previous section, in the DGP the random numbers are generated to form random process in order to obtain the desired distribution of the estimators.

Random numbers are numbers that are realizations of a given random variable and create an unpredictable sequence, whereas in practice genuine random numbers are unachievable. Modern pseudorandom number generators are sufficiently good at mimicking genuine randomness to aid effective simulation experiments. Nevertheless, it is important to note that the apparently random numbers at the core of a simulation are in fact produced by completely deterministic algorithms. Practically pseudorandom numbers cannot be distinguished from real random numbers. In Monte Carlo simulations, where it mentions random numbers, it essentially refers to pseudorandom numbers.

Since pseudorandom numbers are generated by deterministic function with specified seeds (the initial input value in the function), they are exactly reproducible under the same algorithm and seeds. There is a vast literature on the topic of random number generation. Extensive discussions and references can be found in Bratley et al. (1987), Devroye (1986), Gentle (1998) etc. Generally, the basic random numbers are uniformly distributed between 0 and 1 (denote as U_i) . Essentially samples from any other distribution can be transformed from uniform random variables on the unit interval. A very important property of random numbers is that any pair of values generated should be uncorrelated and the value of U_i should not be predictable from $U_1, \ldots, U_{i-1}$.

There are some potential problems in random number generators in terms of the quality of the generated random numbers, such as how well the randomness property of random numbers is maintained; whether the generators can produce sufficiently enough random numbers for large experiments. A poorly designed generator can cause failure of simulation experiments and produce misleading results. In addition, the producing speed of a generator is also a concern.

Glasserman (2004) highlights several considerations in the construction of a random number generator. (ⅰ) *Period length*. Generally pseudorandom number will eventually repeat itself. Other things being equal, generators with longer length are preferred, *i.e.* generators that produce more distinct values before repeating. (ⅱ) *Reproducibility*. A drawback of genuine random process is that it is hard to reproduce. However, research often needs to rerun simulations using exactly the same random variables as before, or the same random variables are used in two or more simulation experiments for comparison purpose or else. This is easily accomplished with pseudorandom number generator by using the same seed. (ⅲ) *Speed*. Since a random number generator may be called thousands or millions of times in a simulation (depends on the number of replications), it must be fast. (ⅳ) Portability. A random number generating algorithm should produce the same sequence of values on all computing platforms. (ⅴ) Randomness. This is the most important consideration and is the hardest one to define or ensure. Random numbers generated with good theoretical properties can be tested by statistical tests for scrutiny. So far this field is sufficiently well developed and many generators in the literature that have survived rigorous tests can be comfortably used.

A Simple Example of Monte Carlo Simulation Experiment

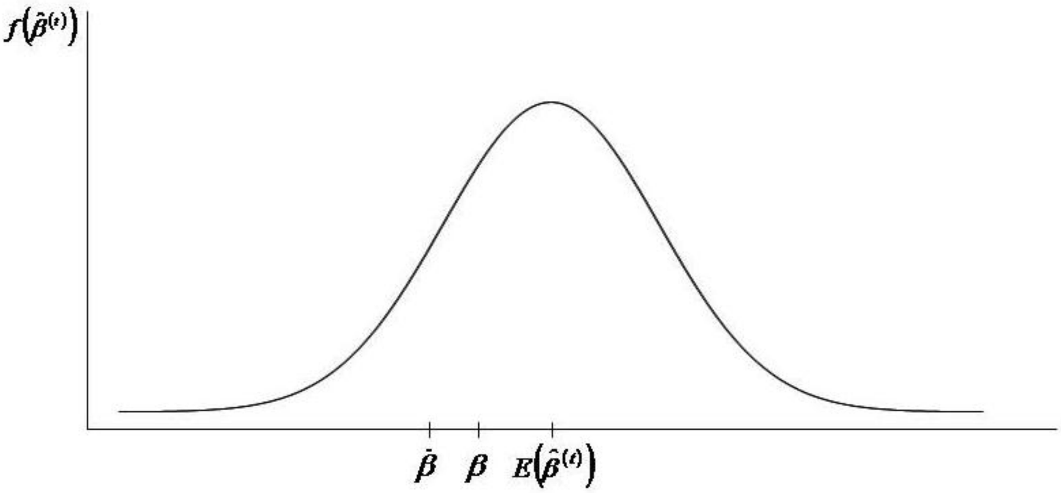
Properties of Consistent Estimator

Consider the typical econometric problem of consistency as shown in Figure A.1. Let the true parameter be β; the estimator $\hat{\beta}^{(t)}$ has a **plim** of $\dot{\beta}$ (*i.e.* $\text{plim}\,\hat{\beta} = \dot{\beta}$) and expectation $E\left(\hat{\beta}^{(t)}\right)$, where *t* is sample size. It implies that

$$E\left(\hat{\beta}^{(t)}\right)\xrightarrow[t\to\infty]{}\dot{\beta} \quad \text{and} \quad Var\left(\hat{\beta}^{(t)}\right)\xrightarrow[t\to\infty]{}0$$

Therefore, $\left[E\left(\hat{\beta}^{(t)}\right)-\dot{\beta}\right]$ is small sample bias, *i.e.* in repeated sampling, the estimator on average may not equal to the true parameter; and $\left(\dot{\beta}-\beta\right)$ is large sample bias, *i.e.* the estimator may not converge to the true parameter even when sample size grows infinitely large. For a consistent estimator, however, $\dot{\beta}=\beta$; as the sample size goes to infinity ($t\to\infty$), $E\left(\hat{\beta}^{(t)}\right)\to\dot{\beta}$ and so $E\left(\hat{\beta}^{(t)}\right)\to\beta$.

Figure A.1 Illustration of the problem associated with consistency



Thus if the estimator $\hat{\beta}^{(t)}$ in the example is consistent, its Bias (BIAS) and Root Mean Square Error (RMSE)[1] have the following property

$$BIAS\left(\hat{\beta}^{(t)}\right)=E\left(\hat{\beta}^{(t)}-\beta\right)\xrightarrow[t\to\infty]{}0$$

$$\text{and} \quad RMSE\left(\hat{\beta}^{(t)}\right)=\sqrt{Var\left(\hat{\beta}^{(t)}\right)+\left[E\left(\hat{\beta}^{(t)}-\beta\right)\right]^2}\xrightarrow[t\to\infty]{}0$$

[1] BIAS indicates that by how much, on average, the estimator under- or overestimates the true value of the parameter; RMSE indicates that by how much, on average, the estimator deviates from the true value of parameter.

Following the simulation example the BIAS and RMSE of $\hat{\beta}^{(t)}$ observed on a simulation basis is

$$BIAS\left(\hat{\beta}^{(t)}\right) = E\left(\hat{\beta}^{(t)} - \beta\right) \approx \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_i^{(t)} - \beta = \overline{\beta} - \beta$$

$$\begin{aligned} RMSE\left(\hat{\beta}^{(t)}\right) &= \sqrt{Var\left(\hat{\beta}^{(t)}\right) + \left[E\left(\hat{\beta}^{(t)} - \beta\right)\right]^2} \\ &\approx \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_i^{(t)} - \beta\right)^2} \\ &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_i^{(t)} - \overline{\beta}\right)^2 + BIAS^2} \end{aligned}$$

Ignoring t, recall also that the variance of OLS estimator $\hat{\beta}$ is given by

$$Var\left(\hat{\beta}\right) = \sigma_u^2\left(X'X\right)^{-1}$$

This suggests that the BIAS and RMSE of OLS estimators are related to sample size and the variance of error terms.

The Simulation Experiment

As specified in the DGP section, the true coefficients of the independent variables are $\beta = (2,2)'$. Set the number of replications as $n = 5000$ and conduct Monte Carlo experiments. Table A.1 and A.2 present the simulation results of the BIAS and RMSE for $\hat{\beta}_1$ as example. Figure A.2 plots the BIAS and RMSE in Table A.1 and A.2 to provide a clearer view. These depict the apparent empirical trends of the BIAS and RMSE of $\hat{\beta}_1$ as sample size and residual variance vary. In particular, RMSE systematically decreases as sample size *T* increases and grows with the increase in σ_u. These suggest that there exists potential systematic relationship between BIAS, RMSE and *T*, σ_u^2.

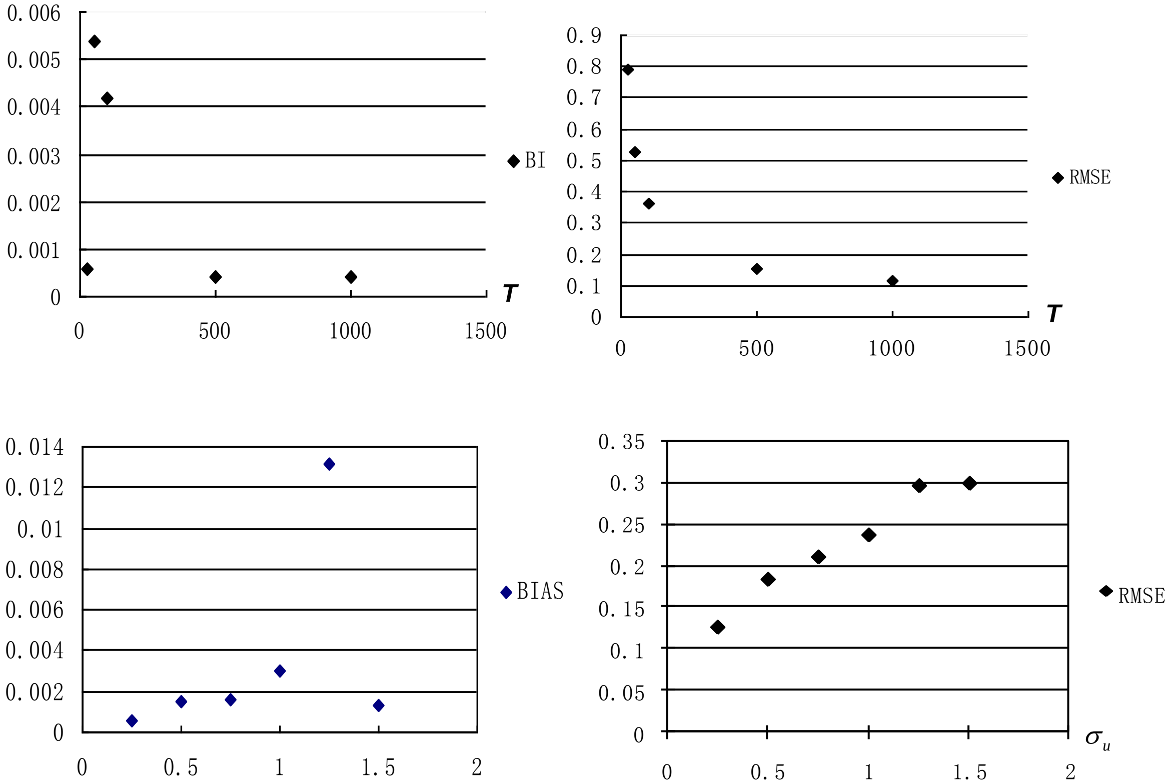
Table A.1 BIAS and RMSE of $\hat{\beta}_1$ with fixed σ_u ($\sigma_u = 1$)

Sample size	BIAS	RMSE
25	-0.0006	0.7918
50	0.0054	0.5243
100	-0.0042	0.3596
500	-0.0004	0.1555
1000	-0.0004	0.1130

Table A.2 BIAS and RMSE of $\hat{\beta}_1$ with fixed sample size (T=200)

σ_u	BIAS	RMSE
0.25	-0.0006	0.1250
0.5	-0.0015	0.1830
0.75	-0.0016	0.2101
1	0.0030	0.2368
1.25	0.0132	0.2965
1.5	0.0013	0.2994

Figure A.2 Plot of BIAS and RMSE of $\hat{\beta}_1$ as T and σ_u vary



Limitations of Monte Carlo Experiments and a Remedy

There is criticism to simulation experiments due to the fact that setting of the problem is stochastically reformulated and the estimators have sampling distributions. This incurs two unavoidable difficulties as noted by Hendry (1984), *imprecision* and *specificity*. Imprecision is incurred by sampling distributions. There is a well developed literature that provides various methods to increase efficiency. Since the techniques to improve efficiency are not a focus for the study in this thesis, they are not discussed here. Elaborate discussions can be found in Glynn and Iglehart (1988), Fishman (1996), Fournié et al. (1997); Asmussen and Binswanger (1997), Avramidis and Wilson (1998), Schmeiser et al. (2001), Glasserman (2004), etc.

The other problem, specificity, refers to the fact that Monte Carlo estimation with only certain choice of parameters can do little more than provide some unknown generality. For instance in the above example, since the estimation of BIAS and RMSE of $\hat{\beta}^{(t)}$ is closely related to sample size T (for clarification, sample is denoted by T hereafter) and error term variance σ_u^2, it must be re-estimated and re-examined as T and σ_u vary. This is particularly problematic in investigating finite sample distributions of testing statistics, e.g. interpolating percentiles for distributions under any sample size T. Unfortunately this can hardly be solved by analytics either. On the *specificity* pitfall in simulation experiments, the *response surface* methodology can act as an effective remedy to extend the Monte Carlo experiment results beyond those obtained under the parameters specified in the DGP through approximation. In the analysis where econometric tests experience complicated asymptotics or the speed of convergence of testing statistic is slow and causes finite sample properties to behave substantially different from asymptotics, the response surface method can provide contributive value to the theories.

Response Surface Analysis

The Concept of Response Surface Analysis

The response surface methodology is a statistical method which explores the relationships between several explanatory variables and one or more response variables. It was introduced by Box and Wilson (1951). They acknowledge that although this method only provides an approximation, it is useful due to the easy estimation and application even when little information is known about the problem. The general procedure of the method is firstly to identify the factors that affect the response variable; once the important factors have been identified, the next step is to determine the settings or functional forms in which these factors result in the optimum value of the response variable.

The technique has gained favour in several subjects such as physics, engineering, chemistry, biology, etc. Hendry (1984) firstly introduced response surface method into the field of econometrics to remedy the limitation of specificity in Monte Carlo experiments. In the application procedure, after identifying the factors that affect the response variable, regressions are used to examine the relationship between the response and the factors, with the response being dependent variable and the factors being explanatory variables. The functional form of the regression is usually determined by the significance of the coefficients of independent variables and a reasonably high value of R^2 which shows the high explanatory power of the factors to the response variable. The early econometric applications of response surface method are by MacKinnon (1991, 1994) where the percentiles of the statistic distributions in several tests under a range of finite sample sizes are examined as well as the finite sample (or numerical) distribution functions of the statistics. The method was well welcomed and later has been applied in a number of studies (c.f. Cheung and Lai, 1995a,b; Sephton, 1995; Carrion et al, 1999; MacKinnon et al., 1999; Cook, 2001; Ericsson and MacKinnon, 2002; Presno and López, 2003, etc.).

Following the simulation example, this section continues to illustrate the response surface method as a remedy to the problem of specificity in the Monte Carlo experiment. As is indicated by the Monte Carlo simulation example that there is potential systematic relationship between BIAS, RMSE and some factors that influence them, *i.e.* T, and σ_u^2. The relationship is particularly obvious for RMSE which can be clearly observed from Figure A.2. Regressing the series of BIAS or RMSE as response variables on the corresponding series of factors can provide important insight to the analysis, *i.e.* whether BIAS and RMSE will eventually disappear as sample size T passes to infinity and thus suggests if the estimator is consistent. Moreover, using the regression estimation results, the analysis can also assist in calculating the magnitude of BIAS or RMSE under any specific parameter values, T, and σ_u^2. The next section will demonstrate the method using RMSE as an example.

The Response Surface Regression

To run response surface regression, a series of RMSE need to be generated by Monte Carlo experiments with different sets of T and σ_u. For example, they can be set as

$$\sigma_u = \{0.25, 0.5, 0.75, 1.0, 1.25, 1.50\}$$

$$T = \{25, 50, 75, 100, 250, 500, 1000\}$$

So there are overall $6 \times 7 = 42$ Monte Carlo experiments with different combinations of σ_u and *T*. Each experiment consists of *n* = 5,000 replications. Take the RMSE of $\hat{\beta}_1$ as example. For each experiment the RMSE of $\hat{\beta}_1$ is recorded.

The next step is to specify the functional form of response surface regression. Generally there is no specific rule on the design of the regression representation, rather it is a more empirical and trial procedure. Basically a functional form is chosen on the significance of estimator coefficients and reasonably high R^2 to show the explanatory power of regressors. Due to the preliminary analysis that the RMSE of an OLS estimator is affected by *T* and σ_u^2, some potential candidates of the explanatory variables for the response surface regression can be $\frac{1}{\sqrt{T}}$, $\frac{1}{T}$, $\frac{1}{\sqrt{T^3}}$, $\frac{\sigma_u}{\sqrt{T}}$, $\frac{\sigma_u}{\sqrt{T^3}}$, $\frac{\sqrt{\sigma_u}}{\sqrt{T^3}}$. After some trials and take the considerations stated above, the model is specified as

$$RMSE\left(\hat{\beta}_1\right) = f\left(T, \sigma_u^2\right) = \alpha + \gamma \frac{1}{\sqrt{T}} + \delta \frac{\sigma_u}{\sqrt{T}} + \varphi \frac{\sigma_u}{T} + e \qquad e \sim NIID\left(0, \sigma_e^2\right) \tag{A.4}$$

Apply the OLS estimation and obtain

$$RMSE\left(\hat{\beta}_1\right) = -0.011 + 1.6698 \frac{1}{\sqrt{T}} + 1.9427 \frac{\sigma_u}{\sqrt{T}} - 8.8595 \frac{\sigma_u}{T} \tag{A.5}$$

standard error	(0.0136)	(0.1426)	(0.1484)	(4.3364)
p-value	(0.422)	(0.000)	(0.000)	(0.048)

$R^2 = 0.975$, (Prob > F) = 0.0000

where the null hypothesis of *F*-test is $\alpha = \gamma = \delta = \varphi = 0$.

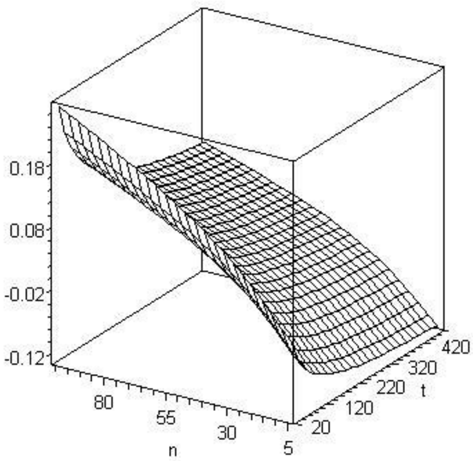
The extremely high value of R^2 indicates that nearly all the variations in RMSE can be explained by the regressors. The constant term is insignificant as expected, which indicates that as *T* increases and passes to infinity, the RMSE of $\hat{\beta}_1$ reduces to zero.

Keeping T fixed, the increasing σ_u will dramatically add value to RMSE; whereas as T grows large, the excessive RMSE caused by σ_u will eventually disappear. Overall, the findings of response surface regression suggest that as sample size goes infinitely large, the RMSE of $\hat{\beta}_1$ will pass to zero despite the influence of σ_u, and thus the consistency of OLS estimator is supported.

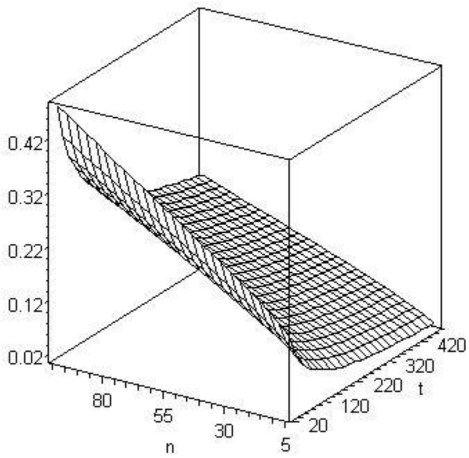
Although the properties of OLS estimators have been analytically well established, they are employed here for the illustration of the methodology. In the main body of this thesis, the same spirit is adopted to investigate the finite sample bias and numerical distributions of the recently developed popular panel unit root tests, since the tests experience poor finite sample performance and also bear certain criticism regarding their asymptotic properties. On the performance of the tests under non-normality assumption, the response surface technique is again employed, due to the complexity of the analytical analysis. The trend of test distortion caused by the increasing degree of non-normality is observed with the assistance of response surface regressions and provides important indications for empirical studies.

Appendix B Graphs and Tables

Figure 2.2 Estimates of finite sample bias of the 1% and 10% percentiles for the three DGPs as sample sizes *N* and *T* vary (different scales for *N* and *T*)

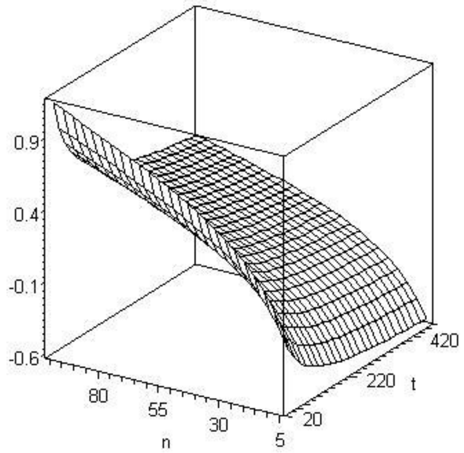


(a) DGP1-1%

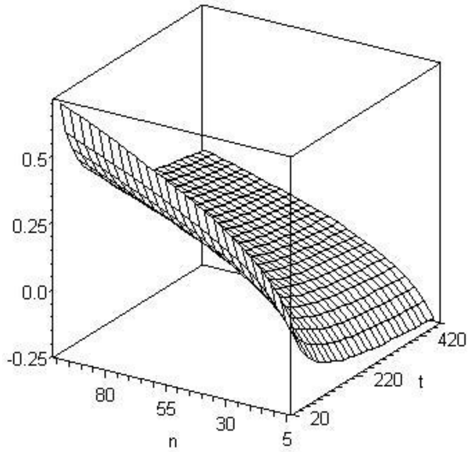


(c) DGP1-10%

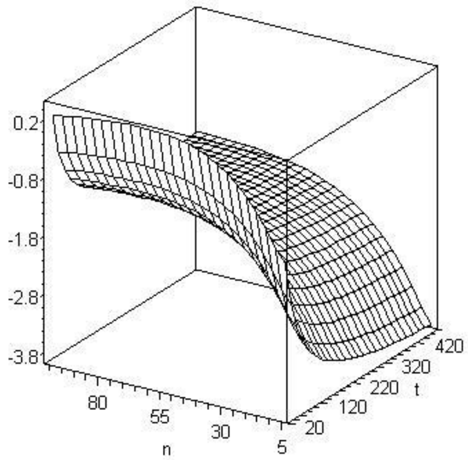
Figure 2.2 Estimates of finite sample bias of the 1% and 10% percentiles for the three DGPs as sample sizes *N* and *T* vary (different scales for *N* and *T*) **(Cont'd)**



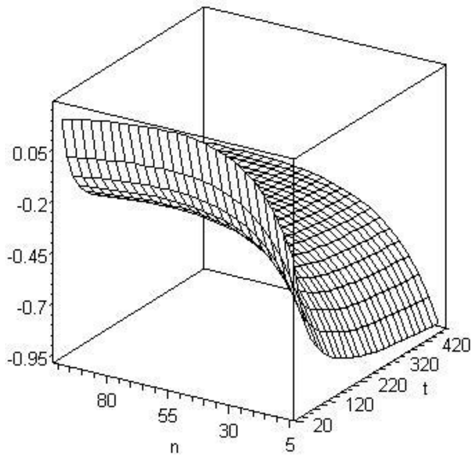
(d) DGP2-1%



(f) DGP2-10%

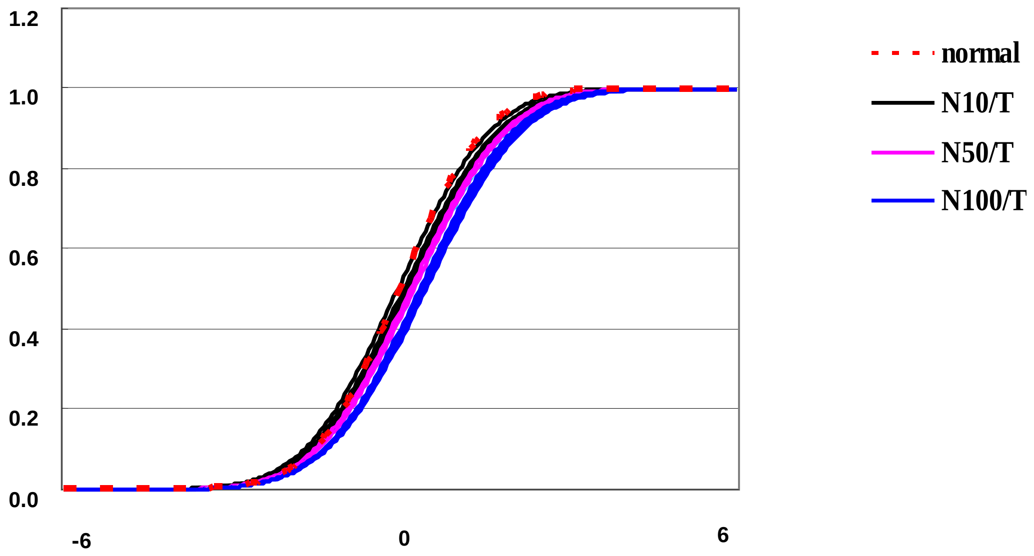


(g) DGP3-1%

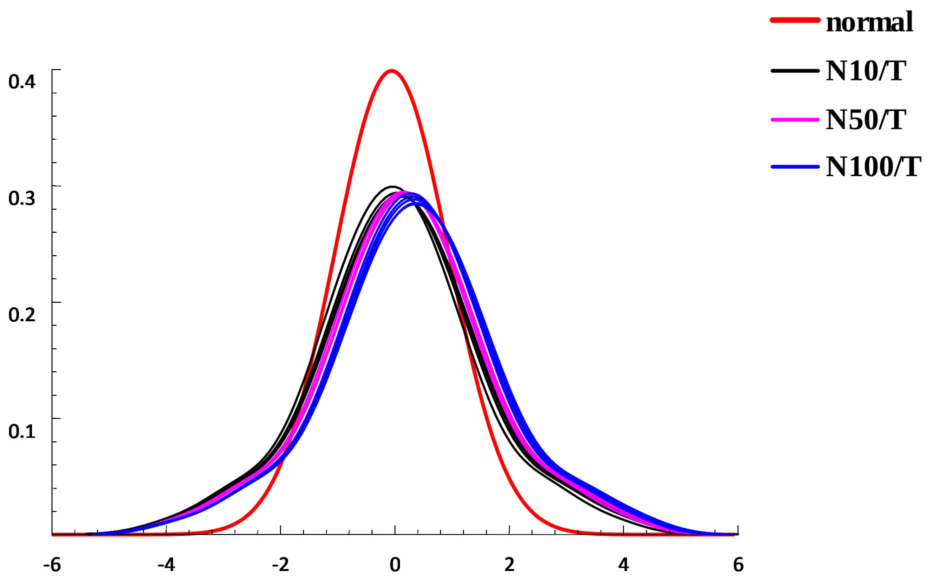


(i) DGP3-10%

Figure 2.4 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP1

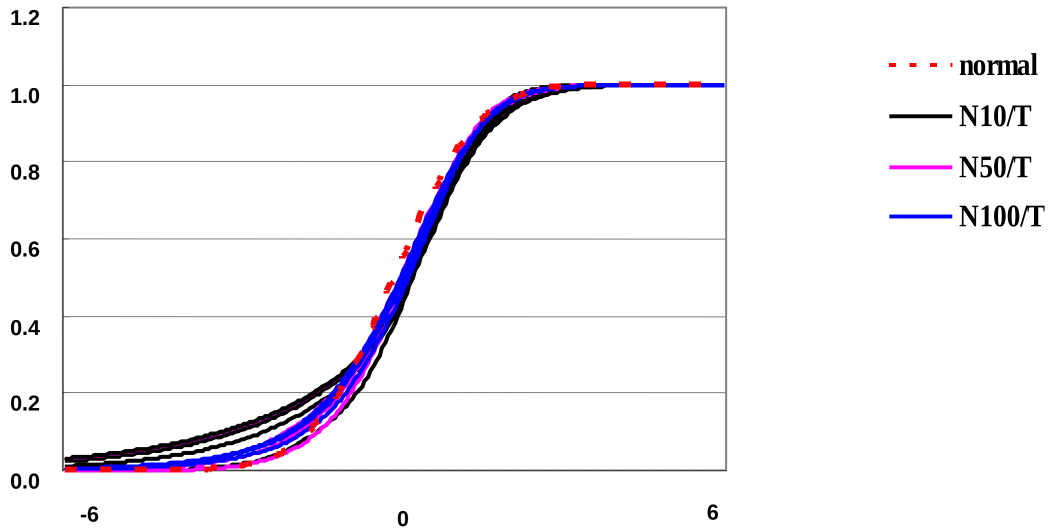


(a) DGP1-CDF

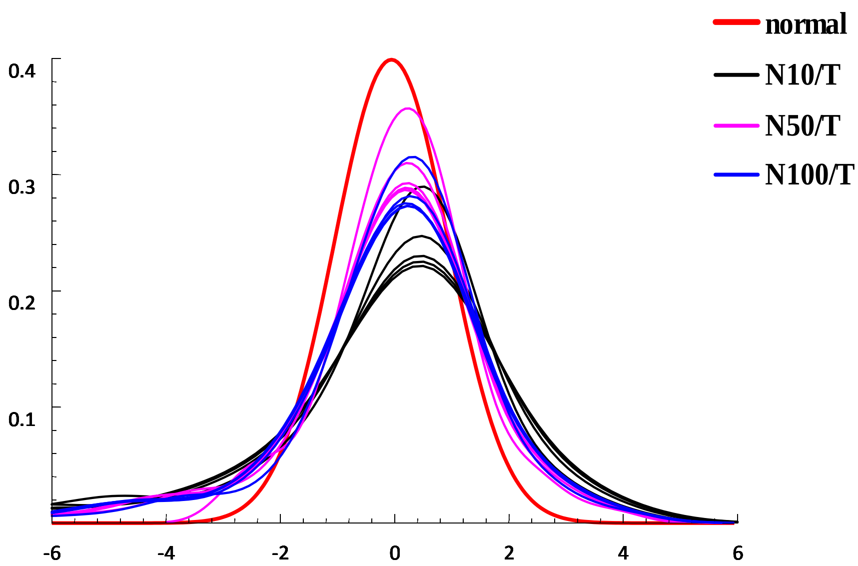


(b) DGP1-PDF

Figure 2.6 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP3



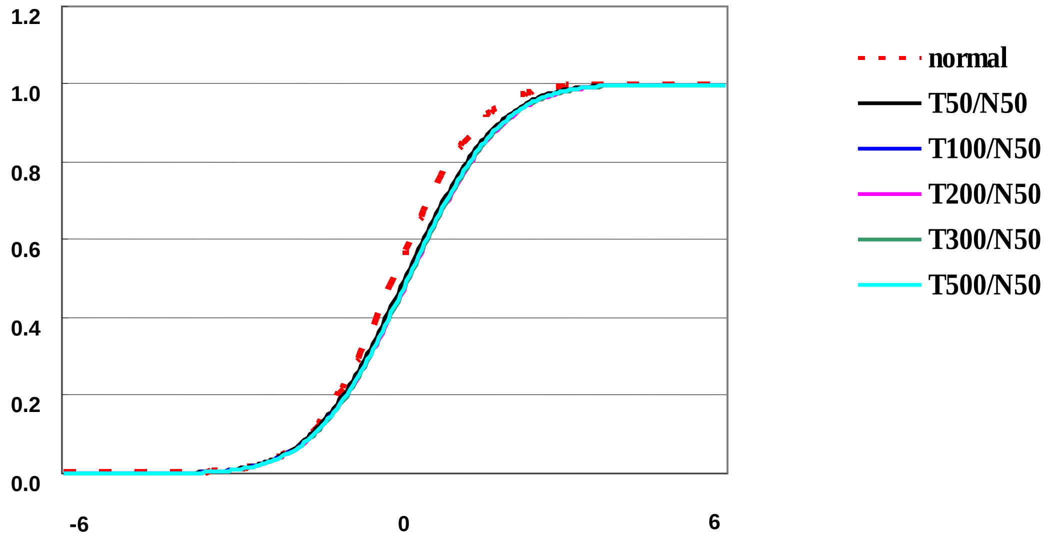
(a) DGP3-CDF



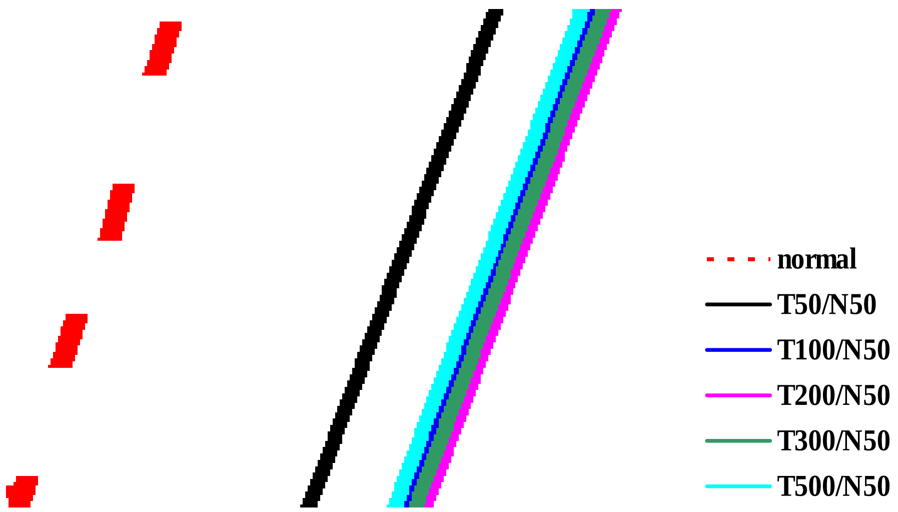
(b) DGP3-PDF

Figure 2.7 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP1, $N = 50$

(b) and **(d)** are enlarged view of **(a)** and **(c)**, respectively



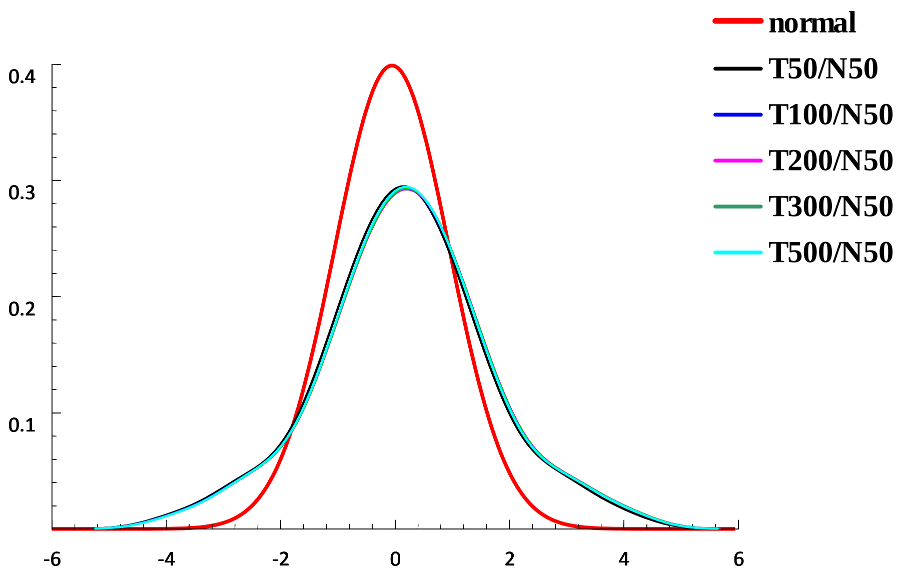
(a) DGP1-CDF



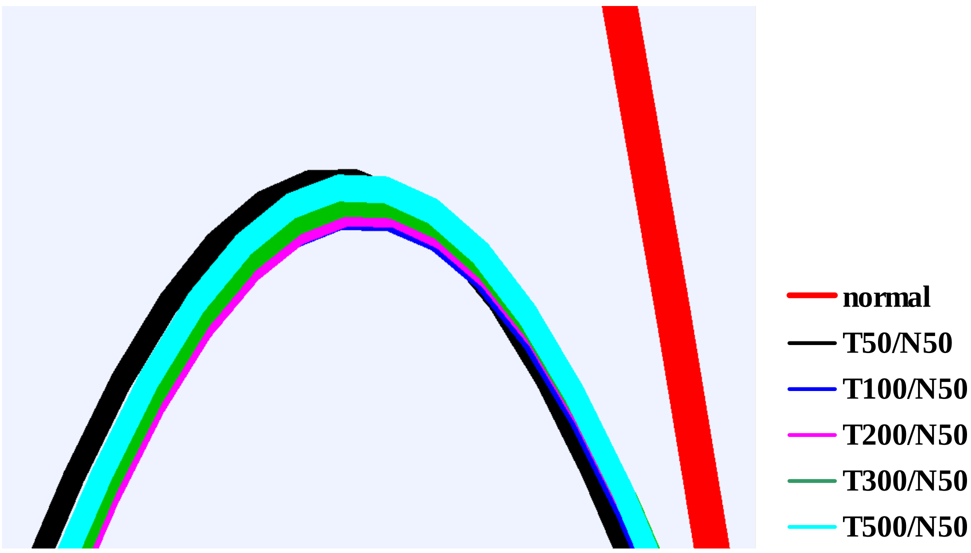
(b) DGP1-CDF-enlarged view of (a)

Figure 2.7 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP1, $N = 50$ **(Cont'd)**

(b) and **(d)** are enlarged view of **(a)** and **(c)**, respectively

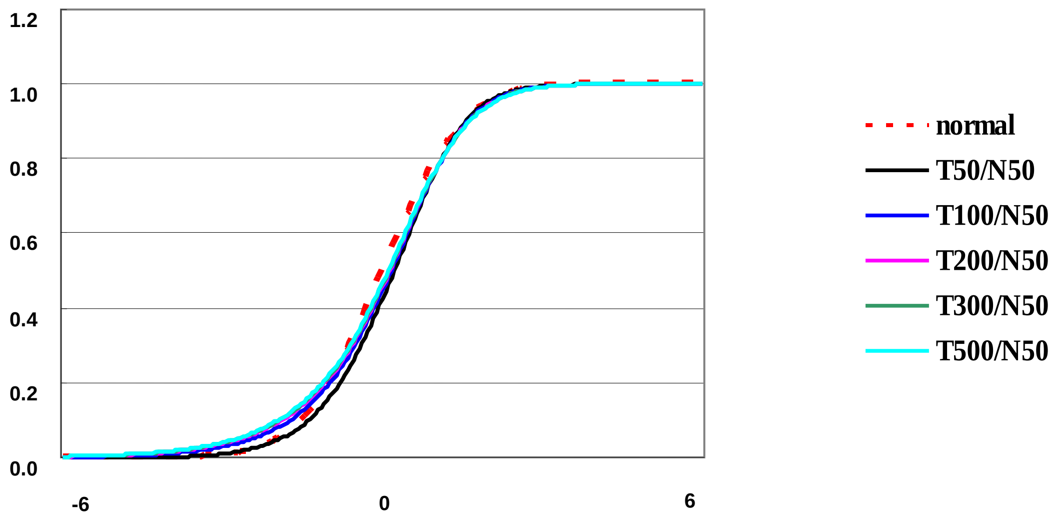


(c) DGP1-PDF

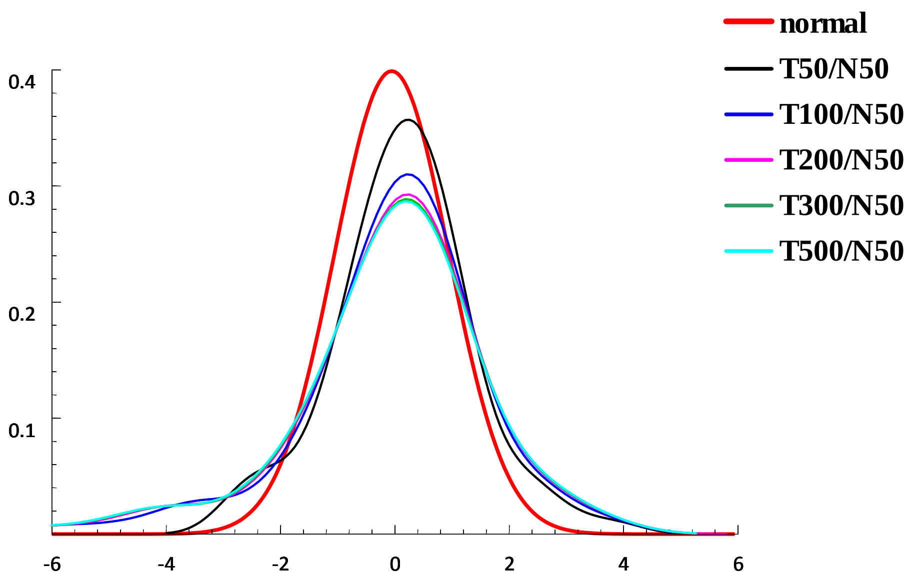


(d) DGP1-PDF-enlarged view of (c)

Figure 2.9 Plots of the numerical cumulative density functions (CDF) and probability density functions (PDF) for DGP3, $N = 50$



(a) DGP3-CDF



(b) DGP3-PDF

Table 2.3 Finite sample critical values intervals of Chang (2002) test and Chang and Song (2005, 2009) test

Chang (2002) test under weak dependence (DGP1)

N	*T*	1%		5%		10%	
10	50	-2.543	-2.397	-1.754	-1.672	-1.353	-1.285
	75	-2.614	-2.468	-1.811	-1.729	-1.415	-1.347
	100	-2.550	-2.404	-1.768	-1.686	-1.364	-1.296
	125	-2.544	-2.398	-1.760	-1.678	-1.361	-1.293
	150	-2.600	-2.454	-1.739	-1.657	-1.325	-1.257
	175	-2.523	-2.377	-1.809	-1.727	-1.392	-1.324
	200	-2.463	-2.317	-1.745	-1.663	-1.338	-1.270
	225	-2.619	-2.473	-1.849	-1.767	-1.418	-1.350
	250	-2.536	-2.390	-1.768	-1.686	-1.361	-1.293
	275	-2.518	-2.372	-1.684	-1.602	-1.299	-1.231
	300	-2.566	-2.420	-1.788	-1.706	-1.400	-1.332
	325	-2.487	-2.341	-1.798	-1.716	-1.396	-1.328
	350	-2.517	-2.371	-1.755	-1.673	-1.363	-1.295
	375	-2.630	-2.484	-1.815	-1.733	-1.392	-1.324
	400	-2.515	-2.369	-1.779	-1.697	-1.361	-1.293
	425	-2.593	-2.447	-1.784	-1.702	-1.377	-1.309
	450	-2.604	-2.458	-1.783	-1.701	-1.349	-1.281
	475	-2.554	-2.408	-1.810	-1.728	-1.392	-1.324
	500	-2.495	-2.349	-1.781	-1.699	-1.371	-1.303
20	50	-2.621	-2.475	-1.802	-1.720	-1.344	-1.276
	75	-2.464	-2.318	-1.757	-1.675	-1.368	-1.300
	100	-2.499	-2.353	-1.702	-1.620	-1.318	-1.250
	125	-2.580	-2.434	-1.757	-1.675	-1.342	-1.274
	150	-2.542	-2.396	-1.717	-1.635	-1.299	-1.231
	175	-2.625	-2.479	-1.764	-1.682	-1.356	-1.288
	200	-2.514	-2.368	-1.704	-1.622	-1.320	-1.252
	225	-2.545	-2.399	-1.752	-1.670	-1.345	-1.277
	250	-2.516	-2.370	-1.738	-1.656	-1.345	-1.277
	275	-2.570	-2.424	-1.721	-1.639	-1.315	-1.247
	300	-2.533	-2.387	-1.746	-1.664	-1.318	-1.250
	325	-2.518	-2.372	-1.724	-1.642	-1.318	-1.250
	350	-2.532	-2.386	-1.690	-1.608	-1.300	-1.232
	375	-2.536	-2.390	-1.732	-1.650	-1.306	-1.238
	400	-2.504	-2.358	-1.726	-1.644	-1.330	-1.262
	425	-2.487	-2.341	-1.663	-1.581	-1.302	-1.234
	450	-2.538	-2.392	-1.756	-1.674	-1.321	-1.253
	475	-2.500	-2.354	-1.728	-1.646	-1.336	-1.268
	500	-2.540	-2.394	-1.719	-1.637	-1.305	-1.237
30	50	-2.642	-2.496	-1.746	-1.664	-1.312	-1.244
	75	-2.470	-2.324	-1.700	-1.618	-1.274	-1.206
	100	-2.573	-2.427	-1.726	-1.644	-1.303	-1.235
	125	-2.453	-2.307	-1.664	-1.582	-1.282	-1.214
	150	-2.526	-2.380	-1.684	-1.602	-1.291	-1.223
	175	-2.510	-2.364	-1.700	-1.618	-1.287	-1.219
	200	-2.479	-2.333	-1.697	-1.615	-1.261	-1.193
	225	-2.463	-2.317	-1.696	-1.614	-1.313	-1.245
	250	-2.483	-2.337	-1.714	-1.632	-1.313	-1.245
	275	-2.531	-2.385	-1.709	-1.627	-1.308	-1.240
	300	-2.464	-2.318	-1.661	-1.579	-1.274	-1.206
	325	-2.498	-2.352	-1.693	-1.611	-1.288	-1.220
	350	-2.490	-2.344	-1.695	-1.613	-1.289	-1.221
	375	-2.464	-2.318	-1.718	-1.636	-1.308	-1.240

Chang (2002) test under weak dependence (DGP1), Cont'd

N	*T*	1%		5%		10%	
30	400	-2.407	-2.261	-1.664	-1.582	-1.247	-1.179
	425	-2.474	-2.328	-1.709	-1.627	-1.312	-1.244
	450	-2.523	-2.377	-1.686	-1.604	-1.276	-1.208
	475	-2.515	-2.369	-1.705	-1.623	-1.282	-1.214
	500	-2.423	-2.277	-1.676	-1.594	-1.233	-1.165
40	50	-2.584	-2.438	-1.797	-1.715	-1.344	-1.276
	75	-2.521	-2.375	-1.704	-1.622	-1.315	-1.247
	100	-2.466	-2.320	-1.702	-1.620	-1.324	-1.256
	125	-2.538	-2.392	-1.712	-1.630	-1.279	-1.211
	150	-2.498	-2.352	-1.706	-1.624	-1.288	-1.220
	175	-2.535	-2.389	-1.734	-1.652	-1.297	-1.229
	200	-2.518	-2.372	-1.721	-1.639	-1.324	-1.256
	225	-2.581	-2.435	-1.743	-1.661	-1.331	-1.263
	250	-2.593	-2.447	-1.744	-1.662	-1.325	-1.257
	275	-2.428	-2.282	-1.696	-1.614	-1.289	-1.221
	300	-2.516	-2.370	-1.739	-1.657	-1.315	-1.247
	325	-2.576	-2.430	-1.759	-1.677	-1.307	-1.239
	350	-2.499	-2.353	-1.699	-1.617	-1.293	-1.225
	375	-2.522	-2.376	-1.731	-1.649	-1.323	-1.255
	400	-2.518	-2.372	-1.727	-1.645	-1.311	-1.243
	425	-2.519	-2.373	-1.707	-1.625	-1.297	-1.229
	450	-2.506	-2.360	-1.718	-1.636	-1.273	-1.205
	475	-2.496	-2.350	-1.706	-1.624	-1.287	-1.219
	500	-2.517	-2.371	-1.704	-1.622	-1.285	-1.217
50	50	-2.492	-2.346	-1.685	-1.603	-1.270	-1.202
	75	-2.498	-2.352	-1.678	-1.596	-1.271	-1.203
	100	-2.492	-2.346	-1.713	-1.631	-1.289	-1.221
	125	-2.421	-2.275	-1.648	-1.566	-1.229	-1.161
	150	-2.519	-2.373	-1.735	-1.653	-1.275	-1.207
	175	-2.506	-2.360	-1.720	-1.638	-1.287	-1.219
	200	-2.458	-2.312	-1.660	-1.578	-1.202	-1.134
	225	-2.495	-2.349	-1.689	-1.607	-1.261	-1.193
	250	-2.513	-2.367	-1.668	-1.586	-1.267	-1.199
	275	-2.510	-2.364	-1.698	-1.616	-1.267	-1.199
	300	-2.487	-2.341	-1.683	-1.601	-1.247	-1.179
	325	-2.506	-2.360	-1.631	-1.549	-1.214	-1.146
	350	-2.459	-2.313	-1.679	-1.597	-1.247	-1.179
	375	-2.436	-2.290	-1.631	-1.549	-1.233	-1.165
	400	-2.422	-2.276	-1.634	-1.552	-1.241	-1.173
	425	-2.434	-2.288	-1.643	-1.561	-1.260	-1.192
	450	-2.506	-2.360	-1.646	-1.564	-1.240	-1.172
	475	-2.460	-2.314	-1.677	-1.595	-1.276	-1.208
	500	-2.442	-2.296	-1.651	-1.569	-1.255	-1.187
60	50	-2.466	-2.320	-1.706	-1.624	-1.267	-1.199
	75	-2.404	-2.258	-1.650	-1.568	-1.230	-1.162
	100	-2.536	-2.390	-1.694	-1.612	-1.273	-1.205
	125	-2.440	-2.294	-1.660	-1.578	-1.268	-1.200
	150	-2.514	-2.368	-1.737	-1.655	-1.297	-1.229
	175	-2.457	-2.311	-1.682	-1.600	-1.255	-1.187
	200	-2.460	-2.314	-1.673	-1.591	-1.271	-1.203
	225	-2.383	-2.237	-1.649	-1.567	-1.222	-1.154
	250	-2.410	-2.264	-1.695	-1.613	-1.261	-1.193
	275	-2.517	-2.371	-1.661	-1.579	-1.251	-1.183
	300	-2.465	-2.319	-1.666	-1.584	-1.241	-1.173
	325	-2.498	-2.352	-1.692	-1.610	-1.239	-1.171

Chang (2002) test under weak dependence (DGP1), Cont'd

N	*T*	1%		5%		10%	
60	350	-2.400	-2.254	-1.679	-1.597	-1.270	-1.202
	375	-2.398	-2.252	-1.662	-1.580	-1.223	-1.155
	400	-2.544	-2.398	-1.690	-1.608	-1.244	-1.176
	425	-2.487	-2.341	-1.665	-1.583	-1.249	-1.181
	450	-2.470	-2.324	-1.665	-1.583	-1.215	-1.147
	475	-2.434	-2.288	-1.646	-1.564	-1.221	-1.153
	500	-2.517	-2.371	-1.637	-1.555	-1.241	-1.173
70	50	-2.486	-2.340	-1.641	-1.559	-1.245	-1.177
	75	-2.462	-2.316	-1.626	-1.544	-1.214	-1.146
	100	-2.354	-2.208	-1.569	-1.487	-1.177	-1.109
	125	-2.373	-2.227	-1.635	-1.553	-1.211	-1.143
	150	-2.484	-2.338	-1.620	-1.538	-1.188	-1.120
	175	-2.383	-2.237	-1.609	-1.527	-1.185	-1.117
	200	-2.429	-2.283	-1.610	-1.528	-1.185	-1.117
	225	-2.299	-2.153	-1.582	-1.500	-1.212	-1.144
	250	-2.410	-2.264	-1.592	-1.510	-1.199	-1.131
	275	-2.356	-2.210	-1.570	-1.488	-1.169	-1.101
	300	-2.415	-2.269	-1.610	-1.528	-1.205	-1.137
	325	-2.359	-2.213	-1.600	-1.518	-1.190	-1.122
	350	-2.334	-2.188	-1.596	-1.514	-1.188	-1.120
	375	-2.405	-2.259	-1.608	-1.526	-1.188	-1.120
	400	-2.341	-2.195	-1.571	-1.489	-1.155	-1.087
	425	-2.391	-2.245	-1.593	-1.511	-1.169	-1.101
	450	-2.286	-2.140	-1.568	-1.486	-1.174	-1.106
	475	-2.476	-2.330	-1.656	-1.574	-1.241	-1.173
	500	-2.378	-2.232	-1.597	-1.515	-1.201	-1.133
80	50	-2.397	-2.251	-1.637	-1.555	-1.226	-1.158
	75	-2.405	-2.259	-1.648	-1.566	-1.236	-1.168
	100	-2.405	-2.259	-1.590	-1.508	-1.207	-1.139
	125	-2.342	-2.196	-1.587	-1.505	-1.160	-1.092
	150	-2.370	-2.224	-1.546	-1.464	-1.152	-1.084
	175	-2.412	-2.266	-1.608	-1.526	-1.192	-1.124
	200	-2.381	-2.235	-1.572	-1.490	-1.171	-1.103
	225	-2.475	-2.329	-1.604	-1.522	-1.213	-1.145
	250	-2.390	-2.244	-1.587	-1.505	-1.176	-1.108
	275	-2.296	-2.150	-1.506	-1.424	-1.145	-1.077
	300	-2.392	-2.246	-1.587	-1.505	-1.180	-1.112
	325	-2.379	-2.233	-1.592	-1.510	-1.194	-1.126
	350	-2.411	-2.265	-1.605	-1.523	-1.177	-1.109
	375	-2.382	-2.236	-1.599	-1.517	-1.163	-1.095
	400	-2.401	-2.255	-1.584	-1.502	-1.174	-1.106
	425	-2.402	-2.256	-1.600	-1.518	-1.190	-1.122
	450	-2.360	-2.214	-1.562	-1.480	-1.154	-1.086
	475	-2.329	-2.183	-1.556	-1.474	-1.149	-1.081
	500	-2.333	-2.187	-1.576	-1.494	-1.196	-1.128
90	50	-2.440	-2.294	-1.587	-1.505	-1.202	-1.134
	75	-2.399	-2.253	-1.597	-1.515	-1.168	-1.100
	100	-2.335	-2.189	-1.593	-1.511	-1.164	-1.096
	125	-2.274	-2.128	-1.537	-1.455	-1.139	-1.071
	150	-2.384	-2.238	-1.554	-1.472	-1.132	-1.064
	175	-2.380	-2.234	-1.553	-1.471	-1.127	-1.059
	200	-2.440	-2.294	-1.610	-1.528	-1.217	-1.149
	225	-2.319	-2.173	-1.527	-1.445	-1.121	-1.053
	250	-2.312	-2.166	-1.560	-1.478	-1.155	-1.087
	275	-2.327	-2.181	-1.536	-1.454	-1.114	-1.046

Chang (2002) test under weak dependence (DGP1), Cont'd

N	*T*	1%		5%		10%	
90	300	-2.326	-2.180	-1.561	-1.479	-1.163	-1.095
	325	-2.282	-2.136	-1.536	-1.454	-1.120	-1.052
	350	-2.293	-2.147	-1.584	-1.502	-1.138	-1.070
	375	-2.356	-2.210	-1.581	-1.499	-1.192	-1.124
	400	-2.312	-2.166	-1.508	-1.426	-1.128	-1.060
	425	-2.402	-2.256	-1.596	-1.514	-1.173	-1.105
	450	-2.315	-2.169	-1.560	-1.478	-1.168	-1.100
	475	-2.349	-2.203	-1.581	-1.499	-1.171	-1.103
	500	-2.306	-2.160	-1.548	-1.466	-1.158	-1.090
100	50	-2.435	-2.289	-1.599	-1.517	-1.185	-1.117
	75	-2.395	-2.249	-1.592	-1.510	-1.201	-1.133
	100	-2.327	-2.181	-1.576	-1.494	-1.175	-1.107
	125	-2.330	-2.184	-1.533	-1.451	-1.134	-1.066
	150	-2.449	-2.303	-1.581	-1.499	-1.164	-1.096
	175	-2.400	-2.254	-1.585	-1.503	-1.167	-1.099
	200	-2.336	-2.190	-1.575	-1.493	-1.151	-1.083
	225	-2.339	-2.193	-1.543	-1.461	-1.133	-1.065
	250	-2.320	-2.174	-1.552	-1.470	-1.137	-1.069
	275	-2.376	-2.230	-1.536	-1.454	-1.136	-1.068
	300	-2.359	-2.213	-1.563	-1.481	-1.148	-1.080
	325	-2.391	-2.245	-1.548	-1.466	-1.144	-1.076
	350	-2.448	-2.302	-1.586	-1.504	-1.154	-1.086
	375	-2.404	-2.258	-1.551	-1.469	-1.130	-1.062
	400	-2.367	-2.221	-1.587	-1.505	-1.161	-1.093
	425	-2.379	-2.233	-1.587	-1.505	-1.160	-1.092
	450	-2.399	-2.253	-1.547	-1.465	-1.108	-1.040
	475	-2.439	-2.293	-1.601	-1.519	-1.178	-1.110
	500	-2.298	-2.152	-1.539	-1.457	-1.119	-1.051

Chang and Song (2005, 2009) test under strong dependence (DGP2)

N	*T*	1%		5%		10%	
10	50	-2.671	-2.525	-1.745	-1.663	-1.298	-1.230
	75	-2.887	-2.741	-1.895	-1.813	-1.410	-1.342
	100	-2.988	-2.842	-1.952	-1.870	-1.491	-1.423
	125	-2.954	-2.808	-1.906	-1.824	-1.436	-1.368
	150	-2.969	-2.823	-1.992	-1.910	-1.500	-1.432
	175	-2.763	-2.617	-1.863	-1.781	-1.420	-1.352
	200	-3.339	-3.193	-2.187	-2.105	-1.630	-1.562
	225	-3.004	-2.858	-2.052	-1.970	-1.563	-1.495
	250	-3.300	-3.154	-2.265	-2.183	-1.716	-1.648
	275	-3.341	-3.195	-2.190	-2.108	-1.654	-1.586
	300	-2.797	-2.651	-1.953	-1.871	-1.523	-1.455
	325	-3.095	-2.949	-2.075	-1.993	-1.542	-1.474
	350	-3.280	-3.134	-2.228	-2.146	-1.689	-1.621
	375	-3.141	-2.995	-2.127	-2.045	-1.610	-1.542
	400	-3.233	-3.087	-2.210	-2.128	-1.668	-1.600
	425	-3.328	-3.182	-2.248	-2.166	-1.709	-1.641
	450	-3.222	-3.076	-2.100	-2.018	-1.600	-1.532
	475	-3.470	-3.324	-2.312	-2.230	-1.743	-1.675
	500	-3.130	-2.984	-2.162	-2.080	-1.622	-1.554
20	50	-2.321	-2.175	-1.535	-1.453	-1.148	-1.080
	75	-2.320	-2.174	-1.589	-1.507	-1.193	-1.125
	100	-2.238	-2.092	-1.548	-1.466	-1.188	-1.120

Chang and Song (2005, 2009) test under strong dependence (DGP2), Cont'd

N	*T*	1%		5%		10%	
	125	-2.469	-2.323	-1.656	-1.574	-1.256	-1.188
	150	-2.645	-2.499	-1.775	-1.693	-1.354	-1.286
	175	-2.545	-2.399	-1.693	-1.611	-1.282	-1.214
	200	-2.659	-2.513	-1.811	-1.729	-1.391	-1.323
	225	-2.758	-2.612	-1.892	-1.810	-1.439	-1.371
	250	-2.950	-2.804	-1.967	-1.885	-1.504	-1.436
	275	-2.833	-2.687	-1.860	-1.778	-1.453	-1.385
	300	-2.802	-2.656	-1.884	-1.802	-1.468	-1.400
	325	-2.750	-2.604	-1.940	-1.858	-1.487	-1.419
	350	-3.272	-3.126	-2.166	-2.084	-1.619	-1.551
	375	-2.719	-2.573	-1.860	-1.778	-1.456	-1.388
	400	-2.979	-2.833	-2.021	-1.939	-1.581	-1.513
	425	-2.785	-2.639	-1.889	-1.807	-1.471	-1.403
	450	-2.825	-2.679	-1.949	-1.867	-1.518	-1.450
	475	-3.200	-3.054	-2.201	-2.119	-1.703	-1.635
	500	-3.183	-3.037	-2.184	-2.102	-1.661	-1.593
30	50	-1.843	-1.697	-1.208	-1.126	-0.902	-0.834
	75	-2.060	-1.914	-1.402	-1.320	-1.076	-1.008
	100	-2.019	-1.873	-1.380	-1.298	-1.044	-0.976
	125	-2.275	-2.129	-1.531	-1.449	-1.156	-1.088
	150	-2.177	-2.031	-1.484	-1.402	-1.122	-1.054
	175	-2.404	-2.258	-1.638	-1.556	-1.243	-1.175
	200	-2.495	-2.349	-1.664	-1.582	-1.265	-1.197
	225	-2.674	-2.528	-1.864	-1.782	-1.403	-1.335
	250	-2.542	-2.396	-1.732	-1.650	-1.340	-1.272
	275	-2.656	-2.510	-1.792	-1.710	-1.382	-1.314
	300	-2.596	-2.450	-1.771	-1.689	-1.382	-1.314
	325	-2.691	-2.545	-1.854	-1.772	-1.428	-1.360
	350	-2.926	-2.780	-1.968	-1.886	-1.487	-1.419
	375	-2.779	-2.633	-1.872	-1.790	-1.462	-1.394
	400	-2.718	-2.572	-1.854	-1.772	-1.458	-1.390
	425	-2.819	-2.673	-1.914	-1.832	-1.490	-1.422
	450	-2.924	-2.778	-2.008	-1.926	-1.524	-1.456
	475	-2.955	-2.809	-2.007	-1.925	-1.577	-1.509
	500	-2.856	-2.710	-1.981	-1.899	-1.517	-1.449
40	50	-1.890	-1.744	-1.236	-1.154	-0.930	-0.862
	75	-1.999	-1.853	-1.329	-1.247	-0.999	-0.931
	100	-2.117	-1.971	-1.394	-1.312	-1.043	-0.975
	125	-1.988	-1.842	-1.374	-1.292	-1.042	-0.974
	150	-2.267	-2.121	-1.501	-1.419	-1.135	-1.067
	175	-2.366	-2.220	-1.591	-1.509	-1.210	-1.142
	200	-2.403	-2.257	-1.632	-1.550	-1.255	-1.187
	225	-2.298	-2.152	-1.588	-1.506	-1.228	-1.160
	250	-2.299	-2.153	-1.600	-1.518	-1.223	-1.155
	275	-2.724	-2.578	-1.842	-1.760	-1.403	-1.335
	300	-2.293	-2.147	-1.573	-1.491	-1.235	-1.167
	325	-2.562	-2.416	-1.758	-1.676	-1.362	-1.294
	350	-2.737	-2.591	-1.855	-1.773	-1.423	-1.355
	375	-2.620	-2.474	-1.796	-1.714	-1.391	-1.323
	400	-2.368	-2.222	-1.639	-1.557	-1.255	-1.187
	425	-2.631	-2.485	-1.849	-1.767	-1.435	-1.367
	450	-2.612	-2.466	-1.822	-1.740	-1.404	-1.336
	475	-2.842	-2.696	-1.898	-1.816	-1.481	-1.413
	500	-2.759	-2.613	-1.856	-1.774	-1.435	-1.367
50	50	-1.740	-1.594	-1.218	-1.136	-0.925	-0.857

Chang and Song (2005, 2009) test under strong dependence (DGP2), Cont'd

N	*T*	1%		5%		10%	
50	75	-1.879	-1.733	-1.263	-1.181	-0.968	-0.900
	100	-1.928	-1.782	-1.313	-1.231	-0.981	-0.913
	125	-2.001	-1.855	-1.345	-1.263	-0.997	-0.929
	150	-1.977	-1.831	-1.362	-1.280	-1.030	-0.962
	175	-2.155	-2.009	-1.424	-1.342	-1.100	-1.032
	200	-2.206	-2.060	-1.484	-1.402	-1.115	-1.047
	225	-2.282	-2.136	-1.560	-1.478	-1.204	-1.136
	250	-2.069	-1.923	-1.419	-1.337	-1.122	-1.054
	275	-2.344	-2.198	-1.623	-1.541	-1.255	-1.187
	300	-2.242	-2.096	-1.548	-1.466	-1.211	-1.143
	325	-2.320	-2.174	-1.627	-1.545	-1.276	-1.208
	350	-2.382	-2.236	-1.615	-1.533	-1.229	-1.161
	375	-2.283	-2.137	-1.585	-1.503	-1.282	-1.214
	400	-2.442	-2.296	-1.688	-1.606	-1.307	-1.239
	425	-2.415	-2.269	-1.700	-1.618	-1.337	-1.269
	450	-2.455	-2.309	-1.696	-1.614	-1.338	-1.270
	475	-2.478	-2.332	-1.732	-1.650	-1.352	-1.284
	500	-2.598	-2.452	-1.722	-1.640	-1.356	-1.288
60	50	-1.891	-1.745	-1.262	-1.180	-0.954	-0.886
	75	-1.812	-1.666	-1.241	-1.159	-0.942	-0.874
	100	-1.802	-1.656	-1.258	-1.176	-0.956	-0.888
	125	-2.035	-1.889	-1.365	-1.283	-1.009	-0.941
	150	-2.025	-1.879	-1.387	-1.305	-1.062	-0.994
	175	-2.112	-1.966	-1.436	-1.354	-1.079	-1.011
	200	-2.225	-2.079	-1.499	-1.417	-1.133	-1.065
	225	-2.174	-2.028	-1.490	-1.408	-1.146	-1.078
	250	-2.344	-2.198	-1.578	-1.496	-1.217	-1.149
	275	-2.303	-2.157	-1.590	-1.508	-1.193	-1.125
	300	-2.259	-2.113	-1.559	-1.477	-1.217	-1.149
	325	-2.391	-2.245	-1.618	-1.536	-1.266	-1.198
	350	-2.322	-2.176	-1.622	-1.540	-1.249	-1.181
	375	-2.231	-2.085	-1.537	-1.455	-1.217	-1.149
	400	-2.265	-2.119	-1.567	-1.485	-1.234	-1.166
	425	-2.387	-2.241	-1.729	-1.647	-1.366	-1.298
	450	-2.447	-2.301	-1.694	-1.612	-1.335	-1.267
	475	-2.472	-2.326	-1.680	-1.598	-1.333	-1.265
	500	-2.598	-2.452	-1.825	-1.743	-1.395	-1.327
70	50	-1.757	-1.611	-1.259	-1.177	-0.952	-0.884
	75	-1.895	-1.749	-1.248	-1.166	-0.948	-0.880
	100	-1.822	-1.676	-1.234	-1.152	-0.933	-0.865
	125	-2.007	-1.861	-1.335	-1.253	-0.993	-0.925
	150	-1.776	-1.630	-1.236	-1.154	-0.940	-0.872
	175	-1.981	-1.835	-1.347	-1.265	-1.003	-0.935
	200	-1.960	-1.814	-1.342	-1.260	-1.018	-0.950
	225	-2.102	-1.956	-1.448	-1.366	-1.090	-1.022
	250	-1.852	-1.706	-1.292	-1.210	-1.000	-0.932
	275	-1.939	-1.793	-1.349	-1.267	-1.058	-0.990
	300	-2.227	-2.081	-1.479	-1.397	-1.127	-1.059
	325	-2.168	-2.022	-1.521	-1.439	-1.189	-1.121
	350	-2.222	-2.076	-1.501	-1.419	-1.146	-1.078
	375	-2.450	-2.304	-1.638	-1.556	-1.253	-1.185
	400	-2.153	-2.007	-1.499	-1.417	-1.176	-1.108
	425	-2.311	-2.165	-1.613	-1.531	-1.259	-1.191
	450	-2.271	-2.125	-1.605	-1.523	-1.264	-1.196
	475	-2.298	-2.152	-1.610	-1.528	-1.269	-1.201

Chang and Song (2005, 2009) test under strong dependence (DGP2), Cont'd

N	*T*	1%		5%		10%	
70	500	-2.148	-2.002	-1.484	-1.402	-1.196	-1.128
80	50	-1.841	-1.695	-1.241	-1.159	-0.964	-0.896
	75	-1.816	-1.670	-1.265	-1.183	-0.956	-0.888
	100	-1.883	-1.737	-1.283	-1.201	-0.989	-0.921
	125	-1.811	-1.665	-1.265	-1.183	-0.970	-0.902
	150	-1.908	-1.762	-1.254	-1.172	-0.952	-0.884
	175	-1.791	-1.645	-1.243	-1.161	-0.935	-0.867
	200	-2.058	-1.912	-1.379	-1.297	-1.044	-0.976
	225	-1.797	-1.651	-1.246	-1.164	-0.963	-0.895
	250	-2.010	-1.864	-1.380	-1.298	-1.059	-0.991
	275	-1.992	-1.846	-1.380	-1.298	-1.058	-0.990
	300	-2.053	-1.907	-1.410	-1.328	-1.096	-1.028
	325	-2.195	-2.049	-1.498	-1.416	-1.170	-1.102
	350	-2.176	-2.030	-1.531	-1.449	-1.172	-1.104
	375	-2.124	-1.978	-1.469	-1.387	-1.143	-1.075
	400	-1.995	-1.849	-1.410	-1.328	-1.136	-1.068
	425	-2.355	-2.209	-1.623	-1.541	-1.257	-1.189
	450	-2.098	-1.952	-1.457	-1.375	-1.144	-1.076
	475	-2.176	-2.030	-1.542	-1.460	-1.216	-1.148
	500	-2.112	-1.966	-1.541	-1.459	-1.224	-1.156
90	50	-1.914	-1.768	-1.310	-1.228	-1.006	-0.938
	75	-1.905	-1.759	-1.325	-1.243	-0.998	-0.930
	100	-1.901	-1.755	-1.314	-1.232	-0.983	-0.915
	125	-1.817	-1.671	-1.253	-1.171	-0.926	-0.858
	150	-1.994	-1.848	-1.338	-1.256	-1.008	-0.940
	175	-1.963	-1.817	-1.358	-1.276	-1.022	-0.954
	200	-1.956	-1.810	-1.349	-1.267	-1.008	-0.940
	225	-1.794	-1.648	-1.244	-1.162	-0.969	-0.901
	250	-1.894	-1.748	-1.301	-1.219	-1.016	-0.948
	275	-2.037	-1.891	-1.401	-1.319	-1.058	-0.990
	300	-2.067	-1.921	-1.451	-1.369	-1.107	-1.039
	325	-2.157	-2.011	-1.470	-1.388	-1.132	-1.064
	350	-2.135	-1.989	-1.457	-1.375	-1.121	-1.053
	375	-2.262	-2.116	-1.589	-1.507	-1.206	-1.138
	400	-2.166	-2.020	-1.470	-1.388	-1.145	-1.077
	425	-2.232	-2.086	-1.556	-1.474	-1.211	-1.143
	450	-2.210	-2.064	-1.532	-1.450	-1.199	-1.131
	475	-2.234	-2.088	-1.573	-1.491	-1.240	-1.172
	500	-2.368	-2.222	-1.644	-1.562	-1.286	-1.218
100	50	-1.934	-1.788	-1.350	-1.268	-1.050	-0.982
	75	-1.915	-1.769	-1.332	-1.250	-1.011	-0.943
	100	-1.858	-1.712	-1.269	-1.187	-0.958	-0.890
	125	-1.877	-1.731	-1.291	-1.209	-0.990	-0.922
	150	-1.720	-1.574	-1.193	-1.111	-0.918	-0.850
	175	-1.858	-1.712	-1.271	-1.189	-0.942	-0.874
	200	-1.863	-1.717	-1.274	-1.192	-0.979	-0.911
	225	-1.965	-1.819	-1.313	-1.231	-0.986	-0.918
	250	-1.945	-1.799	-1.349	-1.267	-1.037	-0.969
	275	-2.022	-1.876	-1.381	-1.299	-1.033	-0.965
	300	-2.138	-1.992	-1.414	-1.332	-1.107	-1.039
	325	-2.134	-1.988	-1.381	-1.299	-1.050	-0.982
	350	-2.113	-1.967	-1.460	-1.378	-1.082	-1.014
	375	-2.106	-1.960	-1.441	-1.359	-1.128	-1.060
	400	-2.131	-1.985	-1.498	-1.416	-1.176	-1.108
	425	-1.991	-1.845	-1.409	-1.327	-1.105	-1.037

Chang and Song (2005, 2009) test under strong dependence (DGP2), Cont'd

N	*T*	1%		5%		10%	
100	450	-2.108	-1.962	-1.460	-1.378	-1.139	-1.071
	475	-2.162	-2.016	-1.504	-1.422	-1.183	-1.115
	500	-2.181	-2.035	-1.504	-1.422	-1.202	-1.134

Chang and Song (2005, 2009) test under long run dependence (DGP3)

N	*T*	1%		5%		10%	
10	50	-3.551	-3.405	-2.205	-2.123	-1.572	-1.504
	75	-4.193	-4.047	-2.507	-2.425	-1.712	-1.644
	100	-4.636	-4.490	-2.735	-2.653	-1.838	-1.770
	125	-4.768	-4.622	-2.813	-2.731	-1.907	-1.839
	150	-4.912	-4.766	-3.046	-2.964	-2.036	-1.968
	175	-5.353	-5.207	-3.263	-3.181	-2.188	-2.120
	200	-5.218	-5.072	-3.162	-3.080	-2.138	-2.070
	225	-5.581	-5.435	-3.483	-3.401	-2.466	-2.398
	250	-5.295	-5.149	-3.249	-3.167	-2.243	-2.175
	275	-5.481	-5.335	-3.388	-3.306	-2.319	-2.251
	300	-5.545	-5.399	-3.417	-3.335	-2.355	-2.287
	325	-5.483	-5.337	-3.540	-3.458	-2.496	-2.428
	350	-5.548	-5.402	-3.539	-3.457	-2.352	-2.284
	375	-5.619	-5.473	-3.479	-3.397	-2.387	-2.319
	400	-5.825	-5.679	-3.740	-3.658	-2.568	-2.500
	425	-5.833	-5.687	-3.723	-3.641	-2.608	-2.540
	450	-5.552	-5.406	-3.517	-3.435	-2.421	-2.353
	475	-6.276	-6.130	-3.995	-3.913	-2.768	-2.700
	500	-5.772	-5.626	-3.615	-3.533	-2.474	-2.406
20	50	-3.118	-2.972	-1.873	-1.791	-1.322	-1.254
	75	-3.547	-3.401	-2.145	-2.063	-1.441	-1.373
	100	-3.991	-3.845	-2.296	-2.214	-1.527	-1.459
	125	-4.195	-4.049	-2.410	-2.328	-1.642	-1.574
	150	-4.544	-4.398	-2.522	-2.440	-1.677	-1.609
	175	-5.078	-4.932	-2.848	-2.766	-1.841	-1.773
	200	-5.099	-4.953	-2.757	-2.675	-1.742	-1.674
	225	-5.440	-5.294	-3.028	-2.946	-1.984	-1.916
	250	-5.385	-5.239	-2.932	-2.850	-1.884	-1.816
	275	-5.460	-5.314	-3.101	-3.019	-1.978	-1.910
	300	-5.586	-5.440	-3.091	-3.009	-2.038	-1.970
	325	-5.520	-5.374	-3.158	-3.076	-2.132	-2.064
	350	-5.783	-5.637	-3.325	-3.243	-2.135	-2.067
	375	-5.768	-5.622	-3.270	-3.188	-2.166	-2.098
	400	-6.293	-6.147	-3.519	-3.437	-2.337	-2.269
	425	-6.238	-6.092	-3.456	-3.374	-2.229	-2.161
	450	-5.880	-5.734	-3.168	-3.086	-2.073	-2.005
	475	-6.444	-6.298	-3.546	-3.464	-2.211	-2.143
	500	-6.294	-6.148	-3.406	-3.324	-2.148	-2.080
30	50	-2.842	-2.696	-1.829	-1.747	-1.337	-1.269
	75	-3.321	-3.175	-1.937	-1.855	-1.354	-1.286
	100	-3.761	-3.615	-2.099	-2.017	-1.449	-1.381
	125	-3.746	-3.600	-2.151	-2.069	-1.469	-1.401
	150	-4.121	-3.975	-2.332	-2.250	-1.514	-1.446
	175	-4.518	-4.372	-2.529	-2.447	-1.729	-1.661
	200	-4.443	-4.297	-2.457	-2.375	-1.621	-1.553
	225	-4.856	-4.710	-2.660	-2.578	-1.797	-1.729
	250	-4.917	-4.771	-2.700	-2.618	-1.846	-1.778

Chang and Song (2005, 2009) test under long run dependence (DGP3), Cont'd

N	*T*	1%		5%		10%	
30	275	-5.215	-5.069	-2.808	-2.726	-1.839	-1.771
	300	-5.042	-4.896	-2.694	-2.612	-1.789	-1.721
	325	-5.064	-4.918	-2.807	-2.725	-1.909	-1.841
	350	-5.227	-5.081	-2.881	-2.799	-1.903	-1.835
	375	-5.480	-5.334	-2.949	-2.867	-2.005	-1.937
	400	-5.488	-5.342	-2.983	-2.901	-2.017	-1.949
	425	-5.732	-5.586	-3.084	-3.002	-2.026	-1.958
	450	-5.652	-5.506	-2.965	-2.883	-2.008	-1.940
	475	-5.978	-5.832	-3.079	-2.997	-2.001	-1.933
	500	-5.874	-5.728	-3.100	-3.018	-2.088	-2.020
40	50	-2.898	-2.752	-1.824	-1.742	-1.329	-1.261
	75	-3.084	-2.938	-1.897	-1.815	-1.350	-1.282
	100	-3.453	-3.307	-2.076	-1.994	-1.436	-1.368
	125	-3.681	-3.535	-2.096	-2.014	-1.467	-1.399
	150	-3.829	-3.683	-2.212	-2.130	-1.522	-1.454
	175	-4.257	-4.111	-2.536	-2.454	-1.765	-1.697
	200	-4.324	-4.178	-2.458	-2.376	-1.701	-1.633
	225	-4.590	-4.444	-2.653	-2.571	-1.821	-1.753
	250	-4.363	-4.217	-2.476	-2.394	-1.698	-1.630
	275	-4.862	-4.716	-2.698	-2.616	-1.875	-1.807
	300	-4.745	-4.599	-2.655	-2.573	-1.851	-1.783
	325	-4.652	-4.506	-2.640	-2.558	-1.837	-1.769
	350	-4.774	-4.628	-2.679	-2.597	-1.861	-1.793
	375	-4.913	-4.767	-2.752	-2.670	-1.927	-1.859
	400	-5.312	-5.166	-2.968	-2.886	-2.063	-1.995
	425	-5.331	-5.185	-2.835	-2.753	-1.934	-1.866
	450	-5.337	-5.191	-2.895	-2.813	-1.941	-1.873
	475	-5.438	-5.292	-2.926	-2.844	-1.961	-1.893
	500	-5.581	-5.435	-2.922	-2.840	-2.002	-1.934
50	50	-2.755	-2.609	-1.835	-1.753	-1.333	-1.265
	75	-2.978	-2.832	-1.815	-1.733	-1.302	-1.234
	100	-3.168	-3.022	-1.966	-1.884	-1.354	-1.286
	125	-3.406	-3.260	-2.008	-1.926	-1.421	-1.353
	150	-3.688	-3.542	-2.148	-2.066	-1.534	-1.466
	175	-4.135	-3.989	-2.460	-2.378	-1.762	-1.694
	200	-3.859	-3.713	-2.215	-2.133	-1.478	-1.410
	225	-4.249	-4.103	-2.460	-2.378	-1.731	-1.663
	250	-4.270	-4.124	-2.435	-2.353	-1.728	-1.660
	275	-4.607	-4.461	-2.529	-2.447	-1.771	-1.703
	300	-4.452	-4.306	-2.428	-2.346	-1.715	-1.647
	325	-4.504	-4.358	-2.607	-2.525	-1.849	-1.781
	350	-4.483	-4.337	-2.475	-2.393	-1.711	-1.643
	375	-4.555	-4.409	-2.603	-2.521	-1.842	-1.774
	400	-4.834	-4.688	-2.699	-2.617	-1.862	-1.794
	425	-4.991	-4.845	-2.799	-2.717	-1.949	-1.881
	450	-4.903	-4.757	-2.746	-2.664	-1.882	-1.814
	475	-5.258	-5.112	-2.869	-2.787	-2.017	-1.949
	500	-5.129	-4.983	-2.830	-2.748	-1.912	-1.844
60	50	-2.846	-2.700	-1.821	-1.739	-1.339	-1.271
	75	-2.929	-2.783	-1.849	-1.767	-1.340	-1.272
	100	-3.129	-2.983	-1.934	-1.852	-1.386	-1.318
	125	-3.323	-3.177	-1.989	-1.907	-1.385	-1.317
	150	-3.462	-3.316	-2.107	-2.025	-1.494	-1.426
	175	-3.659	-3.513	-2.196	-2.114	-1.493	-1.425
	200	-3.712	-3.566	-2.169	-2.087	-1.508	-1.440

Chang and Song (2005, 2009) test under long run dependence (DGP3), Cont'd

N	*T*	1%		5%		10%	
60	225	-4.000	-3.854	-2.237	-2.155	-1.576	-1.508
	250	-3.965	-3.819	-2.265	-2.183	-1.601	-1.533
	275	-4.223	-4.077	-2.426	-2.344	-1.683	-1.615
	300	-4.016	-3.870	-2.259	-2.177	-1.588	-1.520
	325	-4.164	-4.018	-2.420	-2.338	-1.701	-1.633
	350	-4.265	-4.119	-2.452	-2.370	-1.729	-1.661
	375	-4.398	-4.252	-2.479	-2.397	-1.783	-1.715
	400	-4.486	-4.340	-2.613	-2.531	-1.817	-1.749
	425	-4.670	-4.524	-2.657	-2.575	-1.873	-1.805
	450	-4.733	-4.587	-2.640	-2.558	-1.858	-1.790
	475	-4.723	-4.577	-2.709	-2.627	-1.885	-1.817
	500	-4.656	-4.510	-2.700	-2.618	-1.871	-1.803
70	50	-2.780	-2.634	-1.848	-1.766	-1.386	-1.318
	75	-2.947	-2.801	-1.869	-1.787	-1.380	-1.312
	100	-3.086	-2.940	-1.920	-1.838	-1.348	-1.280
	125	-3.192	-3.046	-1.947	-1.865	-1.373	-1.305
	150	-3.380	-3.234	-2.052	-1.970	-1.447	-1.379
	175	-3.702	-3.556	-2.204	-2.122	-1.543	-1.475
	200	-3.609	-3.463	-2.057	-1.975	-1.414	-1.346
	225	-3.635	-3.489	-2.138	-2.056	-1.505	-1.437
	250	-3.942	-3.796	-2.237	-2.155	-1.544	-1.476
	275	-3.899	-3.753	-2.358	-2.276	-1.646	-1.578
	300	-4.011	-3.865	-2.318	-2.236	-1.656	-1.588
	325	-4.354	-4.208	-2.447	-2.365	-1.760	-1.692
	350	-4.283	-4.137	-2.411	-2.329	-1.745	-1.677
	375	-4.129	-3.983	-2.360	-2.278	-1.693	-1.625
	400	-4.253	-4.107	-2.491	-2.409	-1.774	-1.706
	425	-4.402	-4.256	-2.577	-2.495	-1.848	-1.780
	450	-4.350	-4.204	-2.638	-2.556	-1.858	-1.790
	475	-4.666	-4.520	-2.579	-2.497	-1.807	-1.739
	500	-4.591	-4.445	-2.592	-2.510	-1.790	-1.722
80	50	-2.855	-2.709	-1.881	-1.799	-1.394	-1.326
	75	-2.962	-2.816	-1.854	-1.772	-1.341	-1.273
	100	-3.001	-2.855	-1.891	-1.809	-1.352	-1.284
	125	-3.219	-3.073	-1.945	-1.863	-1.401	-1.333
	150	-3.271	-3.125	-1.980	-1.898	-1.397	-1.329
	175	-3.546	-3.400	-2.172	-2.090	-1.516	-1.448
	200	-3.513	-3.367	-2.030	-1.948	-1.439	-1.371
	225	-3.741	-3.595	-2.271	-2.189	-1.592	-1.524
	250	-3.641	-3.495	-2.164	-2.082	-1.516	-1.448
	275	-3.980	-3.834	-2.282	-2.200	-1.593	-1.525
	300	-3.902	-3.756	-2.308	-2.226	-1.639	-1.571
	325	-4.002	-3.856	-2.377	-2.295	-1.681	-1.613
	350	-3.999	-3.853	-2.383	-2.301	-1.707	-1.639
	375	-4.245	-4.099	-2.416	-2.334	-1.692	-1.624
	400	-4.134	-3.988	-2.382	-2.300	-1.693	-1.625
	425	-4.412	-4.266	-2.552	-2.470	-1.817	-1.749
	450	-4.467	-4.321	-2.535	-2.453	-1.818	-1.750
	475	-4.452	-4.306	-2.591	-2.509	-1.853	-1.785
	500	-4.594	-4.448	-2.642	-2.560	-1.880	-1.812
90	50	-2.878	-2.732	-1.868	-1.786	-1.412	-1.344
	75	-2.875	-2.729	-1.846	-1.764	-1.352	-1.284
	100	-3.115	-2.969	-1.853	-1.771	-1.331	-1.263
	125	-3.057	-2.911	-1.834	-1.752	-1.315	-1.247
	150	-3.311	-3.165	-1.962	-1.880	-1.379	-1.311

Chang and Song (2005, 2009) test under long run dependence (DGP3), Cont'd

N	*T*	1%		5%		10%	
	175	-3.442	-3.296	-2.089	-2.007	-1.457	-1.389
	200	-3.397	-3.251	-1.981	-1.899	-1.383	-1.315
	225	-3.393	-3.247	-2.062	-1.980	-1.444	-1.376
	250	-3.651	-3.505	-2.068	-1.986	-1.448	-1.380
	275	-3.580	-3.434	-2.183	-2.101	-1.565	-1.497
	300	-3.700	-3.554	-2.216	-2.134	-1.573	-1.505
	325	-4.008	-3.862	-2.266	-2.184	-1.614	-1.546
	350	-3.941	-3.795	-2.296	-2.214	-1.617	-1.549
	375	-4.042	-3.896	-2.348	-2.266	-1.687	-1.619
	400	-4.162	-4.016	-2.384	-2.302	-1.638	-1.570
	425	-4.108	-3.962	-2.460	-2.378	-1.751	-1.683
	450	-4.259	-4.113	-2.446	-2.364	-1.733	-1.665
	475	-4.246	-4.100	-2.437	-2.355	-1.757	-1.689
	500	-4.310	-4.164	-2.493	-2.411	-1.791	-1.723
100	50	-2.912	-2.766	-1.940	-1.858	-1.461	-1.393
	75	-2.883	-2.737	-1.875	-1.793	-1.374	-1.306
	100	-3.088	-2.942	-1.918	-1.836	-1.400	-1.332
	125	-2.955	-2.809	-1.822	-1.740	-1.299	-1.231
	150	-3.264	-3.118	-1.985	-1.903	-1.414	-1.346
	175	-3.420	-3.274	-2.067	-1.985	-1.461	-1.393
	200	-3.228	-3.082	-1.941	-1.859	-1.375	-1.307
	225	-3.299	-3.153	-2.069	-1.987	-1.449	-1.381
	250	-3.538	-3.392	-2.117	-2.035	-1.526	-1.458
	275	-3.827	-3.681	-2.206	-2.124	-1.559	-1.491
	300	-3.716	-3.570	-2.165	-2.083	-1.557	-1.489
	325	-3.865	-3.719	-2.263	-2.181	-1.577	-1.509
	350	-3.968	-3.822	-2.264	-2.182	-1.594	-1.526
	375	-3.987	-3.841	-2.395	-2.313	-1.725	-1.657
	400	-4.012	-3.866	-2.378	-2.296	-1.665	-1.597
	425	-4.008	-3.862	-2.480	-2.398	-1.778	-1.710
	450	-4.236	-4.090	-2.458	-2.376	-1.750	-1.682
	475	-4.119	-3.973	-2.447	-2.365	-1.724	-1.656
	500	-4.263	-4.117	-2.499	-2.417	-1.747	-1.679

References

Ahn, S.K., Fotopoulos S.B. and L. He (2001), Unit Root Tests with Infinite Variance Errors. Econometric Reviews, 20, 461–483.

Asmussen, S. and K. Binswanger (1997), Simulation of Ruin Probabilities for Subexponential Claims. ASTIN Bulletin, 27, 297-318.

Avramidis, A.N. and J.R. Wilson (1998), Correlation-induction Techniques for Estimating Quantiles in Simulation Experiments. Operations Research, 46, 574-591.

Bai, J. and S. Ng (2002), Determining the Number of Factors in Approximate Factor Models. Econometrica, 70, 191–221.

Bai, J. and S. Ng (2004), A Panic Attack on Unit Roots and Cointegration. Econometrica, 72, 1127–1177.

Baillie, R.T. (1989), Commodity Prices and Aggregate Inflation: Would Commodity Price Rule Be Worthwhile?. Carnegie Rochester Conference Series on Public Policy, 31, 185-240.

Baillie, R.T., Chung, C.F. and M.A. Tieslau (1996), Analyzing Inflation by the Fractionally Integrated ARFIMA-GARCH Model. Journal of Applied Econometrics, 11, 23-40.

References

Ball, L. and S.G. Cecchetti (1990), Inflation and Uncertainty at Short and Long Horizons. Brookings Papers on Economic Activity, 215-254.

Baltagi, B.H. (2008), Econometric Analysis of Panel Data, Fourth Edition. John Wiley & Sons.

Baltagi, B.H. and J.M. Griffin (1983), Gasoline Demand in the OECD: an Application of Pooling and Testing Procedures. European Economic Review, 22, 117-137.

Baltagi, B.H. and J.M. Griffin (1988), A General Index of Technique Change. Journal of Political Economy, 96, 20-41.

Baltagi, B. and C. Kao (2000), Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey. Advances in Econometrics, 15, 7-51.

Baltagi, B.H., Bresson, G. and A. Pirotte (2002), Comparison of Forecast Performance for Homogeneous, Heterogeneous and Shrinkage Estimators: Some Empirical Evidence from US Electricity and Natural-gas Consumption. Economics Letters, 76, 375-382.

Banerjee A. (1999), Panel Data Unit Roots and Cointegration: An overview. Oxford Bulletin of Economics and Statistics, 61, 607–629.

Banerjee, A., Dolado, J. J. and R. Mestre (1998), Error-correction Mechanism Tests for Cointegration in a Single-equation Framework. Journal of Time Series Analysis, 19, 267–283.

Banerjee A., Marcellino M. and C. Osbat (2004), Some Cautions on the Use of Panel Methods for Integrated Series of Macroeconomic Data. Econometrics Journal, 7, 322–340.

Banerjee, A., Marcellino, M. and C. Osbat (2005), Testing for PPP: Should we use Panel Methods?. Empirical Economics, 30, 77–91.

References

Beaulieu, J.J. and J.A. Miron (1993), Seasonal Unit Roots in Aggregate U.S. Data. Journal of Econometrics, 55, 305-328.

Benati, L. (2005), The Inflation Targeting Framework from an Historical Perspective. Bank of England Quarterly Bulletin, Summer 2005, 160-168.

Besley, T. (2008), Inflation and the Global Economy. Speech delivered at the Canada-UK Chamber of Commerce, London, 22 April 2008. www.bankofengland.co.uk/publications/speeches/2008/speech343.pdf.

Borio, C. and A. Filardo (2007), Globalisation and Inflation: New Cross-country Evidence on the Global Determinants of domestic Inflation. Bank for International Settlements Working Papers No. 227.

Boschen, J. and C. Weise (2003), What Starts Inflation: Evidence from the OECD Countries. Journal of Money, Credit and Banking, 35, 324-349.

Box, G.E.P. and M.E. Muller (1958), A Note on the Generation of Random Normal Deviates. Annuals of Mathematical Statistics, 29, 610-611.

Box, G.E.P. and K.B. Wilson (1951), On the Experimental Attainment of Optimum Conditions (with discussion). Journal of the Royal Statistical Society Series B, 13, 1-45.

Boyarchenko, S.I. and S.Z. Levendorskii (2000), Option Pricing for Truncated Lévy Processes. International Journal of Theoretical and Applied Finance, 3, 549-552.

Bratley, P., Fox, B.L. and L. Schrage (1987), A Guide to Simulation, Second Edition. Springer-Verlag, New York.

Breitung, J. and S. Das (2005), Panel Unit Root Tests Under Cross Sectional Dependence. Statistica Neerlandica, 59, 414–433.

Breitung, J. and M.H. Pesaran (2008), Unit Roots and Cointegration in Panels, in Matyas, L. and P. Sevestre (eds.), The Econometrics of Panel Data, Third Edition, Ch 9, 279-322. Springer-Verlag, New York.

Brunner, K. and A.H. Meltzer (1977), The Explanation of Inflation: Some International Evidence. The American Economic Review, 67, 148-154.

Burridge, P. and D. Hristova (2008), Consistent Estimation and Order Selection for Nonstationary Autoregressive Process with Stable Innovations. Journal of Time Series Analysis, 29, 695-718.

Cambanis, S. and S.B. Fotopoulos (1995), Conditional Variance for Stable Random Vectors. Probability and Mathematical Statistics, 15, 195–214.

Callegari, F., Cappuccio N. and D. Lubian (2003), Asymptotic Inference in Time Series Regressions with a Unit Root and Infinite Variance Errors. Journal of Statistical Planning and Inference, 116, 277–303.

Calvo, G. and C. Reinhart (2002), Fear of Floating. Quarterly Journal of Economics, 117, 379–408.

Carr, P., Geman, H., Madan, D. and M. Yor (2002), The Fine Structure of Asset Returns: An Empirical Investigation. Journal of Business, 75, 305-332.

Carrion i Silvestre, J.L., Sansó i Rosselló, A. and M. Artís Ortuňo (1999), Response Surfaces Estimates for the Dickey-Fuller Unit Root Test with Structural Breaks. Economics Letters, 63, 279-283.

Cassese, A. and J.R. Lothian (1982), The Timing of Monetary and Price Changes and the International Transmission of Inflation. Journal of Monetary Economics, 10, 1-23.

References

Chambers, J.M., Mallows, C.L. and B.W. Stuck (1976), A Method for Simulating Stable Random Variables. Journal of the American Statistical Association, 71, 340-344.

Chan, N.H. and L.T. Tran (1989), On the First-order Autoregressive Process with Infinite Variance. Econometric Theory, 5, 354–362.

Chang, Y. (2002), Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. Journal of Econometrics, 110, 261–292.

Chang, Y. (2004), Bootstrap Unit Root Tests in Panels with Cross-sectional Dependency. Journal of Econometrics, 120, 263-293.

Chang, Y. and W. Song (2005), Unit Root Tests for Panels in the Presence of Short-run and Long-run Dependencies: Nonlinear IV Approach with Fixed *N* and Large *T*. Manuscript, Rice University.

Chang, Y. and W. Song (2009), Unit Root Tests for Panels in the Presence of Short-run and Long-run Dependencies: Nonlinear IV Approach with Fixed *N* and Large *T*. Review of Economic Studies, 73, 903-935.

Chang, Y., Park, J.Y. and P.C.B. Phillips (2001), Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors. Econometrics Journal, 4, 1-36.

Chang, Y., Sickles, R. and W. Song (2001), Bootstrapping Unit Root Tests with Covariates. (Mimeo, Rice University).

Charemza, W.W. and D.F. Deadman (1997), New Directions in Econometric Practice: General to Specific Modelling, Cointegration, and Vector Autoregression, Second Edition. Edward Elgar, Lyme, NH.

Charemza, W.W., Hristova, D. and P. Burridge (2005), Is Inflation Stationary? Applied Economics, 37, 901-903.

Cheung, Y.W. and K.S. Lai (1995a), Lag Order and Critical Values of the Augmented Dickey-Fuller Test. Journal of Business and Economic Statistics, 13, 277-280.

Cheung, Y.W. and K.S. Lai (1995b), Lag Order and Critical Values of a Modified Dickey-Fuller Test. Oxford Bulleting of Economics and Statistics, 57, 411-419.

Choi, I. (1994), Residual-Based Tests for the Null of Stationarity with Applications to U.S. Macroeconomic Time Series. Econometric Theory, 10,720-746.

Choi, I. (2001), Unit Root Tests for Panel Data. Journal of International Money and Finance, 20, 249-272.

Ciccarelli, M. and B. Mojon (2005), Global Inflation. European Central Bank Working Paper No. 537.

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

Cook, S. (2001), Finite-sample Critical Values for the Augmented Dickey-Fuller Statistic: a Note on Lag Order. Economics Issues, 6, 31-38.

Corvoisier, S, and B. Mojon (2005), Breaks in the Mean Inflation: How They Happen and What to Do with Them. European Central Bank Working Paper No. 451.

Costantini, M. and C. Lupi (2007), An Analysis of Inflation and Interest Rates. New Panel Unit Root Results in the Presence of Structural Breaks. Economics Letters, 95, 408–414.

Culver, S.E. and D.H. Papell (1997), Is There a Unit Root in the Inflation Rate? Evidence from Sequential Break and Panel Data Models. Journal of Applied Econometrics, 12, 436-444.

Darby, M.R. (1983), The United States as an Exogenous Source of World Inflation, in Darby, M.J., Lothian, J.R., Gandolfi, A.E., Schwartz, A.J., and A.C. Stockman (eds.), The International Transmission of Inflation, 478-490. University of Chicago Press, Chicago.

David, F.N. and N.L. Johnson (1954), Statistical Treatment of Censored Data: Part Ⅰ. Fundamental Formulae. Biometrika, 41, 228-240.

Debelle, G. and J. Wilkinson (2002), Inflation Targeting and the Inflation Process: Some Lessons from an Open Economy. Reserve Bank of Australia Research Discussion Paper No. 2002-01.

De Long, J.B. (1997), America's Peacetime Inflation: The 1970's, in Romer, C.D. and D.H. Romer (eds.), Reducing Inflation: Motivation and Strategy, 247-276. University of Chicago Press, Chicago.

Devroye, L. (1986), Non-Uniform Random Variate Generation. Springer-Verlag, New York.

Dickey, D.A. and W.A. Fuller (1979), Distribution of the Estimates for Autoregressive Time Series With a Unit Root. Journal of the American Statistical Association, 74, 427-431.

Dickey, D.A. and W.A. Fuller (1981), Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. Econometrica, 49, 1057-1052.

DuMouchel, W.H. (1983), Estimating the Stable Index α in Order to Measure Tail Thickness: A Critique. The Annals of Statistics, 11, 1019–1031.

Elliott, G., Rothenberg, T.J. and J.H. Stock (1996), Efficient Tests for an Autoregressive Unit Root. Econometrica, 64, 813-836.

Enders, W. (2004), Applied Econometric Time Series, Second Edition. John Wiley & Sons, New York.

Ericsson, N.R. (1986), Post Simulation Analysis of Monte Carlo Experiments: Interpreting Pesaran's (1974) Study of Non-nested Hypothesis Test Statistics. Review of Economic Studies, 53, 691-707.

Ericsson, N.R. and J.G. MacKinnon (2002), Distributions of Error Correction Tests for Cointegration. Econometrics Journal, 5, 285-318.

Fama, E. (1965), The Behavior of Stock-Market Prices. Journal of Business, 38, 34-105.

Feller, W. (1971), An Introduction to Probability Theory and Its Applications, Volume 2, Second Edition. John Wiley & Sons, New York.

Fisher, I, (1930), The theory of interest. McMillan, New York.

Fisher, R.A. (1932), Statistical Methods for Research Workers, Oliver and Bond, Fourth Edition. Edinburgh.

Fishman, G.S. (1996), Monte Carlo: Concept, Algorithms, and Applications. Springer-Verlag, New York.

Flood, R. and M. Taylor (1996), Exchange Rate Economics: What's Wrong with the Conventional Macro Approach?, in Frankel, J., Galli, G. and A. Giovannini (eds.), The Microstructure of Foreign Exchange Markets. University of Chicago Press, Chicago.

Fotopoulos, S.B. (1998), On the Conditional Variance-Covariance of Stable Random Vectors, II, in Karatzas, I., Rajput, B., and M. Taqqu (eds.), Stochastic Processes and Related Topics, 123-140. Birkhäuser, Boston.

Fournié, E., Lasry, J.M. and N. Touzi (1997), Monte Carlo Methods for Stochastic Volatility Models, in Rogers, L.C.G. and D. Talay (eds.), Numerical Methods in Finance, 146-164. Cambridge University Press, Cambridge.

Franses, P.H. and B. Hobijn (1997), Critical Values for Unit Root Tests in Seasonal Time Series. Journal of Applied Statistics, 24, 25-47.

Gallant, A.R. and G. Tauchen (1996), Which Moments to Match?. Econometric Theory, 12, 657-681.

Garcia, R. and P. Perron (1996), An Analysis of the Real Interest Rate under Regime Shifts. Review of Economics and Statistics, 78, 111-125.

Gentle, J.E. (1998), Random Number Generation and Monte Carlo Method. Springer-Verlag, New York.

Glasserman, P. (2004), Monte Carlo Methods in Financial Engineering. Springer Science + Business Media, New York.

Glynn, P.W. and D.L. Iglehart (1988), Simulation Methods for Queues: An Overview. Queuing Systems: Theory and Applications, 3, 221-256.

Gouriéroux, C. and A. Monfort (1996), Simulation-Based Econometric Methods. Oxford University Press, Oxford.

Greene, W.H. (2007), Econometric Analysis, Sixth Edition. Prentice Hall, New Jersey.

Gujarati, D.N. (2009), Basic Econometrics, Fifth Edition. McGraw-Hill Higher Education, New York.

Hadri, K. (2000), Testing for Stationarity in Heterogeneous Panel Data, Econometrics Journal, 3, 148-161.

Hall, R. (1984), Monetary Strategy with an Elastic Price Standard, in Price stability and Public Policy, 137-159, Federal Reserve Bank of Kansas City.

Hamilton, J.D. (1994), Time Series Analysis. Princeton University Press.

Hammersley, J.M. and D.C. Handscomb (1975), Monte Carlo Methods. Methuen, London.

References

Harbo, I., Johansen, S., Nielsen, B. and A. Rahbek (1998), Asymptotic Inference on Cointegrating Rank in Partial Systems. Journal of Business and Economic Statistics, 16, 388-399.

Harvey, A. and D. Bates (2003), Multivariate Unit Root Tests, Stability and Convergence. DAE Working Paper No.0301, University of Cambridge.

Hasen, B.E. (1995), Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power. Econometric Theory, 11, 1148-1171.

Hendry, D.F. (1984), Monte Carlo Experimentation, in Griliches, Z. and M.D. Intriligator (eds.), Handbook of Econometrics, Volume 2, Ch. 16, 939-976. Elsevier Science, Amsterdam.

Hoffman, P. (1998), The Man Who Loved Only Numbers: The Story of Paul Erdős and the Search for Mathematical Truth, 238-239. Hyperion, New York.

Hurd, C.C. (1985), A Note on Early Monte Carlo Computations and Scientific Meetings. Annals of the History of Computing, 7, 141-155.

Hylleberg, S., Engle, R.F., Granger, C.W.J. and B.S. Yoo (1990), Seasonal Integration and Cointegration. Journal of Econometrics, 44, 215-238.

Im, K.S. and M.H. Pesaran (2003), On the Unit Root Testing Using Nonlinear Instrumental Variables. Working Paper Series, University of Cambridge, England.

Im, K.S., Pesaran, M.H., and Y. Shin (1997), Testing for Unit Roots in Heterogeneous Panels. Manuscript, Department of Applied Economics, University of Cambridge, United Kingdom.

Im, K.S., Pesaran, M.H., and Y. Shin (2003), Testing for Unit Roots in Heterogenous Panels. Journal of Econometrics, 115, 53–74.

Johansen, S. (1992), Testing Weak Exogeneity and the Order of Cointegration in UK Money Demand Data. Journal of Policy Modeling, 14, 313-334.

Koop, G. and M.F.J. Steel, (2001), Bayesian Analysis of Stochastic Frontier Models, in Baltagi, B.H (ed.), A Companion to Theoretical Econometrics, Ch. 24, 520-537. Blackwell Publisher, Massachusetts.

Ng, S. and P. Perron (2001), Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. Econometrica, 69, 1519-1554.

Jönsson, K. (2005a), Cross-sectional Dependency and Size Distortion in a Small-sample Homogeneous Panel Data Unit Root Test. Oxford Bulletin of Economics and Statistics, 67, 369-392.

Jönsson, K. (2005b), Cross-sectional and Serial Correlation in a Small-sample Homogeneous Panel Data Unit Root Test. Applied Economics Letters, 12, 899–905.

Kanter, M. and W.L. Steiger (1974), Regression and Autoregression with Infinite Variance. Advances in Applied Probability, 6, 768-783.

Keane, M.P. (1993), Simulation Estimation for Panel Data Models with Limited Dependent Variables, in Maddala, G.S., Rao, C.R. and H.D. Vinod (eds.), Handbook of Statistics 11: Econometrics, 545-571. North-Holland, Amsterdam.

Kim, Y.S., Rachev, S.T., Bianchi, M.L. and F.J., Fabozzi (2008), Financial Market Models with Lévy Processes and Time-varying Volatility. Journal of Banking & Finance, 32, 1363-1378.

Knight, K. (1989), Consistency of Akaike's Information Criterion for Infinite Variance Autoregressive Processes. The Annals of Statistics, 17, 824-840.

Knight, K. (1993), Estimation in Dynamic Linear Regression Models with Infinite Variance Errors. Econometric Theory, 9, 570-588.

Koponen, I. (1995), Analytic Approach to the Problem of Convergence of Truncated Lévy Flights towards the Gaussian Stochastic Process. Physical Review, 52, 1197-1199.

Lee, C. and C. Chang (2007), Mean Reversion of Inflation Rates in 19 OECD Countries: Evidence from Panel Lm Unit Root Tests with Structural Breaks. Economics Bulletin, 3, 1-15.

Lee, H. and J. Wu (2001), Mean Reversion of Inflation Rates: Evidence from 13 OECD Countries. Journal of Macroeconomics, 23, 477-487.

Lehmer, D.H. (1951), Mathematical Methods in Large-scale Computing methods, Ann. Comp. Lab., 26, 141-146.

Levin, A. and C.F. Lin (1992), Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties. Unpublished manuscript, University of California, San Diego.

MacKeown, P. K. (1997), Stochastic Simulation in Physics. Springer, New York.

MacKinnon, J.G. (1991), Critical Values for Cointegration Tests, in Engle, R.F. and C.W.J. Granger (eds.), Long-run Economic Relationships: Readings in Cointegration, 267-276. Oxford University Press, Oxford.

MacKinnon, J.G. (1994), Approximate Asymptotic Distribution Functions for Unit-root and Cointegration Tests. Journal of Business and Economic Statistics, 12, 167-176.

MacKinnon, J.G. (1996), Numerical Distribution Functions for Unit Root and Cointegration tests. Journal of Applied Econometrics, 11, 601-618.

MacKinnon, J.G. (2000), Computing Numerical Distribution Functions in Econometrics, in A. Pollard, D. Mewhort, and D. Weaver (eds.), High Performance Computing Systems and Applications, 455-470. Kluwer, Amsterdam.

MacKinnon, J.G. and H. White (1985), Some Heteroskedasticity-consistent Covariance Matrix Estimators with Improved Finite Sample Properties. Journal of Econometrics, 29, 305-325.

MacKinnon, J.G., Haug, A.A. and L. Michelis (1999), Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. Journal of Applied Econometrics, 14, 563-577.

Maddala, G.S. and S. Wu (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.

Maddala, G.S., Rao, C.R. and H.D. Vinod (eds.) (1993), Handbook of Statistics 11: Econometrics. North-Holland, Amsterdam.

Maddala, G.S., Trost, R.P., Li, H. and F. Joutz (1997), Estimation of Short-run and Long-run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators. Journal of Business and Economic Statistics, 15, 90-100.

Mandelbrot, B. (1961), Stable Paretian Random Functions and the Multiplicative Variation of Income. Econometria, 29, 517-543.

Mandelbrot, B. (1963), The Variation of Certain Speculative Prices. Journal of Business, 36, 394-419.

Mandelbrot, B. (1967), The Variation of Some Other Speculative Prices. Journal of Business, 40, 394-413.

Mantegna, R.N. and H.E. Stanley (1997), Physics Investigation of Financial Markets. Proceedings of the International School of Physics “Enrico Fermi”. Course CXXXIV, 473-489.

Mariano, R.S. and B.W. Brown (1993), Stochastic Simulation for Inference in Nonlinear Errors-in-Variables Models, in Maddala, G.S., Rao, C.R. and H.D. Vinod (eds.), Handbook of Statistics 11: Econometrics, 611-627. North-Holland, Amsterdam.

McCulloch, J.H. (1996), Numerical Approximation of the Symmetric Stable Distribution and Density, in Adler R., Feldman R. and M.S. Taqqu (eds.), A practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions, 489-500. Birkhauser, Boston.

McCulloch, J.H. (1997), Measuring Tail Thickness to Estimate the Stable Index α : A Critique. Journal of Business and Economic Statistics, 15, 74-81.

McCulloch, J.H. (1998), Numerical Linear Regression with Stable Disturbances, in Adler R., Feldman R. and M.S. Taqqu (eds.), A practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions, 359-376. Birkhauser, Boston.

Metropolis, N. (1987), The Beginning of the Monte Carlo Method. Los Alamos Science, No. 15, 125-130. http://jackman.stanford.edu/mcmc/metropolis1.pdf.

Metropolis, N. and S. Ulam (1949), The Monte Carlo Method. Journal of the American Statistical Association, 44, 335-341.

Moon, H.R. and P. Perron (2004), Testing for Unit Root in Panels with Dynamic Factors. Journal of Econometrics, 122, 81-126.

Moulton, B.R. (1986), Random Group Effects and the Precision of Regression Estimates. Journal of Econometrics, 32, 385-397.

Munnell, A. (1990), Why Has Productivity Growth Declined? Productivity and Public Investment. New England Economic Review, 3-22.

Nabeya, S. and K. Tanaka (1988), Asymptotic Theory of a Test for the Constancy of Regression Coefficients against the Random Walk Alternative. Annals of Statistics, 16, 218-235.

Nelson, C.R. and C.I. Plosser (1982), Trends and Random Walks in Macroeconomic Time Series. Journal of Monetary Economics, 10,139-162.

Nelson, C.R. and G.W. Schwert (1977), Short-term Interest Rates as Predictors of Inflation: On Testing the Hypothesis that the Real Rate of Interest is Constant. American Economic Review, 67,478-486.

Ng, S. and P. Perron (2001), Lag Length Selection and the Construction of Unit Root Test with Good Size and Power. Econometrica, 69, 1519-1554.

Niederreiter, H. (ed.) (2004), Monte Carlo and Quasi-Monte Carlo Methods 2002: Proceedings of a Conference held at the National University of Singapore, Republic of Singapore, 2002. Springer-Verlag, Berlin.

O'Connell, P. (1998), The Overvaluation of Purchasing Power Parity. Journal of International Economics, 44, 1-19.

Otero, J. (2008), Testing for Stationarity in Inflation Rates: A Re-examination of the Evidence from 13 OECD Countries. Oxford Business & Economics Conference Programme, ISBN : 978-0-9742114-7-3.

Pehnelt, G, (2007), Globalisation and Inflation in OECD Countries. Jena Economic Research Papers No.055, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Perron, P. (1989), The Great Crash, The Oil Shock and the Unit Root Hypothesis. Econometrica, 57, 1361-1401.

Perron, P. (1997), Further Evidence from Breaking Trend Functions in Macroeconomic Variables, Journal of Econometrics, 80, 355-385.

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

Pesaran, M.H. (2007), A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. Journal of Applied Econometrics, 22, 265-312.

Pesaran, M.H. and R. Smith (1995), Estimating Long-run Relationships from Dynamic Heterogeneous Panels. Journal of Econometrics, 68, 79-113.

Pesaran, M.H., Shin, Y. and R.J. Smith (2000), Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables. Journal of Econometrics, 97, 293-343.

Phillips, P.C.B. (1987), Time Series Regression with a Unit Root. Econometrica, 55, 277-301.

Phillips, P.C.B. (1990), Time Series Regression with a Unit Root and Infinite Variance errors. Econometric Theory, 6, 44–62.

Phillips, P.C.B. and H.R. Moon (2000), Nonstationary Panel Data Analysis: An Overview of Some Recent Developments. Econometric Reviews, 19, 263-286.

Phillips, P.C.B. and P. Perron (1988), Testing for a Unit Root in Time Series Regression. Biometrika , 75, 335-346.

Phillips, P.C.B. and D. Sul (2003), Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence. Econometrics Journal, 6, 217-259.

Presno, M.J. and A.J. López (2003), Response Surface Estimates of Stationarity Tests with a Structural Break. Economics Letters, 78, 395-399.

Rachev, S.T. and S. Mittnik (2000), Stable Paretian Models in Finance. John Wiley & Sons, Chichester.

Rachev, S.T., Mittnik, S., Fabozzi, F.J., Focardi, S.M. and T. Jašić (2007), Financial Econometrics: From Basics to Advanced Modelling Techniques. John Wiley & Sons, New Jersey.

Rapach, D.E. and M.E. Wohar (2005), Regime Changes in International Real Interest Rates: Are They a Monetary Phenomenon?. Journal of Money, Credit and Banking, 37, 887-906.

Reinhart, C. and K. Rogoff (2004), The Modern History of Exchange Rate Arrangements: A Reinterpretation. The Quarterly Journal of Economics, 119, 1-48.

Roberson, D. and J. Symons (1992), Some Strange Properties of Panel Data Estimators. Journal of Applied Econometrics, 7, 175-189.

Robert, C.P. and G. Casella (2004), Monte Carlo Statistical Methods, Second Edition. Springer, New York.

Rogoff, K.S. (2003), Globalization and Global Disinflation. Federal Reserve Bank of Kansas City, Economic Review, Fourth Quarter, 2003.

Romero-Ávila, D. and C. Usabiaga (2009), The Hypothesis of a Unit Root in OECD Inflation Revisited. Journal of Economics and Business, 61, 153-161.

Rose, A. (1988), Is the Real Interest Rate Stable. Journal of Finance, 43, 1095-1112.

Rubinstein, R.Y. and D.P. Kroese (2007), Simulation and the Monte Carlo Method, Second Edition. John Wiley & Sons, New York.

Sargent, T. (1999), The Conquest of American Inflation. Princeton University Press, Princeton.

Sargan, J.D. and A. Bhargava (1983), Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk. Econometrica, 51, 153-174.

Samorodnitsky, G. and M.S. Taqqu (1994), Stable Non-Gaussian Random Processes. Champman & Hall, London.

Sato, K.I. (1999), Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge.

Schmeiser, B.W., Taaffe, M.R. and J. Wang (2001), Biased Control-variate Estimation. IIE Transactions, 33, 219-228.

Schmidt, P. and R.C. Sickles (1984), Production Frontiers and Panel Data. Journal of Business and Economic Statistics, 2, 367-374.

Sephton, P.S. (1995), Response Surface Estimates of the KPSS Stationarity Test. Economics Letters, 47, 255-261.

Stock, J.H. and M.W. Watson (1988), Testing for Common Trends. Journal of the Amrecian Statistical Association, 83, 1097-1107.

Strauss, J. and T. Yigit (2003), Shortfalls of Panel Unit Root Testing. Economics Letters, 81, 309-313.

Taylor, J. (1985), What Would Nominal GNP Targeting Do to the Business Cycle?. Carnegie-Rochester Conference Series on Public Policy, 22, 61-84.

Taylor, J. (1992), The Great Inflation, the Great Disinflation, and Policies for Future Price Stability, in Blundell-Wignall, A. (ed.), Inflation, Disinflation and Monetary Policy, 9-34. Ambassador Press, Sydney.

Vogelsang, T. and P. Perron (1998), Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time. International Economic Review, 39, 1073-1100.

Woodford, M. (2003), Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press.

Zolotarev, V. (1986), One Dimensional Stable Distributions. American Mathematical Society, Providence, R. I.