UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 104

On Bootstrap Evaluation of Tests for Unit Root and Cointegration

JIANXIN WEI

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

ISSN 1652-9030
ISBN 978-91-554-9069-0
urn:nbn:se:uu:diva-233885

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2, Ekonomikum, Uppsala, Friday, 28 November 2014 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Thomas Holgersson (Linnaeus University).

Abstract

Wei, J. 2014. On Bootstrap Evaluation of Tests for Unit Root and Cointegration. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 104. 23 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9069-0.

This thesis is comprised of five papers that all relate to bootstrap methodology in analysis of non-stationary time series.

The first paper starts with the fact that the Dickey-Fuller unit root test using asymptotic critical value has bad small sample performance. The small sample correction proposed by Johansen (2004) and bootstrap are two effective methods to improve the performance of the test. In this paper we compare these two methods as well as analyse the effect of bias-adjusting through a simulation study. We consider AR(1) and AR(2) models and both size and power properties are investigated.

The second paper studies the asymptotic refinement of the bootstrap cointegration rank test. We expand the test statistic of a simplified VECM model and a Monte Carlo simulation was carried out to verify that the bootstrap test gives asymptotic refinement.

The third paper focuses on the number of bootstrap replicates in bootstrap Dickey-Fuller unit root test. Through a simulation study, we find that a small number of bootstrap replicates are sufficient for a precise size, but, with too small number of replicates, we will lose power when the null hypothesis is not true.

The fourth and last paper of the thesis concerns unit root test in panel setting focusing on the test proposed by Palm, Smeekes and Urbain (2011). In the fourth paper, we study the robustness of the PSU test with comparison with two representative tests from the second generation panel unit root tests. In the last paper, we generalise the PSU test to the model with deterministic terms. Two different methods are proposed to deal with the deterministic terms, and the asymptotic validity of the bootstrap procedure is theoretically checked. The small sample properties are studied by simulations and the paper is concluded by an empirical example.

Keywords: non-stationary time series, unit root test, bootstrap, asymptotic refinement, cointegration, panel unit root test, cross-sectional dependence

Jianxin Wei, Department of Statistics, Uppsala University, SE-75120 Uppsala, Sweden.

ISSN 1652-9030
ISBN 978-91-554-9069-0
urn:nbn:se:uu:diva-233885 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-233885)

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Lyhagen, J. and Wei, J. (2013). Bootstrap versus Bartlett-type correction of the Dickey-Fuller test.

II Wei, J. (2013) A simulation verification of the asymptotic refinement in the bootstrap cointegration rank test.

III Wei, J. (2013) The number of bootstrap replicates in bootstrap Dickey-Fuller unit root test.

IV Liu, X. and Wei, J. (2014) On the robustness of the block bootstrap panel unit root test.

V Liu, X. and Wei, J. (2014) Block bootstrap panel unit root tests with deterministic terms.

Contents

1. Introduction

This thesis consists of five essays in the field of time series econometrics, focusing on the bootstrap methodology in unit root testing and cointegration rank testing. This introduction of the thesis provides an overall orientation concerning the topics covered in the thesis.

1.1 Unit root test

Many economic and financial time series exhibit trending behaviour, such as real gross domestic product (GDP), exchange rates and asset prices. The trend behaviour can be due to either a stochastic trend (unit root process) or a deterministic trend. An important econometric task is determining the most appropriate form of trend in the data. For example, in the autoregressive integrated moving average (ARIMA) framework the data must be transformed into stationary data prior to further analysis. If the data are characterized by a unit root process, first differencing is used whereas if the data are characterized by a deterministic trend, time-trend regression is used. In addition, economic and financial theory often suggests long-run equilibrium relationships among economic variables. If these variables follow unit root processes, cointegration approaches can be used to model these long-run relationships. Hence, testing for unit roots is an important pre-test in cointegration analysis.

1.1.1 Univariate unit root test

Consider an autoregressive (AR) process of order 1,

$$Y_t = \rho Y_{t-1} + \varepsilon_t, \tag{1.1}$$

where ε_t are independently, identically distributed (i.i.d.) with $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) < \infty$. The process exhibits significant different behaviours according to the parameter value ρ. If $|\rho| < 1$, then Y_t is (weakly) stationary. If $\rho = 1$, then Y_t is a unit root process and is not stationary. If $\rho > 1$, then the process is explosive, which usually is not an interesting case in econometric time series. A unit root test is a hypothesis test for testing if Y_t is a unit root process based on the observations $y_1, \ldots, y_T$ with hypotheses $H_0 : \rho = 1$ and $H_1 : \rho < 1$.

The best-known unit root test is the Dickey-Fuller test (DF test) proposed in the seminal work by Dickey and Fuller (1979). Consider the ordinary least square (OLS) estimator of ρ from (1.1),

$$\hat{\rho} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}, \tag{1.2}$$

under the null hypothesis $\rho = 1$, $\hat{\rho}$ is super-consistent and $T(\hat{\rho} - 1)$ converges to a non-standard distribution which can be expressed as a functional of a Wiener process,

$$T(\hat{\rho} - 1) \xrightarrow{d} \frac{\frac{1}{2}[W^2(1) - 1]}{\int_0^1 W^2(r)dr}, \tag{1.3}$$

where $W(r)$ is a standard Wiener process. The DF test uses critical values from the asymptotic distribution (1.3), which can be tabulated by simulations. If deterministic terms are included in the model in (1.1), the asymptotic distribution has to be adjusted according to the types of the deterministic terms, but will still be a Dickey-Fuller-type distribution. For a comprehensive review, see Hamilton (1994).

Said and Dickey (1984) take the general serial correlation into account in the DF test by adding the autocorrelations in the residuals of the fitted AR(1) model.

$$Y_t = \rho Y_{t-1} + z_t, \tag{1.4}$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_{k-1} z_{t-1+k} + \varepsilon_t. \tag{1.5}$$

The models in (4) and (5) are usually rewritten as

$$\Delta Y_t = \alpha Y_{t-1} + \sum_{i=j}^{k-1} \beta_j \Delta Y_{t-j} + \varepsilon_t, \tag{1.6}$$

and testing for unit roots can be carried out in the same way as the DF test. In this case, the standard DF test has been 'augmented' by ΔY_{t-j}, and hence the test is called the augmented Dickey-Fuller (ADF) test. Phillips and Perron (1988) proposed a semi-parametric test using modified ADF test statistics. Besides the framework of the DF tests, some alternative approaches have also been discussed. For example, Sargan and Bhargava (1983) suggested a unit root test in the Durbin-Waston framework. Cochrance (1988) proposed a variance ratio unit root test by measuring the degree of persistence in a time series. There are several problems with these univariate unit root tests, mainly size distortions and low power. Schwert (1989) first presented Monte Carlo evidence to point out the size distortion problems of the commonly used unit root tests. He argued that the distribution of the Dickey-Fuller tests is far different from the distribution reported by Dickey and Fuller if the underlying distribution contains a moving-average (MA) component. DeJong et al. (1992) argued that the unit root tests have low power when compared with plausible trend-stationary alternatives and concluded that tests with higher power need to be developed. For an excellent review, see Maddala and Kim (1998).

1.1.2 Panel unit root test

Panel unit root tests are extensions of the simple time series case by utilizing the cross-sectional dimension. The main motivation behind panel unit root tests is the desire to increase the power of unit root tests by increasing the sample size.

One question arises, naturally, when panel data are used for unit root testingd, -namely what is the relationship between cross-sectional units? The answer to this question roughly divides the panel unit root tests into two categories. The so-called first-generation panel unit root tests assume that the cross-sectional units are independent. Two representative tests belonging to this category are the tests proposed in Levin et al. (2002) and Im et al. (2003), usually denoted by the LLC test and the IPS test, respectively. Both tests assume cross-sectional independence, but allow for heterogeneous individual deterministic terms and heterogeneous serial correlation of the residuals. Both methods test the same null hypothesis

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_N = 1, \tag{1.7}$$

but differ with regard to the alternative hypothesis. The LLC test assumes a homogeneous first order autoregressive parameter and considers the alternative hypothesis

$$H_1 : \rho_1 = \rho_2 = \cdots = \rho_N = \rho < 1. \tag{1.8}$$

For the LLC test, a pooled t-statistic is constructed based on the estimation of ρ. The IPS test relaxes the assumption that $\rho_i = \rho$ for all i. Instead of pooling the data, the IPS test uses separate unit root tests for the N cross-section units and combines the evidence from the N tests. The statistic of the IPS test is based on the standardized average of individual ADF statistics.

Cross-sectional independence is an unrealistic assumption in the majority of macroeconomic applications of unit root tests, since co-movements of economies are often observed, for example concerning purchasing power parity. If we apply the tests belonging to the first generation directly to data with cross-sectional dependence, the results will have size distortions and low power, as pointed out in Banerjee et al. (2004). This led to the development of the so-called second-generation panel unit root tests, which relax this unrealistic assumption and allow for cross-sectional dependence. Dealing with cross-sectional dependence is complicated and two major approaches can be distinguished according to how cross-sectional dependence is treated. The first approach introduces cross-sectional dependence by imposing few or no restrictions on the error covariance matrix. Under this assumption, several tests have been developed, notably by Chang (2002; 2004) and Breitung and Das (2005).

The second approach, which is more popular, relies on a factor structure. The most influential method based on this approach is the PANIC procedure (panel analysis of non-stationarity in idiosyncratic and common components), proposed by Bai and Ng (2004). They consider a factor model:

$$Y_{i,t} = D_{i,t} + \lambda_i' F_t + e_{i,t}, \tag{1.9}$$

where $D_{i,t}$ is the deterministic term containing either a constant or a linear trend, F_t is an $m \times 1$ vector of common factors which are the cause of the cross-sectional dependence, λ_i is a vector of factor loadings, and $e_{i,t}$ is the idiosyncratic error term. The unit root can enter either into the common factors F_t or into the idiosyncratic errors $e_{i,t}$, and $Y_{i,t}$ is said to be non-stationary if F_t and/or $e_{i,t}$ is non-stationary. In the PANIC procedure, $\hat{F}_t$ and $\hat{e}_{it}$ are consistently estimated by the principal components method. Then the presence of a unit root is tested separately for these two parts. For the common factor component $\hat{F}_t$, the pooled DF test is used when only one factor is included. When more than one factor are included, a sequence of common trends tests is applied. For the idiosyncratic errors $\hat{e}_{it}$, either the combining p-values test or the type of test devised by Moon and Perron (2004) can be used. Other popular test methods which allow for cross-sectional dependence include those devised by Moon and Perron (2004), Pesaran (2007), and Phillips and Sul (2003).

1.2 Cointegration rank test

A time series, Y_t, is integrated of order d if d differences are required to make it stationary. If two or more time series are individually integrated of order d, but a linear combination of them is integrated of a lower order, the vector series is said to be cointegrated. Consider the p-dimensional vector error correction model (VECM),

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-1} + \Phi D_t + \varepsilon_t, \quad t = 1, ..., T, \tag{1.10}$$

where ε_t are independent $N_p(0, \Omega)$, and α and β are $p \times r$ matrices of full rank r, which is the cointegration rank, i.e. the number of linearly independent cointegration vectors. In applied work the cointegration rank is determined by a sequence of tests, $\mathscr{H}(r)$ against $\mathscr{H}(p)$ $\big(\mathscr{H}(r)|\mathscr{H}(p)\big)$ starting with $r = 0$. The procedure is sometimes called the 'top-to-bottom' procedure:

$$\underbrace{(\text{rank} = 0)}_{\mathscr{H}(0)} \subset \cdots \subset \underbrace{(\text{rank} \leq r)}_{\mathscr{H}(r)} \subset \cdots \subset \underbrace{(\text{rank} \leq p)}_{\mathscr{H}(p)}.$$

If $\mathscr{H}(0)$ is rejected, we continue to test $\mathscr{H}(1)$ until we accept a null hypothesis, $\mathscr{H}(r)$, and the cointegration rank is then determined as r. The Johansen trace test statistic for testing $\mathscr{H}(r)|\mathscr{H}(p)$ is $-T\sum_{i=r+1}^{p} \log(1-\hat{\lambda}_i)$, where $1 > \hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_p > 0$ are the eigenvalues of a corresponding matrix; for details, see, e.g., Johansen (1995). Under the assumption that (a) rank $= r \leq p$ and (b) the number of unit roots is $p-r$, the test statistic has the asymptotic distribution,

$$tr\left\{\int_0^1 (dB)F'\left[\int_0^1 FF'du\right]^{-1}\int_0^1 F(dB)'\right\}, \qquad (1.11)$$

where B and F are $p-r$ dimensional standard Brownian motions where F is adjusted according to the types of deterministic terms. If there are no deterministic terms, then $F=B$.

1.3 Bootstrap

The bootstrap method is a resampling mechanism designed to provide information about the sampling distribution of a statistic of interest. It is becoming a more and more appealing method due to its broad application and the increasing capability of computers.

Suppose that we have an i.i.d. sample, $\mathbf{z} = (z_1, z_2, ..., z_n)$, from a distribution, F_0. $T(\mathbf{z})$ is the statistic of interest and we want to perform an inference based on the distribution of $T(\mathbf{z})$. Let $G_n(x,F_0) = P\big(T(\mathbf{z}) \leq x\big)$ denote the distribution function of $T(\mathbf{z})$ (note that $G_n(x,F_0)$ is a function of F_0 and the sample size n). It is usually very difficult to derive an explicit form of G_n. One way to approximate G_n is to use the asymptotic distribution $G_\infty(x,F_0)$, i.e. the limiting distribution when the sample size n tends to infinity.

Bootstrap is another approach to approximating $G_n(x,F_0)$. The bootstrap approximation can be denoted as $G_n(x,\hat{F}_n)$, where $\hat{F}_n$ is the empirical cumulative distribution function, which can be considered as an approximation of F_0. $G_n(x,\hat{F}_n)$ can be simulated by the following algorithm:

1. draw the i.i.d. bootstrap samples $\mathbf{z}^* = \{z_1^*, z_2^*, ..., z_n^*\}$ from $\mathbf{z} = \{z_1, z_2, ..., z_n\}$;
2. calculate the bootstrap statistics $T^* = T(\mathbf{z}^*)$;
3. repeat steps 1 and 2 B times to obtain $T_1^*, T_2^*, ..., T_B^*$ and then let $G_n(x,\hat{F}_n) = \frac{1}{B}\sum_{b=1}^{B}\mathbf{1}(T_b^* \leq x)$, where $\mathbf{1}()$ is the indicator function.

A bootstrap procedure is consistent if the bootstrap distribution $G_n(x,\hat{F}_n)$ converges to the asymptotic distribution, $G_\infty(x,F_0)$, as n tends to infinity. More precisely, the bootstrap is (weakly) consistent if $\sup_x |G_n(x,\hat{F}_n) - G_n(x,F_0)| \xrightarrow{p} 0$, where $\xrightarrow{p}$ denotes convergence in probability. Consistency is considered as the basic requirement for a valid bootstrap procedure. If the asymptotic distribution is pivotal, i.e. free of nuisance parameters, the bootstrap usually gives a more accurate approximation in order than the asymptotic distribution. This is usually verified using an Edgeworth expansion. To illustrate this, suppose that $EZ_i = \mu$ and that $Var(Z_i) = \sigma^2$, and then we have the following expansion:

$$P(\sqrt{n}\frac{\bar{\mathbf{z}}-\mu}{\sigma} \leq t) = \Phi(t) + \frac{1}{\sqrt{n}}h_1(t,F_0) + \frac{1}{n}h_2(t,F_0) + O(\frac{1}{n^{3/2}}), \qquad (1.12)$$

where $\Phi(t)$ is the distribution function of the standard normal, and h_1 and h_2 are functionals of F_0. Hence, the accuracy of the asymptotic approximation is $O(\frac{1}{\sqrt{n}})$. The bootstrap version of the above expansion can be expressed as:

$$P(\sqrt{n}\frac{\bar{\mathbf{z}}^* - \bar{\mathbf{z}}}{se(\mathbf{z})} \leq t) = \Phi(t) + \frac{1}{\sqrt{n}} h_1(t, \hat{F}_n) + \frac{1}{n} h_2(t, \hat{F}_n) + O(\frac{1}{n^{3/2}}). \quad (1.13)$$

Taking the difference between (1.12) and (1.13) yields the following:

$$(\sqrt{n}\frac{\bar{\mathbf{z}} - \mu}{\sigma} \leq t) - P(\sqrt{n}\frac{\bar{\mathbf{z}}^* - \bar{\mathbf{z}}}{se(\mathbf{z})} \leq t) = \frac{1}{\sqrt{n}}\big(h_1(t, F_0) - h_1(t, \hat{F}_n)\big) + O(\frac{1}{n}) = O(\frac{1}{n}), \quad (1.14)$$

where it can be shown that $\hat{F}_n - F_0 = O(\frac{1}{\sqrt{n}})$. In summary, when we bootstrap a pivotal statistic, the accuracy of the approximation is $O(\frac{1}{n})$, and the gain in the accuracy is called the asymptotic refinement. For a review of the bootstrap theory, see, e.g., Horowitz (2001).

1.4 Bootstrap unit root test

When bootstrapping time series data, the i.i.d. bootstrap is not applicable because it does not capture the serial correlation, and consequently the bootstrap consistency will not hold. Instead, we can either bootstrap the independent component from the model and reconstruct the bootstrap sample, for example with the sieve bootstrap, or adjust the bootstrap method to capture the serial correlation, for example with the block bootstrap.

1.4.1 Univariate bootstrap unit root test

For simplicity, consider the AR model $Y_t = \rho Y_{t-1} + \varepsilon_t$, where ε_t are mutually uncorrelated. With the observations $y_1, ..., y_T$, a bootstrap unit root test can be conducted as follows:

1. calculate the DF test statistic and denote it as G;
2. obtain the centred errors, $\hat{\varepsilon}_t = y_t - y_{t-1} - \frac{1}{T}\sum_1^T (y_i - y_{i-1})$;
3. obtain the bootstrap errors, $\varepsilon_1^*, \ldots, \varepsilon_T^*$ through an i.i.d. bootstrap on the centred errors from step 2, i.e. sample with replacement T times from $\{\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_T\}$;
4. construct the bootstrap sample $y_1^*, \ldots, y_T^*$ with

   $$y_t^* = y_{t-1}^* + \varepsilon_t^*, t = 1, ...T,$$

   starting from $y_0^* = y_0 = 0$;
5. calculate the bootstrap test statistic G^* from the bootstrap sample in the same way as step 1;

6. set the number of bootstrap replicates B and repeat step 3 to step 5 B times to obtain B bootstrap statistics $\{G_1^*, ..., G_B^*\}$;
7. calculate the P-value of the bootstrap test with

$$P = B^{-1} \sum_{b=1}^{B} I\{G_b^* < G\}. \quad (1.15)$$

Several variations of the bootstrap algorithm for unit root testing have been developed in the literature. For all the bootstrap tests, the most important issue is that the bootstrap distribution must mimic the distribution under the null hypothesis. In the context of bootstrap unit root testing, this is achieved by imposing the unit root parameter $\rho = 1$ when constructing the bootstrap sample as in step 4.

With respect to obtaining the bootstrap errors in step 2, two different approaches can be distinguished, the difference-based approach and the residual-based approach. In the difference-based approach, the bootstrap errors are obtained with the restriction $\rho = 1$, i.e. $\hat{\varepsilon}_t = y_t - y_{t-1}$. This approach is used in, for example, Psaradakis (2001) and Park (2003). In the residual-based approach, on the other hand, the bootstrap errors are obtained by estimating ρ without restriction, i.e. $\hat{\varepsilon}_t = y_t - \hat{\rho} y_{t-1}$. This approach is used in Paparoditis and Politis (2003). In general, both approaches do not affect the consistency of the bootstrap. However, as argued in Paparoditis and Politis (2005), the residual-based approach will have a better power property, especially when the true DGP is from the alternative hypothesis.

If serial correlation is considered, we write the model as $Y_t = \rho Y_{t-1} + u_t$, where u_t captures the serial correlation. In this more general case, different bootstrap techniques have been developed, among which the sieve bootstrap and the block bootstrap are the ones most commonly used. The sieve bootstrap fits an autoregressive model to $\hat{u}_t$ in order to filter out the "i.i.d." errors from which the bootstrap errors are sampled. Using this bootstrap method, Psaradakis (2001) proposed a difference-based DF sieve bootstrap unit root test, while Chang and Park (2003) considered a difference-based ADF sieve bootstrap test. Palm et al. (2008) extended the two tests as residual-based tests and showed that the consistency of the bootstrap tests still holds. The block bootstrap resamples blocks of residuals, $\hat{u}_t$, to preserve the correlation structure. This bootstrap method is used in Paparoditis and Politis (2003), for example.

1.4.2 Bootstrap panel unit root test

In contrast to the univariate bootstrap unit root test, the errors to be sampled from in panel data are a matrix and, hence, the bootstrap algorithm has to be considered in both the time dimension and the cross-section unit dimension. Denote the estimated error terms by $\hat{\varepsilon}_{it}$, $t = 1, ..., T, i = 1, ..., N$. Maddala and

Wu (1999) proposed resampling $\hat{\varepsilon}_{it}$ with the cross-section index fixed; i.e. they resample $\hat{\varepsilon}_{.t} = (\hat{\varepsilon}_{1t}, ..., \hat{\varepsilon}_{Nt})'$ to obtain $\hat{\varepsilon}^*_{.t}$. This resampling scheme has the advantage that the cross-sectional dependence is preserved. Using this resampling scheme, Chang (2004) proposed a sieve bootstrap unit root test in which each time series, y_i, is approximated by an AR(p_i) process and all of the N regression equations are stacked in one regression equation. This test allows for cross-sectional dependence, but the form of the cross-sectional dependence is restricted. In particular, common factors are not allowed and the dependence only occurs through the covariance matrix of the error terms.

Kapetanios (2008) proposed a different bootstrap scheme which resamples in the cross-sectional dimension instead of the time dimension; i.e. they resamples with the time index fixed. This approach needs the assumption of cross-sectional independence.

In Palm et al. (2011), Palm, Smeekes and Urbain (PSU) proposed a block bootstrap panel unit root test which is herein referred to as the PSU test and which allows for a rather general complicated cross-sectional dependence structure. Since the last two papers of the thesis are motivated by this test, we introduce the PSU test here in greater detail. The authors of the test consider the model,

$$y_t = \Lambda F_t + w_t, \tag{1.16}$$

where

$$F_t = \Phi F_{t-1} + f_t, \tag{1.17}$$

$$w_t = \Theta w_{t-1} + v_t, \tag{1.18}$$

with v_t and f_t generated by

$$\begin{pmatrix} v_t \\ f_t \end{pmatrix} = \Psi(L)\varepsilon_t = \begin{bmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ \Psi_{21}(L) & \Psi_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{v,t} \\ \varepsilon_{f,t} \end{bmatrix}, \tag{1.19}$$

where ε_t are i.i.d. with $E\varepsilon_t = 0, E\varepsilon_t\varepsilon_t' = \Sigma$ and $E|\varepsilon_t|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. A rather general cross-sectional dependence structure can be accounted for in this model setting. Remarkably, correlation of the common factor and the idiosyncratic errors is allowed. In addition, a dynamic dependence is captured by $\Psi(L)$, which is not allowed in all the other tests that preceded the PSU test. The PSU test is not interested in which component has the unit root, and instead simply tests if $y_{i,t}$ has a unit root for all i. To achieve this, an auxiliary regression, $\Delta y_{it} = \rho_i y_{i,t-1} + u_{it}$, is considered. From this regression, the pooled and group-mean statistics are constructed following the ideas in the LLC and IPS tests. The bootstrap is applied to the residuals from the auxiliary regression. Palm et al. (2011) suggest a block bootstrap with the cross-section index fixed in order to preserve both the cross-sectional dependence and the temporal correlation. The consistency of the bootstrap test has been validated theoretically, and simulation studies, making comparison with the LLC and

IPS tests, show that the PSU tests have generally acceptable size and power under the DGP of a wide range of cross-sectional dependence.

2. Summary of the papers

2.1 Paper I: Bootstrap versus Bartlett-type Correction of the Dickey-Fuller Test

Consider the AR(k) model in the form,

$$\Delta x_t = \pi x_{t-1} + \beta t^d + \sum_{i=1}^{k-1} \gamma_i \Delta x_{t-i} + \sum_{i=0}^{d-1} \beta_i t^i + \varepsilon_t,$$

where $t = 1, 2, \ldots, T$, k is the number of lags in levels, d determines the forms of the deterministic components, and the error term ε_t is i.i.d. $N(0, \sigma^2)$. We are interested in the properties of the likelihood ratio test in testing the joint hypothesis $H_0 : \pi = \beta = 0$. The null hypothesis ensures that the order of the trend is the same under the null as under the alternative. The test statistic can be expressed by

$$-T(\frac{RSS_U}{RSS_R}),$$

where RSS_U and RSS_R are the residual sums of squares from the unrestricted and the restricted model, respectively. Under H_0 the test statistic has an asymptotic distribution of the Dickey-Fuller-type. When using the asymptotic critical value, there is a size distortion and for small T, the inference made is not reliable. There are mainly two methods for achieving more accurate finite sample inferences, the Bartlett-type correction (see, e.g., Johansen (2004)), and the bootstrap (see, e.g., Paparoditis and Politis (2003)). Estimates of the short-run dynamics $\hat{\gamma}_i$ are involved in both of the two methods and it is well known that the OLS estimator is biased in the AR model. Tjostheim and Paulsen (1983) and van Giersbergen (2005) derived the bias of the OLS estimator of the parameters of an AR(k) process. The purpose of Paper 1 was to compare the Bartlett-type correction and the bootstrap approach, as well as to analyze if using the bias-adjusted estimator would give improvement in both the Bartlett correction and the bootstrap. To be specific, we investigated both the size and the power properties of the following five test methods:

- tests using asymptotic critical values,
- Bartlett-type correction,
- Bartlett-type correction using a bias-adjusted estimator,
- the bootstrap test,
- the bootstrap test using a bias-adjusted estimator.

The results show that, in terms of size, Bartlett correction is sensitive to the short-run parameter values. For some parameter values it corrects the size well, while for some other parameter values it will over-correct and lead to an undersized test. This is in contrast to the bootstrap. The bootstrap is less sensitive to the parameter values and works well with an empirical size close to the nominal. Bias-adjusting leads to a minor improvement. In terms of power, we analyze both the size-adjusted power and the raw power. For the size-adjusted power, there is no clear winner since the curves cross. However, when π is close to the null, the bootstrap gives better power than the Bartlett correction. There is little difference between the two bootstrap methods. The bootstrap also gives higher raw power than the Bartlett-type correction. In conclusion, the bootstrap outperforms the Bartlett correction in general and no improvement is achieved by using the bias-corrected estimates.

2.2 Paper II: A Simulation Verification of the Asymptotic Refinement in the Bootstrap Cointegration Rank Test

The original Johansen (1995) cointegration rank test suffers from size distortion when asymptotic critical values are used. Since the trace statistic is asymptotically pivotal, it is expected that the bootstrap will give asymptotic refinement. We consider a simple VAR model with no-short run dynamics and no deterministic terms,

$$\Delta X_t = \Pi X_{t-1} + \varepsilon_t, \quad t = 1, ..., T,$$

with $X_0 = 0$, and where Π is a $p \times p$ matrix and ε_t is i.i.d. $N(0, \Omega)$. We show that, in this simple framework, the error of the asymptotic approximation is of the stochastic order $O_p(T^{-\frac{1}{2}})$. To be specific,

$$R_T = R - T^{-\frac{1}{2}} R_1 + o_p(T^{-\frac{1}{2}}),$$

where R_T is the test statistic of $\mathscr{H}(0)|\mathscr{H}(p)$, R represents the random variable which follows the asymptotic distribution, and R_1 involves a p-dimensional Brownian motion and a normal random vector, U. When the asymptotic critical value is used, the difference between the rejection probability and the nominal size is $O(T^{-\frac{1}{2}})$. If the bootstrap can give asymptotic refinement, the error in the rejection probability of using the bootstrap critical value will be $o(T^{-\frac{1}{2}})$. We conducted an intensive Monte Carlo experiment in testing $\mathscr{H}(0)$ against $\mathscr{H}(2)$ in the simple VAR(2) model. We ran the simulation for one million replicates and obtained $\hat{\alpha}_{A,T}$ and $\hat{\alpha}_{B,T}$, which are the estimated empirical rejection probabilities of the test using asymptotic critical values and the bootstrap test, respectively. We find that $\sqrt{T}(\hat{\alpha}_{A,T} - 0.05)$ converges to a non-zero constant and $\sqrt{T}(\hat{\alpha}_{B,T} - 0.05)$ converges to zero, and this corroborates the conjecture that the bootstrap gives asymptotic refinement in the cointegration rank test.

2.3 Paper III: The Number of Bootstrap Replicates in the Bootstrap Dickey-Fuller Unit Root Test

This paper studies the bootstrap unit root test, focusing on the number of bootstrap replicates. We consider the simple AR(1) model $Y_t = \rho Y_{t-1} + \varepsilon_t$ and the bootstrap unit root test proposed by Park (2003). Through a simulation study, we find that a small number of bootstrap replicates are sufficient to obtain an empirical rejection probability close to the nominal level. However, a small number of bootstrap replicates will lose power when the data generating process (DGP) is from the alternative hypothesis. The study in this paper helps to reduce the computational cost in the intensive simulation that we applied in Paper II.

2.4 Paper IV: On the Robustness of the Block Bootstrap Panel Unit Root Test

This paper studies the robustness of the block bootstrap panel unit root test proposed by Palm et al. (2011)(PSU). Relying on a block bootstrap approach, the PSU test can deal with a rather general cross-sectional dependence structure, including the popular common factor framework. PSU proved the consistency of the block bootstrap tests but their simulation study is incomplete in that the small sample performance of the PSU tests is still unclear in the following cases:

1 when the idiosyncratic errors and the common factors are dependent;
2 when the idiosyncratic errors are of integrated order $I(1)$ while the common factors are of integrated order $I(0)$;
3 when there are more than one common factor.

In addition, in their simulation study the PSU test was only compared with the LLC and IPS tests, both of which belong to the first-generation panel unit root tests. Prior to Paper IV, no comparison with tests from the second generation had been performed and such a comparison is of interest to practitioners. This circumstance motivated us to conduct a further investigation on the small sample properties of the PSU test under a more general and complicated DGP and in comparison with two other panel unit root tests from the second generation.

The tests proposed by Bai and Ng (2004) and Chang (2004) were chosen as the two reference tests in the comparison, denoted by the BN test and the Chang test, respectively. The models from both tests are special cases of the model from the PSU test. The robustness of the PSU test was investigated from two aspects. First, the PSU test was applied on data generated by the specific models considered in Bai and Ng (2004) and Chang (2004), to compare the small sample performance of the PSU test with that of the BN test and that of the Chang test. Second, all the three tests were applied on data generated according to the DGP of the PSU test with more general cross-sectional dependence than that considered in the simulation of the PSU test. We added a correlation between the idiosyncratic errors and the common factors. The extent of this correlation had two levels, denoted by Gen 1 and Gen 2. Gen 1 considered the correlation between the two components only through the covariance matrix Σ. Gen 2 extended the correlation through the lag polynomials $\Psi(L)$. Furthermore, in the DGPs with a factor component, all the four possible combinations of integrated order were considered, and both one and two common factors were included.

The conclusions of the study can be summarized as follows.

1 Using the specific DGPs of Chang (2004) and Bai and Ng (2004), the PSU test, especially τ_{gm}, has generally as good size and power properties as the Chang and the BN tests, except in the case with negative moving-average coefficients.

2 Using the DGP in PSU with a general cross-sectional dependence structure, the PSU test exhibits robustness with good size and power properties in the cases when the idiosyncratic errors and the common factor are of the same integrated order. However, when the two components are of a different integrated order, the PSU test is oversized.

3 The group-mean test τ_{gm} is more robust than the pooled test τ_p. In both the DGP of Chang (2004) and that of Bai and Ng (2004), when the ARMA coefficients are randomly chosen from $U[-0.8, 0.8]$, τ_p exhibits severe size distortion.

2.5 Paper V: Block Bootstrap Panel Unit Root Tests with Deterministic Terms

In this paper we made an extension of the PSU test by including deterministic terms in the model from PSU. Dealing with deterministic terms in a bootstrap test is not trivial, since the consistency of the bootstrap is not guaranteed by default. In this paper, we proposed two approaches by which the PSU tests can be extended to the model with deterministic terms. Consider the panel data $x_t = (x_{1,t}, ..., x_{N,t})$ for $t = 1, ..., T$ generated by

$$x_t = d_t^m + y_t, \tag{2.1}$$

where $d_t^m = (d_{1,t}^m, ..., d_{N,t}^m)'$ for $m = 0, 1, 2$ is the deterministic component. d_t can be zero (without deterministic terms), a constant or a constant and a linear time trend. Specifically, for each i, $d_{i,t}^m = \beta_i^{m'} z_t^m$, where

$$z_t^m = \begin{cases} 0 & \text{if} \quad m = 0, \\ 1 & \text{if} \quad m = \mu, \\ (1,t)' & \text{if} \quad m = \tau, \end{cases} \quad \text{and} \quad \beta_i^m = \begin{cases} 0 & \text{if} \quad m = 0, \\ \beta_{1i} & \text{if} \quad m = \mu, \\ (\beta_{1i}, \beta_{2i})' & \text{if} \quad m = \tau. \end{cases} \tag{2.2}$$

y_t is the stochastic component with the same structure as in Palm et al. (2011). In the first approach, which is called the conventional approach in the paper, the statistics of the PSU test are modified by estimating the unit root parameter ρ_i from the auxiliary regression with deterministic terms. To be more specific, for the model $m = \mu$, ρ_i is estimated from

$$\Delta x_{it} = \alpha_i + \rho_i x_{i,t-1} + e_{it}, \tag{2.3}$$

and for the model $m = \tau$, ρ_i is estimated from

$$\Delta x_{it} = \alpha_i + \beta_i t + \rho_i x_{i,t-1} + e_{it}. \tag{2.4}$$

Then the pooled statistic τ_p and the group-mean statistic τ_{gm} are constructed as in Palm et al. (2011). The bootstrap algorithm is modified accordingly by generalizing the block bootstrap algorithm proposed by Paparoditis and Politis (2003) to the panel case and we proved the consistency of this bootstrap algorithm.

The second approach is a two-stage approach. The first stage is to detrend the data. We considered three detrending methods, namely ordinary least square (OLS) detrending, generalized least square (GLS), devised by Elliott et al. (1996), and recursive detrending (REC), devised by Taylor (2002). After the first step, the deterministic component is removed and the same statistics as in Palm et al. (2011) are used. Although the test statistics are the same, the bootstrap algorithm must be adjusted according to the detrending methods. The adjustment of the bootstrap algorithm is based on the methods proposed in Smeekes (2009). We theoretically proved the consistency of the bootstrap procedure of the OLS detrending method under the hypothesis that both the common factors and the idiosyncratic errors have a unit root.

A simulation study was performed to study the small sample properties of our tests. The simulation included both the conventional approach and the two-stage approach, and in the latter approach all the three detrending methods were considered. The simulation results show that all the tests have acceptable size properties. The two-stage detrending methods have a better power property than the conventional methods. Within the two-stage method, GLS detrending has the highest power.

3. Acknowledgements

I want to express my first and foremost gratitude to my supervisor, Rolf Larsson. I truly appreciate all the time, ideas, and funding you have contributed throughout my study. Thank you for always giving me feedback on the next day and tolerating me missing deadlines. I also would like to thank my second supervisor, Lars Forsberg. Thank you for your encouragement and all the rewarding discussions. My thanks also goes to Professor Johan Lyhagen. Thank you for sharing your idea with me and giving me fruitful advice. A special thanks goes to Professor Fan Yang Wallentin. Thank you for introducing me to Sweden in the first place. Your kind help and useful suggestions make my life in Sweden easier. I am also grateful to the other committee members, Rauf Ahmad, Inger Persson and Bo Wallentin for reading and commenting my thesis. In addition, I want to extend my acknowledgement to Professor Dietrich von Rossen at SLU. Thank you for all the inspiring discussions regarding science and other interesting aspects, and also thank you for organizing the PhD seminars.

It is a pleasure to work in the Department of Statistics. I would like to take this opportunity to express my thanks to all the current and former colleagues at the department. Without your support, this thesis would not have been possible. In particular, thanks to Lisbeth Hansson for helping me with all the problems I have met in teaching. Thanks to Davoud Emamjomeh for all the IT support and for helping me buying my little red car, which makes my life much more colorful. Thanks to Eva Enefjord for helping me with all the administrative formalities. Moreover, I would like to show my gratitude to my fellow PhD students. I enjoy all the relaxing after work events with you guys. Special thanks go to Xijia Liu and Xingwu Zhou. I will never forget the interesting discussions we have in our office and the exciting pingpong times in the gym. And to Xijia, you are a great coauthor.

Finally, my deepest thanks goes to my family. I am especially indebted to my parents. No words can express my gratitude for your unconditional love and support. I have saved my warmest thanks to my wife, Ying Li. Thank you for your kindness, support and unwavering love. I cannot imagine how this journey would be without your company. And Yichen, our son, you always bring me joy and remind me to look toward the bright side in life.

Jianxin Wei
Uppsala, September 2014

References

Bai, J., Ng, S. (2002). Determining the number of factors in approximate factor models, *Econometrica* **70**: 191–221.

Bai, J., Ng, S. (2004). A PANIC attack on unit roots and cointegration, *Econometrica* **72**: 1127–177.

Bai, J., Ng, S. (2010). Panel unit root tests with cross-section dependence: a further investigation, *Econometric Theory* **26**: 1088–1114.

Banerjee, A., Marcellino, M., Osbat, C. (2004). Some cautions on the use of panel methods for integrated series of macro-economic data, *Econometrics Journal* **7**: 322–340.

Breitung, J., Das, S. (2005). Panel unit roots under cross-sectional dependence, *Statistica Neerlandica* **4**: 414–433.

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., Park, J.Y. (2003). A sieve bootstrap for the test of a unit root, *Journal of Time Series Analysis* **24**: 379–400.

Cochrane, J.H. (1988). How big is the random walk in DNP? *Journal of Political Economy* **96**: 893–920.

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

Dickey, D.A., Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* **74**: 472-31.

Elliott, T., Rothenberg, T.J., Stock, J.H. (1996). Efficient tests for an autoregressive unit root, *Econometrica* **64**: 813–836.

Gengenbach, C., Palm, F.C., Urbain, J.P. (2010). Panel unit root tests in the presence of cross-sectional dependencies: comparison and implications for modelling, *Econometric Reviews* **29**: 111–145.

van Giersbergen, N.P.A. (2005). On the effect of deterministic terms on the bias in stable AR models, *Economics Letters* **89**: 75–82.

Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press, Princeton, N.J.

Horowitz, J.L. (2011). The bootstrap, *Handbook of Econometrics* Vol.5, Chapeter 52, Elsevier, Amsterdam, pp. 3159–3228.

Im, K.S., Pesaran, M.H., Shin, Y. (2003). Testing for unit roots in heterogeneous panels, *Journal of Econometrics* **115**: 53–74.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S. (2004). A small sample correction of the Dickey-Fuller test, *New Directions in Macromodelling*, Welfe, A., pp. 49-68.

Kapetanios, G. (2008). A bootstrap procedure for panel data sets with many cross-sectional units, *Econometrics Journal* **11**: 377–395.

Levin, A., Lin, C.F., Chu, C.S.J. (2010). Unit root tests in panel data: asymptotic and finite-sample properties, *Journal of Econometrics* **108**: 1–24.

Maddala, G.S., Kim, I.M. (1998). *Unit Roots, Cointegration and Structural Change*, Cambridge University Press, Cambridge.

Maddala, G.S, Wu, S. (1999). A comparative study of panel data unit root tests and a simplified test, *Oxford Bulletin of Economics and Statistics* **61**: 631–652.

Moon, R., Perron, P. (2004). Testing for unit root in panels with dynamic factors, *Journal of Econometrics* **122**: 81–126.

Palm, F.C., Smeekes, S., Urbain, J.P. (2008). Bootstrap unit-root tests: comparison and extensions, *Journal of Time Series Analysis* **29**: 371–401.

Palm, F.C., Smeekes, S., Urbain, J.P. (2011). Cross-sectional dependence robust block bootstrap panel unit root tests, *Journal of Econometrics* **163**: 85–104.

Paparoditis, E., Politis, D.N. (2003). Residual-based block bootstrap for unit root testing, *Econometrica* **71**: 813–815.

Paparoditis, E., Politis, D.N. (2005). Bootstrapping unit root tests for autoregressive time series, *Journal of the American Statistical Association* **100**: 545–53.

Park, J.Y. (2003). Bootstrap unit root tests, *Econometrica* **71**: 1845–1895.

Perron, P., Ng, S. (1996). Useful modifications to unit root tests with dependent errors and their local asymptotic properties, *Review of Economic Studies* **63**: 435–465.

Pesaran, M.H. (2007). A simple panel unit root test in the presence of cross-section dependence, *Journal of Applied Econometrics* **22**: 265–312.

Phillips, P.C.B., Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross-section dependence, *Econometrics Journal* **6**: 217–259.

Phillips, P.C.B., Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika* **75**: 335–46.

Politis, D.N., Romano, J.P., Wolf, M. (1999). *Subsampling,* Springer-Verlag, New York.

Psaradakis, Z. (2001). Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors, *Journal of Time Series Analysis* **22**: 577–94.

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

Sargan, J.D., Bhargava, A. (1983). Testing for residuals from least squares regression being generated by Gaussian random walk, *Econometrica* **51**: 153–174.

Schwert, G.W. (1989). Testing for unit roots: a Monte Carlo investigation, *Journal of Business and Economic Statistics* **7**: 147–159.

Smeekes, S. (2009). Detrending bootstrap unit root tests, *Econometric Reviews* **32**: 869-891.

Taylor, R. (2002). Regression-based unit root tests with recursive mean adjustment for seasonal and nonseasonal time series, *Journal of Business and Economic Statistics* **20**: 269–81.

Tjostheim, D., Paulsen, J. (1983). Bias of some commonly-used time series estimates, *Biometrika* **70**: 389–399.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 104

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-233885

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014