Bootstrapping Cointegrating Regressions[1]

Yoosoon Chang
Department of Economics
Rice University

Joon Y. Park
School of Economics
Seoul National University

Kevin Song
Department of Economics
Yale University

Abstract

In this paper, we consider bootstrapping cointegrating regressions. It is shown that the method of bootstrap, if properly implemented, generally yields consistent estimators and test statistics for cointegrating regressions. We do not assume any specific data generating process, and employ the sieve bootstrap based on the approximated finite-order vector autoregressions for the regression errors and the first differences of the regressors. In particular, we establish the bootstrap consistency for OLS method. The bootstrap method can thus be used to correct for the finite sample bias of the OLS estimator and to approximate the asymptotic critical values of the OLS-based test statistics in general cointegrating regressions. The bootstrap OLS procedure, however, is not efficient. For the efficient estimation and hypothesis testing, we consider the procedure proposed by Saikkonen (1991) and Stock and Watson (1993) relying on the regression augmented with the leads and lags of differenced regressors. The bootstrap versions of their procedures are shown to be consistent, and can be used to do inferences that are asymptotically valid. A Monte Carlo study is conducted to investigate the finite sample performances of the proposed bootstrap methods.

This version: February 18, 2002

Key words and phrases: cointegrating regression, sieve bootstrap, efficient estimation and hypothesis testing, AR approximation.

1. Introduction

The bootstrap has become a standard tool for the econometric analysis. Roughly, the purpose of using the bootstrap methodology is twofold: To find the distributions of statistics whose asymptotic distributions are unknown or dependent upon nuisance parameters, and to obtain refinements of the asymptotic distributions that are closer to the finite sample distributions of the statistics. It is well known that the bootstrap statistics have the same asymptotic distributions as the corresponding sample statistics for a very wide, if not all, class of models, and therefore, the unknown or nuisance parameter dependent limit distributions can be approximated by the bootstrap simulations. Moreover, if properly implemented to pivotal statistics, the bootstrap simulations indeed provide better approximations to the finite sample distributions of the statistics than their asymptotics. See Horowitz (2002) for an excellent nontechnical survey on the subject.

The purpose of this paper is to develop the bootstrap theory for cointegrating regressions. The bootstrap can be potentially more useful for models with nonstationary time series than for the standard models with stationary time series, since the statistical theories for the former are generally nonstandard and depend, often in a very complicated manner, upon various nuisance parameters. Nevertheless, the bootstrap theories for the former are much less developed compared to those for the latter. Virtually all of the published works on the theoretical aspects of the nonstationary bootstrap consider simple unit root models. See, e.g., Basawa *et al.* (1991), Chang (2000), Chang and Park (2002b) and Park (2000, 2002). The bootstrap cointegrating regression has been studied only by simulations as in Li and Maddala (1997). The bootstrap method, however, is used quite frequently and extensively by empirical researchers to approximate the distributions of the statistics in more general models with nonstationary time series.

We consider the sieve bootstrap to resample from the cointegrating regressions. The method does not assume any specific data generating processes, and the data are simply fitted by a VAR of order increasing with the sample size. The bootstrap samples are then constructed using the fitted VAR from the resampled innovations. We show in the paper that under such a scheme the bootstrap becomes consistent for both the usual OLS and the efficient OLS by Saikkonen (1991) and Stock and Watson (1993). The sieve bootstrap can therefore be used to reduce the finite sample bias of the OLS estimator, and also to find the asymptotic critical values of the tests based on the OLS estimator. The bootstrapped OLS estimator, however, is inefficient just as is the sample OLS estimator. The bootstrap does not improve the efficiency. To attain efficiency, we need to bootstrap the efficient OLS estimator. The sieve bootstrap can be naturally implemented to do resampling for the efficient estimator, which itself relies on the idea of sieve estimation of the cointegrating regression. We show in the paper that the sieve bootstrap is generally consistent for the efficient OLS method.

Though we focus on the prototype multivariate cointegration model in the paper for concreteness, the theory we derive here can be used to analyze more general cointegrated models. The immediate extensions are the cointegrating regressions with more flexible deterministic trends including those allowing for structural breaks. Cointegrating regressions with shifts in the coefficients can also be analyzed using the methods developed in the paper.

Moreover, our theory extends further to the cointegrating regressions represented as error correction models, seemingly unrelated cointegrated models and panels with cointegration in individual units. The sieve bootstrap proposed here can be applied to all such models only with some obvious modifications. The estimation of and testing for the cointegration parameters can therefore be performed or refined by the sieve bootstrap. It is well expected that the sieve bootstrap is consistent for the aforementioned models if implemented properly as suggested in the paper, and it can be proved rigorously, if necessary, using the theory established in the paper.

The rest of the paper is organized as follows. Section 2 introduces the model, assumptions and some preliminary results. The multivariate cointegrating regression with a detailed specification for data generating process is given, and a strong invariance principle, which can be used for both the sample and bootstrap asymptotics, is introduced and discussed. The standard asymptotic results for the cointegrating regressions are then derived. In Section 3, the sieve bootstrap procedure is presented, and the relevant bootstrap asymptotics are developed. The bootstrap consistency is established there. Section 4 summarizes our simulation study, and Section 5 concludes the paper. All the mathematical proofs are given in Section 6.

A word on notation. Following the standard convention, we use the superscript "$*$" to signify whatever is related to the bootstrap samples and dependent upon the realization of the samples. The usual notations for the modes of convergence such as $\to_{a.s.}$, $\to_p$ and $\to_d$ are used without additional references, and the notation $=_d$ denotes the equality in distribution. The stochastic order symbols o_p and O_p are also used frequently. Moreover, we often use such standard notations with the superscript "$*$" to imply their bootstrap counterparts. We use the notation $|\cdot|$ to denote the Euclidean norm for vectors and matrices, i.e., $|x|^2 = x'x$ and $|A|^2 = \operatorname{tr} A'A$ for a vector x and a matrix A. For matrices, we also use the operator norm $\|\cdot\|$, i.e., $\|A\| = \max_x |Ax|/|x|$ for a vector x and a matrix A which are of conformable dimensions. For a matrix A, $\operatorname{vec}(A)$ denotes a column vector which stacks row vectors of A.

2. The Model, Assumptions and Preliminary Results

2.1 The Model and Assumptions

We consider the regression model given by

$$
\begin{aligned}
y_t &= \Pi'x_t + u_t \\
x_t &= x_{t-1} + v_t
\end{aligned}
\tag{1}
$$

for $t = 1, 2, \ldots$, where

$$
w_t = (u_t', v_t')'
$$

is an $(\ell+m)$-dimensional stationary vector process. Under this specification, the model introduced in (1) becomes a multivariate cointegrating regression. For the subsequent development of our theory, we let x_0 be any random variable which is stochastically bounded,

and let (w_t) be a linear process given by

$$w_t = \Psi(L)\varepsilon_t \tag{2}$$

where

$$\Psi(z) = \sum_{k=0}^{\infty} \Psi_k z^k$$

We make the following assumptions.

Assumption 2.1 *We assume*
(a) (ε_t) are iid random variables such that $\mathbf{E}\varepsilon_t = 0$, $\mathbf{E}\varepsilon_t\varepsilon_t' = \Sigma > 0$ and $\mathbf{E}|\varepsilon_t|^a < \infty$ for some $a \geq 4$.
(b) $\det \Psi(z) \neq 0$ for all $|z| \leq 1$, and $\sum_{k=0}^{\infty} |k|^b|\Psi_k| < \infty$ for some $b \geq 1$.

The conditions in Assumption 2.1 are not necessary for our results in this section. In particular, the iid assumption on the innovation (ε_t) is not required. It is introduced here just to make our forthcoming bootstrap procedure more meaningful. All the subsequent theoretical results may be obtained under weaker martingale difference assumptions on (ε_t). Also, 1-summability of (Ψ_k) is assumed to simplify the proofs, and can be weakened to 1/2-summability. Yet, a wide class of cointegrated models, including Gaussian error correction models considered in Johansen (1988, 1991), can be represented as the model specified in (1) and (2) with (ε_t) and (Ψ_k) satisfying the conditions in Assumption 2.1.

2.2 Invariance Principles

For the iid random vectors (ε_t), we define

$$W_n(r) =_d n^{-1/2} \sum_{t=1}^{[nr]} \varepsilon_t$$

Then the invariance principle for the iid random vectors (ε_t) holds, i.e.,

$$W_n \to_d W \tag{3}$$

where W is a vector Brownian motion with variance Σ. In particular, we have

Lemma 2.2 *Let $\mathbf{E}|\varepsilon_t|^a < \infty$ for some $a > 2$. Then we may define on a common probability space W_n and W such that*

$$\mathbf{P}\left\{ \sup_{0 \leq r \leq 1} |W_n(r) - W(r)| > n^{-1/2}c_n \right\} \leq K n\, c_n^{-a}\, \mathbf{E}|\varepsilon_t|^a$$

for any numerical sequence (c_n), $c_n = n^{1/a+\delta}$ for some $\delta > 0$, where K is an absolute constant depending only on a and $\ell+m$.

Lemma 2.2 follows immediately from the strong approximation result established in Einmahl (1987). For any given $a > 2$, we may choose δ such that $0 < \delta < 1/2 - 1/a$ to show

$$\sup_{0 \leq r \leq 1} |W_n(r) - W(r)| = o_p(1)$$

and therefore, the invariance principle (3) follows directly from Lemma 2.2. The use of strong approximation for the proof of the invariance principle (3) is very useful in our context, since it can also be directly applied to derive the corresponding invariance principle for the bootstrap samples. This will be shown in the next section.

The invariance principle for (ε_t) directly carries over to the one for (w_t). We have from the Beveridge-Nelson decompostion that

$$w_t = \Psi(1)\varepsilon_t + (\bar{w}_{t-1} - \bar{w}_t)$$

with $\bar{w}_t = \sum_{k=0}^{\infty} \bar{\Psi}_k \varepsilon_{t-k}$ and $\bar{\Psi}_k = \sum_{i=k+1}^{\infty} \Psi_i$, and therefore,

$$
\begin{aligned}
B_n(r) &= n^{-1/2} \sum_{t=1}^{[nr]} w_t \\
&= n^{-1/2} \Psi(1) \sum_{t=1}^{[nr]} \varepsilon_t + n^{-1/2}(\bar{w}_0 - \bar{w}_{[nr]}) \\
&= \Psi(1) W_n(r) + R_n(r)
\end{aligned}
\tag{4}
$$

where

$$\sup_{0 \leq r \leq 1} |R_n(r)| = o_p(1)$$

since $(\bar{w}_n)$ is well-defined and stationary under our 1-summability condition on (Ψ_k). The reader is referred to Phillips and Solo (1992) for more details. It now follows immediately from (4) that

$$B_n \to_d B \tag{5}$$

where B is an $(\ell+m)$-dimensional vector Brownian motion with variance - $= \Psi(1)\Sigma\Psi(1)'$. We call - the longrun variance of (w_t). If we let $\Gamma(k) = \mathbf{E}w_t w_{t+k}'$ be the autocovariance function of (w_t), then we may define the longrun variance of (w_t) as - $= \sum_{k=-\infty}^{\infty} \Gamma(k)$. Correspondingly, we denote by Δ the one-way longrun variance of w_t, i.e., $\Delta = \sum_{k=0}^{\infty} \Gamma(k)$. We let $B = (B_1', B_2')'$ and partition - $= (-_{ij})$ and $\Delta = (\Delta_{ij})$ into cell matrices for $i, j = 1, 2$, conformably with $w_t = (u_t', v_t')'$.

2.3 Inference on Parameters

As is well known, the parameter Π in the multivariate cointegrating regression (1) can be consistently estimated by the OLS estimator $\hat{\Pi}_n$, whose limiting distribution is given by

$$n(\hat{\Pi}_n - \Pi) \to_d \left(\int_0^1 B_2 B_2' \right)^{-1} \left(\int_0^1 B_2 dB_1' + \Delta_{21} \right)$$

as $n \to \infty$. Though the OLS estimator $\hat{\Pi}_n$ is super-consistent, it is asymptotically biased and inefficient when (u_t) and (v_t) are correlated. Moreover, the tests based on the OLS estimator $\hat{\Pi}_n$ are generally invalid. For the test of the hypothesis

$$\mathrm{H}_0: \ \lambda(\Pi) = 0 \tag{6}$$

where $\lambda : \mathbf{R}^{\ell m} \to \mathbf{R}^{\kappa}$ is continuously differentiable with first-order derivative $\Lambda = \partial\lambda/\partial\pi', \pi = \mathrm{vec}\,\Pi$, we may consider the Wald-type statistic

$$\hat{T}_n = \hat{\lambda}'_n \left(\hat{\Lambda}_n (M_n^{-1} \quad \text{-}_{11}) \hat{\Lambda}'_n \right)^{-1} \hat{\lambda}_n \tag{7}$$

where $\hat{\lambda}_n = \lambda(\hat{\Pi}_n)$, $\hat{\Lambda}_n = \Lambda(\hat{\Pi}_n)$ and $M_n = \sum_{t=1}^n x_t x'_t$. For the practical implementation, of course, the longrun variance -$_{11}$ of (u_t) must be estimated. The statistic has the limiting distribution

$$\hat{T}_n \to_d \tau' Q^{-1} \tau \tag{8}$$

where $\tau = \Lambda(M^{-1} \quad I_\ell)(\int_0^1 B_2 \quad dB_1 + \delta_{21})$ and $Q = \Lambda(M^{-1} \quad \text{-}_{11})\Lambda'$ using the notations $\Lambda = \Lambda(\Pi)$, $\delta_{21} = \mathrm{vec}\,\Delta_{21}$ and $M = \int_0^1 B_2 B'_2$. The asymptotic distribution of T_n is thus nonstandard and dependent upon various nuisance parameters. See, e.g., Park and Phillips (1988) for the distribution theories for $\hat{\Pi}_n$ and $\hat{T}_n$.

For the efficient estimation of Π, we consider the procedure suggested by Saikkonen (1991) and Stock and Watson (1993), which is based on the regressions augmented with the leads and lags of the first-differenced regressors $\triangle x_t$. Under Assumption 2.1, we may write

$$u_t = \sum_{k=-\infty}^{\infty} \Pi'_k v_{t-k} + \eta_t \tag{9}$$

where (η_t) is uncorrelated with $(v_t) = (\triangle x_t)$ at all leads and lags with variance -$_{11\cdot2}$ = -$_{11}$ $-$ -$_{12}$ $_{22}^{-1}$-$_{21}$, and $\sum_{k=-\infty}^{\infty} |\Pi_k| < \infty$. The reader is referred to Saikkonen (1991) and the references cited therein for the representation in (9).

We are therefore led to consider the regression

$$y_t = \Pi' x_t + \sum_{|k| \leq p} \Pi'_k \triangle x_{t-k} + \eta_{pt} \tag{10}$$

where $\eta_{pt} = \eta_t + \sum_{|k|>p} \Pi'_k \triangle x_{t-k}$. Since (Π_k) are absolutely summable, we may well expect that the error (η_{pt}) will become close to (η_t) if we let the number p of included leads and lags of the differenced regressors increase appropriately as the sample size n grows. Indeed, Saikkonen (1991) shows that if we let $p \to \infty$ such that $p^3/n \to 0$ and $n^{1/2} \sum_{|k|>p} |\Pi_k| \to 0$, then the regression (10) is asymptotically equivalent to

$$y_t = \Pi' x_t + \eta_t \tag{11}$$

In particular, the OLS estimator $\tilde{\Pi}_n$ of Π in regression (10), which we call the efficient OLS estimator, has the asymptotics given by

$$n(\tilde{\Pi}_n - \Pi) \to_d \left(\int_0^1 B_2 B'_2 \right)^{-1} \int_0^1 B_2 d(B_1 - \text{-}_{12} \text{ }_{22}^{-1} B_2)' \tag{12}$$

which is precisely the same as the asymptotics of the OLS estimator of Π from regression (11).

Note that $B_1 - \Omega_{12}\Omega_{22}^{-1}B_2$ has reduced variance $\Omega_{11\cdot2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$ relative to the variance Ω_{11} of B_1. The asymptotic variance of $\tilde{\Pi}_n$ is therefore smaller than that of $\hat{\Pi}_n$. Moreover, $B_1 - \Omega_{12}\Omega_{22}^{-1}B_2$ is independent of B_2, and consequently the limiting distribution of $\tilde{\Pi}_n$ is mixed normal, which is quite a contrast to the nonstandard limit theory of $\hat{\Pi}_n$. As a result, the usual chi-square test is valid if we use a test statistic based on $\tilde{\Pi}_n$. Indeed, for testing the hypothesis (6), we may use

$$\tilde{T}_n = \tilde{\lambda}_n' \left(\tilde{\Lambda}_n (M_{nn}^{-1} \quad \Omega_{11\cdot2}) \tilde{\Lambda}_n' \right)^{-1} \tilde{\lambda}_n \tag{13}$$

where $\tilde{\lambda}_n = \lambda(\tilde{\Pi}_n)$ and $\tilde{\Lambda}_n = \Lambda(\tilde{\Pi}_n)$ are defined similarly as $\hat{\lambda}_n$ and $\hat{\Lambda}_n$ introduced in (7), and M_{nn} is the matrix defining the sample covariance of $\tilde{\Pi}_n$, which corresponds to M_n for $\hat{\Pi}_n$ given also in (7). Then it follows that

$$\tilde{T}_n \to_d \chi^2_\kappa \tag{14}$$

As before, $\Omega_{11\cdot2}$ should be consistently estimated for practical applications.

In what follows, we reestablish the asymptotics for the augmented regression (10) using a weaker condition on the expansion rate for p, relative to the one required by Saikkonen (1991). We assume

Assumption 2.3 *Let $p \to \infty$ and $p = o(n^{1/2})$ as $n \to \infty$.*

Our condition on p is substantially weaker than the one used in Saikkonen (1991). Our upper bound for p is extended to $o(n^{1/2})$ compared to his $o(n^{1/3})$. Moreover, we do not impose any restriction on the minimum rate at which p must grow. Thus we may allow for the logarithmic rate that is usually imposed for the common order selection rules such as AIC and BIC. The logarithmic rate, however, is not allowed in Saikkonen (1991) unless (Π_k) decreases at a geometric rate. The rate for p here is therefore valid for more general stationary process (w_t), as can be seen clearly in our proof for the following lemma.

Lemma 2.4 *Under* Assumptions 2.1 and 2.3, *we have* (12) *and* (14) *as $n \to \infty$.*

The asymptotics established in the above lemma will be used as a basis for the subsequent development of our bootstrap asymptotics.

3. Bootstrap Procedures and Their Asymptotics

3.1 Bootstrap Procedures

In this section, we introduce the bootstrap procedures for our cointegrating regression model (1) and develop their asymptotics. For our bootstrap procedures introduced below, we may

use various consistent estimates for Π. Therefore, we use the generic notation Π_n to denote any estimate of Π that is n-consistent. More explicitly, we let

$$\Pi_n = \hat{\Pi}_n, \tilde{\Pi}_n, \Pi_0$$

where $\hat{\Pi}_n$ and $\tilde{\Pi}_n$ are the OLS estimators of Π in regressions (1) and (10), respectively, and Π_0 denotes the hypothesized or estimated value of Π under the restriction given by the hypothesis (6). Other estimates of Π, which are asymptotically equivalent to any of these estimators, can also be used.

The cointegrating regression model (1) with the specification (2) of its stationary component as a linear process can be bootstrapped using the standard sieve method. The method of sieve bootstrap requires to fit the linear process (w_t) to a finite order VAR with the order increasing as the sample size grows. We may rewrite (w_t) as a VAR

$$\Phi(L)w_t = \varepsilon_t$$

with $\Phi(z) = I - \sum_{k=1}^{\infty} \Phi_k z^k$, since $\Psi(L)$ in (2) is invertible, and therefore it is reasonable to approximate (w_t) as a finite order VAR

$$w_t = \Phi_1 w_{t-1} + \cdots + \Phi_q w_{t-q} + \varepsilon_{qt} \tag{15}$$

with $\varepsilon_{qt} = \varepsilon_t + \sum_{|k|>q} \Phi_k w_{t-k}$. The order q of the approximated VAR is set to increase at a controlled rate of n, as we will specify below. In practice, it can be chosen by one of the commonly used order selection rules such as AIC and BIC.

Assumption 3.1 *Let $q \to \infty$ and $q = o(n^{1/2})$ as $n \to \infty$.*

The required condition on the expansion rate for q in Assumption 3.1 is identical to that for p given in Assumption 2.3.

Outlined below are the necessary steps for obtaining the bootstrap samples (x_t^*) and (y_t^*) for (x_t) and (y_t), respectively.

Step 1: Fit regression (1) to obtain $(\hat{u}_t)$, i.e.,

$$y_t = \Pi_n' x_t + \hat{u}_t$$

and define $\hat{w}_t = (\hat{u}_t', v_t')'$, where $v_t = \triangle x_t$. As mentioned above, we may use various estimates Π_n of Π here.

Step 2: Apply the sieve estimation method to the VAR (15) of $(\hat{w}_t)$ to get the fitted values $(\hat{\varepsilon}_{qt})$ of (ε_{qt}), i.e.,

$$\hat{w}_t = \hat{\Phi}_1 \hat{w}_{t-1} + \cdots + \hat{\Phi}_q \hat{w}_{t-q} + \hat{\varepsilon}_{qt} \tag{16}$$

Obtain (ε_t^*) by resampling the centered fitted residuals

$$\left(\hat{\varepsilon}_{qt} - \frac{1}{n} \sum_{t=1}^{n} \hat{\varepsilon}_{qt} \right)_{t=1}^{n}$$

and construct the bootstrap samples (w_t^*) recursively using

$$w_t^* = \hat{\Phi}_1 w_{t-1}^* + \cdots + \hat{\Phi}_q w_{t-q}^* + \varepsilon_t^*$$

given the initial values $w_t^* = w_t$ for $t = 0, \ldots, 1 - q$. This step amounts to the usual sieve bootstrap for $(\hat{w}_t)$.

Step 3: Define $w_t^* = (u_t^{*\prime}, v_t^{*\prime})'$ analogously as $w_t = (u_t', v_t')'$. Obtain the bootstrap samples (x_t^*) by integrating (v_t^*), i.e., $x_t^* = x_0^* + \sum_{k=1}^{t} v_t^*$, with $x_0^* = x_0$ and generate the bootstrap samples (y_t^*) from

$$y_t^* = \Pi_n' x_t^* + u_t^*$$

The estimate Π_n of Π here need not be the same as the one used in Step 1.

The bootstrap samples (x_t^*) and (y_t^*) can be used to simulate the distributions of various statistics as explained below.

Discussions on some practical issues arising from the implementation of our bootstrap method are in order. First, the choice of an estimator Π_n for Π should be made in Steps 1 and 3. The choice in Step 3 is not important, as long as we regard the chosen estimate as the true parameter for the bootstrap samples. The choice of Π_n for Π we use in Step 1, however, is very important, as it would affect the subsequent bootstrap procedure in Steps 2 and 3. In particular, the order selection rule applied to determine the model in Step 2 can be very sensitive to the choice of Π_n made in Step 1. Naturally, we recommend to use the best possible estimate, i.e., the most efficient estimate available incorporating all restrictions for the hypothesis testing. This is what is observed by Chang and Park (2002b) for the bootstrap unit root tests. Li and Maddala (1997) also find that using the true value in Step 1 yields the best result for the bootstrap hypotheses tests. Second, the initializations for the generations of (w_t^*) and (x_t^*) are necessary respectively in Steps 2 and 3. The choices $w_t^* = w_t$ for $t = 0, \ldots, 1 - q$ and $x_0^* = x_0$ make the results conditional on these initial values of the samples. To reduce or eliminate such dependencies, we may generate sufficiently large number of (w_t^*) and retain only last n drawings, or include a constant so that the estimate Π_n of Π is independent of the initial value of (x_t^*).

We consider the bootstrap version of regression (1) given as

$$y_t^* = \Pi_n' x_t^* + u_t^* \tag{17}$$

which yields the bootstrap estimator for the usual OLS, viz.,

$$\hat{\Pi}_n^* = \left(\sum_{t=1}^{n} x_t^* x_t^{*\prime} \right)^{-1} \left(\sum_{t=1}^{n} x_t^* y_t^{*\prime} \right)$$

We also look at the bootstrap version of regression (10), which we write as

$$y_t^* = \Pi_n' x_t^* + \sum_{|k| \leq p} \Pi_k^{*\prime} \triangle x_{t-k}^* + \eta_{pt}^* \tag{18}$$

where

$$\eta_{pt}^* = \sum_{|k|>p} \Pi_k^{*\prime} v_{t-k}^* + \eta_t^*$$

The efficient bootstrap OLS estimator, i.e., the OLS estimator from regression (18), is given by

$$\tilde{\Pi}_n^* = \left(\sum_{t=1}^n x_t^* x_t^{*\prime} - \left(\sum_{t=1}^n x_t^* v_{pt}^{*\prime} \right) \left(\sum_{t=1}^n v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \left(\sum_{t=1}^n v_{pt}^* x_t^{*\prime} \right) \right)^{-1}$$
$$\cdot \left(\sum_{t=1}^n x_t^* y_t^{*\prime} - \left(\sum_{t=1}^n x_t^* v_{pt}^{*\prime} \right) \left(\sum_{t=1}^n v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \left(\sum_{t=1}^n v_{pt}^* y_t^{*\prime} \right) \right)$$

where

$$v_{pt}^* = (\triangle x_{t+p}^{*\prime}, \ldots, \triangle x_{t-p}^{*\prime})'$$

In both regressions (17) and (18), we denote by Π_n the estimate of Π used to generate the bootstrap samples (y_t^*) from (x_t^*) and (u_t^*). Note that in regression (18) the coefficients of the leads and lags of the differenced regressors depend on the realized samples, which we signify by attaching the superscript "$*$" to the coefficients Π_k's. The test statistics

$$\hat{T}_n^* = \hat{\lambda}_n^{*\prime} \left(\hat{\Lambda}_n^* (M_n^{*-1} \quad \hat{-}_{11}) \hat{\Lambda}_n^{*\prime} \right)^{-1} \hat{\lambda}_n^*$$
$$\tilde{T}_n^* = \tilde{\lambda}_n^{*\prime} \left(\tilde{\Lambda}_n^* (M_{nn}^{*-1} \quad \hat{-}_{11\cdot 2}) \tilde{\Lambda}_n^{*\prime} \right)^{-1} \tilde{\lambda}_n^*$$

which are constructed from the bootstrap OLS and efficient estimators, $\hat{\Pi}_n^*$ and $\tilde{\Pi}_n^*$, analogously as as their sample counterparts $\hat{T}_n$ and $\tilde{T}_n$ defined in (7) and (13). For the definitions of the bootstrap statistics $\hat{T}_n^*$ and $\tilde{T}_n^*$, we may use the longrun variance estimate obtained from each bootstrap sample, say $-_{11}^*$ and $-_{11\cdot 2}^*$, in the places of the sample estimates $\hat{-}_{11}$ and $\hat{-}_{11\cdot 2}$. This would not affect any of our subsequent results, since we only consider the first order asymptotics.

3.2 Bootstrap Asymptotics

The asymptotic theories of the estimators $\hat{\Pi}_n^*$ and $\tilde{\Pi}_n^*$ can be developed similarly as those for $\hat{\Pi}_n$ and $\tilde{\Pi}_n$. To develop their asymptotics, we first establish the bootstrap invariance principle for (ε_t^*). We have

Lemma 3.2 *Under* Assumptions 2.1 *and* 3.1,

$$\mathbf{E}^* |\varepsilon_t^*|^a = O_p(1)$$

as $n \to \infty$.

Roughly, Lemma 3.2 allows us to regard the bootstrap samples (ε_t^*) as iid random variables with finite a-th moment, given a sample realization. Following the usual convention, we

use the notation $\mathbf{P}^*$ to denote the bootstrap probability conditional on the samples. The notation $\mathbf{E}^*$ signifies the expectation taken with respect to $\mathbf{P}^*$, for each realization of the samples.

To develop the bootstrap asymptotics, it is convenient to introduce the bootstrap stochastic order symbols o_p^* and O_p^*, which correspond respectively to o_p and O_p. Let $\epsilon > 0$ be given arbirarily, and $M > 0$ be chosen accordingly. For a bootstrap statistic T_n^*, we define

$$T_n^* = o_p^*(1) \ \text{ if and only if } \ \mathbf{P}\left\{\mathbf{P}^*\{|S_n^*| < \epsilon\} > \epsilon\right\} < \epsilon$$

for all large n, and

$$T_n^* = O_p^*(1) \ \text{ if and only if } \ \mathbf{P}\left\{\mathbf{P}^*\{|S_n^*| > M\} > \epsilon\right\} < \epsilon$$

for all large n. They all satisfy the usual rules applicable for o_p and O_p. See Chang and Park (2002b) for more details. In particular, whenever $T_n^* \to_{d^*} T$ in $\mathbf{P}$, we have $T_n^* = O_p^*(1)$ and $T_n^* + o_p^*(1) \to_{d^*} T$ in $\mathbf{P}$. Note also that $\mathbf{E}^*|T_n^*| = o_p(1)$ and $O_p(1)$ imply $T_n^* = o_p^*(1)$ and $O_p^*(1)$, respectively.

Define

$$W_n^*(r) =_{d^*} n^{-1/2} \sum_{t=1}^{n} \varepsilon_t^*$$

analogously as W_n introduced earlier in Section 2, and let W be a vector Brownian motion with variance Σ. It follows from the strong approximation result in Lemma 2.2 applied to the bootstrap samples (ε_t^*) that we may choose W_n^* satisfying

$$\mathbf{P}^*\left\{\sup_{0 \leq r \leq 1} |W_n^*(r) - W(r)| > n^{-1/2}c_n\right\} \leq Kn\, c_n^{-a}\, \mathbf{E}^*|\varepsilon_t^*|^a$$

for (c_n) and K given exactly as in Lemma 2.2. Note in particular that K does not depend on the realization of the samples. However, due to the result in Lemma 3.2, we have

$$\sup_{0 \leq r \leq 1} |W_n^* - W| = o_p^*(1)$$

as long as $a > 2$, and consequently

$$W_n^* \to_{d^*} W \ \text{ in } \mathbf{P}$$

An 'in probability' version of the bootstrap invariance principle is therefore established for the bootstrap samples (ε_t^*).

The corresponding invariance principle for (w_t^*) can also be derived easily. We let

$$\hat{\Phi}(1) = 1 - \sum_{k=1}^{q} \hat{\Phi}_k$$

and

$$\hat{\Psi}(1) = \hat{\Phi}(1)^{-1}$$

Then we may deduce after some algebra that

$$
\begin{aligned}
w_t^* &= \hat{\Psi}(1)\varepsilon_t^* + \hat{\Psi}(1)\sum_{i=1}^{q}\left(\sum_{j=i}^{q}\hat{\Phi}_j\right)(w_{t-i}^* - w_{t-i+1}^*) \\
&= \hat{\Psi}(1)\varepsilon_t^* + (\bar{w}_{t-1}^* - \bar{w}_t^*)
\end{aligned}
$$

where $\bar{w}_t^* = \hat{\Psi}(1)\sum_{i=1}^{q}(\sum_{j=i}^{q}\hat{\Phi}_j)w_{t-i+1}^*$. We therefore have

$$
\begin{aligned}
B_n^*(r) &= n^{-1/2}\sum_{t=1}^{[nr]}w_t^* \\
&= n^{-1/2}\hat{\Psi}(1)\sum_{t=1}^{[nr]}\varepsilon_t^* + n^{-1/2}(\bar{w}_0^* - \bar{w}_{[nr]}^*) \\
&= \hat{\Psi}(1)W_n^*(r) + R_n^*(r)
\end{aligned}
$$

It is thus well expected that the invariance principle holds for (w_t^*) if $\hat{\Psi}(1) \to_p \Psi(1)$ and $\sup_{0 \leq r \leq 1}|R_n^*(r)| = o_p^*(1)$. Let B be the vector Brownian motion introduced in (5). Then we may indeed show that

Theorem 3.3 *Under* Assumptions 2.1 *and* 3.1, *we have*

$$
B_n^* \to_{d^*} B \text{ in } \mathbf{P}
$$

as $n \to \infty$.

Let

$$
z_t^* = z_0^* + \sum_{i=1}^{t}w_i^*
$$

Then we have

Lemma 3.4 *Under* Assumptions 2.1 *and* 3.1, *we have*

$$
\frac{1}{n^2}\sum_{t=1}^{n}z_t^* z_t^{*\prime} \to_{d^*} \int_0^1 BB' \text{ in } \mathbf{P}
$$

$$
\frac{1}{n}\sum_{t=1}^{n}z_{t-1}^* w_t^{*\prime} \to_{d^*} \int_0^1 BdB' + \Delta \text{ in } \mathbf{P}
$$

as $n \to \infty$.

The limit distributions of the bootstrap OLS estimator and the associated statistic can now be obtained easily from Lemma 3.4, and are given in the following theorem. Let τ and Q be defined as in (8).

Theorem 3.5 *Under* Assumptions 2.1 *and* 3.1, *we have*

$$n(\hat{\Pi}_n^* - \Pi_n) \to_{d^*} \left(\int_0^1 B_2 B_2' \right)^{-1} \left(\int_0^1 B_2 dB_1' + \Delta_{21} \right) \ \text{in} \ \mathbf{P}$$

and

$$\hat{T}_n^* \to_{d^*} \tau' Q^{-1} \tau \ \text{in} \ \mathbf{P}$$

as $n \to \infty$.

The bootstrap is therefore consistent for the OLS regression. The bootstrap estimator $\hat{\Pi}_n^*$ has the same limiting distribution as $\hat{\Pi}_n$. Therefore, in particular, the bootstrap can be used to remove the asymptotic bias in $\hat{\Pi}_n$, since the bootstrap distribution has the same asymptotic bias as the sample distribution. The bias corrected OLS estimator defined by

$$\hat{\Pi}_n^c = \hat{\Pi}_n - \mathbf{E}^*(\hat{\Pi}_n^* - \Pi_n)$$

does not have asymptotic bias, and is thus expected to have smaller bias in finite samples as well. Hypothesis testing can also be done using the OLS estimator, if the critical values are obtained from the bootstrap distribution. For example, consider the test of the hypothesis (6). If we let $c^*(\alpha)$ be the bootstrap critical value given by

$$\mathbf{P}^* \left\{ \hat{T}_n^* \leq \hat{c}^*(\alpha) \right\} = \alpha$$

for a prescribed size α, then we have

$$\mathbf{P} \left\{ \hat{T}_n \leq \hat{c}^*(\alpha) \right\} \to \alpha$$

as $n \to \infty$. Therefore, the bootstrap test has the asymptotically correct size.

We now consider the asymptotics for the bootstrap efficient estimator $\tilde{\Pi}_n^*$ obtained from (18). Denote by $f^*(\lambda)$ and $\Gamma^*(k)$, respectively, the spectral density and autocovariance function of (w_t^*). For all bootstrap samples yielding spectral density bounded away from zero and absolutely summable autocovariance function, we would have the representation

$$u_t^* = \sum_{k=-\infty}^{\infty} \Pi_k^{*\prime} v_{t-k}^* + \eta_t^* \tag{19}$$

where $\sum_{k=-\infty}^{\infty} |\Pi_k^*| < \infty$ and (η_t^*) are uncorrelated with (v_t^*) at all leads and lags.

As shown in Lemma A3 in Appendix, we have

$$\sup_\lambda \|f^*(\lambda) - f(\lambda)\| = o_p^*(1)$$

and

$$\sum_{k=-\infty}^{\infty} \Gamma^*(k) = \sum_{k=-\infty}^{\infty} \Gamma(k) + o_p^*(1)$$

It therefore follows that

$$\mathbf{P}^* \left\{ f^*(\lambda) > \epsilon \right\} \to_p 1$$

for some $\epsilon > 0$ and

$$\mathbf{P}^* \left\{ \sum_{k=-\infty}^{\infty} |\Gamma^*(k)| < \infty \right\} \to_p 1$$

since $f(\lambda)$ is bounded away from zero and $\Gamma(k)$ is absolutely summable. We may therefore deduce that the probability of realization of samples which allow for the representation (19) approaches to one as the sample size increases. The bootstrapped augmented regression (18) is therefore well justified.

Similarly as for the asymptotic equivalence between the sample regressions (10) and (11), we have the bootstrap asymptotic equivalence between the regression (18) and

$$y_t^* = \Pi_n' x_t^* + \eta_t^* \tag{20}$$

for the inference on Π_n. We define

$$N_n^* = \frac{1}{n} \sum_{t=1}^{n} x_t^* \eta_t^{*\prime}, \quad M_n^* = \frac{1}{n^2} \sum_{t=1}^{n} x_t^* x_t^{*\prime}$$

and

$$
\begin{aligned}
N_{nn}^* &= \frac{1}{n} \sum_{t=1}^{n} x_t^* \eta_{pt}^{*\prime} - \left(\frac{1}{n} \sum_{t=1}^{n} x_t^* v_{pt}^{*\prime} \right) \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* \eta_{pt}^{*\prime} \right) \\
&= N_n^* + P_n^* \\
M_{nn}^* &= \frac{1}{n^2} \sum_{t=1}^{n} x_t^* x_t^{*\prime} - \frac{1}{n} \left(\frac{1}{n} \sum_{t=1}^{n} x_t^* v_{pt}^{*\prime} \right) \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* x_t^{*\prime} \right) \\
&= M_n^* + Q_n^*
\end{aligned}
$$

Then we have

Lemma 3.6 *Under* Assumptions 2.1, 2.3 *and* 3.1, *we have*

$$P_n^*, \ Q_n^* = o_p^*(1)$$

as $n \to \infty$.

Therefore, the bootstrap asymptotics for the regressions (18) and (20) are the same. Consequently it follows that

Theorem 3.7 *Under* Assumptions 2.1, 2.3 and 3.1, *we have*

$$n(\tilde{\Pi}_n^* - \Pi_n) \to_{d*} \left(\int_0^1 B_2 B_2' \right)^{-1} \int_0^1 B_2 d(B_1 - \text{-}_{12} \text{-}_{22}^{-1} B_2)' \ \text{in } \mathbf{P}$$

and

$$\tilde{T}_n^* \to_{d*} \chi_\kappa^2 \ \text{in } \mathbf{P}$$

as $n \to \infty$.

The limit distribution of the bootstrap efficient estimator $\tilde{\Pi}_n^*$ is equivalent to that of its sample counterpart $\tilde{\Pi}_n$, just as in the case for the OLS estimator. The associated statistics $\tilde{T}_n^*$ and $\tilde{T}_n$ also have the identical asymptotic distribution. The bootstrap consistency of the test statistic for cointegrating regression (1) therefore extends to the dynamic cointegrating regression (10) augmented with the leads and lags of the differenced regressors.

We may also consider the bias corrected estimator

$$\tilde{\Pi}_n^c = \tilde{\Pi}_n - \mathbf{E}^*(\tilde{\Pi}_n^* - \Pi_n)$$

for the efficient estimator $\tilde{\Pi}_n$. Though it has no asymptotic bias, the correction may still reduce the bias in finite samples. Moreover, the bootstrap critical value for the test $\tilde{T}_n$ can be found

$$\mathbf{P}^* \left\{ \tilde{T}_n^* \leq \tilde{c}^*(\alpha) \right\} = \alpha$$

for the size α test. The bootstrap test based on $b^*(\alpha)$ would then have the correct asymptotic size, since

$$\mathbf{P} \left\{ \tilde{T}_n \leq \tilde{c}^*(\alpha) \right\} \to \alpha$$

as $n \to \infty$. Note that $\tilde{T}_n$ is asymptotically pivotal, unlike $\hat{T}_n$ whose limiting distribution depends on various nuisance parameters. Therefore, the bootstrap test based on $\tilde{T}_n$ may provide an asymptotic refinement. In this case, the bootstrap test utilizing the bootstrap critical values $\tilde{c}^*(\alpha)$ would yield rejection rates closer to the nominal values, compared to the test relying on the asymptotic chi-square critical values.

4. Simulations

In this section, we conduct a set of simulations to investigate finite sample performances of the bootstrap procedures in the context of cointegrating regressions. Our simulations are based on the simple bivariate cointegrating regression model

$$\begin{aligned} y_t &= \pi x_t + u_t \\ x_t &= x_{t-1} + v_t \end{aligned}$$

where the regression error (u_t) and the innovation (v_t) for the regressor (x_t) are generated as stationary AR(1) processes given by

$$\begin{aligned} u_t &= \varphi_1 v_{t-1} + \varepsilon_{1t} \\ v_t &= \varphi_2 v_{t-1} + \varepsilon_{2t} \end{aligned}$$

The innovations (ε_t), $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, are drawn from bivariate normal distribution with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Our simulation setup is entirely analogous to that of Stock and Watson (1993). In our simulations, we set $\varphi_1 = 0.6$, $\varphi_2 = 0.3$, $\rho = 0.5$, and $\pi = 0, 0.1$.

We consider two least squares regressions

$$y_t = \hat{\pi}_n x_t + \hat{u}_t \tag{21}$$
$$y_t = \tilde{\pi}_n x_t + \tilde{\pi}_{1n} \triangle x_t + \tilde{\pi}_{2n} \triangle x_{t-1} + \tilde{\eta}_t \tag{22}$$

and compare the finite sample performances of the estimators $\hat{\pi}_n$ and $\tilde{\pi}_n$ and the test statistics $\hat{T}_n$ and $\tilde{T}_n$ based respectively on $\hat{\pi}_n$ and $\tilde{\pi}_n$. Note that our simulation setup introduces nonzero asymptotic correlation between the regressor and the regression error in regression (21), causing the asymptotic bias of $\hat{\pi}_n$ and the invalidity of $\hat{T}_n$. The problems are however expected to vanish, at least asymptotically, in regression (22). As explained earlier, $\tilde{\pi}_n$ has no asymptotic bias, and the test using $\tilde{T}_n$ is asymptotically valid. For our simulation setup, we have in regression (22)

$$\pi_1 = \rho, \quad \pi_2 = \varphi_1 - \rho\varphi_2$$

and

$$\eta_t = \varepsilon_{1t} - \rho\varepsilon_{2t}$$

Notice that (η_t) defined as above are uncorrelated with (ε_{2t}) at all leads and lags, and therefore are orthogonal also to (v_t) at all leads and lags, which is required for the validity of the model (22) as an efficient cointegrating regression introduced in (10).

To examine the effectiveness of the bootstrap bias correction, the finite sample performances of $\hat{\pi}_n$ and $\tilde{\pi}_n$ are compared to those of the bias corrected estimators $\hat{\pi}_n^c$ and $\tilde{\pi}_n^c$, in terms of bias, variance and mean square error (MSE). For the sake of comparisons, the scaled biases, variances and MSE's for the estimators $\hat{\pi}_n$, $\tilde{\pi}_n$, $\hat{\pi}_n^c$ and $\tilde{\pi}_n^c$ are computed. For the test statistics, we compute the rejection probabilities under both the null and the alternative hypotheses, for which we set the value of π respectively at 0 and 0.1. Rejection probabilities are obtained using the critical values from χ_1^2 distribution for the sample test statistics, $\hat{T}_n$ and $\tilde{T}_n$, and from the bootstrap distribution for the bootstrap tests, $\hat{T}_n^*$ and $\tilde{T}_n^*$. The simulation results are presented in Tables 1, 2 and 3, respectively for the models with no deterministic components, with drift and with a linear trend. We use the superscripts "μ" and "τ" to signify that the associated estimators/tests are computed from the models with drift and with a linear trend. Samples of sizes $n = 25, 50, 100$ and 200 are considered. For each case, 5000 samples are simulated, and for each of the simulated samples, 1000 bootstrap repetitions are carried out to compute bootstrap estimators and test statistics.

The simulation results are largely consistent with the theory derived in the paper. The bias of $\hat{\pi}_n$ is quite noticeable and does not vanish as the sample size increases. It turns out, however, that the bootstrap is quite effective in reducing the bias of $\hat{\pi}_n$. The bootstrap reduces the bias of $\hat{\pi}_n$ substantially in all cases that we consider in our simulations, and the bias reduction becomes more effective as the sample size increases. In some of the reported cases, the bias corrected estimate has bias as small as approximately 2% of that of $\hat{\pi}_n$. On the other hand, the bootstrap bias correction has little impact on the sampling variation.

The sample variances of $\hat{\pi}_n^c$ are roughly comparable to, or even slightly larger than, those of $\hat{\pi}_n$. The bootstrap also reduces the bias in $\tilde{\pi}_n$, albeit the magnitudes of improvements are not comparable to those for $\hat{\pi}_n$. Clearly, $\tilde{\pi}_n$ is asymptotically unbiased, and thus there is not as much room for improvement in $\tilde{\pi}_n$ as in $\hat{\pi}_n$. The bootstrap does not reduce the sampling variations for $\tilde{\pi}_n$, just as for $\hat{\pi}_n$, and $\tilde{\pi}_n^c$ has no smaller sample variances. For both estimators, the bootstrap bias corrections become relatively more important for models with a mean or a linear trend.

It appears that the bootstrap generates quite precise critical values for both tests $\hat{T}_n$ and $\tilde{T}_n$, even when the sample sizes are small. In particular, the performance of $\hat{T}_n^*$ is quite satisfactory. The test $\hat{T}_n$ is invalid, and naturally, yields the rejection probabilities under the null that are very distinct from its nominal test sizes. However, its bootstrap version $\hat{T}_n^*$ has quite accurate null rejection probabilities even for small samples. As expected, the test $\tilde{T}_n$ based on the efficient estimator $\tilde{\pi}_n$ performs much better than its OLS counterpart $\hat{T}_n$ in terms of the rejection probabilities under the null. Its bootstrap version $\tilde{T}_n^*$ also improves the finite sample performances of $\tilde{T}_n$. While the null rejection probabilities of both bootstrap tests $\hat{T}_n^*$ and $\tilde{T}_n^*$ approach to the nominal values as the size of samples increase, the bootstrap efficient test $\tilde{T}_n^*$ does better in small samples than its OLS counterpart $\hat{T}_n^*$. This is more so for models with a mean or a linear trend. Due to the presence of non-uniform size distortions, the finite sample power comparisons between $\hat{T}_n$ and $\tilde{T}_n$ and their bootstrap counterparts $\hat{T}_n^*$ and $\tilde{T}_n^*$ are not clear. It just seems that they are all roughly comparable.

5. Conclusion

In this paper we consider the bootstrap for cointegrating regressions. We introduce the sieve bootstrap based on a VAR of order increasing with the sample size, and establish its consistency and asymptotic validity for two procedures: the usual OLS and the efficient OLS relying on the regressions augmented with the leads and lags of the differenced regressors. For the usual OLS, the bootstrap can thus be employed to correct for biases in the estimated parameters, and to compute the critical values of the tests. With the bootstrap bias correction, the OLS estimator becomes asymptotically unbiased. Moreover, the OLS-based tests become asymptotically valid, if the bootstrap critical values are used. The bootstrap OLS method, however, is not efficient. For the efficient inference, we should base our bootstrap procedure on the efficient OLS method. The sieve bootstrap proposed in the paper appears to improve upon the efficient OLS method in finite samples. It generally reduces the finite sample biases for the estimators and yields sizes that are closer to the nominal sizes of the tests.

The theory and method developed in the paper can be used to analyze more general cointegrated models. The models with deterministic trends and/or structural breaks can be analyzed similarly. The error correction models, seemingly unrelated and panel cointegration models are other examples, to which our theory and method are readily applicable. Indeed, the required modifications of the theory and method for such extensions are minimal, and can be applied only with some obvious adjustments. The theory discussed in the paper is concerned only with the consistency of the bootstrap. For the pivotal statistics,

however, it might well be the case that the bootstrap provides the asymptotic refinements as well. Our simulation study in the paper is in fact somewhat indicative of this possibility. Moreover, it appears that the theoretical illustration for the asymptotic refinement is possible using the techniques developed in Park (2000) to establish the bootstrap refinement for the unit root tests.

6. Mathematical Proofs

6.1 Useful Lemmas and Their Proofs

We consider the regression

$$w_t = \tilde{\Phi}_1 w_{t-1} + \cdots + \tilde{\Phi}_q w_{t-q} + \tilde{\varepsilon}_{qt} \tag{23}$$

which is the fitted version of regression (15). Since Π_n is consistent, it is well expected that the fitted regression (23) with (w_t) should be asymptotically equivalent to the fitted regression (16) with $(\hat{w}_t)$ introduced in Step 1 of our bootstrap procedure.

Lemma A1 *Under* Assumptions 2.1 *and* 3.1, *we have as* $n \to \infty$

$$\hat{\Phi}_k = \tilde{\Phi}_k + O_p(n^{-1/2})$$

uniformly in $1 \le k \le q$. *Moreover*,

$$\max_{1 \le t \le n} |\hat{\varepsilon}_{qt} - \tilde{\varepsilon}_{qt}| = O_p(n^{-1/2})$$

as $n \to \infty$.

Proof of Lemma A1 It follows immediately from the definition of $(\hat{w}_t)$ that

$$\max_{0 \le i,j \le q} \left| \frac{1}{n} \sum_{t=1}^n \hat{w}_{t-i} \hat{w}'_{t-j} - \frac{1}{n} \sum_{t=1}^n w_{t-i} w'_{t-j} \right| \le \left(\max_{1 \le t \le n} |\hat{w}_t - w_t| \right) \left(\frac{1}{n} \sum_{t=1}^n |\hat{w}_t| + \frac{1}{n} \sum_{t=1}^n |w_t| \right)$$

However, we have

$$\max_{1 \le t \le n} |\hat{w}_t - w_t| = O_p(n^{-1/2})$$

because

$$|\hat{w}_t - w_t| = \left| \begin{pmatrix} y_t - \Pi'_n x_t \\ v_t \end{pmatrix} - w_t \right| = \left| \begin{pmatrix} (\Pi - \Pi_n)' x_t \\ 0 \end{pmatrix} \right| \le |\Pi - \Pi_n| \, |x_t|$$

and

$$|\Pi_n - \Pi| = O_p(n^{-1}), \quad \max_{1 \le t \le n} |x_t| = O_p(n^{1/2})$$

Since

$$\frac{1}{n} \sum_{t=1}^n |w_t|, \ \frac{1}{n} \sum_{t=1}^n |\hat{w}_t| = O_p(1)$$

we have

$$\max_{0 \le i,j \le q} \left| \frac{1}{n} \sum_{t=1}^{n} \hat{w}_{t-i} \hat{w}'_{t-j} - \frac{1}{n} \sum_{t=1}^{n} w_{t-i} w'_{t-j} \right| = O_p(n^{1/2})$$

The rest of the proof is rather straightforward, and we omit the details. ∎

Let

$$w_{\kappa t} = (w'_{t-1}, \dots, w'_{t-\kappa})'$$

and define

$$M_{\kappa\kappa} = \mathbf{E} w_{\kappa t} w'_{\kappa t}$$

Moreover, denote by f the spectral density of (w_t). Then we have

Lemma A2 *Under* Assumption 2.1, *we have*

$$\left\| M_{\kappa\kappa}^{-1} \right\| \le \frac{1}{2\pi} \left(\inf_\lambda \| f(\lambda) \| \right)^{-1}$$

for all $\kappa \ge 1$.

Proof of Lemma A2 Let $c_\kappa \in \mathbf{R}^\kappa$ be an eigenvector associated with the smallest eigenvalue $\lambda_{\min}$ of $M_{\kappa\kappa}$. Define

$$\varphi_\kappa(\lambda) = (e^{i\lambda}, \dots, e^{i\kappa\lambda})'$$

and use "¯" to denote its cojugate. It follows that

$$
\begin{aligned}
\lambda_{\min} &= c'_\kappa M_{\kappa\kappa} c_\kappa \\
&= \int_{-\pi}^{\pi} c'_\kappa \left(\varphi_\kappa(\lambda) \bar{\varphi}_\kappa(\lambda)' \quad f(\lambda) \right) c_\kappa \, d\lambda \\
&\ge \int_{-\pi}^{\pi} \left(\inf_\lambda \| \varphi_\kappa(\lambda) \bar{\varphi}_\kappa(\lambda)' \quad f(\lambda) \| \right) d\lambda \\
&= \left(\inf_\lambda \| f(\lambda) \| \right) \int_{-\pi}^{\pi} \| \varphi_\kappa(\lambda) \bar{\varphi}_\kappa(\lambda)' \| d\lambda \\
&= 2\pi \left(\inf_\lambda \| f(\lambda) \| \right)
\end{aligned}
$$

However, we have

$$\left\| M_{\kappa\kappa}^{-1} \right\| = \lambda_{\min}^{-1}$$

from which the stated result can be deduced immediately. ∎

Lemma A3 *Under* Assumptions 2.1 *and* 3.1, *we have*

$$\sup_\lambda |f^*(\lambda) - f(\lambda)| = o_p^*(1)$$

and

$$\sum_{k=-\infty}^{\infty} \Gamma^*(k) = \sum_{k=-\infty}^{\infty} \Gamma(k) + o_p^*(1)$$

as $n \to \infty$.

Proof of Lemma A3 Given Lemma A1, the stated results are just straightforward extensions of Lemma A2 in Chang and Park (2002b). Here we only obtain 'in probability' versions, instead of 'almost sure' versions, since the results in Lemma A1 hold only in probability. ■

Lemmma A4 *Under* Assumptions 2.1 *and* 3.1, *we have*

$$\mathbf{E}^* \left| \sum_{t=1}^{n} (w_{t-i}^* w_{t-j}^{*\prime} - \Gamma^*(i-j)) \right|^2 = O_p(n)$$

uniformly in i *and* j, *as* $n \to \infty$.

Proof of Lemma A4 Once again, the stated result follows exactly as in Lemma A4 in Chang and Park (2002b), due to Lemma A1, under some obvious modifications to deal with multiple time series. ■

Lemmma A5 *Under* Assumptions 2.1, 2.3 *and* 3.1, *we have*

$$\mathbf{E}^* \left\| \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \right\| = O_p(1)$$

$$\mathbf{E}^* \left| \sum_{t=1}^{n} x_t^* v_{pt}^{*\prime} \right| = O_p(np^{1/2})$$

as $n \to \infty$.

Proof of Lemma A5 The proof is essentially identical to that of Lemma 3.3 in Chang and Park (2002b). ■

Lemma A6 *Under* Assumptions 2.1, 2.3 *and* 3.1, *we have*

$$\left| \sum_{t=1}^{n} v_{pt}^* \eta_{pt}^{*\prime} \right| = O_p^*(n^{1/2} p^{1/2})$$

and

$$\sum_{t=1}^{n} x_t^* (\eta_t^* - \eta_{pt}^*)' = o_p^*(n)$$

as $n \to \infty$.

Proof of Lemma A6 Define $(\hat{\Psi}_{pk})$ such that

$$\eta_{pt}^* - \eta_t^* = \sum_{|k|>p} \hat{\Pi}_k' v_{t-k}^* = \sum_{|k|>p} \hat{\Psi}_{pk} \varepsilon_{t-k}^*$$

Note that

$$\sum_{|k|>p} |\hat{\Psi}_{pk}| \leq \left(\sum_{k=1}^{\infty} |\hat{\Psi}_k|\right) \left(\sum_{|k|>p} |\hat{\Pi}_k|\right)$$

as one may easily deduce.

To show the first part, we write for $1 \leq i \leq p$

$$\sum_{t=1}^{n} v_{t-i}^* \eta_{pt}^{*\prime} = \sum_{t=1}^{n} v_{t-i}^* \eta_t^{*\prime} + \sum_{t=1}^{n} v_{t-i}^* (\eta_{pt}^* - \eta_t^*)'$$

It is easy to see

$$\mathbf{E}^* \left|\sum_{t=1}^{n} v_{t-i}^* \eta_t^{*\prime}\right|^2 = O_p(n)$$

uniformly in $1 \leq i \leq p$. Therefore, it suffices to show that

$$\sum_{t=1}^{n} v_{t-i}^* (\eta_{pt}^* - \eta_t^*)' = o_p^*(n^{1/2}) \tag{24}$$

uniformly in $1 \leq i \leq p$. However, we have as in the proof of Lemma 3.1 in Chang and Park (2002a)

$$\begin{aligned}
\sum_{t=1}^{n} w_{t-i}^* \left(\sum_{|j|>p} \hat{\Psi}_{pj} \varepsilon_{t-j}^*\right)' &= \left(\sum_{|k|>p} |\hat{\Psi}_{k-1}||\hat{\Psi}_{pk}|\right) O_p^*(n) + \left(\sum_{i=0}^{\infty} \sum_{|j|>p} |\hat{\Psi}_i||\hat{\Psi}_{pj}|\right) O_p^*(n^{1/2}) \\
&= \left(\sum_{|k|>p} |\hat{\Psi}_{pk}|\right) O_p^*(n^{1/2}) \\
&= \left(\sum_{|k|>p} |\hat{\Pi}_k|\right) O_p^*(n^{1/2})
\end{aligned}$$

uniformly in $1 \leq i \leq p$, and (24) follows immediately.

To prove the second part, we define

$$xi_t^* = \sum_{i=1}^{t} \varepsilon_i^* \tag{25}$$

so that

$$z_t^* = \hat{\Psi}(1)\xi_t^* + (\bar{w}_0^* - \bar{w}_t^*)$$

It follows that

$$\sum_{t=1}^{n} z_t^* (\eta_{pt}^* - \eta_t^*)' = \hat{\Psi}(1) \sum_{t=1}^{n} \xi_t^* (\eta_{pt}^* - \eta_t^*)' + \bar{w}_0^* \sum_{t=1}^{n} (\eta_{pt}^* - \eta_t^*)' - \sum_{t=1}^{n} \bar{w}_t^* (\eta_{pt}^* - \eta_t^*)' \tag{26}$$

We have

$$\sum_{t=1}^{n}(\eta_{pt}^{*} - \eta_{t}^{*}) = \sum_{|k|>p} \hat{\Psi}_{pk} \sum_{t=1}^{n} \varepsilon_{t-k}^{*}$$

$$= \left(\sum_{|k|>p} |\hat{\Psi}_{pk}|\right) O_{p}^{*}(n^{1/2})$$

$$= \left(\sum_{|k|>p} |\hat{\Pi}_{k}|\right) O_{p}^{*}(n^{1/2}) \tag{27}$$

We also have similarly as in the proof of the first part

$$\sum_{t=1}^{n} \bar{w}_{t}^{*}(\eta_{pt}^{*} - \eta_{t}^{*})' = \left(\sum_{|k|>p} |\hat{\Pi}_{k}|\right) O_{p}^{*}(n) \tag{28}$$

Moreover, we have as in the proof of Lemma 3.1 in Chang and Park (2002a)

$$\sum_{t=1}^{n} \xi_{t}^{*}(\eta_{pt}^{*} - \eta_{t}^{*})' = \left(\sum_{|k|>p} |\hat{\Psi}_{pk}|\right) O_{p}^{*}(n) + \left(\sum_{|k|>p} |\hat{\Psi}_{pk}|\right) O_{p}^{*}(n^{1/2})$$

$$= \left(\sum_{|k|>p} |\hat{\Pi}_{k}|\right) O_{p}^{*}(n) \tag{29}$$

The second part can now be easily deduced from (26) and (27) - (29). The proof is therefore complete. ∎

6.2 Proofs of Lemmas and Theorems

Proof of Lemma 2.2 The stated result follows from Einmahl (1987). In particular, he shows that his Equation 1.3 holds for all δ when $2 < s < 4$, and for $\delta \geq K_{sn}^{\gamma}$ with $\gamma < 1/(2s-4)$ when $s \geq 4$, in his notation. In either case, his δ is greater than our $n^{1/a+\delta}$ with any $\delta > 0$ as long as n is sufficiently large. His result is therefore applicable as we formulate here. ∎

Proof of Lemma 2.4 We write

$$n(\tilde{\Pi}_{n} - \Pi) = \left(\frac{1}{n^{2}}\sum_{t=1}^{n} x_{t}x_{t}' - Q_{n}\right)^{-1} \left(\frac{1}{n}\sum_{t=1}^{n} x_{t}\eta_{t} - P_{n}\right)$$

where

$$P_{n} = \frac{1}{n}\sum_{t=1}^{n} x_{t}(\eta_{t} - \eta_{pt})' + \left(\frac{1}{n}\sum_{t=1}^{n} x_{t}v_{pt}'\right) \left(\frac{1}{n}\sum_{t=1}^{n} v_{pt}v_{pt}'\right)^{-1} \left(\frac{1}{n}\sum_{t=1}^{n} v_{pt}\eta_{pt}'\right)$$

$$Q_{n} = \frac{1}{n}\left(\frac{1}{n}\sum_{t=1}^{n} x_{t}v_{pt}'\right) \left(\frac{1}{n}\sum_{t=1}^{n} v_{pt}v_{pt}'\right)^{-1} \left(\frac{1}{n}\sum_{t=1}^{n} v_{pt}x_{t}'\right)$$

where in turn

$$v_{pt} = (v'_{t+p}, \ldots, v'_{t-p})'$$

To get the stated result, it suffices to show that

$$P_n, \; Q_n = o_p(1)$$

under Assumptions 2.1 and 2.3.

In the subsequent proof, we use

$$\left\| \frac{1}{n} \sum_{t=1}^{n} x_t v'_{pt} \right\| = O_p(p^{1/2}) \tag{30}$$

$$\left\| \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt} v'_{pt} \right)^{-1} \right\| = O_p(1) \tag{31}$$

which are the multivariate extensions of the results established in Chang and Park (2002a). The required extensions are straightforward and the details are omitted. It follows immediately from (30) and (31) that

$$Q_n = n^{-1} O_p(p^{1/2}) O_p(1) O_p(p^{1/2}) = O_p(n^{-1}p)$$

since

$$\|Q_n\| \le \frac{1}{n} \left\| \frac{1}{n} \sum_{t=1}^{n} x_t v'_{pt} \right\| \left\| \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt} v'_{pt} \right)^{-1} \right\| \left\| \frac{1}{n} \sum_{t=1}^{n} v_{pt} x'_t \right\|$$

as one may easily see.

We now show that

$$P_n = \left(\sum_{|k|>p} |\Pi_k| \right) O_p(1) + O_p(n^{-1/2}p)$$

We first write

$$\begin{aligned}
\|P_n\| &\le \left\| \frac{1}{n} \sum_{t=1}^{n} x_t (\eta_{pt} - \eta_t)' \right\| + \left\| \frac{1}{n} \sum_{t=1}^{n} x_t v'_{pt} \right\| \left\| \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt} v'_{pt} \right)^{-1} \right\| \left\| \frac{1}{n} \sum_{t=1}^{n} v_{pt} \eta'_{pt} \right\| \\
&= A_n + B_n
\end{aligned}$$

We may show as in the proof of Lemma 3.1 of Chang and Park (2002a) that

$$A_n = \left(\sum_{|k|>p} |\Pi_k| \right) O_p(1)$$

Moreover, we have

$$
\begin{aligned}
\left| \sum_{t=1}^{n} v_{t-i} \eta'_{pt} \right| &\leq \left| \sum_{t=1}^{n} v_{t-i} \left(u_t - \sum_{|j| \leq p} \Pi'_k v_{t-j} \right)' \right| \\
&\leq \left| \sum_{t=1}^{n} v_{t-i} u'_t \right| + \sum_{j=-\infty}^{\infty} |\Pi_j| \left| \sum_{t=1}^{n} v_{t-i} v'_{t-j} \right| \\
&= O_p(n^{1/2})
\end{aligned}
$$

uniformly in i for $|i| \leq p$. It therefore follows that

$$
B_n = O_p(p^{1/2}) O_p(1) O_p(n^{-1/2} p^{1/2}) = O_p(n^{-1/2} p)
$$

as was to be shown. ∎

Proof of Lemma 3.2   Given the result in Lemma A1, the proof is the trivial extension of the proof of Lemma 3.2 in Park (2002). The details are, therefore, omitted. ∎

Proof of Theorem 3.3   To derive the bootstrap invariance principle for (w_t^*) from that of (ε_t^*), we need to show

$$
\hat{\Phi}(1) \to_p \Phi(1) \tag{32}
$$

and

$$
\mathbf{P}^* \left\{ \max_{1 \leq t \leq n} \left| n^{-1/2} \bar{w}_t^* \right| > \epsilon \right\} = o_p(1) \tag{33}
$$

for any $\epsilon > 0$.

Let $\tilde{\Phi}(1)$ be defined exactly as $\hat{\Phi}(1)$ using the fitted coefficients $(\tilde{\Phi}_k)$ in regression (23). It follows immediately from Lemma A1 that

$$
\hat{\Phi}(1) = \tilde{\Phi}(1) + O_p(n^{-1/2} q)
$$

Moreover, we may deduce as in the proof of Lemma 3.5 in Chang and Park (2002a) that

$$
\tilde{\Phi}(1) = \Phi(1) + O_p(n^{-1/2} q) + o(q^{-b})
$$

using the result in Shibata (1981). We therefore have

$$
\hat{\Phi}(1) = \Phi(1) + o_p(1)
$$

and obtain (32). The proof of (33) is essentially identical to the proof of Theorem 3.3 in Park (2002).

Proof of Lemma 3.4   Set $z_0^* = 0$ for simplicity. The required modification to allow for nonzero z_0^* is trivial. The first part follows immediately, since

$$\frac{1}{n^2} \sum_{t=1}^n z_t^* z_t^{*\prime} =_{d^*} \int_0^1 B_n^* B_n^{*\prime} + \frac{1}{n^2} z_n^* z_n^{*\prime}$$

and

$$n^{-1/2} z_n^* = O_p^*(1)$$

for large n.

To prove the second part, we let (ξ_t^*) be defined as in (25) so that we have

$$\sum_{t=1}^n z_{t-1}^* w_t^{*\prime} = \hat{\Psi}(1) \sum_{t=1}^n \xi_{t-1}^* \varepsilon_t^{*\prime} \hat{\Psi}(1)' + \sum_{t=1}^n w_t^* \bar{w}_t^{*\prime}$$
$$- z_n^* \bar{w}_n^{*\prime} + \bar{w}_0^* \sum_{t=1}^n \varepsilon_t^{*\prime} \hat{\Psi}(1)' - \sum_{t=1}^n \bar{w}_{t-1}^* \varepsilon_t^{*\prime} \hat{\Psi}(1)'$$

It is straightforward to show

$$z_n^* \bar{w}_n^{*\prime}, \ \bar{w}_0^* \sum_{t=1}^n \varepsilon_t^{*\prime} \hat{\Psi}(1)', \ \sum_{t=1}^n \bar{w}_{t-1}^* \varepsilon_t^{*\prime} \hat{\Psi}(1)' = O_p^*(n^{1/2})$$

and we therefore have

$$\sum_{t=1}^n z_{t-1}^* w_t^{*\prime} = \hat{\Psi}(1) \sum_{t=1}^n \xi_{t-1}^* \varepsilon_t^{*\prime} \hat{\Psi}(1)' + \sum_{t=1}^n w_t^* \bar{w}_t^{*\prime} + o_p^*(1) \tag{34}$$

for large n.

It follows that

$$\hat{\Psi}(1) \sum_{t=1}^n \xi_{t-1}^* \varepsilon_t^{*\prime} \hat{\Psi}(1)' \to_{d^*} \int_0^1 B dB'$$

by the bootstrap invariance principle and Kurtz and Protter (1991). Moreover, it can be deduced analogously as in Lemma A4 that

$$\mathbf{E}^* \left| \sum_{t=1}^n (w_t^* \bar{w}_t^{*\prime} - \mathbf{E}^* w_t^* \bar{w}_t^{*\prime}) \right|^2 = O_p(n^2)$$

and we have

$$\frac{1}{n} \sum_{t=1}^n w_t^* \bar{w}_t^{*\prime} = \mathbf{E}^* w_t^* \bar{w}_t^{*\prime} + O_p^*(n^{-1/2})$$

However,

$$\mathbf{E}^* w_t^* \bar{w}_t^{*\prime} = \sum_{k=0}^\infty \Gamma^*(k) = \sum_{k=0}^\infty \Gamma(k) + o_p^*(1)$$

from which, together with (34), the stated result follows immediately. ∎

Proof of Theorem 3.5 The results can easily be derived from Lemma 3.4 using the bootstrap invariance principle and continuous mapping theorem. ∎

Proof of Lemma 3.6 Notice that

$$|Q_n^*| \leq \frac{1}{n} \left| \frac{1}{n} \sum_{t=1}^{n} x_t^* v_{pt}^{*\prime} \right| \left\| \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \right\| \left| \frac{1}{n} \sum_{t=1}^{n} v_{pt} x_t^{*\prime} \right|$$

It therefore follows that

$$Q_n^* = n^{-1} O_p^*(p^{1/2}) O_p^*(1) O_p^*(p^{1/2}) = O_p^*(n^{-1}p)$$

from Lemma A5. This shows that $Q_n^* = o_p^*(1)$.

Moreover, we have

$$|P_n^*| \quad \leq \quad \left| \frac{1}{n} \sum_{t=1}^{n} x_t^* (\eta_{pt}^* - \eta_t^*)' \right| + \left| \frac{1}{n} \sum_{t=1}^{n} x_t^* v_{pt}^{*\prime} \right| \left\| \left(\frac{1}{n} \sum_{t=1}^{n} v_{pt}^* v_{pt}^{*\prime} \right)^{-1} \right\| \left| \frac{1}{n} \sum_{t=1}^{n} v_{pt} \eta_{pt}^{*\prime} \right|$$

and it follows from Lemmas A5 and A6 that

$$P_n = o_p^*(1) + O_p^*(p^{1/2}) O_p^*(1) O_p^*(n^{-1/2} p^{1/2}) = o_p^*(1)$$

as required to be shown. ∎

Proof of Theorem 3.7 Due to Lemma 3.6, we have

$$n(\tilde{\Pi}_n^* - \Pi_n) = \left(\frac{1}{n^2} \sum_{t=1}^{n} x_t^* x_t^{*\prime} \right)^{-1} \frac{1}{n} \sum_{t=1}^{n} x_t^* \eta_t^{*\prime} + o_p^*(1)$$

The bootstrap asymptotic distribution of $\tilde{\Pi}_n^*$ can now be easily deduced from the bootstrap invariance principle and Kurtz and Protter (1991). The bootstrap asymptotic distribution of $\tilde{T}_n^*$ may similarly be obtained. ∎

References

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

Chang, Y. and J.Y. Park (2002a). "On the asymptotics of ADF tests for unit roots," forthcoming in *Econometric Reviews*.

Chang, Y. and J.Y. Park (2002b). "A sieve bootstrap for the test of a unit root," forthcoming in *Journal of Time Series Analysis*.

Einmahl, U. (1987). "A useful estimate in the multidimensional invariance principle," *Probability Theory and Related Fields* 76: 81-101.

Horowitz, J. (2002). "The bootstrap," forthcoming in *Handbook of Econometrics* Vol. 5, Elsevier, Amsterdam.

Johansen, S. (1988). "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control* 12: 231-254.

Johansen, S. (1991). "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica* 59: 1551-1580.

Kurtz, T.G. and P. Protter (1991). "Weak limit theorems for stochastic integrals and stochastic differential equations," *Annals of Probability* 19: 1035-1070.

Li, H. and J.S. Maddala (1997). "Bootstrapping cointegrating regressions," *Journal of Econometrics* 80: 297-348.

Park, J.Y. (2000). "Bootstrap unit root tests," Mimeographed, School of Economics, Seoul National University.

Park, J.Y. (2002). "An invariance principle for sieve bootstrap in time series," forthcoming in *Econometric Theory*.

Park, J.Y. and P.C.B. Phillips (1988). "Statistical inference in regressions with integrated processes: Part 1," *Econometric Theory* 4: 468-497.

Phillips, P.C.B. and V. Solo (1992). "Asymptotics for linear processes," *Annals of Statistics* 20: 971-1001.

Saikkonen, P. (1991). "Asymptotically efficient estimation of cointegration regressions," *Econometric Theory* 7: 1-21.

Shibata, R. (1980). "Asymptotically efficient selection of the order of the model for estimating parameters of a linear process," *Annals of Statistics* 8: 147-164.

Stock, J.H. and M.W. Watson (1993). "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica* 61: 783-820.

Table 1.1: Finite Sample Performances of the Estimators

n	estimator	$n$bias	n^2var	n^2MSE
25	$\hat{\pi}_n$	2.516	4.826	11.157
	$\hat{\pi}_n^c$	0.576	6.315	6.647
	$\tilde{\pi}_n$	0.025	2.981	2.982
	$\tilde{\pi}_n^c$	0.026	2.981	2.981
50	$\hat{\pi}_n$	2.705	6.213	13.528
	$\hat{\pi}_n^c$	0.365	7.305	7.439
	$\tilde{\pi}_n$	0.015	2.489	2.489
	$\tilde{\pi}_n^c$	0.016	2.490	2.490
100	$\hat{\pi}_n$	2.746	6.573	14.112
	$\hat{\pi}_n^c$	0.197	7.201	7.240
	$\tilde{\pi}_n$	0.015	2.224	2.224
	$\tilde{\pi}_n^c$	0.014	2.229	2.229
200	$\hat{\pi}_n$	2.723	6.457	13.871
	$\hat{\pi}_n^c$	0.060	6.775	6.779
	$\tilde{\pi}_n$	-0.005	2.165	2.165
	$\tilde{\pi}_n^c$	-0.004	2.169	2.169

Table 1.2: Finite Sample Performances of the Test Statistics

		sizes			powers		
n	test	1% test	5% test	10% test	1% test	5% test	10% test
25	$\hat{T}_n$	0.058	0.216	0.345	0.583	0.806	0.885
	$\hat{T}_n^*$	0.011	0.053	0.100	0.404	0.558	0.660
	$\tilde{T}_n$	0.046	0.110	0.178	0.485	0.626	0.691
	$\tilde{T}_n^*$	0.007	0.045	0.090	0.296	0.481	0.592
50	$\hat{T}_n$	0.055	0.195	0.324	0.893	0.977	0.990
	$\hat{T}_n^*$	0.008	0.048	0.097	0.727	0.873	0.934
	$\tilde{T}_n$	0.026	0.080	0.131	0.811	0.892	0.919
	$\tilde{T}_n^*$	0.012	0.051	0.097	0.744	0.856	0.902
100	$\hat{T}_n$	0.047	0.174	0.296	0.997	1.000	1.000
	$\hat{T}_n^*$	0.010	0.047	0.094	0.981	0.997	0.999
	$\tilde{T}_n$	0.018	0.066	0.117	0.982	0.994	0.997
	$\tilde{T}_n^*$	0.010	0.049	0.096	0.973	0.991	0.996
200	$\hat{T}_n$	0.039	0.161	0.270	1.000	1.000	1.000
	$\hat{T}_n^*$	0.009	0.045	0.094	1.000	1.000	1.000
	$\tilde{T}_n$	0.012	0.054	0.105	1.000	1.000	1.000
	$\tilde{T}_n^*$	0.009	0.045	0.096	1.000	1.000	1.000

Table 2.1: Finite Sample Performances of the Estimators

n	estimator	$n$bias	n^2var	n^2MSE
25	$\hat{\pi}_n^{\mu}$	3.288	11.692	22.504
	$\hat{\pi}_n^{\mu c}$	0.912	14.649	15.481
	$\tilde{\pi}_n^{\mu}$	0.184	7.171	7.205
	$\tilde{\pi}_n^{\mu c}$	0.074	7.233	7.238
50	$\hat{\pi}_n^{\mu}$	3.398	13.270	24.816
	$\hat{\pi}_n^{\mu c}$	0.463	15.670	15.885
	$\tilde{\pi}_n^{\mu}$	0.069	5.226	5.231
	$\tilde{\pi}_n^{\mu c}$	0.018	5.241	5.241
100	$\hat{\pi}_n^{\mu}$	3.494	13.384	25.590
	$\hat{\pi}_n^{\mu c}$	0.277	14.820	14.897
	$\tilde{\pi}_n^{\mu}$	0.045	4.587	4.589
	$\tilde{\pi}_n^{\mu c}$	0.019	4.596	4.596
200	$\hat{\pi}_n^{\mu}$	3.517	13.540	25.912
	$\hat{\pi}_n^{\mu c}$	0.146	14.255	14.276
	$\tilde{\pi}_n^{\mu}$	0.028	4.312	4.312
	$\tilde{\pi}_n^{\mu c}$	0.017	4.322	4.322

Table 2.2: Finite Sample Performances of the Test Statistics

n	test	sizes			powers		
		1% test	5% test	10% test	1% test	5% test	10% test
25	$\hat{T}_n^{\mu}$	0.020	0.142	0.261	0.265	0.547	0.676
	$\hat{T}_n^{\mu*}$	0.017	0.059	0.106	0.227	0.384	0.498
	$\tilde{T}_n^{\mu}$	0.044	0.117	0.180	0.257	0.415	0.505
	$\tilde{T}_n^{\mu*}$	0.008	0.043	0.090	0.093	0.248	0.369
50	$\hat{T}_n^{\mu}$	0.031	0.140	0.257	0.623	0.841	0.905
	$\hat{T}_n^{\mu*}$	0.011	0.055	0.106	0.498	0.693	0.795
	$\tilde{T}_n^{\mu}$	0.026	0.084	0.135	0.591	0.736	0.808
	$\tilde{T}_n^{\mu*}$	0.012	0.052	0.102	0.479	0.672	0.767
100	$\hat{T}_n^{\mu}$	0.029	0.135	0.231	0.967	0.994	0.997
	$\hat{T}_n^{\mu*}$	0.010	0.054	0.105	0.924	0.978	0.990
	$\tilde{T}_n^{\mu}$	0.017	0.067	0.121	0.942	0.974	0.984
	$\tilde{T}_n^{\mu*}$	0.011	0.051	0.099	0.921	0.967	0.980
200	$\hat{T}_n^{\mu}$	0.026	0.125	0.220	1.000	1.000	1.000
	$\hat{T}_n^{\mu*}$	0.011	0.052	0.100	1.000	1.000	1.000
	$\tilde{T}_n^{\mu}$	0.014	0.060	0.115	0.999	1.000	1.000
	$\tilde{T}_n^{\mu*}$	0.010	0.050	0.103	0.999	1.000	1.000

Table 3.1: Finite Sample Performances of the Estimators

n	estimator	$n$bias	n^2var	n^2MSE
25	$\hat{\pi}_n^{\tau}$	5.925	19.070	54.170
	$\hat{\pi}_n^{\tau c}$	2.460	27.327	33.379
	$\tilde{\pi}_n^{\tau}$	1.273	17.658	19.280
	$\tilde{\pi}_n^{\tau c}$	0.230	19.762	19.815
50	$\hat{\pi}_n^{\tau}$	6.408	22.298	63.360
	$\hat{\pi}_n^{\tau c}$	1.538	29.023	31.389
	$\tilde{\pi}_n^{\tau}$	0.592	11.893	12.244
	$\tilde{\pi}_n^{\tau c}$	0.037	12.256	12.257
100	$\hat{\pi}_n^{\tau}$	6.609	22.742	66.418
	$\hat{\pi}_n^{\tau c}$	0.907	26.866	27.688
	$\tilde{\pi}_n^{\tau}$	0.286	9.585	9.667
	$\tilde{\pi}_n^{\tau c}$	0.013	9.675	9.675
200	$\hat{\pi}_n^{\tau}$	6.660	24.429	68.783
	$\hat{\pi}_n^{\tau c}$	0.467	26.910	27.128
	$\tilde{\pi}_n^{\tau}$	0.168	8.592	8.620
	$\tilde{\pi}_n^{\tau c}$	0.034	8.611	8.612

Table 3.2: Finite Sample Performances of the Test Statistics

		sizes			powers		
n	test	1% test	5% test	10% test	1% test	5% test	10% test
25	$\hat{T}_n^{\tau}$	0.051	0.212	0.351	0.204	0.503	0.661
	$\hat{T}_n^{\tau *}$	0.029	0.071	0.123	0.120	0.243	0.355
	$\tilde{T}_n^{\tau}$	0.056	0.135	0.203	0.161	0.298	0.388
	$\tilde{T}_n^{\tau *}$	0.010	0.042	0.091	0.035	0.131	0.228
50	$\hat{T}_n^{\tau}$	0.060	0.246	0.387	0.522	0.790	0.879
	$\hat{T}_n^{\tau *}$	0.016	0.057	0.114	0.271	0.490	0.612
	$\tilde{T}_n^{\tau}$	0.025	0.080	0.134	0.329	0.516	0.612
	$\tilde{T}_n^{\tau *}$	0.010	0.050	0.100	0.221	0.430	0.551
100	$\hat{T}_n^{\tau}$	0.065	0.235	0.375	0.927	0.987	0.995
	$\hat{T}_n^{\tau *}$	0.010	0.056	0.108	0.753	0.905	0.951
	$\tilde{T}_n^{\tau}$	0.015	0.064	0.116	0.815	0.904	0.938
	$\tilde{T}_n^{\tau *}$	0.009	0.049	0.099	0.770	0.886	0.927
200	$\hat{T}_n^{\tau}$	0.062	0.229	0.362	1.000	1.000	1.000
	$\hat{T}_n^{\tau *}$	0.012	0.056	0.107	0.998	1.000	1.000
	$\tilde{T}_n^{\tau}$	0.014	0.057	0.109	0.997	0.999	0.999
	$\tilde{T}_n^{\tau *}$	0.011	0.050	0.101	0.996	0.999	0.999