Novel Panel Cointegration Tests Emending for Cross-Section Dependence with N Fixed[1][2]

Kaddour Hadri — Management School, Queen's University Belfast

Eiji Kurozumi — Department of Economics, Hitotsubashi University

Yao Rao — Management School, The University of Liverpool

April 11, 2014

Abstract

In this paper, we propose new cointegration tests for single equations and panels. In both cases, the asymptotic distributions of the tests, which are derived with N fixed and $T \to \infty$, are shown to be standard normals. The effects of serial correlation and cross-sectional dependence are mopped out via long-run variances. An effective bias correction is derived which is shown to work well in finite samples; particularly when N is smaller than T. Our panel tests are robust to possible cointegration across units.

JEL classification: C12, C15, C22, C23.

Keywords: cointegration, panel cointegration, cross-section dependence, bias correction, DOLS, FCLT.

[1] Correspondence: Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan; e-mail: kurozumi@stat.hit-u.ac.jp.

[2] Kurozumi's research was supported by JSPS KAKENHI Grant Numbers 23243038 and 25285067.

1. Introduction

Since the seminal work of Engle and Granger (1987), we observe a continuous and prolific stream of publications on estimating and testing long-run relationships amongst non-stationary economic variables. This literature can be divided into two branches: Single equation and system approaches (for an relevant review of system equations cointegration tests cf. Hubrich, Lütkepohl and Saikkonen, 2001).

This paper deals mainly with the former type and in this context, we start by proposing a new single equation test for the null hypothesis of cointegration based on the sample autocovariance. Our test can be seen as an extension of the stationarity test proposed by Harris, McCabe and Leybourne (2003, hereafter HML). Analogous tests for the null of cointegration in the literature include Hansen (1992), Quintos and Phillips (1993), Shin (1994), Jansson (2005) and Kurozumi and Arai (2008). Hansen (1992) derives tests for parameter instability with $I(1)$ processes. He shows that these tests can be viewed as tests of the null hypothesis of cointegration against the alternative of no cointegration. Similarly, Quintos and Phillips (1993) develop tests for parameter constancy in cointegrating regressions. Their approach delivers a test of the null hypothesis of cointegration against particular directions of departure from the null hypothesis. The test proposed by Shin (1994), which is a residual based test, for testing the null hypothesis of cointegration against the alternative of no cointegration, is based on the approach adopted by Kwiatkowski, Phillips, Schmidt, and Shin (1992). Jansson (2005) offers a feasible point optimal test of the null hypothesis of cointegration whose local asymptotic power function is showed to be close to the limiting Gaussian power envelope. Finally, a locally best invariant and unbiased (LBIU) test is proposed in Kurozumi and Arai (2008), which also discuss the relative merits and demerits of the tests of Shin (1994), Jansson (2005) and theirs.

In addition to single equation tests for cointegration, panel cointegration tests have been developed in the literature since late 90's. Early contributions are given by Kao (1999) and Pedroni (1999, 2004) among others, in which individuals are assumed to be independent. On the other hand, Banerjee, Marcellino and Osbat (2004) have shown through simulations that panel cointegration tests have severely distorted size in presence of cross units cointegration. In many empirical applications this is likely the case, because of economic links across regions and units. Another source of size distorsion is the likely presence of cross-sectional dependence. A succession of panel cointegration tests accounting for cross-sectional dependence have been proposed by Marcellino and Osbat (2004), Gengenbach, Palm and Urbain (2006), Westerlund (2008), Westerlund and Edgerton (2008), Hanck (2009), Chang and Nguyen (2012), and Bai and Carrion-i-Silvestre (2013) among others. Most of these tests assume the null hypothesis of no cointegration against the alternative of cointegration, while Chang and Nguhen (2012) also consider the case where the null of hypothesis is such that a part of individuals are not cointegrated. In addition to the theoretical development, panel unit root/cointegration tests have been implemented in empirical analysis. For example, the PPP hypothesis have been investigated by many papers, including Banerjee, Marcellino and Osbat (2005), while Westerlund (2008) considers testing the Fisher effect using panel cointegration tests. See also Banerjee and Wagner (2009) for the theoretical overview and empirical examples.

Contrary to tests for the null of no panel cointegration, there are only a few papers about tests for the null of the existence of panel cointegration. McCoskey and Kao (1999) propose LBIU tests and Westerlund (2005) consider CUSUM-based panel cointegration tests, but they assume cross-sectional independence. Westerlund and Edgerton (2007) implement the bootstrap method to McCoskey and Kao's (1999) test to correct for cross-sectional dependence. All the above panel cointegration tests are residula-based tests. Few panel cointegration tests adopted the Johansen methodology. These include Larsson, Lyhagen and Löthgren (2001), Larsson and Lyhagen (2007) and Groen and Kleibergen (2003). The last two papers extended Larsson, Lyhagen and Löthgren (2001) to allow for cross-sectional dependence. Choi (2013) offers a recent review of tests for panel cointegration. In this paper, we propose tests for the null hypothesis of panel cointegration by using the same method as Harris, Leybourne and McCabe (2005, hereafter HLM), which consider the null hypothesis of panel stationarity by pooling individual autocovariances. Like HLM (2005), our test statistic has a negative bias, which makes our test conservative. To avoid this problem, we propose an effective bias correction to improve the finite sample properties of our test.

It is widely known that panel unit root/cointegration tests are more powerful than single equation tests and Baltagi (2008) and Breitung and Pesaran (2008) provide comprehensive surveys on the subject. On the other hand, this improved power comes, in general, at a price in terms of a more involved asymptotic theory dealing with two indices simultaneously[3] and the need to emend for likely occurrence of cross-sectional dependence. Instead of using the joint asymptotics to obtain a test whose null limiting distribution is free of nuisance parameters, we use a simpler asymptotic theory where N is fixed and $T \to \infty$. This is due to the fact that the limiting distribution of the statistic of each unit is a standard normal distribution. Therefore, our panel cointegration tests are valid for any N.

The remainder of the paper has the following structure. In Section 2, we review the autocovariance based test proposed by HLM (2003) and HLM (2005). The new univariate cointegration test is analyzed in the following section. Section 4 investigates the novel panel cointegration tests. The finite sample property of our tests are investigated in Section 5. Finally, Section 6 offers some concluding remarks, and all proofs are collected in the Appendix.

2. Review of the Autocovariance Based Test

In this section, we briefly review stationarity tests based on the autocovariance proposed by Harris, McCabe and Leybourne (2003, hereafter HML) and Harris, Leybourne and McCabe (2005, hereafter HLM). Let us consider the following local level model[4]:

$$y_t = \mu + z_t \quad \text{for} \quad t = 1, 2, \cdots, T,$$

[3] Cf. Phillips and Moon (1999) for a theoretical exposition and Hadri, Larsson and Rao (2012) for a discussion of the different limit theories including the limit theory where T is fixed and N is allowed to go to infinity.

[4] HML (2003) and HLM (2005) allowed for deterministic regressors in addition to a constant but we restrict our attention to the local level model in order to simplify the explanation.

and suppose that we want to test for the null hypothesis that z_t is stationary whereas it is a unit root process under the alternative. HML (2003) note the differences in the convergence order of the sample autocovariance under the null and the alternative hypotheses,

$$\frac{1}{T-K}\sum_{t=K+1}^{T}\hat{z}_t\hat{z}_{t-K} \quad \overset{p}{\longrightarrow} \quad E[(y_t-\mu)(y_{t-K}-\mu_K)] \equiv C_K \quad \text{under the null hypothesis}$$

$$\frac{1}{(T-K)^2}\sum_{t=K+1}^{T}\hat{z}_t\hat{z}_{t-K} \quad \overset{d}{\longrightarrow} \quad \int_0^1 \tilde{B}^2(r)dr \quad \text{under the alternative}$$

for a given lag order K, where $\hat{z}_t = y_t - \bar{y}$ and $\tilde{B}(r)$ is a demeaned Brownian motion. Although it seems inconvenient to use the sample autocovariance as a test statistic because it converges to a fixed value C_K, HML (2003) notice that $C_K \to 0$ as $K \to \infty$ and thus the central limit theorem (CLT) for the sample autocovariance with a suitable normalization is expected to hold as K goes to infinity. In fact, they showed that

$$\frac{\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}\hat{z}_t\hat{z}_{t-K}}{\hat{\omega}_{zz}} \overset{d}{\longrightarrow} N(0,1) \quad \text{under the null hypothesis,} \tag{1}$$

where $\hat{\omega}_{zz}^2$ is the kernel estimator of the long-run variance based on $\hat{z}_t\hat{z}_{t-K}$, whereas the left-hand side diverges to infinity under the alternative. They also proposed a test for heteroskedastic cointegration using a similar principle.

The above stationarity test based on the autocovariance was extended to a panel stationarity test by HLM (2005). For a panel data model given by

$$y_{it} = \mu_i + z_{it} \quad \text{for} \quad i = 1, 2, \cdots, N \quad \text{and} \quad t = 1, 2, \cdots, T,$$

we have the regression residuals normalized by the standard deviation; that is,

$$\tilde{z}_{i,t} = \frac{\hat{z}_{i,t}}{\hat{\sigma}_{i,z}}, \quad \text{where } \hat{\sigma}_{i,z} \text{ is the sample standard deviation of } \hat{z}_{i,t}.$$

Then, the test statistic for panel stationarity is constructed by pooling the sample autocovariances across cross-sections, which is given by

$$\hat{S}_K = \frac{\tilde{C}_K}{\hat{\omega}_a}, \quad \text{where} \quad \tilde{C}_K = \frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}\tilde{a}_{K,t} \quad \text{with} \quad \tilde{a}_{K,t} = \sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-K}$$

and $\hat{\omega}_a^2$ is the long-run variance estimator based on $\tilde{a}_{K,t}$. HLM (2005) showed that $\hat{S}_K \overset{d}{\longrightarrow} N(0,1)$ under the null hypothesis whereas it diverges to infinity under the alternative.

Although the size of the above test can be controlled at least asymptotically, HLM (2005) showed that $\hat{S}_K$ suffers from under-size distortion in finite samples because of the negative bias of the test statistic. Since $\hat{z}_{i,t} = z_{i,t} - \bar{z}_i$, we can see that

$$\frac{1}{\sqrt{T-K}}\sum_{t=K+1}^{T}\tilde{z}_{i,t}\tilde{z}_{i,t-K} = \frac{1}{\hat{\sigma}_{i,z}^2\sqrt{T-K}}\sum_{t=K+1}^{T}z_{i,t}z_{i,t-K} - \frac{1}{\hat{\sigma}_{i,z}^2\sqrt{T-K}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}z_{i,t}\right)^2 + o_p\left(\frac{1}{\sqrt{T}}\right),$$

and thus the negative bias comes from the second term on the right-hand side of the above equation. Note that this negative bias accumulates when we pool the sample autocovariances, so that the panel stationarity test tends to be severely undersized as N gets larger. Because the expectation of $(T^{-1/2}\sum_{t=1}^{T} z_{i,t})^2$ is approximated by the long-run variance of its limiting distribution, HLM (2005) proposed the following bias corrected version of the test statistic:

$$\tilde{S}_K = \frac{\tilde{C}_K + \tilde{b}}{\hat{\omega}_a} \quad \text{where} \quad \tilde{b} = \frac{1}{\sqrt{T-K}}\sum_{i=1}^{N}\frac{\hat{\omega}_{i,z}^2}{\hat{\sigma}_{i,z}^2}$$

with $\hat{\omega}_{i,z}^2$ being the long-run variance estimator based on $\hat{z}_{i,t}$. Because the bias term is negligible when T is large, we still have $\tilde{S}_K \xrightarrow{d} N(0,1)$ under the null hypothesis.

3. Univariate Cointegration Test

3.1. Model and assumptions

We start with a univariate cointegrating regression model given by

$$y_t = \beta' X_t + u_t \quad \text{for} \quad t = 1, 2, \cdots, T, \tag{2}$$

where $X_t = [1, x_t']'$ (constant case) or $X_t = [1, t, x_t']'$ (trend case), y_t and x_t are 1- and p_x-dimensional processes with

$$x_t = x_{t-1} + v_t \quad \text{and} \quad u_t = \rho u_{t-1} + u_t^*.$$

We make the following assumption for u_t^* and v_t:

Assumption 1 *(a) $[u_t^*, v_t']'$ is a vector linear process given by*

$$\begin{bmatrix} u_t^* \\ v_t \end{bmatrix} = \sum_{j=0}^{\infty} \Phi_j \varepsilon_{t-j} \quad \textit{with} \quad \sum_{j=0}^{\infty} j^2 \|\Phi_j\| < \infty,$$

where $\{\varepsilon_t\}$ is an (p_x+1)-dimensional i.i.d. sequence with mean 0 and variance given by Σ_ε, which is positive definite, and has the finite fourth order moments.
(b) The spectral density of $[u_t^*, v_t']'$, denoted by $f(\lambda) \equiv (2\pi)^{-1}\Phi(e^{-i\lambda})\Sigma_\varepsilon \Phi'(e^{i\lambda})$, is nonsingular and $f(\lambda) \geq \alpha I_{p_x+1}$ for some $\alpha > 0$ for all $\lambda \in [0, \pi]$.

Assumption 1 implies that $[u_t^*, v_t']'$ is stationary and that there is no cointegrating relation among x_t. The 2-summability of $\{\Phi_j\}$ is stronger than usual but we need this condition to derive the bias later. The assumption on the spectral density will be used to derive the leads and lags expression. We also note that, since $\{\varepsilon_t\}$ is an *i.i.d.* sequence with the finite fourth order moments, exercise 2.13 of Brillinger (1981) implies that $[u_t^*, v_t']'$ satisfies Assumption 2.6.2 of Brillinger (1981). That is, the fourth order cumulants of $[u_t^*, v_t']'$, which are denoted by $\kappa_{ijkl}(m_1, m_2, m_3)$, satisfy

$$\sum\sum\sum_{m_1,m_2,m_3=-\infty}^{\infty} |\kappa_{ijkl}(m_1, m_2, m_3)| < \infty.$$

The testing problem we consider is given by

$$H_0 : \ |\rho| < 1 \quad \text{vs.} \quad H_1 : \ \rho = 1.$$

That is, y_t is cointegrated with x_t under the null hypothesis whereas they are not cointegrated under the alternative. Note that under the null hypothesis, $[u_t, v_t']'$ also satisfies the same conditions as given by Assumption 1.

Since it is known that $\tilde{D}_T(\hat{\beta}_{ols} - \beta)$ converges in distribution where $\hat{\beta}_{ols}$ is obtained by regressing y_t on X_t and $\tilde{D}_T = diag\{\sqrt{T}, T\ I_{p_x}\}$ (constant case) or $\tilde{D}_T = diag\{\sqrt{T}, T\sqrt{T}, T\ I_{p_x}\}$ (trend case), we can see that the same weak convergence holds as given by (1) with $\hat{z}_t$ replaced by $\hat{u}_t$. However, such a test suffers from under-size distortion as discussed in the previous section; therefore, we need to construct the bias corrected version of the test. In the case of cointegration model (2), we have

$$\begin{aligned} \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \hat{u}_t \hat{u}_{t-K} &= \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} u_t u_{t-K} \\ &\quad - \frac{1}{\sqrt{T-K}} \sum_{t=1}^{T} u_t X_t' \left(\sum_{t=1}^{T} X_t X_t' \right)^{-1} \sum_{t=1}^{T} X_t u_t + o_p\left(\frac{1}{\sqrt{T}}\right), \end{aligned} \tag{3}$$

so that the second term in (3) corresponds to the bias term. Under Assumption 1 and the null hypothesis, we can show that

$$\tilde{D}_T^{-1} \sum_{t=1}^{T} X_t X_t' \tilde{D}_T^{-1} \xrightarrow{d} \int_0^1 \tilde{B}(r)\tilde{B}'(r)dr \quad \text{and} \quad \tilde{D}_T^{-1} \sum_{t=1}^{T} X_t u_t \xrightarrow{d} \int_0^1 \tilde{B}(r)dB_u(r) + \lambda_{xu},$$

where $\tilde{B}(r) = [1, B'(r)]'$ (constant case) or $\tilde{B}(r) = [1, r, B'(r)]'$ (trend case) with $B(r)$ being a p_x-dimensional Brownian motion, $B_u(r)$ is a 1-dimensional Brownian motion, and λ_{xu} is the so called one-sided long run variance. As a result, the expectation of the bias term approximately becomes

$$E\left[\left(\int_0^1 \tilde{B}(r)dB_u(r) + \lambda_{xu} \right)' \left(\int_0^1 \tilde{B}(r)^{-1}\tilde{B}'(r)dr \right)^{-1} \left(\int_0^1 \tilde{B}(r)dB_u(r) + \lambda_{xu} \right) \right]. \tag{4}$$

In this case, the problem is that $B(r)$ is correlated with $B_u(r)$ in general, so that it is too difficult to evaluate the above expectation in general. Exception is the case when u_t is independent of v_t, so that $B(r)$ is independent of $B_u(r)$ and $\lambda_{xu} = 0$. In such a special case, (4) reduces to $\omega_u^2(p_c + p_x)$ where $p_c = 1$ or 2 depending on constant or trend case while ω_u^2 is the long-run variance of u_t, which can be estimated using $\hat{u}_t$. In other words, if $v_t = \Delta x_t$ is uncorrelated with the regression error u_t for all the leads and lags, then we can evaluate expectation (4).

In order to establish such a reasonable relation, we exploit the dynamic ordinary least squares (DOLS) technique[5] considered by Phillips and Loretan (1991), Saikkonen (1991) and

[5] We also considered the fully modified (FM) regression proposed by Phillips and Hansen (1990). However, it can be shown that the tedious bias still remains even if the FM method is applied and thus we do not pursue the FM technique in this paper.

Stock and Watson (1993). Under Assumption 1 and the null hypothesis, we have the following leads and lags expression by Theorem 8.3.1 of Brillinger (1981):

$$u_t = \sum_{j=-\infty}^{\infty} \pi_j' v_{t-j} + \eta_t, \tag{5}$$

where $E[v_s \eta_t] = 0$ for all s and t, and the transfer function associated with $\{\pi_j\}$ is given by $f_{uv}(\lambda) f_{vv}^{-1}(\lambda)$ with $f_{uv}(\lambda)$ and $f_{vv}(\lambda)$ being the corresponding blocks of $f(\lambda)$. Then, the assumption of the 2-summability of $\{\Phi_j\}$ implies that $\{\pi_j\}$ is also 2-summable. In addition, because $[u_t, v_t']'$ is a linear process with *i.i.d.* innovations, η_t can be expressed as

$$\eta_t = \sum_{j=-\infty}^{\infty} \phi_j \xi_{t-j} \quad \text{with} \quad \sum_{j=-\infty}^{\infty} |j|^2 |\phi_j| < \infty, \tag{6}$$

where $\{\xi_t\}$ is an independent sequence with mean 0, variance σ_ξ^2 and the finite fourth order moments. By replacing u_t in (2) with (5), we have

$$y_t = \beta' X_t + \sum_{j=-\infty}^{\infty} \pi_j' v_{t-j} + \eta_t.$$

By truncating infinite leads and lags at $j = \pm M$, we obtain the DOLS regression as follows:

$$y_t = \beta' X_t + \sum_{j=-M}^{M} \pi_j' v_{t-j} + \eta_t^*, \quad \text{for } t = M+1, \cdots, T-M, \tag{7}$$

where $\eta_t^* = \eta_t + \sum_{j>|M|} \pi_j' v_{t-j}$. Note that the truncation points can be different at the leads and the lags; in fact, the finite sample performance with the different truncation points could be better in some cases as investigated by Hayakawa and Kurozumi (2008) and Choi and Kurozumi (2012). In this paper, the same truncation points are used only for notational convenience.

In the following, we consider constructing a test statistic based on regression (7) and thus for notational convenience, we re-define $T = T - 2M$ and denote the effective sample period $t = M+1, \cdots, T-M$ as $t = 1, \cdots, T$.

As discussed in Saikkonen (1991), the truncation point M must diverge to infinity at a suitable rate and we make the following assumption on the divergence rate of M:

Assumption 2 *As $T \to \infty$,*

$$\frac{M^4}{T} \to 0, \tag{8}$$

$$\sqrt{T} \sum_{|j|>M} \|\pi_j\| \to 0. \tag{9}$$

Conditions (8) and (9) gives the upper and lower bounds for the divergence rate of M, respectively. Note that Saikkonen (1991) assumed $M^3/T \to 0$, which is weaker than (8) and sufficient to guarantee the asymptotic normality of π_j for a given j. The stronger assumption 2 is required in order to evaluate the bias term in our cointegration test. Note that, as shown by Kejriwal and Perron (2008), we can relax Assumption 2 as far as the efficient estimation of β is concerned.

3.2. Cointegration test with DOLS regressions

We construct the test statistic following HML (2003). Let $\hat{\eta}_t^*$ be the regression residuals from DOLS regression (7) and the standardized version[6] is given by

$$\tilde{\eta}_t^* = \frac{\hat{\eta}_t^*}{\hat{\sigma}_\eta}, \quad \text{where} \quad \hat{\sigma}_\eta^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{\eta}_t^{*2}.$$

Then, the test statistic for the null of cointegration is given by

$$\hat{S}_K = \frac{\tilde{C}_K}{\hat{\omega}_a} \quad \text{where} \quad \tilde{C}_K = \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \tilde{a}_{K,t} \quad \text{with} \quad \tilde{a}_{K,t} = \tilde{\eta}_t^* \tilde{\eta}_{t-K}^*,$$

and $\hat{\omega}_a^2$ is the long-run variance estimator of $\tilde{a}_{K,t}$ with the Bartlett kernel given by

$$\hat{\omega}_a^2 = \hat{\gamma}_{a,0} + 2\sum_{j=1}^{J}\left(1 - \frac{j}{J+1}\right)\hat{\gamma}_{a,j} \quad \text{where} \quad \hat{\gamma}_{a,j} = \frac{1}{T-K}\sum_{t=K+j+1}^{T} \tilde{a}_{K,t}\tilde{a}_{K,t-j} \tag{10}$$

and J is the bandwidth of order $o(T^{1/2})$.

We would like to show that the functional central limit theorem (FCLT) holds for $\tilde{C}_K$, but we cannot directly apply theorems in HML (2003) because they assume a causal linear process for the stochastic term z_t whereas η_t in our model is not a causal but a linear process with leads and lags of the innovations $\{\xi_t\}$. Then, we first have to establish the Beveridge–Nelson (B–N) decomposition for $\eta_t \eta_{t-K}$. In the following, the coefficients and the lag polynomials depend on K but we suppress it for notational convenience.

Lemma 1 *For $\{\eta_t\}$ satisfying (6), we have*

$$\eta_t \eta_{t-K} = \sum_{j=1}^{\infty} G_j \xi_t \xi_{t-j} - \Delta \tilde{r}_t - \Delta^+ \tilde{r}_t^+ + r_{1,t} + r_{2,t} + r_{3,t}, \tag{11}$$

[6] Exactly speaking, it is not necessary for the residuals to be standardized as far as the univariate case is concerned; the standardization is required only for the panel cointegration test in order for the test statistic to be scale invariant. We standardize them in the univariate case just because the univariate cointegration test can be seen as a special case of the panel cointegration test.

where $\Delta = 1 - L$ *and* $\Delta^+ = 1 - L^{-1}$ *with* L *being the lag operator,* $G_j = G_{1,j} + G_{2,j}$ *with*

$$G_{1,j} = \sum_{\ell=1-(j\wedge K)}^{K-1} \phi_\ell \phi_{j+\ell-K} \text{ and } G_{2,j} = \begin{cases} \sum_{\ell=1}^{K-j-1} \phi_{\ell-K}\phi_{j+\ell}, & (j = 1, \cdots, K-2), \\ 0, & (j > K+2), \end{cases}$$

$\tilde{r}_t = \tilde{r}_{1,t} + \tilde{r}_{2,t}$ *with*

$$\begin{aligned}
\tilde{r}_{1,t} &= \sum_{j=1}^{\infty} \tilde{G}_{1,j}(L)\xi_t\xi_{t-j} \quad \text{where } \tilde{G}_{1,j}(L) = \sum_{\ell=0}^{K-2} \tilde{G}_{1,\ell}L^\ell \text{ with } \tilde{G}_{1,\ell} = \sum_{i=\ell+1}^{K-1} \phi_i\phi_{i+j-K}, \\
\tilde{r}_{2,t} &= \sum_{j=1}^{K-2} \tilde{G}_{2,j}(L)\xi_t\xi_{t-j} \quad \text{where } \tilde{G}_{2,j}(L) = \sum_{\ell=0}^{K-j-2} \tilde{G}_{2,\ell}L^\ell \text{ with } \tilde{G}_{2,\ell} = \sum_{i=\ell+1}^{K-j-1} \phi_{i+j}\phi_{i-K}, \\
\tilde{r}_t^+ &= \sum_{j=2}^{\infty} \tilde{G}_j^+(L)\xi_t\xi_{t-j} \quad \text{where } \tilde{G}_j^+(L) = \sum_{\ell=2-(j\wedge K)}^{0} \tilde{G}_\ell^+ L^\ell \text{ with } \tilde{G}_\ell^+ = \sum_{i=1-(j\wedge K)}^{\ell-1} \phi_i\phi_{i+j-K},
\end{aligned}$$

$$r_{1t} = \sum_{j=1}^{K-1} \phi_j\phi_{j-K}\xi_{t-j}^2, \quad r_{2t} = \sum_{|j|\geq K}^{\infty} \sum_{\ell=-\infty}^{\infty} \phi_j\phi_\ell\xi_{t-j}\xi_{t-K-\ell}, \quad r_{3t} = \sum_{j=-K+1}^{K-1} \sum_{\ell=-\infty}^{-K} \phi_j\phi_\ell\xi_{t-j}\xi_{t-K-\ell}.$$

Lemma 1 implies that $\eta_t\eta_{t-K}$ can be decomposed into the first term on the right-hand side of (11) plus the remaining terms, the former of which is a martingale difference array. In order to establish the FCLT for the partial sum process of $\eta_t\eta_{t-K}$, we make the following assumption on the divergence rate of K.

Assumption 3 *The lag order* K *diverges to infinity at a rate of* T^δ *for* $1/4 \leq \delta < 1$.

The divergence rate of K is related with the establishment of Lemma 5(ii) in the appendix, the proof of which implies that if, in general, $\{\phi_i\}$ is j-summable, then K could be T^δ for $1/(2j) \leq \delta < 1$. Since $\{\phi_i\}$ is 2-summable in our case, we make Assumption 3. Note that Assumptions 2 and 3 imply that $M/K \to 0$, which is required in the proofs of the lemmas and theorems.

From expression (11), the FCLT for a sequence of martingale difference arrays can be applied to the first term on the right-hand side of (11) by the following Lemma 2 while the differencing operators $\Delta = 1 - L$ and $\Delta^+ = 1 - L^{-1}$ avoid from accumulating the effect of $\tilde{r}_t$ and $\tilde{r}_t^+$. Intuitively, the partial sums of the remaining terms $r_{1,t}$, $r_{2,t}$ and $r_{3,t}$ become negligible because they include ϕ_j for $j \geq K$, which converges to zero sufficiently rapidly.

Lemma 2 *Suppose that Assumptions 1 and 3 hold. Under the null hypothesis, the following FCLT holds as* $T \to \infty$:

$$\frac{1}{\sqrt{T-K}} \sum_{t=1}^{[Tr]} \eta_t\eta_{t-K} \Rightarrow B(r), \tag{12}$$

where $[a]$ is the largest integer less than a, $0 \leq r \leq 1$, $\Rightarrow$ signifies weak convergence of the associated probability measures, and $B(r)$ is a Brownian motion with the variance $\omega_a^2 \equiv \sigma_\xi^4 \lim_{K\to\infty} \sum_{j=1}^{\infty} G_j^2$.

Note that (12) holds only when $K \to \infty$ at a suitable rate; otherwise, the left-hand side apparently goes to infinity.

We are now in a position to apply Lemma 2 to the residuals in DOLS regression (7). Since

$$\hat{\eta}_t^* = \eta_t - (\hat{\beta} - \beta)' X_t - (\hat{\Pi} - \Pi)' V_t + e_t,$$

where $\hat{\beta}$ and $\hat{\Pi}$ are the estimators of β and Π in (7) with $\Pi = [\pi_M, \pi_{M-1}, \cdots, \pi_{-M}]$, $V_t = [v'_{t-M}, v'_{t-M+1}, \cdots, v'_{t+M}]'$, and $e_t = \sum_{|j|>M} \pi'_j v_{t-j}$, we have

$$\frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \hat{\eta}_t^* \hat{\eta}_{t-K}^* = \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \eta_t \eta_{t-K} + \frac{1}{\sqrt{T-K}} \left(R_{\beta,T} + R_{\Pi,T} + R_T \right), \quad (13)$$

where

$$R_{\beta,T} = (\hat{\beta}-\beta)' \sum_{t=K+1}^{T} X_t X_{t-K} (\hat{\beta}-\beta) - (\hat{\beta}-\beta)' \sum_{t=K+1}^{T} X_{t-K} \eta_t - (\hat{\beta}-\beta)' \sum_{t=K+1}^{T} X_t \eta_{t-K}, \quad (14)$$

$$R_{\Pi,T} = (\hat{\Pi}-\Pi)' \sum_{t=K+1}^{T} V_t V_{t-K} (\hat{\Pi}-\Pi) - (\hat{\Pi}-\Pi)' \sum_{t=K+1}^{T} V_{t-K} \eta_t - (\hat{\Pi}-\Pi)' \sum_{t=K+1}^{T} V_t \eta_{t-K}, \quad (15)$$

$$\begin{aligned} R_T \;=\; & \sum_{t=K+1}^{T} e_t e_{t-K} + \sum_{t=K+1}^{T} (\eta_t e_{t-K} + \eta_{t-K} e_t) + (\hat{\beta} - \beta)' \sum_{t=K+1}^{T} (X_t V_{t-K} + X_{t-K} V_t)(\hat{\Pi} - \Pi) \\ & -(\hat{\beta} - \beta)' \sum_{t=K+1}^{T} (X_t e_{t-K} + X_{t-K} e_t) - (\hat{\Pi} - \Pi)' \sum_{t=K+1}^{T} (V_t e_{t-K} + V_{t-K} e_t). \quad (16) \end{aligned}$$

The following theorem is obtained by applying Lemma 2 to the first term on the right-hand side of (13) whereas the remaining terms are shown to be negligible by directly applying the results of Saikkonen (1991), so that $\hat{C}_K \xrightarrow{d} N(0, \omega_a^2)$ under the null hypothesis. The consistency of $\hat{\omega}_a^2$ is also proved similarly to HML (2003). On the other hand, the test statistic diverges to infinity as proved by HML (2003) and then we omit the details.

Theorem 1 *Suppose that Assumptions 1, 2 and 3 hold. Under the null hypothesis, as $T \to \infty$,*

$$\hat{S}_K \to N(0, 1),$$

whereas under the alternative, it diverges to infinity.

From Theorem 1, we can test for the null hypothesis of cointegration using the same test statistic as HML (2003) using the DOLS regression residuals, even though they are not causal but expressed as the leads and lags of the innovations.

3.3. Bias correction of the cointegration tests

As explained in the previous section, the cointegration test based on the autocovariance suffers from under-size distortion and we need to construct the bias corrected version of the test statistic as suggest by HLM (2005). Because the first term on the right-hand side of (13) is the leading term, we define the bias of (13) as the expectation of the remaining terms up to $O_p(T^{-1/2})$. It is shown in the proof of Lemma 3 that the bias appears only from $R_{\beta,T}$ in (13) while $R_{\Pi,T}$ and R_T can be negligible.

Lemma 3 *The bias of* (13), $-b$, *is given by*

$$-b = -\frac{(p_c + p_x)\omega_\eta^2}{\sqrt{T-K}\sigma_\eta^2}, \quad \text{where } p_c = 1 \text{ (constant case) or } p_c = 2 \text{ (trend case).}$$

From the result of Lemma 3, the bias corrected version of the test statistic is defined by

$$\tilde{S}_K = \frac{\tilde{C}_K + \tilde{b}}{\hat{\omega}_a} \quad \text{where} \quad \tilde{b} = \frac{(p_c + p_x)\hat{\omega}_\eta^2}{\sqrt{T-K}\hat{\sigma}_\eta^2}$$

with $\hat{\omega}_\eta^2$ is the long-run variance estimator based on $\hat{\eta}_t^*$ with the Bartlett kernel defined as (10) with $\tilde{a}_{K,t}$ replaced by $\hat{\eta}_t^*$. Then, we have the following corollary:

Corollary 1 *Suppose that Assumptions 1, 2 and 3 hold. Under the null hypothesis, as* $T \to \infty$,

$$\hat{S}_K \to N(0,1),$$

whereas under the alternative, it diverges to infinity.

4. Panel Cointegration Test

In the case of panel cointegration, model (2) becomes

$$y_{i,t} = \beta_i' X_{i,t} + u_{i,t} \quad \text{for} \quad i = 1, 2, \cdots, N \quad \text{and} \quad t = 1, 2, \cdots, T, \tag{17}$$

where $X_{i,t} = [1, x_{i,t}']'$ (constant case) or $X_{it} = [1, t, x_{i,t}']'$ (trend case), $y_{i,t}$ and x_{it} are 1- and $p_{i,x}$- dimensional processes with

$$x_{i,t} = x_{i,t-1} + v_{i,t} \quad \text{and} \quad u_{i,t} = \rho_i u_{i,t-1} + u_{i,t}^*.$$

Note that the specification of the non-stochastic term and the dimension of the $I(1)$ regressors can be different for individuals.

Let $u_{a,t}^* = [u_{1,t}^*, u_{2,t}^*, \cdots, u_{N,t}^*]'$ and $v_{a,t} = [v_{1,t}', v_{2,t}', \cdots, v_{N,t}']'$ are N- and $p_{a,x} \equiv (p_{1,x} + p_{2,x} + \cdots + p_{N,x})$-dimensional vectors, respectively. In the case of panel cointegration, we make the following assumption:

Assumption 4 *(a)* $[u^{*\prime}_{a,t}, v'_{a,t}]'$ *is a vector linear process given by*

$$\begin{bmatrix} u^*_{a,t} \\ v_{a,t} \end{bmatrix} = \sum_{j=0}^{\infty} \Phi_{a,j}\varepsilon_{a,t-j} \quad \textit{with} \quad \sum_{j=0}^{\infty} j^2 \|\Phi_{a,j}\| < \infty,$$

where $\{\Phi_{a,j}\}$ *is a set of* $(N + p_{a,x}) \times p_a$ *coefficients* (p_a *is not necessarily equal to* $N + p_{a,x}$) *and* $\{\varepsilon_{a,t}\}$ *is a* p_a*-dimensional i.i.d. sequence with mean 0 and variance given by* $\Sigma_{a,\varepsilon}$*, which is positive definite, and has the finite fourth order moments.*
(b) The marginal distribution of $[u^*_{i,t}, v'_{i,t}]$ *satisfies Assumption 1 for* $i = 1, 2, \cdots, N$.

As in the univariate case, we do not allow for cointegration among regressors in each individual regression (17) by Assumption 4(b). On the other hand, it is possible for some $x_{i,t}$ to be cointegrated with $x_{j,t}$ with $i \neq j$. In this case, because $x_{i,t}$ and $x_{j,t}$ are driven by common stochastic trends, the dimension of $\varepsilon_{a,t}$ could be less than $N + p_{a,x}$ and hence $\Phi_{a,j}$ are not necessarily square matrices but the column dimension becomes smaller than the row dimension. We also note that the cross-sectional dependence in $u_{i,t}$ is allowed through the off-diagonal elements of $\Phi_{a,j}$ and $\Sigma_{a,\varepsilon}$ and that common factors can be included in u_{it} as far as they can be expressed as linear processes. Assumption 4(b) implies that $[u^*_{i,t}, v'_{i,t}]'$ can be expressed as in Assumption 1 using a $(p_{i,c} + p_{i,x})$-dimensional *i.i.d.* sequence $\{\varepsilon_{i,t}\}$ and that $\varepsilon_{i,s}$ are independent of $\varepsilon_{j,t}$ for $s \neq t$. The latter property will be used to establish the joint convergence of the individual test statistics.

For example, when x_{it} is one-dimensional and common for all i and the errors are linear processes given by

$$\Delta x_{i,t} = v_{i,t} = \sum_{j=0}^{\infty} \phi^v_j \varepsilon^v_{t-j} \quad \text{for all } i \quad \text{and} \quad u^*_{i,t} = \sum_{j=0}^{\infty} \phi^u_{i,j}\varepsilon^u_{i,t-j},$$

where $\{\varepsilon^v_t\}$ is independent of $\{\varepsilon^u_{i,t}\}$, we can see that

$$\begin{bmatrix} u^*_{1,t} \\ \vdots \\ u^*_{N,t} \\ v_{1,t} \\ \vdots \\ v_{N,t} \end{bmatrix} = \sum_{j=0}^{\infty} \begin{bmatrix} \phi^u_{1,j} & & 0 \\ & \ddots & \\ & & \phi^u_{N,j} & \\ & & & \phi^v_j \\ & & & \vdots \\ 0 & & & \phi^v_j \end{bmatrix} \begin{bmatrix} \varepsilon^u_{1,t-j} \\ \vdots \\ \varepsilon^u_{N,t-j} \\ \varepsilon^v_{t-j} \end{bmatrix} \quad \text{with} \quad \Sigma_{a,\varepsilon} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} & \\ \vdots & \ddots & \vdots & 0 \\ \sigma_{N1} & \cdots & \sigma_{NN} & \\ & 0 & & \sigma_{vv} \end{bmatrix}.$$

In this case, $\varepsilon_{a,t}$ are $(N + 1)$-dimensional vectors and $\Phi_{a,j}$ are $2N \times (N + 1)$ matrices.

The null hypothesis in the panel case is that all the individuals are cointegrated whereas at least one individual is not cointegrated under the alternative. That is,

$$H_0: \ \rho_i < 1 \text{ for all } i \quad \text{vs.} \quad H_1: \ \rho_i = 1 \text{ for } i = 1, \cdots, N_1 \text{ with } 1 \leq N_1 \leq N.$$

Note that because the cross-sectional dimension N is fixed in our model, we can reject the null hypothesis even if only one individual is not cointegrated. However, it is not difficult to expect that the test against small N_1 is less powerful than that against large N_1.

As in the univariate case, individual regression (17) is augmented by the leads and lags of the first differences of the $I(1)$ regressors and we obtain the DOLS regression given by

$$y_{i,t} = \beta_i' X_{i,t} + \sum_{j=-M}^{M} \pi_{i,j}' v_{i,t-j} + \eta_{i,t}^* \tag{18}$$

where $\eta_{i,t}^*$ is defined as before and the standardized regression residuals are defined as

$$\tilde{\eta}_{i,t}^* = \frac{\hat{\eta}_{i,t}^*}{\hat{\sigma}_{i,\eta}} \quad \text{where} \quad \hat{\sigma}_{i,\eta}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{\eta}_{i,t}^{*2}.$$

Note that the truncation point M can be different over cross-sections but we proceed with the same M for notational convenience.

In this case, the test statistic for panel cointegration is given by

$$\hat{S}_K = \frac{\tilde{C}_K}{\hat{\omega}_a} \quad \text{where} \quad \tilde{C}_K = \frac{1}{\sqrt{T-K}} \sum_{t=K+1}^{T} \tilde{a}_{K,t} \quad \text{with} \quad \tilde{a}_{K,t} = \sum_{i=1}^{N} \tilde{\eta}_{i,t}^* \tilde{\eta}_{i,t-K}^*,$$

while the bias corrected version of the test statistic is defined as

$$\tilde{S}_K = \frac{\tilde{C}_K + \tilde{b}}{\hat{\omega}_a} \quad \text{where} \quad \tilde{b} = \frac{1}{\sqrt{T-K}} \sum_{i=1}^{N} \frac{(p_{i,c} + p_{i,x})\hat{\omega}_{i,\eta}^2}{\hat{\sigma}_{i,\eta}^2}.$$

Theorem 2 *Suppose that Assumptions 4, 2 and 3 hold. Under the null hypothesis, as $T \to \infty$,*

$$\hat{S}_K,\ \tilde{S}_K \to N(0,1),$$

whereas under the alternative, they diverge to infinity.

As discussed in the introduction, the advantage of using HLM (2005) test is that we do not have to rely on the joint limit theorem in order to obtain a test statistic whose null limiting distribution is free of nuisance parameter. This is because the test statistic in the univariate case has the limiting normal distribution. As a result, we can apply our test even for panel data with small N.

5. Monte Carlo Simulations

5.1. Single cointegration tests

In this section, we first investigate the finite sample performance of the single cointegration tests proposed in this paper. The data generating process is given by

$$y_t = \alpha' c_t + \beta x_t + u_t, \qquad x_t = x_{t-1} + v_t,$$

$$u_t = \phi u_{t-1} + \varepsilon_t^u, \qquad v_t = \psi v_{t-1} + \varepsilon_t^v,$$

where $c_t = 1$ in the constant case while $c_t = [1, t]'$ in the trend case, $\alpha = 0$, $\beta = 1$, and $[\varepsilon_t^u, \varepsilon_t^v]' \sim i.i.d.N(0, \Sigma)$ with $vech(\Sigma) = [1, 0.5, 1]'$. The initial values of u_t and v_t are $u_0 = v_0 = 0$. To control serial correlation in v_t, we set $\psi = 0$, 0.4, 0.8. Under the null hypothesis of cointegration, ϕ must be less than 1 in absolute value and then we consider three cases; $\phi = 0$, 0.4, 0.8, while the alternative of no cointegration corresponds to the case with $\phi = 1$.

Throughout the simulations, the bandwidth J for the long-run variance estimation and the leads-lags truncation parameter M are set to

$$J = \left[12(T/100)^{1/4}\right] \quad \text{and} \quad M = \left[2(T/100)^{1/5}\right].$$

We also investigate the effect of the lag order K on our tests based on the autocovariance because the finite sample performance will crucially depend on K. We consider

$$K = \left[(aT)^{\delta}\right] \quad \text{for } a = 1,\ 2 \text{ and } 3 \text{ and } \delta = 1/4,\ 1/2 \text{ and } 3/4,$$

so that 9 lag orders are used in the simulations. Note that HML (2003) and HLM (2005) recommended $K = O(T^{1/2})$. The number of replications is 5,000 and the significance level is set to 0.05. All computations are conducted using the GAUSS matrix language.

For the purpose of comparison, we also calculate the single cointegration test statistic proposed by Shin (1994), which is one of the most frequently used test in applications, and the LBIU test by Kurozumi and Arai (2008), which can control the empirical size well even when the errors are strongly serially correlated. The leads-lags truncation parameter for the Shin's test is the same as the above, while the semiparametric correction is used for the LBIU test; see Kurozumi and Arai (2008) for details.

Table 1 reports the rejection frequencies of the tests. The Shin's test can well control the size of the test when the serial correlation is not strong but when $\phi = 0.8$, it suffers from over-size distortion. On the other hand, the empirical size of the LBIU test is close to 0.05 even in the case of strong serial correlation. For the autocovariance based tests, the columns $\hat{S}_k(a)$ and $\tilde{S}_k(a)$ correspond to the case when $K = [(aT)^{1/2}]$ for $a = 1$, 2 and 3 are used; the cases with the other lag orders are omitted to save space.[7] From the table, we observe that the tests with no bias correction, $\hat{S}_k(a)$, tend to be conservative because of the effect of the negative bias, whereas the empirical sizes of the bias corrected versions, $\tilde{S}_k(a)$, are close to the nominal one except for the case where $a = 1$ and $\phi = \psi = 0.8$. Overall, the finite sample performance of our test under the null hypothesis is better than that of the Shin's test and as good as the LBIU test.

With respect to power, the Shin's test seems more powerful than the LBIU test by observing the results for $\psi = 0$ and 0.4, in which case the sizes of these two tests are close to the nominal one. On the other hand, the bias corrected versions of the autocovariance based test are more powerful than the LBIU test and relatively comparable to the Shin's test. Taking into account the finite sample performance under both the null and the alternative hypotheses, we recommend using our bias corrected test with $K = [(2T)^{1/2}]$.

[7] Roughly speaking, the autocovariance based tests with $K = [(aT)^{1/4}]$ and $[(aT)^{3/4}]$ result in the over-size distortion when the errors are strongly serially correlated.

5.2. Panel cointegration tests

We next investigate the finite sample performance of the panel cointegration tests. The data generating process is similar to the single equation case and it is given by

$$y_{it} = \alpha_i' c_t + \beta_i x_{it} + e_{i,t}, \qquad x_{i,t} = x_{i,t} + v_{i,t}, \qquad e_{i,t} = u_{i,t} + \lambda_i f_t,$$

$$u_{i,t} = \phi_i u_{i,t-1} + \varepsilon_{i,t}^u, \qquad v_{i,t} = \psi_i v_{i,t-1} + \varepsilon_{i,t}^v,$$

where $\alpha_i = 0$, $\beta_i = 1$, and $[\varepsilon_{i,t}^u, \varepsilon_{i,t}^v]' \sim i.i.d.N(0, \Sigma)$ with $vech(\Sigma) = [1, 0.5, 1]'$ with $u_0 = v_0 = 0$. The error terms $e_{i,t}$ consist of the idiosyncratic errors $u_{i,t}$ and the common components $\lambda_i f_t$ with common factor f_t and loading factors λ_i. The idiosyncratic errors, $u_{i,t}$, and the driving force of the I(1) regressors, $v_{i,t}$, are correlated for the same individual i, but they are cross-sectionally independent. We set $\lambda_i = 0$ for the case of no cross-sectional correlation while $\lambda_i \sim U(0, 1)$ and $f_t \sim i.i.d.N(0, 1)$ for the case of cross-sectional dependence. To see the effect of serial correlation on the tests under the null hypothesis, we consider three cases; $\phi_i \sim U(-0.4, 0.4)$ and $\psi_i \sim U(-0.4, 0.4)$ (mild serial correlation), $\phi_i \sim U(-0.8, 0.8)$ and $\psi_i \sim U(-0.8, 0.8)$ (diversified serial correlation), and $\phi_i \sim U(0.7, 0.9)$ and $\psi_i \sim U(0.7, 0.9)$ (strong serial correlation). Under the alternative hypothesis that not all the individuals are cointegrated, we generate $\phi_i = 1$ for $i = 1, \cdots, N_1$ and ϕ_i for $i = N_1 + 1, \cdots, N$ are the same as in the null hypothesis.

Table 2 summarizes the results for the case of no cross-sectional dependence and mild serial correlation. For the purpose of comparison, we also calculate the test statistic proposed by McCoskey and Kao (1998) and report the rejection frequencies on the column denoted by MK.[8] From the rows of $N_1/N = 0$ in the table, we can see that the MK test suffers from over-size distortion in almost all the cases.[9] As theoretically expected from the previous section, our tests with no bias correction tends to under-reject the null hypothesis. In particular, the empirical size is almost zero when $N = 100$. On the other hand, the bias corrected versions work well, which indicates the effectiveness of our bias correction, except for the case where $T = 100$ and $N \geq 50$. Because of the size distortion there is no need to discuss the power of the MK test and the statistics $\hat{S}_k(a)$. The powers of $\tilde{S}_k(a)$ is generally good, it increases with the sample size T and the ratio N_1/N as expected.

Table 3 reports the case of diversified serial correlation. In this case, the empirical size of $\tilde{S}_k(2)$ is close to the nominal one in many cases whereas $\tilde{S}_k(3)$ tend to over-reject the null hypothesis when $T = 100$, although its performance improves as T gets larger. On the other hand, when the serial correlation is positively strong, the control of the empirical size becomes more difficult as in Table 4. However, we observe that the empirical size of $\tilde{S}_k(2)$ is still close to 0.05 in many cases even in the presence of strong serial correlation.

[8] We used the asymptotic mean and variance to construct the test statistic by McCoskey and Kao (1998). Because they reported the mean and variance only in the constant case, we calculate those values in the trend case by simulations.

[9] Our preliminary simulations show that the size of the MK test becomes closer to the nominal one when all the errors are independent.

The rejection frequencies in the case of the cross-sectional dependence are reported in Tables 5-7. The relative performance is preserved but it seems that the sizes of $\tilde{S}_k(a)$ become better in this case.

All of the above results correspond to the case where $x_{i,t}$ is cross-sectionally independent. However, in practical analysis, they may be correlated and moreover, it is possible for x_{it} to be cointegrated with $x_{j,t}$. Then, we consider the same data generating process as before except that

$$x_{i,t} = \gamma_i(x_t + w_{i,t}), \qquad x_t = x_{t-1} + v_t,$$

where $v_t = \psi v_{t-1} + \varepsilon_t^v$, $w_{i,t} \sim i.i.d.N(0,1)$ and $\gamma_i \sim U(0.5, 1.5)$. That is, we consider the case where $x_{i,t}$ are driven by the common stochastic trend, which implies that $[x_{1,t}, \cdots, x_{N,t}]'$ are cointegrated with cointegrating rank $N-1$.

Again, the overall property of $\tilde{S}_k(a)$ is preserved but they suffer from over-size distortion when $T = 100$ and N is not small as in Tables 8–13. In summary, our bias corrected tests with $K = [(2T)^{1/2}]$ or $K = [(2T)^{1/3}]$ are recommended in practical analysis on the basis of our extensive simulations.

6. Conclusion

In this paper we have proposed tests assuming a null hypothesis of cointegration. Contrary to the single equation cointegration tests in the literature where the limiting distributions are non-standard, we show that our tests have a standard normal asymptotic distribution. Our tests are transposed to panel data cointegration tests allowing for cross-section dependence and serial correlation. We prove for N fixed and $T \to \infty$ that the limiting distributions of our statistics are standard normals. We have derived a bias correction which is shown to work well in finite sample via Monte Carlo simulations, particularly when T is larger than N. Finally, our tests are robust to the likely presence of cointegration across units which is often the case in macroeconomic data.

Appendix

In this appendix, $\bar{c}$ signifies a generic positive constant that may differ from place to place.

Proof of Lemma 1

Using expression (6), we decompose $\eta_t \eta_{t-K}$ into 5 parts as follows:

$$\begin{aligned}
\eta_t \eta_{t-K} &= \sum_{j=-\infty}^{\infty} \phi_j \xi_{t-j} \sum_{\ell=-\infty}^{\infty} \phi_\ell \xi_{t-K-\ell} \\
&= \sum_{j=1}^{K-1} \sum_{\ell=0}^{\infty} g_t(j,\ell) + \sum_{j=1}^{K-1} \sum_{\ell=1-K}^{-1} g_t(j,\ell) + \sum_{j=1-K}^{0} \sum_{\ell=1-K}^{\infty} g_t(j,\ell) \\
&\quad + \sum_{|j| \geq K}^{\infty} \sum_{\ell=-\infty}^{\infty} g_t(j,\ell) + \sum_{j=1-K}^{K-1} \sum_{\ell=-\infty}^{-K} g_t(j,\ell)
\end{aligned}$$

$$\equiv \quad C_{1,t} + C_{2,t} + C_{3,t} + r_{2,t} + r_{3,t}, \quad \text{say,} \tag{19}$$

where $g_t(j,\ell) = \phi_j \phi_\ell \xi_{t-j} \xi_{t-K-\ell}$.

For $C_{1,t}$, we can see that

$$C_{1,t} = \sum_{j=1}^{K-1} \sum_{\ell=0}^{j-1} g_t(j,\ell) + \sum_{j=1}^{K-1} \sum_{\ell=j}^{\infty} g_t(j,\ell).$$

The first term becomes

$$\begin{aligned}
\sum_{j=1}^{K-1} \sum_{\ell=0}^{j-1} g_t(j,\ell) &= [g_t(K-1,0)] + [g_t(K-2,0) + g_t(K-1,1)] \\
&\qquad + \cdots + [g_t(1,0) + g_t(2,1) + \cdots + g_t(K-1,K-2)] \\
&= \sum_{j=1}^{K-1} \sum_{\ell=0}^{j-1} g_t(K-j+\ell,\ell) \\
&= \sum_{j=1}^{K-1} \sum_{\ell=K-j}^{K-1} g_t(\ell, j+\ell-K) \\
&= \sum_{j=1}^{K-1} \sum_{\ell=K-j}^{K-1} \phi_\ell \phi_{j+\ell-K} \xi_{t-\ell} \xi_{t-\ell-j}, \qquad (20)
\end{aligned}$$

where the third equality holds by re-defining ℓ as $K-j+\ell$. Similarly, we have

$$\begin{aligned}
\sum_{j=1}^{K-1} \sum_{\ell=j}^{\infty} g_t(j,\ell) &= \sum_{j=0}^{\infty} \sum_{\ell=1}^{K-1} g_t(\ell, j+\ell) \\
&= \sum_{j=K}^{\infty} \sum_{\ell=1}^{K-1} g_t(\ell, j+\ell-K) \\
&= \sum_{j=K}^{\infty} \sum_{\ell=1}^{K-1} \phi_\ell \phi_{j+\ell-K} \xi_{t-\ell} \xi_{t-\ell-j}, \qquad (21)
\end{aligned}$$

where the second equality is obtained by re-defining j as $j+K$.

Similarly, we have

$$\begin{aligned}
C_{2,t} &= \sum_{j=1}^{K-1} \sum_{\ell=1-K}^{-1} g_t(j,\ell) \\
&= \sum_{j=1}^{K-1} \sum_{\ell=1}^{K-1} \phi_j \phi_{\ell-K} \xi_{t-j} \xi_{t-\ell} \\
&= \sum_{j=1}^{K-1} \phi_j \phi_{j-K} \xi_{t-j}^2 + \sum_{j=1}^{K-2} \sum_{\ell=j+1}^{K-1} (\phi_j \phi_{\ell-K} + \phi_\ell \phi_{j-K}) \xi_{t-j} \xi_{t-\ell}
\end{aligned}$$

$$= \quad r_{1,t} + \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} (\phi_\ell \phi_{j+\ell-K} + \phi_{j+\ell}\phi_{\ell-K})\xi_{t-\ell}\xi_{t-\ell-j}. \tag{22}$$

Then, we have, from (20)–(22),

$$C_{1,t} + C_{2,t} = r_{1,t} + \sum_{j=1}^{\infty} \sum_{\ell=1}^{K-1} \phi_\ell \phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j} + \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} \phi_{j+\ell}\phi_{\ell-K}\xi_{t-\ell}\xi_{t-\ell-j}. \tag{23}$$

For $C_{3,t}$,

$$\begin{aligned}
C_{3,t} &= \sum_{j=1-K}^{0} \sum_{\ell=1-K}^{\infty} g_t(j,\ell) \\
&= \sum_{j=1-K}^{0} \sum_{\ell=1-K}^{j} g_t(j,\ell) + \sum_{j=1-K}^{0} \sum_{\ell=j+1}^{\infty} g_t(j,\ell) \\
&= \sum_{j=1}^{K} \sum_{\ell=1}^{j} g_t(-j+\ell, \ell-K) + \sum_{j=1}^{\infty} \sum_{\ell=1}^{K} g_t(\ell-K, j+\ell-K) \\
&= \sum_{j=1}^{K} \sum_{\ell=1-j}^{0} g_t(\ell, j+\ell-K) + \sum_{j=1}^{\infty} \sum_{\ell=1-K}^{0} g_t(\ell, j+\ell) \\
&= \sum_{j=1}^{K} \sum_{\ell=1-j}^{0} g_t(\ell, j+\ell-K) + \sum_{j=K+1}^{\infty} \sum_{\ell=1-K}^{0} g_t(\ell, j+\ell-K) \\
&= \sum_{j=1}^{\infty} \sum_{\ell=(1-j)\vee(1-K)}^{0} g_t(\ell, j+\ell-K) \\
&= \sum_{j=1}^{\infty} \sum_{\ell=(1-j)\vee(1-K)}^{0} \phi_\ell \phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j},
\end{aligned} \tag{24}$$

and then from (23) and (24), we have

$$\begin{aligned}
&C_{1,t} + C_{2,t} + C_{3,t} \\
&= \quad r_{1,t} + \sum_{j=1}^{\infty} \sum_{\ell=(1-j)\vee(1-K)}^{K-1} \phi_\ell \phi_{j+\ell-K}\xi_{t-\ell}\xi_{t-\ell-j} + \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} \phi_{j+\ell}\phi_{\ell-K}\xi_{t-\ell}\xi_{t-\ell-j} \\
&= \quad r_{1,t} + C_{a,t} + C_{b,t}, \quad \text{say.}
\end{aligned} \tag{25}$$

We next apply the B–N decomposition to $C_{a,t}$ and $C_{b,t}$. For $C_{a,t}$, we consider three cases where $\ell = 0$, $\ell \geq 1$ and $\ell \leq -1$. For $\ell = 0$, we have

$$C_{a,t} = \sum_{j=1}^{\infty} \phi_0 \phi_{j-K}\xi_t\xi_{t-j}, \tag{26}$$

while for $\ell \geq 1$,

$$
\begin{aligned}
C_{a,t} &= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}L^\ell\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}[1-(1-L^\ell)]\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta\sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\sum_{i=0}^{\ell-1}L^i\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta\sum_{j=1}^{\infty}\sum_{i=0}^{K-2}\left(\sum_{\ell=i+1}^{K-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j} \\
&= \sum_{j=1}^{\infty}\sum_{\ell=1}^{K-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta\tilde{r}_{1,t}. \qquad (27)
\end{aligned}
$$

For $\ell \leq -1$, it is sufficient to consider the case where $j \geq 2$. For $j = 2, \cdots, K-1$, the summand of $C_{a,t}$ becomes

$$
\begin{aligned}
&\sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}L^\ell\xi_t\xi_{t-j} \\
&= \sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}[1-(1-L^\ell)]\xi_t\xi_{t-j} \\
&= \sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}\sum_{i=\ell+1}^{0}L^i\xi_t\xi_{t-j} \\
&= \sum_{\ell=1-j}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{i=2-j}^{0}\left(\sum_{\ell=1-j}^{i-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j}, \qquad (28)
\end{aligned}
$$

where $\Delta^+ = (1-L^{-1})$ and we used the relation $(1-L^\ell) = (1-L^{-1})(1+L^{-1}+\cdots+L^{\ell+1})$ for $\ell < 0$, while for $j \geq K$, it can be expressed as

$$
\begin{aligned}
&\sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}L^\ell\xi_t\xi_{t-j} \\
&= \sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}\sum_{i=\ell+1}^{0}L^i\xi_t\xi_{t-j} \\
&= \sum_{\ell=1-K}^{-1}\phi_\ell\phi_{j+\ell-K}\xi_t\xi_{t-j} - \Delta^+\sum_{i=2-K}^{0}\left(\sum_{\ell=1-K}^{i-1}\phi_\ell\phi_{j+\ell-K}\right)L^i\xi_t\xi_{t-j}. \qquad (29)
\end{aligned}
$$

From (28) and (29), $C_{a,t}$ for $\ell \leq -1$ becomes

$$\begin{aligned} C_{a,t} &= \sum_{j=2}^{\infty} \sum_{\ell=1-(j\wedge K)}^{-1} \phi_\ell \phi_{j+\ell-K} \xi_t \xi_{t-j} - \Delta^+ \sum_{j=2}^{\infty} \sum_{i=2-(j\wedge K)}^{0} \left(\sum_{\ell=1-(j\wedge K)}^{i-1} \phi_\ell \phi_{j+\ell-K} \right) L^i \xi_t \xi_{t-j} \\ &= \sum_{j=2}^{\infty} \sum_{\ell=1-(j\wedge K)}^{-1} \phi_\ell \phi_{j+\ell-K} \xi_t \xi_{t-j} - \Delta^+ \tilde{r}_t^+. \end{aligned} \tag{30}$$

Combining (26), (27) and (30), we can see that

$$\begin{aligned} C_{a,t} &= \sum_{j=1}^{\infty} \sum_{\ell=1-(j\wedge K)}^{K-1} \phi_\ell \phi_{j+\ell-K} \xi_t \xi_{t-j} - \Delta \tilde{r}_{1,t} - \Delta^+ \tilde{r}_t^+ \\ &= \sum_{j=1}^{\infty} G_{1,j} \xi_t \xi_{t-j} - \Delta \tilde{r}_{1,t} - \Delta^+ \tilde{r}_t^+. \end{aligned} \tag{31}$$

In exactly the same way, we have

$$\begin{aligned} C_{b,t} &= \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} \phi_{j+\ell} \phi_{\ell-K} \xi_t \xi_{t-j} - \Delta \sum_{j=1}^{K-2} \sum_{\ell=1}^{K-j-1} \phi_{j+\ell} \phi_{\ell-K} \sum_{i=0}^{\ell-1} L^i \xi_t \xi_{t-j} \\ &= \sum_{j=1}^{K-2} G_{2,j} \xi_t \xi_{t-j} - \Delta \sum_{j=1}^{K-2} \sum_{i=0}^{K-j-2} \left(\sum_{\ell=i+1}^{K-j-1} \phi_{j+\ell} \phi_{\ell-K} \right) L^i \xi_t \xi_{t-j} \\ &= \sum_{j=1}^{K-2} G_{2,j} \xi_t \xi_{t-j} - \Delta \tilde{r}_{2,t}. \end{aligned} \tag{32}$$

Combining (19), (25), (31) and (32), we obtain (11).■

Proof of Lemma 2

From (11) in Lemma 1, we can see that

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \eta_t \eta_{t-K} &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \sum_{j=1}^{\infty} G_j \xi_t \xi_{t-j} + \frac{1}{\sqrt{T}} \left(\tilde{r}_0 - \tilde{r}_{[Tr]} - r_1^+ + r_{[Tr]}^+ \right) \\ &\quad + \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} (r_{1,t} + r_{2,t} + r_{3,t}). \end{aligned} \tag{33}$$

We will show that the FCLT holds for the first term on the right-hand side while the other terms are negligible, using the following lemma:

Lemma 4 *For $\{\phi_j\}_{j=-\infty}^{\infty}$ satisfying the condition given by (6), (i)* $\sum_{|j|\geq K}^{\infty} |\phi_j| = o\left(\frac{1}{K^2}\right)$ *and* $\sum_{|j|\geq K}^{\infty} |\phi_j|^2 = o\left(\frac{1}{K^4}\right)$*, (ii)* $\sum_{j=1}^{\infty} |G_j| < \infty$*, (iii)* $\sum_{j=1}^{\infty}\sum_{\ell=0}^{K-2} |\tilde{G}_{1,\ell}| < \infty$*, (iv)* $\sum_{j=1}^{K-2}\sum_{\ell=0}^{K-j-2} |\tilde{G}_{2,\ell}| < \infty$*, (v)* $\sum_{j=2}^{\infty}\sum_{\ell=2-(j\wedge K)}^{0} |\tilde{G}_{\ell}^{+}| < \infty$*. The relations (ii)–(v) hold uniformly over K.*

Proof of Lemma 4: (i) is shown by

$$\begin{aligned}
\sum_{|j|\geq K}^{\infty} |\phi_j| &\leq \frac{1}{K^2}\sum_{|j|\geq K}^{\infty} |j|^2|\phi_j| = o\left(\frac{1}{K^2}\right), \\
\sum_{|j|\geq K}^{\infty} |\phi_j|^2 &\leq \frac{1}{K^4}\sum_{|j|\geq K}^{\infty} |j|^4|\phi_j|^2 = o\left(\frac{1}{K^4}\right).
\end{aligned}$$

For (ii)-(v), we have

$$\begin{aligned}
\sum_{j=1}^{\infty} |G_j| &\leq \sum_{j=1}^{\infty} |G_{1,j}| + \sum_{j=1}^{K-2} |G_{2j}| \\
&\leq \sum_{j=1}^{\infty}\sum_{\ell=1-(j\wedge K)}^{K-1} |\phi_\ell||\phi_{j+\ell-K}| + \sum_{j=1}^{K-2}\sum_{\ell=1}^{K-j-1} |\phi_{j+\ell}||\phi_{\ell-K}| \\
&\leq \sum_{\ell=1-K}^{K-1} |\phi_\ell| \sum_{j=1}^{\infty} |\phi_{j+\ell-K}| + \sum_{\ell=1}^{K-2} |\phi_{\ell-K}| \sum_{j=1}^{K-2} |\phi_{j+\ell}| \\
&\leq \left(\sum_{\ell=-\infty}^{\infty} |\phi_\ell|\right)^2 + \sum_{\ell=-\infty}^{\infty} |\phi_\ell| \sum_{j=1}^{\infty} |\phi_j| < \infty.
\end{aligned}$$

$$\begin{aligned}
\sum_{j=1}^{\infty}\sum_{\ell=0}^{K-2} |\tilde{G}_{1,\ell}| &\leq \sum_{j=1}^{\infty}\sum_{\ell=0}^{K-2}\sum_{i=\ell+1}^{K-1} |\phi_i||\phi_{i+j-K}| \\
&\leq \sum_{\ell=0}^{K-2}\sum_{i=\ell+1}^{K-1} |\phi_i| \sum_{j=-\infty}^{\infty} |\phi_j| \\
&= \sum_{i=1}^{K-1} i|\phi_i| \sum_{j=-\infty}^{\infty} |\phi_j| < \infty.
\end{aligned}$$

$$\sum_{j=1}^{K-2}\sum_{\ell=0}^{K-j-2} |\tilde{G}_{2,\ell}| \leq \sum_{j=1}^{K-2}\sum_{\ell=0}^{K-j-2}\sum_{i=\ell+1}^{K-j-1} |\phi_{i+j}||\phi_{i-K}|$$

$$
\begin{aligned}
&= \sum_{j=1}^{K-2} \sum_{i=1}^{K-j-1} i|\phi_{i+j}||\phi_{i-K}| \\
&\leq \sum_{i=1}^{K-2} |\phi_{i-K}| \sum_{j=1}^{K-2} (i+j)|\phi_{i+j}| \\
&\leq \sum_{i=-\infty}^{\infty} |\phi_i| \sum_{j=1}^{\infty} j|\phi_j| < \infty.
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=2}^{\infty} \sum_{\ell=2-(j\wedge K)}^{0} |\tilde{G}_\ell^+| &\leq \sum_{j=2}^{\infty} \sum_{\ell=2-(j\wedge K)}^{0} \sum_{i=1-(j\wedge K)}^{\ell-1} |\phi_i||\phi_{i+j-K}| \\
&\leq \sum_{\ell=2-K}^{0} \sum_{i=1-K}^{\ell-1} |\phi_i| \sum_{j=2}^{\infty} |\phi_{i+j-K}| \\
&\leq \sum_{i=1-K}^{-1} |i||\phi_i| \sum_{j=-\infty}^{\infty} |\phi_j| < \infty. \blacksquare
\end{aligned}
$$

Note that the absolute summability in Lemma 4(ii)–(v) implies the square summability of the corresponding terms. Using Lemma 4, we show that all the term on the right hand side of (33), except for the first term, are negligible.

Lemma 5 *For $\tilde{r}_t$, $\tilde{r}_t^+$, $r_{1,t}$, $r_{2,t}$ and $r_{3,t}$ in* (33)*, (i)* $\sup_{0\leq r\leq 1} \left| \frac{1}{\sqrt{T}} \tilde{r}_{[Tr]} \right| = o_p(1)$ *and* $\sup_{0\leq r\leq 1} \left| \frac{1}{\sqrt{T}} \tilde{r}_{[Tr]}^+ \right| = o_p(1)$*. (ii)* $\sup_{0\leq r\leq 1} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} r_{i,t} \right| = o_p(1)$ *for $i = 1, 2$ and 3.*

Proof of Lemma 5: (i) We first note that $\tilde{r}_t = \tilde{r}_{1,t} + \tilde{r}_{2,t}$ as defined in Lemma 1. Since

$$
P\left(\sup_{0\leq r\leq 1} \left| \frac{1}{\sqrt{T}} \tilde{r}_{i,t} \right| \geq \varepsilon \right) \leq TP\left(\frac{1}{\sqrt{T}} |\tilde{r}_{i,t}| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^4 T} E[\tilde{r}_{i,t}^4]
$$

for $i = 1$ and 2, it is sufficient to prove that $E[\tilde{r}_{i,t}^4] < \infty$ for $i = 1$ and 2. Noting that non-zero terms of $E[\tilde{r}_{i,t}^4]$ are related to the products among $E[\xi_t^2]$, $E[\xi_t^3]$ and $E[\xi_t^4]$, all of which are bounded by assumption, we can see that

$$
\begin{aligned}
E[\tilde{r}_{1,t}^4] &\leq \bar{c} \left(\sum_{j=1}^{\infty} \sum_{\ell=0}^{K-2} |\tilde{G}_{1,\ell}| \right)^4 < \infty, \\
E[\tilde{r}_{2,t}^4] &\leq \bar{c} \left(\sum_{j=1}^{K-2} \sum_{\ell=0}^{K-j-2} |\tilde{G}_{2,\ell}| \right)^4 < \infty
\end{aligned}
$$

uniformly over K by Lemma 4(iii) and (iv). The second statement of (i) for $\tilde{r}_t^+$ is proved in exactly the same manner.

(ii) For $i = 1$, we first show that $E[r_{1,t}] = o(1/K^2)$. From the definition of $r_{1,t}$, we have

$$E[|r_{1,t}|] \leq \sigma_\xi^2 \sum_{j=-\infty}^{\infty} |\phi_j||\phi_{j-K}|. \tag{34}$$

Noting that

$$\begin{aligned} \sum_{K=-\infty}^{\infty} |K|^2 \sum_{j=-\infty}^{\infty} |\phi_j||\phi_{j-K}| &\leq 2 \sum_{K=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (|j-K|^2 + |j|^2)|\phi_j||\phi_{j-K}| \\ &\leq 4 \sum_{K=-\infty}^{\infty} |\phi_K| \sum_{j=-\infty}^{\infty} |j|^2|\phi_j| < \infty \end{aligned}$$

because of the 2-summability of $\{\phi_j\}$, we can see that $|K|^2 \sum_{j=-\infty}^{\infty} |\phi_j||\phi_{j-K}|$ is a convergence sequence over K. In other words, $K^2 \sum_{j=-\infty}^{\infty} |\phi_j||\phi_{j-K}|$ is $o(1)$ as $K \to \infty$ and then from (34), $E[|r_{1,t}|] = o(1/K^2)$ uniformly over t.

Using this result, since

$$\sup_{0\leq r\leq 1} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} r_{1,t} \right| \leq \frac{1}{\sqrt{T}} \sum_{t=1}^{T} |r_{1,t}|,$$

we obtain

$$E\left[\sup_{0\leq r\leq 1} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} r_{1,t} \right|\right] \leq \frac{1}{\sqrt{T}} \sum_{t=1}^{T} E\left[|r_{1,t}|\right] = o\left(\frac{\sqrt{T}}{K^2}\right) = o(1).$$

For $i = 2$, by Cauchy-Schwarz inequality, we have

$$\begin{aligned} E[|r_{2,t}|] &\leq \left\{ E\left[\left(\sum_{|j|\geq K}^{\infty} \phi_j \xi_{t-j}\right)^2\right] E\left[\left(\sum_{\ell=-\infty}^{\infty} \phi_\ell \xi_{t-K-\ell}\right)^2\right]\right\}^{1/2} \\ &= \left\{ \sigma_\xi^4 \sum_{|j|\geq K}^{\infty} \phi_j^2 \sum_{\ell=-\infty}^{\infty} \phi_\ell^2 \right\}^{1/2} \\ &= o\left(\frac{1}{K^2}\right) \end{aligned}$$

by Lemma 4(i). Then, we have $E[\sup_r |T^{-1/2} \sum_{t=1}^{[Tr]} r_{2,t}|] = o(1)$ in exactly the same manner as the proof for $i = 1$.

The case with $i = 3$ is shown similarly and we omit the proof.■

From Lemma 5, the rest we have to show is that the FCLT holds for the first term on the right-hand side of (33). From Theorem 27.14 of Davidson (1994), it is sufficient to show that

$$\frac{\sum_{t=1}^{T} m_t^2}{\sum_{t=1}^{T} E[m_t^2]} \xrightarrow{p} 1, \tag{35}$$

$$\frac{\max_{1\leq t\leq T} |m_t|}{(\sum_{t=1}^{T} E[m_t^2])^{1/2}} \xrightarrow{p} 0, \tag{36}$$

$$\lim_{T\to\infty} \frac{\sum_{t=1}^{[Tr]} E[m_t^2]}{\sum_{t=1}^{T} E[m_t^2]} \to r \quad \forall\, 0 \leq r \leq 1, \tag{37}$$

where $m_t = \sum_{j=1}^{\infty} G_j \xi_t \xi_{t-j}$.

The condition (35) holds if we show that $T^{-1}\sum_{t=1}^{T}(m_t^2 - E[m_t^2])) \xrightarrow{p} 0$, which is proved using Chebyshev inequality by showing that

$$\begin{aligned} E\left[\left\{\frac{1}{T}\sum_{t=1}^{T}(m_t^2 - E[m_t^2])\right\}^2\right] &= \frac{1}{T^2}\sum_{t=1}^{T} E\left[(m_t^2 - E[m_t^2])^2\right] \\ &\quad + \frac{2}{T^2}\sum_{s=1}^{T-1}\sum_{t=s+1}^{T} E\left[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}^2])\right] \\ &\to 0. \end{aligned} \tag{38}$$

For the first term on the right-hand side of (38),

$$\begin{aligned} \frac{1}{T^2}\sum_{t=1}^{T} E\left[\left(m_t^2 - E[m_t^2]\right)^2\right] &= \frac{1}{T^2}\sum_{t=1}^{T} E\left[\left\{\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} G_i G_j \left(\xi_t^2 \xi_{t-i}\xi_{t-j} - \sigma_\xi^2 E[\xi_{t-i}\xi_{t-j}]\right)\right\}^2\right] \\ &\leq \frac{\bar{c}}{T}\left(\sum_{j=1}^{\infty} |G_j|\right)^4 \to 0. \end{aligned} \tag{39}$$

For the second term, note that for $s > 0$,

$$
\begin{aligned}
& E[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}^2])] \\
= & \sum_{i_1=1}^{\infty}\sum_{j_1=1}^{\infty}\sum_{i_2=1}^{\infty}\sum_{j_2=1}^{\infty} G_{i_1}G_{i_2}G_{i_3}G_{i_4} E\left[(\xi_t^2\xi_{t-i_1}\xi_{t-j_1} - \sigma_\xi^2 E[\xi_{t-i_1}\xi_{t-j_1}])\right. \\
& \left. (\xi_{t-s}^2\xi_{t-s-i_2}\xi_{t-s-j_2} - \sigma_\xi^2 E[\xi_{t-s-i_2}\xi_{t-s-j_2}])\right].
\end{aligned}
$$

The expectation becomes

$$
\begin{aligned}
& E\left[(\xi_t^2\xi_{t-i_1}\xi_{t-j_1} - \sigma_\xi^2 E[\xi_{t-i_1}\xi_{t-j_1}])(\xi_{t-s}^2\xi_{t-s-i_2}\xi_{t-s-j_2} - \sigma_\xi^2 E[\xi_{t-s-i_2}\xi_{t-s-j_2}])\right] \\
= & E\left[\sigma_\xi^2(\xi_{t-i_1}\xi_{t-j_1} - E[\xi_{t-i_1}\xi_{t-j_1}])(\xi_{t-s}^2 - \sigma_\xi^2)\xi_{t-s-i_2}\xi_{t-s-j_2}\right] \\
& + E\left[\sigma_\xi^2(\xi_{t-i_1}\xi_{t-j_1} - E[\xi_{t-i_1}\xi_{t-j_1}])\sigma_\xi^2(\xi_{t-s-i_2}\xi_{t-s-j_2} - E[\xi_{t-s-i_2}\xi_{t-s-j_2}])\right].
\end{aligned}
$$

Since $\{\xi_t\}$ is an independent sequence, the first expectation takes non-zero values when i) $i_1 = j_1 = s$ and $i_2 = j_2$, ii) $i_1 = s + i_2$ and $j_1 = s + j_2$, (iii) $i_1 = s + j_2$ and $j_1 = s + j_2$, while for the second expectation, it is sufficient to consider either iv) $i_1 = s + i_2$ and $j_1 = s + j_2$ or (v) $i_1 = s + j_2$ and $j_1 = s + j_2$. Therefore, we can see that

$$
\left|E[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}^2])]\right| \le \bar{c}\left[G_s^2\sum_{j_2=1}^{\infty} G_{j_2}^2 + \left(\sum_{i_2=1}^{\infty}|G_{s+i_2}||G_{i_2}|\right)^2\right],
$$

and thus,

$$
\begin{aligned}
& \left|\frac{1}{T^2}\sum_{s=1}^{T-1}\sum_{t=s+1}^{T} E\left[(m_t^2 - E[m_t^2])(m_{t-s}^2 - E[m_{t-s}])^2\right]\right| \qquad (40) \\
\le & \frac{\bar{c}}{T}\left[\sum_{s=1}^{T-1} G_s^2\sum_{j_2=1}^{\infty} G_{j_2}^2 + \sum_{s=1}^{T-1}\left(\sum_{i_2=1}^{\infty}|G_{s+i_2}||G_{i_2}|\right)^2\right] \le \frac{\bar{c}}{T}\left[\left(\sum_{j_2=1}^{\infty} G_{j_2}^2\right)^2 + \left(\sum_{i_2=1}^{\infty}|G_{i_2}|\right)^4\right] \to 0.
\end{aligned}
$$

Then, (38) holds from (39) and (40).

To prove (36), we note that $E[m_t^2] = \sigma_\xi^4\sum_{j=1}^{\infty} G_j^2 < \infty$ and then the denominator of (36) is $O(\sqrt{T})$. On the other hand,

$$
\begin{aligned}
P\left(\max_{1\le t\le T}\frac{1}{\sqrt{T}}|m_t| \ge \varepsilon\right) & \le T\ P\left(\frac{1}{\sqrt{T}}|m_t| \ge \varepsilon\right) \\
& \le \frac{1}{\varepsilon^4 T}E[m_t^4] = O\left(\frac{1}{T}\right),
\end{aligned}
$$

because $E[m_t^4]$ is bounded uniformly in t, T and M. Therefore, we obtain (36).

Finally, we can see that (37) holds even in finite samples because of stationarity of m_t.■

Proof of Lemma 3

Let $\tilde{D}_T = diag\{\sqrt{T}, T\ I_{p_x}\}$ (constant case) or $\tilde{D}_T = \{\sqrt{T}, T\sqrt{T}, I_{p_x}\}$ (trend case), $\|B\| = [tr(B'B)]^{1/2}$ and $\|B\|_1 = \sup\{\|Bx\| : \|x\| \leq 1\}$ for a matrix B. We will show that only $R_{\beta,T}$ yields the non-zero bias whereas $R_{\Pi,T}$ and R_T are negligible, using the following lemma:

Lemma 6 *Suppose that Assumptions 1, 2 and 3 hold. Under the null hypothesis, as* $T \to \infty$, *(i)* $\tilde{D}_T^{-1}(\hat{\beta} - \beta) \xrightarrow{d} \left(\int_0^1 \tilde{B}(r)\tilde{B}'(r)dr\right)^{-1} \int_0^1 \tilde{B}(r)dB_\eta(r)$, *(ii)* $\|\hat{\Pi} - \Pi\|^2 = O_p(M/T)$, *(iii)* $\tilde{D}^{-1}\sum_{t=K+1}^T X_{t-K}\eta_t \xrightarrow{d} \int_0^1 \tilde{B}(r)dB_\eta(r)$ *and* $\tilde{D}^{-1}\sum_{t=K+1}^T X_t\eta_{t-K} \xrightarrow{d} \int_0^1 \tilde{B}(r)dB_\eta(r)$, *(iv)* $\|T^{-1/2}\sum_{t=K+1}^T V_{t-K}\eta_t\| = \|T^{-1/2}\sum_{t=K+1}^T V_t\eta_{t-K}\| = O_p(M^{1/2})$, *(v)* $\|\sum_{t=K+1}^T \eta_t e_{t-K}\| = \|\sum_{t=K+1}^T \eta_{t-K}e_t\| = o_p(1)$, *(vi)* $\tilde{D}^{-1}\sum_{t=K+1}^T X_tX_{t-K}'\tilde{D}^{-1} \xrightarrow{d} \int_0^1 \tilde{B}(r)\tilde{B}'(r)dr$, *(vii)* $\|\tilde{D}_T^{-1}\sum_{t=K+1}^T X_tV_{t-K}'\| = \|\tilde{D}_T\sum_{t=K+1}^T X_{t-K}V_t'\| = O_p(M^{1/2})$, *(viii)* $\|\tilde{D}_T\sum_{t=K+1}^T X_te_{t-K}\| = \|\tilde{D}_T\sum_{t=K+1}^T X_{t-K}e_t\| = o_p(1)$, *(ix)* $\|T^{-1/2}\sum_{t=K+1}^T V_tV_{t-K}'\| = O_p(M)$, *(x)* $\|\left(T^{-1}\sum_{t=K+1}^T V_tV_t'\right)^{-1} - \Gamma_x^{-1}\|_1 = O_p(M/\sqrt{T})$, *(xi)* $\|\sum_{t=K+1}^T V_te_{t-K}\| = \|\sum_{t=K+1}^T V_{t-K}e_t\| = o_p(M^{1/2})$, *(xii)* $\|\sum_{t=K+1}^T e_te_{t-K}\| = o_p(1)$, *where* $\tilde{B}(r) = [1, B'(r)]'$ *(constant case) or* $\tilde{B}(r) = [1, r, B'(r)]'$ *(trend case) with* $B(r)$ *being a* p_x*-dimensional Brownian motion with the variance given by* $\lim_{T\to\infty} E[T^{-1/2}x_T]$, $B_\eta(r)$ *is a one-dimensional Brownian motion independent of* $B(r)$ *with the variance given by* $\omega_\eta^2 = \lim_{T\to\infty} E[(T^{-1/2}\sum_{t=1}^T \eta_t)^2]$ *and* $\Gamma_x = E[V_tV_t']$.

Proof of Lemma 6: All the results, except for (v), (ix) and (xi), are obtained by Saikkonen (1991) using the FCLT with K going to infinity slower than T. For (v), we can see that

$$\left|\sum_{t=K+1}^T \eta_t e_{t-K}\right| \leq \sup_{1\leq t\leq T}|e_t| \sum_{t=1}^T |\eta_t|.$$

Note that $\sum_{t=1}^T |\eta_t| = O_p(T)$ while

$$\begin{aligned} P\left(\sup_{1\leq t\leq T}|e_t| \geq \varepsilon\right) &\leq TP(|e_t| \geq \varepsilon) \\ &\leq \frac{T}{\varepsilon^4}E[e_t^4] \\ &\leq \frac{\bar{c}T}{\varepsilon^4}\left(\sum_{|j|\geq K}\|\pi_j\|\right)^4 \end{aligned}$$

$$= \frac{\bar{c}T}{\varepsilon^4}o\left(\frac{1}{T^2}\right) = o\left(\frac{1}{T}\right), \quad (41)$$

where the second last equality is obtained by (9).We thus obtain (v).

For (ix), each block element is expressed as $T^{-1/2}\sum_{t=K+1}^{T} v_{t-i}v'_{t-K-j}$ for $i,j=-M,\cdots,M$. Since $(t-i)-(t-K-j)=K-i+j\geq K-2M$, we can see that the time difference diverges to infinity at a rate of K because $M/K\to 0$ by Assumptions 2 and 3. Because the conditions for the FCLT given by HML (2003) are satisfied, we can see that each element is $O_p(1)$, which implies (ix).

(xi) is proved by noting that

$$E\left[\left\|\sum_{t=K+1}^{T} V_t e_{t-K}\right\|\right] \leq \sup|e_t| \sum_{t=1}^{T} E[\|V_t\|] = o_p(\sqrt{M}),$$

because $\sup_t |e_t| = o_p(1/T)$ by (41).■

We first evaluate $R_{\beta,T}$. Using Lemma 6(i), (iii) and (vi), we have

$$R_{\beta,T} \xrightarrow{d} -\left(\int_0^1 \tilde{B}(r)dB_\eta(r)\right)'\left(\int_0^1 \tilde{B}(r)\tilde{B}'(r)dr\right)^{-1}\left(\int_0^1 \tilde{B}(r)dB_\eta(r)\right). \tag{42}$$

Since $\int_0^1 \tilde{B}(r)dB_\eta(r)|\tilde{B}(\cdot) \sim N\left(0, \omega_\eta^2 \int_0^1 \tilde{B}(r)\tilde{B}'(r)dr\right)$, we can see that the right-hand side of (42) is distributionally equal to $-\omega_\eta^2$ times a chi-square distribution with (p_c+p_x) degrees of freedom. As a result, $E[R_{\beta,T}]$ can be approximated by $-\omega_\eta^2(p_c+p_x)$.

For $R_{\Pi,T}$, the first term becomes

$$\begin{aligned}\left\|(\hat{\Pi}-\Pi)'\sum_{t=K+1}^{T} V_t V_{t-K}(\hat{\Pi}-\Pi)\right\| &\leq \left\|\hat{\Pi}-\Pi\right\|^2 \left\|\sum_{t=K+1}^{T} V_t V_{t-K}\right\| \\ &= O_p\left(\frac{M^2}{\sqrt{T}}\right) = o_p(1),\end{aligned}$$

using Lemma 6 (ii) and (ix) and Assumption 2.

For the second term of $R_{\Pi,T}$, since it can be shown that

$$\left\|\sqrt{T}(\hat{\Pi}-\Pi) - \left(\frac{1}{T}\sum_{t=1}^{T} V_t V_t'\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} V_t \eta_t\right)\right\| \leq O_p\left(\sqrt{\frac{M}{T}}\right),$$

while

$$\left\|\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\eta_t V_t'\right)\left[\left(\frac{1}{T}\sum_{t=1}^{T} V_t V_t'\right)^{-1} - \Gamma_x^{-1}\right]\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} V_{t-K}\eta_t\right)\right\| = O_p\left(\frac{M^2}{\sqrt{T}}\right) = o_p(1)$$

by Lemma 6 (iv) and (x), it is sufficient to evaluate

$$\begin{aligned}&\left|E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\eta_t V_t'\right)\Gamma_x^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} V_{t-K}\eta_t\right)\right]\right| \\ &\leq \sup|\Gamma_x^{-1}(i,j)|\frac{1}{T}\sum_{t=1}^{T}\sum_{\ell=t-T}^{t-1}\left|E\left[V_t' V_{t-K-\ell}\eta_t\eta_{t-\ell}\right]\right|.\end{aligned} \tag{43}$$

To evaluate the right-hand side of (43), we express η_t using ε_t such that

$$\eta_t = \sum_{j=-\infty}^{\infty} \psi_j' \varepsilon_{t-j}, \quad \text{where} \quad \sum_{j=-\infty}^{\infty} |j|^2 \|\psi_j\| \leq \infty$$

with $\{\psi_j\}_{j=-\infty}^{\infty}$ is a sequence of $p_x + 1$-dimensional coefficient vectors, because

$$\eta_t = u_t - \sum_{j=-\infty}^{\infty} \pi_j' v_{t-j} \quad \text{with} \quad u_t = \sum_{j=0}^{\infty} \Phi_{1,j} \varepsilon_{t-j} \quad \text{and} \quad v_t = \sum_{j=0}^{\infty} \Phi_{2,j} \varepsilon_{t-j},$$

where Φ_j is partitioned into $\Phi_j = [\Phi_{1,j}', \Phi_{2,j}']'$. Then, by focusing on the term $v_t' v_{t-K-\ell}$ in $V_t' V_{t-K-\ell}$, we can see that

$$\begin{aligned}
&\tilde{R}_{\Pi,\ell} \equiv E\left[v_t' v_{t-K-\ell} \eta_t \eta_{t-\ell}\right] \\
&= E\left[\left(\sum_{j_1=0}^{\infty} \Phi_{2,j_1} \varepsilon_{t-j_1}\right)' \left(\sum_{j_2=0}^{\infty} \Phi_{2,j_2} \varepsilon_{t-K-\ell-j_2}\right) \left(\sum_{i_1=-\infty}^{\infty} \psi_{i_1}' \varepsilon_{t-i_1}\right) \left(\sum_{i_2=-\infty}^{\infty} \psi_{i_2}' \varepsilon_{t-\ell-i_2}\right)\right].
\end{aligned}$$

We note that the expectation takes non-zero values when (i) $j_1 = K + \ell + j_2$, $i_1 = \ell + i_2$ and $i_2 \neq K + j_2$, (ii) $i_1 = j_1$, $i_2 = K + j_2$ and $j_1 \neq K + \ell + j_2$, (iii) $i_1 = K + \ell + j_2$, $i_2 = j_1 - \ell$ and $j_1 \neq K + \ell + j_2$, and (iv) $i_1 = K + \ell + j_2$, $i_2 = K + j_2$ and $j_1 = K + \ell + j_2$.

In case (i), for $\ell \geq 0$, the sum of $\tilde{R}_{\Pi,\ell}$ becomes

$$\begin{aligned}
\left|\sum_{\ell=0}^{\infty} \tilde{R}_{\Pi,\ell}\right| &\leq \bar{c} \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{j_2}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\| \|\psi_{i_2}\| \\
&\leq \bar{c} \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \left(\sum_{j=0}^{\infty} \|\Phi_{2,j}\|\right) \left(\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2}\|\right)^2 \\
&\leq \bar{c} \sum_{j_2=K}^{\infty} (j_2 - K + 1) \|\Phi_{2,j_2}\| = o\left(\frac{1}{K}\right), \qquad (44)
\end{aligned}$$

because $\{\Phi_j\}$ is 2-summable.

On the other hand, for $\ell = -1, -2, \cdots, -K$, we have

$$\begin{aligned}
\left|\sum_{\ell=-K}^{-1} \tilde{R}_{\Pi,\ell}\right| &\leq \bar{c} \sum_{\ell=-K}^{[-K/2]} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\|\|\Phi_{j_2}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\|\|\psi_{i_2}\| \\
&\quad +\bar{c} \sum_{\ell=[-K/2]+1}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\|\|\Phi_{j_2}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\|\|\psi_{i_2}\| \\
&\leq \bar{c} \left(\sum_{j_2=0}^{\infty} \|\Phi_{j_2}\|\right)^2 \sum_{\ell=-K}^{[-K/2]} \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\|\|\psi_{i_2}\| \\
&\quad +\bar{c} \left(\sum_{\ell=[-K/2]+1}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\|\right) \left(\sum_{j_2=0}^{\infty} \|\Phi_{j_2}\|\right) \left(\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2}\|\right)^2 \\
&= o\left(\frac{1}{K}\right) + o\left(\frac{1}{K}\right), \qquad (45)
\end{aligned}$$

where the last relation holds because

$$\begin{aligned}
\sum_{K=-\infty}^{\infty} |K|^2 \sum_{i_2=-\infty}^{\infty} \|\psi_{i_2-K}\|\|\psi_{i_2}\| &\leq 2 \sum_{K=-\infty}^{\infty} \sum_{i_2=-\infty}^{\infty} (|i_2|^2 + |i_2 - K|^2)\|\psi_{i_2-K}\|\|\psi_{i_2}\| \\
&\leq 4 \sum_{K=-\infty}^{\infty} \|\psi_K\| \sum_{i_2=-\infty}^{\infty} |i_2|^2 \|\psi_{i_2}\| < \infty,
\end{aligned}$$

which implies $|K|^2 \sum_{i_2=-\infty}^{\infty} \|\psi_{i_2-K}\|\|\psi_{i_2}\| = o(1)$ or, equivalently, $\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2-K}\|\|\psi_{i_2}\| = o(1/K^2)$, while

$$\sum_{\ell=[-K/2]+1}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \leq \left[\frac{K}{2}\right] \sum_{j_2=[K/2]}^{\infty} \|\Phi_{2,j_2}\| = o\left(\frac{1}{K}\right)$$

because of 2-summability of $\{\Phi_{2,j}\}$.

For $\ell \leq -K-1$,

$$\begin{aligned}
\left|\sum_{\ell=-\infty}^{-K+1} \tilde{R}_{\Pi,\ell}\right| &\leq \bar{c} \sum_{\ell=-\infty}^{-K-1} \sum_{j_1=0}^{\infty} \|\Phi_{2,j_1}\|\|\Phi_{j_1-K-\ell}\| \sum_{i_2=-\infty}^{\infty} \|\psi_{\ell+i_2}\|\|\psi_{i_2}\| \\
&\leq \bar{c} \sum_{\ell=-\infty}^{-K-1} \sum_{j_1=0}^{\infty} \|\Phi_{j_1-K-\ell}\| \left(\sum_{j_1=0}^{\infty} \|\Phi_{2,j_1}\|\right) \left(\sum_{i_2=-\infty}^{\infty} \|\psi_{i_2}\|\right)^2 \\
&\leq \bar{c} \sum_{j_1=K}^{\infty} j_1 \|\Phi_{j_1}\| = o\left(\frac{1}{K}\right). \qquad (46)
\end{aligned}$$

From (44)–(46), we have $\left|\sum_{\ell=-\infty}^{\infty} \tilde{R}_{\Pi,\ell}\right| = o(1/K)$ in case (i).

In case (ii), we first note that

$$E[v_s \eta_t] = \sum_{j=0}^{\infty} \Phi_{2,j} \Sigma_\varepsilon \psi_{j+\ell} = 0 \quad \forall \ell = 0, \pm 1, \pm 2, \cdots. \tag{47}$$

Then, we have for $\ell \geq 0$,

$$\begin{aligned}
\left| \sum_{\ell=0}^{\infty} \tilde{R}_{\Pi,\ell} \right| &= \left| \sum_{\ell=0}^{\infty} \sum_{\mathrm{j}_2=0}^{\infty} \sum_{\mathrm{j}_1=0}^{\infty}{}_{j_1 \neq K+\ell+j_2} \psi'_{j_1} \Sigma_\varepsilon \Phi'_{2,j_1} \Phi_{2,j_2} \Sigma_\varepsilon \psi_{K+j_2} \right| \\
&= \left| \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \psi'_{K+\ell+j_2} \Sigma_\varepsilon \Phi'_{2,K+\ell+j_2} \Phi_{2,j_2} \Sigma_\varepsilon \psi_{K+j_2} \right| \\
&\leq \bar{c} \sum_{\ell=0}^{\infty} \sum_{j_2=0}^{\infty} \|\psi_{K+\ell+j_2}\| \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{2,j_2}\| \psi_{K+j_2}\| \\
&\leq \bar{c} \left(\sum_{j_2=0}^{\infty} j_2 \|\psi_{K+j_2}\| \right) \left(\sum_{j_2=0}^{\infty} j_2 \|\Phi_{2,K+j_2}\| \right) \left(\sum_{j_2=0}^{\infty} \|\Phi_{2,j_2}\| \right) \left(\sum_{j_2=0}^{\infty} \|\psi_{K+j_2}\| \right) \\
&= o\left(\frac{1}{K^2} \right),
\end{aligned}$$

where the second equality holds using (47).

Similarly for $\ell = -1, \cdots, -K$,

$$\begin{aligned}
\left| \sum_{\ell=-K}^{-1} R_{\Pi,t} \right| &\leq \bar{c} \sum_{\ell=-K}^{-1} \sum_{j_2=0}^{\infty} \|\psi_{K+\ell+j_2}\| \|\Phi_{2,K+\ell+j_2}\| \|\Phi_{2,j_2}\| \psi_{K+j_2}\| \\
&\leq \bar{c} \left(\sum_{\ell=-K}^{-1} \sum_{j_2=0}^{\infty} \|\Phi_{2,K+\ell+j_2}\| \right) \left(\sum_{j_2=-\infty}^{\infty} \|\psi_{j_2}\| \right) \left(\sum_{j_2=0}^{\infty} \|\Phi_{2,j_2}\| \right) \left(\sum_{j_2=0}^{\infty} \|\psi_{K+j_2}\| \right) \\
&\leq \bar{c} \left(\sum_{j_2=0}^{K-1} (j_2+1) \|\Phi_{2,j_2}\| + K \sum_{j_2=K}^{\infty} \|\Phi_{2,j_2}\| \right) \left(\sum_{j_2=0}^{\infty} \|\psi_{K+j_2}\| \right) = o\left(\frac{1}{K} \right),
\end{aligned}$$

while for $\ell \leq -K-1$,

$$\begin{aligned}
\left| \sum_{\ell=-\infty}^{-K-1} R_{\Pi,t} \right| &= \left| \sum_{\ell=-\infty}^{-K-1} \sum_{\mathrm{j}_2=0}^{\infty} \sum_{\mathrm{j}_1=0}^{\infty}{}_{j_2 \neq j_1-K-\ell} \psi'_{j_1} \Sigma_\varepsilon \Phi'_{2,j_1} \Phi_{2,j_2} \Sigma_\varepsilon \psi_{K+j_2} \right| \\
&= \left| \sum_{\ell=-\infty}^{-K-1} \sum_{j_1=0}^{\infty} \psi'_{j_1} \Sigma_\varepsilon \Phi'_{2,j_1} \Phi_{2,j_1-K-\ell} \Sigma_\varepsilon \psi_{j_1-\ell} \right|
\end{aligned}$$

$$\leq \bar{c}\left(\sum_{j_1=0}^{\infty}\|\psi_{j_1}\|\right)\left(\sum_{j_1=0}^{\infty}\|\Phi_{2,j_1}\|\right)\left(\sum_{j_1=0}^{\infty}j_1\|\Phi_{2,j_1}\|\right)\left(\sum_{j_1=K+1}^{\infty}(j_1-K)\|\psi_{j_1}\|\right)$$
$$= o\left(\frac{1}{K}\right).$$

We then have $\left|\sum_{\ell=-\infty}^{\infty}\tilde{R}_{\Pi,\ell}\right| = o(1/K)$ in case (ii).

In exactly the same way, we have the same order in cases (iii) and (iv), so that $\left|\sum_{\ell=-\infty}^{\infty}\tilde{R}_{\Pi,\ell}\right| = o(1/K)$ in general. Then, we can see that the right-hand side of (43) is $o(M/K) = o(1)$ by Assumptions 2 and 3, so that the second term of $R_{\Pi,T}$ is $o_p(1)$. Similarly, we can show that the third term of $R_{\Pi,T}$ is $o_p(1)$.

Using Lemma 6, it is not difficult to see that $R_T = o_p(1)$. As a result, we obtain the bias.■

Proof of Theorem 2

As given by Lemma 1, we can apply the B–N decomposition to each $\eta_{i,t}\eta_{i,t-K}$. We can also see from Theorem 1 that $\eta_{i,t}\eta_{i,t-K}$ is the dominate term in $\hat{\eta}^*_{i,t}\hat{\eta}^*_{i,t-K}$while the other terms are negligible and the bias becomes as given in Lemma 3 for each i. The rest we have to show is that the FCLT holds for $\sum_{i=1}^{N}\eta_{i,t}\eta_{i,t-K}$. Note that because $\eta_{i,t}$ is obtained by linear transformations of $\varepsilon_{i,t}$, $\eta_{i,t}$ is independent of $\eta_{j,s}$ for all i, j and $s \neq t$. Thus, we can see that $\sum_{i=1}^{N}\eta_{i,t}\eta_{i,t-K}$ is a martingale difference sequence with respect to the sigma-field constructed from $\eta_{1,t}, \eta_{2,t-1}, \cdots, \eta_{2,t}, \eta_{2,t-1}, \cdots, \eta_{N,t}, \eta_{N,t-1}, \cdots$. Because $G_{i,j}$ for $i = 1, \cdots, N$ satisfy Lemma 4(ii), we can see that the conditions of the FCLT given by Theorem 27.14 of Davidson (1994) are satisfied as in the proof of Theorem 1. We then have the theorem.■

References

[1] Bai, J. and J. L. Carrion-i-Silvestre (2013). Testing Panel Cointegration with Unobservable Dynamic Common Factors That Are Correlated with the Regressors. *Econometrics Journal* 16, 222-249.

[2] Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*. Chichester, Wiley.

[3] Banerjee, A., M. Marcellino and C. Osbat (2004). Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data. *Econometrics Journal* 7, 322-340.

[4] Banerjee, A., M. Marcellino and C. Osbat (2005). Testing for PPP: Should We Use Panel Methods? *Empirical Economics* 30, 77-91.

[5] Banerjee, A. and M. Wagner (2009). Panel Methods to Test for Unit Roots and Cointegration, in T. C. Mills and K. Patterson ed., *Palgrave Handbook of Econometrics, Vol.2*. Palgrave Macmillan, New York.

[6] Breitung, J. and M. H. Pesaran (2008). Unit Roots and Cointegration in Panels. In L. Matyas and P. Sevestre (eds), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice (3rd edition)*. Berlin, Spsringer-Verlag.

[7] Brillinger, D. R. (1981). *Time Series Data Analysis and Theory*. San Francisco, Holden-Day.

[8] Chang, Y. and C. M. Nguyen (2012). Residual Based Tests for Cointegration in Dependent Panels. *Journal of Econometrics* 167, 504-520.

[9] Choi, I. (2013). Panel Cointegration. Sogang University, Working Papers number 1208.

[10] Choi, I. and E. Kurozumi (2012). Model Selection Criteria for the Leads-and-Lags Cointegrating Regression. *Journal of Econometrics* 169, 224-238.

[11] Davidson, J. (1994). *Stochastic Limit Theory*. Oxford, Oxford University Press.

[12] Engle, R. F. and C. W. J. Granger (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55, 251-276.

[13] Gengenbach, C., F. C. Palm and J.-P. Urbain (2006). Cointegration Testing in Panels with Common Factors. *Oxford Bulletin of Economics and Statistics* 68, 683-719.

[14] Groen, J.J.J., and F. Kleibergen. (2003). Likelihood-Based Cointegration Analysis in Panels of Vector Error-Correction Models, *Journal of Business & Economic Statistics*, 21, 295–318.

[15] Hanck, C. (2009). Cross-Sectional Correlation Robust Tests for Panel Cointegration. *Journal of Applied Statistics* 36, 817-833.

[16] Hadri K., R. Larsson and Y. Rao (2012). Testing for Stationarity with a Break in Heterogeneous Panels Where the Time Dimension is Finite. *Bulletin of Economic Research* 64, 123-148.

[17] Hansen, B. E. (1992). Tests for Parameter Instability in Regressions with I(1) Processes. *Journal of Business and Economic Statistics* 10, 321-335.

[18] Harris, D., B. McCabe and S. Leybourne (2003). Some Limit Theory for Autocovariances Whose Order Depends on Sample Size. *Econometric Theory* 19, 829-864.

[19] Harris, D., S. Leybourne and B. McCabe (2005). Panel Stationarity Tests for Purchasing Power Parity with Cross-Sectional Dependence. *Journal of Business and Economic Statistics* 23, 395-409.

[20] Hayakawa, K. and E. Kurozumi (2008). The Role of "Leads" in the Dynamic OLS Estimation of Cointegrating Regression Models. *Mathematics and Computers in Simulation* 79, 555-560.

[21] Hubrich, K., H. Lütkepohl, and P. Saikkonen (2001). A Review of Systems Cointegrating Tests. *Econometric Reviews* 20, 247-318.

[22] Jansson, M. (2005). Point Optimal Tests of the Null Hypothesis of Cointegration. *Journal of Econometrics* 124, 187-201.

[23] Kao, C. (1999). Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. *Journal of Econometrics* 90, 1-44.

[24] Kejriwal, M. and P. Perron (2008). Data Dependent Rules for Selection of the Number of Leads and Lags in the Dynamic OLS Cointegrating Regression. *Econometric Theory* 24, 1425-1441.

[25] Kurozumi, E. and Y. Arai (2008). Test for the Null Hypothesis of Cointegration with Reduced Size Distortion. *Journal of Time Series Analysis* 29, 476-500.

[26] Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159-178.

[27] Larsson, R., and J. Lyhagen. (2007). Inference in Panel Cointegration Models with Long Panels, *Journal of Business & Economic Statistics*, 25, 473–483.

[28] Larsson, R., J. Lyhagen, and M. Löthgren. (2001). Likelihood-Based Cointegration Tests in Heterogeneous Panels, *Econometric Journal*, 4, 109–142.

[29] McCoskey, S. and C. Kao (1999). A Residual-Based Test for the Null of Cointegration in Panel Data. *Econometric Reviews* 17, 57-84.

[30] Pedroni, P. (1999). Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. *Oxford Bulletin of Economics and Statistics* 61, 653-670.

[31] Pedroni, P. (2004). Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis. *Econometric Theory* 20, 597-625.

[32] Phillips, P. C. B. and B. E. Hansen (1990). Statistical Inference in Instrumental Variables Regression With I(1) Processes. *Review of Economic Studies* 57, 99-125.

[33] Phillips, P. C. B. and M. Loretan (1991). Estimating Long-Run Economic Equilibria. *Review of Economic Studies* 58, 407-436.

[34] Phillips, P. C. B. and H. R. Moon (1999). Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica* 67, 1057-1111.

[35] Quintos, C. E. and P. C. B. Phillips (1993). Parameter Constancy in Cointegrating Regressions. *Empirical Economics* 18, 675-706.

[36] Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegration Regressions. *Econometric Theory* 7, 1-21.

[37] Shin, Y. (1994). A Residual- Based Test of the Null of Cointegration against the Alternative of No Cointegration. *Econometric Theory* 10, 91-115.

[38] Stock, J. H. and M. W. Watson (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica* 61, 783-820.

[39] Westerlund, J. (2005). A Panel CUSUM Test for the Null of Cointegration. *Oxford Bulletin of Economics and Statistics* 67, 231-262.

[40] Westerlund, J. (2008). Panel Cointegration Tests of the Fisher Effect. *Journal of Applied Econometrics* 23, 193-233.

[41] Westerlund, J. and D. L. Edgerton (2007). A Panel Bootstrap Cointegration Test. *Economics Letters* 97, 185-190.

[42] Westerlund, J. and D. L. Edgerton (2008). A Simple Test for Cointegration in Dependent Panels with Structural Breaks. *Oxford Bulletin of Economics and Statistics* 70, 665-704.