# BLOCK BOOTSTRAP PANEL UNIT ROOT TESTS WITH DETERMINISTIC TERMS

XIJIA LIU† AND JIANXIN WEI†

Abstract. In this paper, we generalize the block bootstrap panel unit root tests proposed by Palm, Smeekes and Urbain (2011, PSU) in the sense of considering the drift terms and time trend in the model. We consider two different methods to deal with the deterministic terms. One is to modify the test statistics by adding the deterministic terms into the auxiliary Dickey-Fuller regression and adjust the bootstrap algorithm regarding the model specification. Another is applying a certain detrending method on the observations and then performing the block bootstrap panel unit root tests on the detrended residuals, i.e. a two stage method. For both of them, we theoretically checked the asymptotic validity of the bootstrap tests under the main null hypothesis. The results of simulation studies indicate that the tests based on the modified test statistics and detrending methods all have acceptable size in general. The two stage method has better power properties than the modified test statistics, and the GLS detrending method has the best performance among others.

## 1. Introduction

It is well known that unit root tests suffer from a poor power problem especially when the sample size is small. In order to overcome this problem, doing unit root tests in a panel setting is introduced so that more powerful tests can be expected by collecting more information from cross section units. After the seminal work of Levin, Lin and Chu (2002, LLC) and Im, Pesaran and Shin (2003, IPS), panel unit root tests became a very active research topic in time series econometrics. However, these papers are based on the assumption that all cross section units are uncorrelated, and the exciting power improvement will be broken down by relaxing this convenient but unrealistic assumption. For example, O'Connell (1998) shows that the tests in LLC and IPS have serious size distortions when cross sectional dependence is present in the panel.

In order to solve this problem, the so called second generation panel unit root tests emerged. Several methods have been proposed. For example, Chang (2002) proposes a nonlinear instrumental variable approach to deal with the cross section dependence; Breitung and Das (2005) propose a robust version of the Dickey-Fuller t-statistic under a contemporaneous structure. However, factor models became the most popular tool to address the cross sectional dependence in panel unit root tests. For example, Moon and Perron (2004) and Pesaran (2007) consider modeling cross section dependence by common factors and test unit roots on the idiosyncratic components. Apart from this, another influential study is the PANIC procedure proposed by Bai and Ng (2004), in which both the unit roots in idiosyncratic components and the cross sectional cointegrations are considered. Based on this framework, they perform the unit root tests on idiosyncratic errors and common factors separately.

---

*Key words and phrases.* Panel unit root test, block bootstrap, robust, cross section dependence and deterministic terms.

The authors wish to gratefully thank Stephen Smeekes for the useful help and valuable suggestions. †Uppsala University. Address correspondence to Xijia Liu, Department of Statistics, Uppsala University, Box 513, 751 20 Uppsala, Sweden; e-mail: xijia.liu@statistik.uu.se.

As argued in PSU, most of the second generation panel unit root tests can be applied in common factor structures and contemporaneous dependence, however the other types of cross sectional dependence are seldom considered. For this motivation, based on the block bootstrap method, PSU proposed a robust panel unit root test which can work not only in the common factors structure but also in a wider range of other plausible dynamic dependence. More specifically, two assumptions in the commonly used factor models are relaxed. One is that they allow the dependence between factors and idiosyncratic components; another one is that they relax the assumption of weak dependence between idiosyncratic components and allow for a wide array of possible dependence both through the long run covariance and the lag polynomials. In one word, PSU contains almost all the form of dependence in the literature as a special case. In addition, they also avoid several nuisance parameter estimation problems, for example, the number of common factors and the nuisance parameters due to the serial correlation, and which make the tests able to be easily implemented.

There are several open problems in PSU, for example, they did not consider the asymptotic properties for large $N$. From the empirical applications point of view, however, another more important issue is that PSU only considered the simplest model, random walk without drift terms. Therefore, in this paper, we focus on the generalization of PSU in the sense of considering the deterministic terms. Two cases are considered: one considers an intercept term in the model and another one includes both the intercept term and a time trend. Two methods of addressing the deterministic terms are considered in this paper. First, we follow the traditional method in unit root tests, i.e. to modify the test statistics with respect to the model specification and complete the generalization of the block bootstrap algorithm proposed in Paparoditis and Politis (2003, PP). Second, we consider the so called two stage method, i.e. we do hypothesis tests in two steps. In the first step, we apply certain detrending methods to get the detrended residuals. Based on the detrended residuals, we apply the same test statistics which are employed in PSU to do hypothesis tests. For a proper adjustment of the bootstrap algorithm, we find that Smeekes (2009) has discussed the role of detrending in sieve bootstrap unit root tests which is based on the Augmented Dickey-Fuller test. Following his main conclusions, we suggest a bootstrap algorithm for the tests which is based on the detrended residuals. For the validity of these bootstrap tests, we theoretically checked the consistency under the main null hypothesis. The consistency of the bootstrap test under the remaining null hypothesis and the power properties of those tests are investigated by simulation studies.

The rest of the paper is organized as follows. In Section 2, we describe the basic framework. The conventional way will be discussed in Section 3. In a parallel way, the tests which are based on the two stage method will be discussed in Section 4. Section 5 reports the results of a simulation study to illustrate the performance in different methods. At last we give the conclusion and discussion. All proofs and tables are included in the Appendix.

## 2. Basic framework

As mentioned before, we consider generalizing the block bootstrap panel unit root tests into a more complicated model which contains the deterministic terms, thus our framework will be built based on PSU. First, about the model, let $x_t = (x_{1,t}, ..., x_{N,t})$ for $t = 1, ..., T$ be generated by

$$x_t = d_t^m + y_t \tag{1}$$

where $d_t^m = \left(d_{1,t}^m, ..., d_{N,t}^m\right)'$ for $m = 0, 1, 2$ is the deterministic part. For each $i$, $d_{i,t}^m = \beta_i^{m\prime} z_t^m$, where

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

Then $d_t^m$ can be expressed as $d_t^m = \beta^m z_t^m$ where $\beta^m = (\beta_1^m, ..., \beta_N^m)'$. For future reference, we also partition $\beta^\tau$ in another way and define $\beta^\tau = (\beta_{1\cdot}, \beta_{2\cdot})$ where $\beta_{j\cdot} = (\beta_{j1}, ..., \beta_{jN})'$ and $j = 1, 2$. $y_t$ is the stochastic part and follows the same model as in PSU. That is $y_t$ is generated as

$$y_t = \Lambda F_t + w_t$$

where the factor loadings, $\Lambda = (\lambda_1, ..., \lambda_d)'$, the common factors, $F_t = (F_{1,t}, ..., F_{d,t})'$ and the idiosyncratic error, $w_t = (w_{1,t}, ..., w_{N,t})'$. Let the factor and idiosyncratic components be generated by

$$\begin{aligned} F_t &= \Phi F_{t-1} + f_t \\ w_t &= \Theta w_{t-1} + v_t \end{aligned}$$

where $\Phi = diag\left(\phi_1, ..., \phi_d\right)$ and $\Theta = diag\left(\theta_1, ..., \theta_N\right)$. Additionally $v_t$ and $f_t$ are generated by

$$\begin{pmatrix} v_t \\ f_t \end{pmatrix} = \psi\left(L\right)\varepsilon_t = \begin{bmatrix} \psi_{11}\left(L\right) & \psi_{12}\left(L\right) \\ \psi_{21}\left(L\right) & \psi_{22}\left(L\right) \end{bmatrix} \begin{bmatrix} \varepsilon_{v,t} \\ \varepsilon_{f,t} \end{bmatrix}$$

where $\psi\left(z\right) = \sum_{j=0}^{\infty} \psi_j z^j$ and $\psi_0 = I$. We also need an assumption on $\psi(z)$ and $\epsilon_t$

**Assumption 1.**

(i) $\det\left(\psi\left(z\right)\right) \neq 0$ for all $\{z \in \mathbb{C} : |z = 1|\}$ and $\sum_{j=0}^{\infty} j\left|\psi_j\right| < \infty$.
(ii) $\varepsilon_t$ is i.i.d. with $E\varepsilon_t = 0$, $E\varepsilon_t\varepsilon_t' = \Sigma$ and $E\left|\varepsilon_t\right|^{2+\epsilon} < \infty$ for some $\epsilon > 0$.

The null hypothesis is $H_0$: $y_{i,t}$ has a unit root for all $i = 1, ..., N$. As discussed in PSU, this can occur in four different settings.

(A) $\theta_i = \phi_j = 1$ for all $i = 1, ..., N$ and $j = 1, ..., d$: both the common factors and the idiosyncratic components have a unit root.
(B) $|\theta_i| < 1$ for all $i = 1, ..., N$, $\phi_j = 1$ for all $j = 1, ..., d$: the common factors have a unit root while the idiosyncratic components are stationary.
(C) $|\phi_j| < 1$ for all $j = 1, ..., d$, $\theta_i = 1$ for all $i = 1, ..., N$: the common factors are stationary while the idiosyncratic components have a unit root.
(D) $|\phi_j| < 1$ for some $j$, and $|\theta_i| < 1$ for some $i$ or all $i = 1, ..., N$: part of common factors have a unit root.

We consider two different alternative hypotheses:

$H_1^a$: $y_{i,t}$ is stationary for all $i = 1, ..., N$. This implies that $|\theta_i| < 1$ for all $i = 1, ..., N$ and $|\phi_j| < 1$ for all $j = 1, ..., d$.
$H_1^b$: $y_{i,t}$ is stationary for a portion of the units. This implies that $|\phi_j| < 1$ for all $j = 1, ..., d$, while $|\theta_i| < 1$ for some $i$.

## 3. Conventional method

In this section, we consider to follow the conventional method in unit root tests, i.e. modify the test statistics corresponding to different model specifications. Then we complete the generalization of the block bootstrap algorithm of PP (2003) in the panel settings.

3.1. **Test statistics.** For model 0, we apply the same test statistics as in PSU.

$$\tau_p = T\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} x_{i,t-1}\Delta x_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} x_{i,t-1}^2} \tag{2}$$

and

$$\tau_{gm} = \frac{1}{N}\sum_{i=1}^{N} T\frac{\sum_{t=1}^{T} x_{i,t-1}\Delta x_{it}}{\sum_{t=1}^{T} x_{i,t-1}^2}. \tag{3}$$

For model 1, we use $\mu$ to indicate the test statistics for the model with intercept term. We define the pooled statistics as the Dickey-Fuller coefficient statistic from the pooled regression of $\Delta x_{it}$ on $(1, x_{i,t-1})'$. That is

$$\tau_p^\mu = T\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(x_{i,t-1} - \overline{x}_{i,-1}\right)\left(\Delta x_{it} - \Delta\overline{x}_i\right)}{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(x_{i,t-1} - \overline{x}_{i,-1}\right)^2} \tag{4}$$

We define the group mean statistic as the average of the DF coefficient statistics from the individual regressions of $\Delta x_{it}$ on $(1, x_{i,t-1})'$ for each $i = 1, ..., N$.

$$\tau_{gm}^\mu = \frac{1}{N}\sum_{i=1}^{N} T\frac{\sum_{t=1}^{T}\left(x_{i,t-1} - \overline{x}_{i,-1}\right)\left(\Delta x_{it} - \Delta\overline{x}_i\right)}{\sum_{t=1}^{T}\left(x_{i,t-1} - \overline{x}_{i,-1}\right)^2}. \tag{5}$$

For model 2, we use $\tau$ to indicate the test statistics for the model with time trend term. Similarly, we define the pooled statistics as the DF coefficient statistic from the pooled regression of $\Delta x_{it}$ on $(1, t, x_{i,t-1})'$. That is

$$\tau_p^\tau = T\frac{\sum_{i=1}^{N}\left(\sum_{t=1}^{T}\Delta x_{it}^\dagger x_{i,t-1}^\dagger\sum_{t=1}^{T} t^{\dagger 2} - \sum_{t=1}^{T}\Delta x_{it}^\dagger t^\dagger\sum_{t=1}^{T} x_{i,t-1}^\dagger t^\dagger\right)}{\sum_{i=1}^{N}\left(\sum_{t=1}^{T} x_{i,t-1}^{\dagger 2}\sum_{t=1}^{T} t^{\dagger 2} - \left(\sum_{t=1}^{T} x_{i,t-1}^\dagger t^\dagger\right)^2\right)} \tag{6}$$

where $\Delta x_{it}^\dagger = \Delta x_{it} - \Delta\overline{x}_i$, $x_{i,t-1}^\dagger = x_{i,t-1} - \overline{x}_{i,-1}$ and $t^\dagger = t - \frac{T-1}{2} \doteq t - \frac{T}{2}$. And the group mean:

$$\tau_{gm}^\tau = \frac{1}{N}\sum_{i=1}^{N} T\frac{\sum_{t=1}^{T}\Delta x_{it}^\dagger x_{i,t-1}^\dagger\sum_{t=1}^{T} t^{\dagger 2} - \sum_{t=1}^{T}\Delta x_{it}^\dagger t^\dagger\sum_{t=1}^{T} x_{i,t-1}^\dagger t^\dagger}{\sum_{t=1}^{T} x_{i,t-1}^{\dagger 2}\sum_{t=1}^{T} t^{\dagger 2} - \left(\sum_{t=1}^{T} x_{i,t-1}^\dagger t^\dagger\right)^2} \tag{7}$$

**Remark 1.** *We emphasize that the null hypothesis of our tests is always one of these three settings, or each cross section unit contains a unit root for all models, even we call them the conventional method. (Conventionally, for different models the unit root test has different null hypotheses. For example, the null hypothesis should be that the autoregressive coefficient is equal to 1 and the coefficient of time trend is fixed as 0 when we consider the model with time trend.) Adopting the same idea as PSU, we may treat those regressions by which the test statistics are constructed from the auxiliary regressions.*

Based on the results from PSU, the following lemma which describes the limiting distribution of test statistics can be easily formulated.

**Lemma 1.** *Let $x_t$ be generated under $H_0$ setting (A) and let Assumption 1 hold. Then as $T \to \infty$*

$$\tau_p^\mu \xrightarrow{d} \frac{\sum_{i=1}^N \left(\int B_i dB_i - B_i(1)\int B_i + \frac{1}{2}(\omega_{0i} + \omega_i)\right)}{\sum_{i=1}^N \left(\int B_i^2 - \left(\int B_i\right)^2\right)},$$

$$\tau_{gm}^\mu \xrightarrow{d} \frac{1}{N}\sum_{i=1}^N \frac{\int B_i(r)\,dB_i(r) - B_i(1)\int B_i(r)\,dr + \frac{1}{2}(\omega_{0i} + \omega_i)}{\int B_i(r)^2\,dr - \left(\int B_i(r)\,dr\right)^2},$$

$$\tau_p^\tau \xrightarrow{d} \frac{\sum_{i=1}^N \left(\int B_i dB_i - B_i(1)\int B_i - 12\left(\frac{1}{2}B_i(1) - \int B_i\right)\left(\int rB_i - \frac{1}{2}\int B_i\right) + \frac{1}{2}(\omega_{0i} + \omega_i)\right)}{\sum_{i=1}^N \left(\left(\int B_i^2 - 12\left(\int B_i\right)^2\right) - \left(\int rB_i - \frac{1}{2}\int B_i\right)^2\right)}$$

*and*

$$\tau_{gm}^\tau \xrightarrow{d} \frac{1}{N}\sum_{i=1}^N \frac{\int B_i dB_i - B_i(1)\int B_i - 12\left(\frac{1}{2}B_i(1) - \int B_i\right)\left(\int rB_i - \frac{1}{2}\int B_i\right) + \frac{1}{2}(\omega_{0i} + \omega_i)}{\int B_i^2 - \left(\int B_i\right)^2 - 12\left(\int rB_i - \frac{1}{2}\int B_i\right)^2}$$

*where $B_i$ is the ith element of $B(r)$ and $\omega_i\,(\omega_{0,i})$ is the $(i,i)$th element of $\Omega\,(\Omega_0)$. $\Omega$ and $\Omega_0$ follow the definitions in PSU, $\Omega = \Gamma'\psi(1)\Sigma\psi'(1)\Gamma$, $\Omega_0 = \sum_{j=0}^\infty \Gamma'\psi_j\Sigma\psi_j'\Gamma$ and $\psi(1) = \sum_{j=1}^\infty \psi_j$.*

3.2. **Bootstrap algorithm.** PSU generalized the residual based block bootstrap (RBB) theory proposed by PP into a multivariate (or panel) setting. PP constructs the RBB algorithm based on the true DGP which is

$$x_t = \beta + x_{t-1} + \epsilon_t \tag{8}$$

which contains two cases: one is $\beta = 0$, i.e. random walk and $\beta \neq 0$, i.e. random walk with non-zero drift term. However PSU only generalizes the algorithm for the first case into the panel setting. In this sense, we will complete this generalization of the RBB theory. The bootstrap algorithm is presented as follows and the corresponding invariance principle will be discussed later.

**Bootstrap algorithm:**

**1:** Calculate the centered residuals. Let

$$\widehat{u}_{it} = (x_{it} - \widetilde{\rho}_i x_{i,t-1}) - \frac{1}{T-1}\sum_{t=2}^T (x_{it} - \widetilde{\rho}_i x_{i,t-1})$$

where for model 0 and 1, $\widetilde{\rho}_i$ is the OLS estimate of regression of $x_{it}$ on $x_{i,t-1}$; for model 2, $\widetilde{\rho}_i$ is the OLS estimate of the autoregressive coefficient of the regression of $x_{it}$ on $(1, x_{i,t-1})'$. Then let $\widehat{u}_t = (\widehat{u}_{1t}, ..., \widehat{u}_{Nt})'$.

**2:** Generate the bootstrap residuals $u_t^*$. Choose a block length $b < T$. Draw $i_0, ..., i_{k-1}$ i.i.d. from the uniform distribution on $1, 2, ..., T-b$, where $k = \lfloor (T-2)/b \rfloor + 1$ is the number of blocks. Then let

$$u_t^* = \widehat{u}_{i_m + s}$$

where $m = \lfloor (t-2)/b \rfloor$ and $s = t - mb - 1$.

**3:** Construct the bootstrap sample. Let $x_1^* = x_1$ and construct $x_t^*$ recursively as

$$x_t^* = \widehat{\beta} + x_{t-1}^* + u_t^*$$

where for model 0 and 1, we set $\widehat{\beta} = 0$; for model 2, let $\widehat{\beta} = \left(\widehat{\beta}_1, ..., \widehat{\beta}_N\right)'$ where for each $i$, $\widehat{\beta}_i$ is the OLS estimate of the intercept term of the regression of $x_{it}$ on $(1, x_{i,t-1})'$.

**4:** Calculate the corresponding test statistics from the bootstrap sample.

**5:** Repeat Step 2-4 $B$ times, obtaining bootstrap test statistics $\tau_k^{*vb}$ where $b = 1, ..., B$, $k = p$, $gm$ and $v = \mu$, $\tau$, and select the bootstrap critical value $c_\alpha^*$ as

$$c_\alpha^* = \max\left\{c : B^{-1}\sum_{b=1}^{B} 1_{\{\tau_k^{*vb}<c\}} \leq \alpha\right\}$$

or equivalently as the $\alpha$-quantile of the ordered $\tau_k^{*vb}$ statistics. Reject the null of a unit root if $\tau_k^v$, calculated from Equation (4) if $k = p$ and $v = \mu$ or Equation (5) if $k = gm$ and $v = \mu$ or Equation (6) if $k = p$ and $v = \tau$ or Equation (7) if $k = gm$ and $v = \tau$, is smaller than $c_\alpha^*$, where $\alpha$ is the significance level of the test.

For the asymptotic validity of the bootstrap test, we also need the following assumption on the block length.

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

3.3. **Asymptotic validity of bootstrap tests.** PSU has proved the bootstrap invariance principle for the model which is a random walk without drift term and the convergence of some moments that appear in the asymptotic distribution, thus the asymptotic validity of bootstrap tests which are based on test statistics $\tau_p^\mu$ and $\tau_{gm}^\mu$ can be proved directly from those results. For the validity of bootstrap tests which are based on model 2, however, PSU did not provide the corresponding bootstrap invariance principle for the true model which is a random walk with drift. Based on the bootstrap algorithm in this section, it turns out that the bootstrap invariance principle is always satisfied no matter if the true model includes a drift term or not. This fact is summarized in the following lemma.

**Lemma 2.** *Under $H_0$ setting (A) and let Assumption 1 and 2 hold. Then no matter if $x_t$ is generated by model 0, 1 or 2, as $T \to \infty$,*

$$S_T^*(r) = T^{-1/2}\sum_{t=1}^{\lfloor Tr \rfloor} u_t^* \xrightarrow{d^*} B(r) \text{ in probability.}$$

After proving Lemma 2, in fact, the asymptotic validity of bootstrap tests has been proved. Since we have the bootstrap invariance principle for model 2, and the convergence of the corresponding moments is checked in PSU, we can directly prove that the limiting distributions of the bootstrap tests are exactly the same as the limiting distribution in Lemma 1.

## 4. Two stage method

In this section, we deal with the deterministic terms in another way. Simply speaking, we do hypothesis testing in two stages. In the first stage, we apply different detrending methods to get the detrending residuals, and then apply the tests in PSU on these residuals in the second step. Note that we only prove the theoretical results for time trend case, so $d_{i,t}$ is always expressed as $\beta_i' z_t$, where $z_t = (1, t)'$ and $\beta_i = (\beta_{1i}, \beta_{2i})'$ in this section. In addition, we only proved the validity of bootstrap tests under the main null hypothesis setting (A) and based on the OLS detrended residuals. The validity of bootstrap tests which are based on GLS, will be investigated by simulation studies. Next, we will illustrate the detrending methods which will be considered in this paper. Then the limiting distribution of the test statistics which are based on detrended residuals will be derived. At last, for the feasibility of hypothesis tests, we introduce the modified block bootstrap algorithm and theoretically check the validity of bootstrap tests which are based on the OLS detrending method under the main null hypothesis.

### 4.1. **Detrending methods and the corresponding asymptotic results.**

4.1.1. *Full sample OLS and GLS detrending methods.* We consider the full sample detrending method, OLS and GLS detrending, i.e. for each $i$, we apply the OLS or GLS method to estimate the coefficients of deterministic terms and calculate the detrended residuals. Define the detrended residuals as $x_t^d = x_t - \widehat{\beta} z_t$. By Equation (1), the detrended residuals can be expressed as

$$x_t^d = y_t - \left(\widehat{\beta} - \beta\right) z_t$$

For the full sample OLS detrending method, we have that for each $i$

$$\left(\widehat{\beta}_i - \beta_i\right)' = \left(\sum_{t=1}^{T} z_t' y_{it}\right)\left(\sum_{t=1}^{T} z_t z_t'\right)^{-1}$$

then

$$\widehat{\beta} - \beta = \left(\sum_{t=1}^{T} y_t z_t'\right)\left(\sum_{t=1}^{T} z_t z_t'\right)^{-1}$$

For the GLS detrending method, we follow the idea of Elliott, Rothenberg and Stock (1996, ERS) and define the following transformation

$$y_{\alpha t} = y_t - \left(1 - \alpha T^{-1}\right) y_{t-1}$$

and

$$z_{\alpha t} = z_t - \left(1 - \alpha T^{-1}\right) z_{t-1}$$

Then we have the following GLS estimation of the deterministic terms. For each $i$

$$\left(\widehat{\beta}_i - \beta_i\right)' = \left(\sum_{t=1}^{T} z_{\bar{c}t}' y_{\bar{c}it}\right)\left(\sum_{t=1}^{T} z_{\bar{c}t} z_{\bar{c}t}'\right)^{-1}$$

then

$$\widehat{\beta} - \beta = \left(\sum_{t=1}^{T} y_{\bar{c}t} z_{\bar{c}t}'\right)\left(\sum_{t=1}^{T} z_{\bar{c}t} z_{\bar{c}t}'\right)^{-1}$$

Following the recommendation by ERS, we set $\bar{c} = 7$ for the intercept only case and $\bar{c} = 13.5$ for the linear time trend case.

**Remark 2.** *In ERS, the coefficient of transformation $\bar{c}$ is decided by the asymptotic power envelope of unit root test in the univariate case. In this paper, however, we did not really find the proper $\bar{c}$ for the panel setting and just naively applied these values in our study. Of course, one could expect better power properties by considering an optimized $\bar{c}$ for the panel setting, however we will not go into this problem in this paper.*

4.1.2. *Asymptotic properties of the test statistics based on detrending residuals.* Next, we will derive the limiting distribution of test statistics which are based on detrended residuals. Under the null hypothesis (A), we have

$$x_t^d = \sum_{j=1}^{t} u_j - \left(\widehat{\beta} - \beta\right) z_t$$

where $u_t = \Delta y_t = \Gamma' \Psi (L) \varepsilon_t$ and $\Gamma = (I_N, \Lambda)'$. First, we have the following invariance principle for detrended residuals $x_t^d$:

**Lemma 3.** *Let $x_t$ be generated under $H_0$ setting (A) and let Assumption 1 hold. Then as $T \to \infty$,*

$$T^{-1/2} x^d_{\lfloor Tr \rfloor} \xrightarrow{d} B^d (r)$$

*where $B^d (r) = B (r) - \left(\int_0^1 B (s)\, ds, \int_0^1 sB (s)\, ds\right) (4 - 6r, 12r - 6)'$. Where using the same notation in PSU, we define $B (r) = \Gamma' \Psi (1) W (r)$, and $W (r)$ denotes an $(N + d)$-dimensional standard Brownian motion.*

Secondly, we define $u_t^d = \Delta x_t = \Delta y_t - \left(\widehat{\beta} - \beta\right) \Delta z_t$. Under the null hypothesis

$$u_t^d = u_t - \left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)$$

Since

$$\widehat{\beta}_{2i} - \beta_{2i} = \frac{\sum_{t=1}^{T} y_{it} \left(t - \bar{t}\right)}{\sum_{t=1}^{T} \left(t - \bar{t}\right)^2} \doteq \frac{\sum_{t=1}^{T} y_{it} \left(t - T/2\right)}{T^3/12}$$

we have

$$\widehat{\beta}_{2\cdot} - \beta_{2\cdot} \doteq \frac{\sum_{t=1}^{T} \left(t - T/2\right) y_t}{T^3/12}$$

Then we define

$$\Omega^d = \lim_{T\to\infty} T^{-1} E \left(\sum_{t=1}^{T} u_t^d\right) \left(\sum_{t=1}^{T} u_t^d\right)'$$

and

$$\Omega_0^d = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E\left(u_t^d u_t^{d\prime}\right).$$

**Lemma 4.** *Let $x_t$ be generated under $H_0$ setting (A) and let Assumption 1 hold. Then*

$$\Omega^d = \lim_{T\to\infty} T^{-1} E \left(\sum_{t=1}^{T} u_t^d\right) \left(\sum_{t=1}^{T} u_t^d\right)' = \Omega/5$$

*and*

$$\Omega_0^d = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E u_t^d u_t^{d\prime} = \Omega_0.$$

Then under the main null hypothesis, the limiting distributions for $\tau_p$ and $\tau_{gm}$ which are based on OLS detrended residuals can be constructed. The main results are summarized in the following theorem.

**Theorem 1.** *Let $x_t$ be generated under $H_0$ setting $(A)$ and let Assumption 1 hold. Then as $T \to \infty$,*

$$\tau_p^d = \frac{\sum_{i=1}^N \left[\int_0^1 B_i^d(r)\, dB_i^d(r) + \frac{1}{2}\left(\omega_i^d - \omega_{0,i}^d\right)\right]}{\sum_{i=1}^N \int_0^1 B_i^d(r)^2\, dr} + o_p(1)$$

*and*

$$\tau_{gm}^d = \frac{1}{N}\sum_{i=1}^N \frac{\int_0^1 B_i^d(r)\, dB_i^d(r) + \frac{1}{2}\left(\omega_i^d - \omega_{0,i}^d\right)}{\int_0^1 B_i^d(r)^2\, dr} + o_p(1)$$

*where $B_i^d(r)$ is the ith element of $B^d(r)$ and $\omega_i^d \left(\omega_{0,i}^d\right)$ is the $(i,i)$th element of $\Omega^d \left(\Omega_0^d\right)$.*

4.2. **Bootstrap tests.** In Smeekes (2009), the role of detrending in bootstrap unit root tests is discussed and he concludes that the detrending procedure should be performed three times. The first one is that we need to detrend the observations before estimating the residuals which will be re-sampled. The second is that we need to detrend the bootstrap sample before calculating the bootstrap statistics. At last, we need to calculate the test statistics based on detrended residuals. Furthermore, the last two detrending procedures must be identical so that the bootstrap tests are asymptotically valid. The first one could be different from the last two, but must satisfy some conditions of convergence rate. We apply these ideas to get the following modified block bootstrap algorithm for our panel unit root tests, and prove the asymptotic validity later.

4.2.1. *Bootstrap algorithm:*

**1:** Calculate the detrended residuals: $x_t^d = x_t - \widetilde{\beta} z_t$ where $\widetilde{\beta}$ is an estimate of $\beta$ in some way.

**2:** Calculate the centered residuals: for each $i$,

$$\widehat{u}_{it}^d = \left(x_{it}^d - \widehat{\rho}_i x_{i,t-1}^d\right) - \frac{1}{T-1}\sum_{t=2}^T \left(x_{it}^d - \widehat{\rho}_i x_{i,t-1}^d\right)$$

where

$$\widehat{\rho}_i = \frac{\sum_{t=2}^T x_{it}^d x_{i,t-1}^d}{\sum_{t=2}^T x_{i,t-1}^{d2}}$$

Let $\widehat{u}_t^d = \left(\widehat{u}_{1t}^d, ..., \widehat{u}_{Nt}^d\right)$.

**3:** Generate the bootstrap residuals $u_t^*$. Choose a block length $b < T$. Draw $i_0, ..., i_{k-1}$ i.i.d. from the uniform distribution on $1, 2, ..., T-b$, where $k = \lfloor (T-2)/b \rfloor + 1$ is the number of blocks. Let

$$u_t^{d*} = \widehat{u}_{i_m+s}^d$$

where $m = \lfloor (t-2)/b \rfloor$ and $s = t - mb - 1$.

**4:** Generate the bootstrap sample

$$y_t^* = \begin{cases} x_1^d & \text{for } t = 1 \\ y_{t-1}^* + u_t^{d*} & \text{for } t > 1 \end{cases}$$

and

$$x_t^* = y_t^* + \beta^* z_t$$

**5:** Perform some method to get the estimate of $\beta^*$ and calculate the detrended residual.

$$x_t^{*d} = x_t^* - \widehat{\beta}^* z_t = y_t^* - \left(\widehat{\beta}^* - \beta^*\right) z_t$$

**6:** Based on the bootstrap samples, calculate the test statistics $\tau_p^{d*}$ and $\tau_{gm}^{d*}$ by Equation (2) and (3).

**7:** Repeat Step 2-6 $B$ times, obtaining bootstrap test statistics $\tau_k^{d*}$ where $b = 1, ..., B$ and $k = p$, $gm$, and select the bootstrap critical value $c_\alpha^*$ as

$$c_\alpha^* = \max\left\{c : B^{-1} \sum_{b=1}^{B} 1_{\{\tau_k^{d*} < c\}} \leq \alpha\right\}$$

or equivalently as the $\alpha$-quantile of the ordered $\tau_k^{d*}$ statistics. Reject the null of a unit root if $\tau_k^d$, calculated from Equation (2) if $k = p$ or Equation (3) if $k = gm$, is smaller than $c_\alpha^*$, where $\alpha$ is the significance level of the test.

**Remark 3.** *As mentioned before, the detrending method of Step 1 could be different from the other two. Indeed, the conclusion of Smeekes (2009) and our proof all indicate that the only condition is that* $\widetilde{\beta} - \beta = \left(O_p\left(T^{1/2}\right), O_p\left(T^{-1/2}\right)\right)$, *and OLS and GLS detrending methods all satisfy it.*

**Remark 4.** *In this study, we also find that the residual based block bootstrap algorithm for random walk with non zero drift term can be applied instead of the detrending procedure in Step 1. More specifically, Step 1 and 2 can be replaced by Step 1 of the bootstrap algorithm in Section 3, and the consistency of bootstrap tests can be proven similarly.*

**Remark 5.** *As discussed in Smeekes (2009), we can simply set* $\beta^* = 0$ *in Step 4, since the tests we consider are invariant with respect to the true deterministic components in the (bootstrap) DGP. However, we should emphasize that the detrending procedure in Step 5 is necessary even though we actually did not include the deterministic terms when we construct the bootstrap sample, otherwise the bootstrap test will not be valid any more.*

4.2.2. *Asymptotic validity of bootstrap tests.* In this section, we will theoretically check the validity of bootstrap tests. First, we have the following bootstrap invariance principle.

**Lemma 5.** *Let* $x_t$ *be generated under* $H_0$ *setting* (*A*) *and let Assumption 1 hold. Then as* $T \to \infty$,

$$T^{-1/2} x_{\lfloor Tr \rfloor}^{*d} \xrightarrow{d^*} B^d(r) \text{ in probability.}$$

Lemma 5 indicates that the bootstrap partial sum process indeed correctly reproduces the non bootstrap partial sum. Next, define $u_t^{*d} = \Delta x_t^{*d} = u_t^{d*} - \left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)$. Check the moments of the bootstrap series corresponding to the moments of the original series. The results are summarized in the following lemma.

**Lemma 6.** *Let $x_t$ be generated under $H_0$ setting $(A)$ and let Assumption 1 hold. Then*

$$\begin{aligned}\Omega^{d*} &= T^{-1}\left[E^*\left(\sum_{t=1}^{T}u_t^{*d}\right)\left(\sum_{t=1}^{T}u_t^{*d}\right)' - E^*\left(\sum_{t=1}^{T}u_t^{*d}\right)E^*\left(\sum_{t=1}^{T}u_t^{*d}\right)'\right] \\ &= \Omega^d + o_p(1)\end{aligned}$$

*and*

$$\Omega_0^{d*} = T^{-1}\sum_{t=1}^{T}\left(E^*\left(u_t^{*d}u_t^{*d\prime}\right) - E^*u_t^{*d}E^*u_t^{*d\prime}\right) = \Omega_0^d + o_p(1)$$

Based on Lemma 5 and 6, we can construct the following asymptotic results for the test statistics which are based on bootstrap samples.

**Theorem 2.** *Let $x_t$ be generated under $H_0$ setting $(A)$ and let Assumption 1 hold. Then as $T \to \infty$,*

$$\tau_p^{d*} \xrightarrow{d^*} \frac{\sum_{i=1}^{N}\left[\int_0^1 B_i^d(r)\,dB_i^d(r) + \frac{1}{2}\left(\omega_i^d - \omega_{0,i}^d\right)\right]}{\sum_{i=1}^{N}\int_0^1 B_i^d(r)^2\,dr}\textit{ in probability}$$

*and*

$$\tau_{gm}^{d*} \xrightarrow{d^*} \frac{1}{N}\sum_{i=1}^{N}\frac{\int_0^1 B_i^d(r)\,dB_i^d(r) + \frac{1}{2}\left(\omega_i^d - \omega_{0,i}^d\right)}{\int_0^1 B_i^d(r)^2\,dr}\textit{ in probability.}$$

Finally, Theorem 1 and 2 jointly imply that the bootstrap tests are consistent asymptotically.

## 5. Simulation studies

5.1. **Monte Carlo design.** We consider the following DGP for the simulation study,

$$x_t = d_t + y_t,$$

where $d_t$ is the deterministic component and $y_t$ is the stochastic component. The deterministic component $d_t = (d_{1t}, d_{2t}, \ldots, d_{Nt})$ has the form $d_{it} = a_i + b_i t$, $i = 1, 2, \ldots, N$, where the parameters $a_i$ and $b_i$ are randomly chosen from $a_i \sim U[2,4]$, $b_i \sim U[0.25, 0.75]$. The stochastic component $y_t$ is of the form

$$y_t = \Lambda F_t + w_t,$$

which is generated the same way as in PSU. To be specific,

$$F_t = \phi F_{t-1} + f_t,$$
$$w_{i,t} = \theta_i w_{i,t-1} + v_{i,t}.$$

Furthermore,

$$v_t = A_1 v_{t-1} + \varepsilon_{1,t} + B_1\varepsilon_{1,t-1},$$
$$f_t = \alpha_2 f_{t-1} + \varepsilon_{2,t} + \beta_2\varepsilon_{2,t-1},$$

where $\varepsilon_{2,t} \sim N(0,1)$, and $\varepsilon_{1,t} \sim N(0,\Sigma)$ with $\Sigma$ generated by the following,

1. Generate an $N \times N$ matrix $U$ with each element from $U[0,1]$. Construct $H = U(U'U)^{-\frac{1}{2}}$.
2. Generate $N$ eigenvalues $\zeta_1, \ldots, \zeta_N$ with $\zeta_1 = r$, $\zeta_N = 1$ and $\zeta_i \sim U[r,1]$ for $i = 2, \ldots, N-1$.

3. Let $Z = diag(\zeta_1, \ldots, \zeta_N)$ and $\Sigma = HZH'$.

The contemporaneous dependence is controlled by $r$ and we consider both $r = 1$ (no contemporaneous dependence, $\Sigma = I$) and $r = 0.1$ (contemporaneous dependence, $\Sigma \neq I$). The dynamic dependence is controlled by the parameters $A_1$ and $B_1$ for which we consider two different options,

1. No dynamic dependence: $A_1 = B_1 = 0$.
2. Dynamic autoregressive moving-average cross-sectional dependence: $A_1$ and $B_1$ are non-diagonal.
   As in PSU, we let

$$A_1 = \begin{pmatrix} \xi_1 & \xi_1\eta_1 & \xi_1\eta_1^2 & \cdots & \xi_1\eta_1^{N-1} \\ \xi_2\eta_2 & \xi_2 & \xi_2\eta_2 & \cdots & \xi_2\eta_2^{N-2} \\ \vdots & & \ddots & & \vdots \\ \xi_N\eta_N^{N-1} & \xi_N\eta_N^{N-2} & \xi_N\eta_N^{N-3} & \cdots & \xi_N\eta_N \end{pmatrix}$$

   where $\xi_i, \eta_i \sim U[-0.5, 0.5]$. The moving-average parameter $B_1$ is generated much the same way as $\Sigma$. Let $M = HLH'$ where $H = U(U'U)^{-\frac{1}{2}}$, with $U$ an $N \times N$ matrix of $U[0, 1]$ variables, and $L = diag(l_1, \ldots, l_N)$ with $l_1 = 0.1, l_N = 1$ and $l_2, \ldots, l_{N-1} \sim U[0.1, 1]$. We then let $B_1 = 2M - I_N$. Note that the invertibility is not guaranteed.

The parameters $\alpha_2$ and $\beta_2$ for the common factor are taken in accordance with the setting for the idiosyncratic components. So if $A_1 = B_1 = 0$ then the same holds for $\alpha_2$ and $\beta_2$. If $A_1$ and $B_1$ are non-diagonal, then we let $\alpha_2, \beta_2 \sim U[-0.5, 0.5]$. (Note that PSU does not specify how to generate $\alpha_2$ and $\beta_2$.)

In order to compare with the result in PSU, we consider the same five settings.

I. No common factor, unit root for all idiosyncratic components: $\lambda_i = 0, \theta_i = 1$ for all $i = 1, ..., N$.
II. Unit root in common factor and idiosyncratic components: $\phi = 1, \theta_i = 1$ for all $i = 1, \ldots, N$ and the factor loadings $\lambda_i \sim U[-1, 3]$.
III. Unit root in common factor, stationary idiosyncratic components: $\phi = 1, \theta_i \sim U[0.8, 1]$ and $\lambda_i \sim U[-1, 3]$.
IV. No common factor, stationary idiosyncratic components: $\theta_i \sim U[0.8, 1]$ and $\lambda_i = 0$ for all $i = 1, \ldots, N$.
V. Stationary common factor and idiosyncratic components: $\phi = 0.95$, $\theta_i \sim U[0.8, 1]$ and $\lambda_i \sim U[-1, 3]$.

In accordance with PSU, within each setting, we consider different types of cross-sectional dependence. For all the settings and all types of cross-sectional dependence, we choose the combinations of $T = 25, 50, 100$ and $N = 5, 25, 50$. The results are based on 2000 simulations and the Warp-Speed bootstrap[1] is used to obtain the estimates for the rejection probabilities of the bootstrap tests. The nominal size is 0.05 and the approximate 95% confidence interval for the estimated rejection probability is about (0.042,0.058). In the simulation study the tests to be evaluated are: the conventional method, OLS detrending and GLS detrending, denoted by CON, OLS and GLS respectively. In addition, for the two stage method, the performance of the tests based on a recursive detrending is also assessed, denoted by REC. Recursively adjusting the deterministic terms reduces the downward small sample bias induced by using the full sample of data, see Taylor (2002). At last, the case of mis-specification by using the test statistics from

[1]In a Warp-Speed bootstrap, only one bootstrap replication is drawn for each simulation and the bootstrap distribution is the empirical distribution of the 2000 bootstrap statistics from 2000 simulations.

TABLE 1. Size properties without common factors (setting I).

| | | | | | $\tau_p^\tau$ | | | | | $\tau_{gm}^\tau$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1, B_1, \Sigma$ | $T$ | $N$ | CON | OLS | GLS | REC | MIS | CON | OLS | GLS | REC | MIS |
| Panel A: No short-run dependence | | | | | | | | | | | | |
| $A_1 = 0,$ | 25 | 5 | 0.6 | 5.6 | 6.4 | 0.0 | 0.0 | 0.4 | 3.3 | 3.7 | 0.0 | 0.0 |
| $B_1 = 0,$ | 25 | 25 | 0.0 | 0.6 | 1.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.8 | 0.0 | 0.0 |
| $r = 1$ | 25 | 50 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 50 | 5 | 1.0 | 9.5 | 8.0 | 0.4 | 0.0 | 0.6 | 9.9 | 8.5 | 0.3 | 0.0 |
| | 50 | 25 | 0.0 | 3.4 | 5.2 | 0.0 | 0.0 | 0.0 | 5.5 | 6.0 | 0.0 | 0.0 |
| | 50 | 50 | 0.0 | 2.7 | 1.9 | 0.0 | 0.0 | 0.0 | 3.9 | 3.9 | 0.0 | 0.0 |
| | 100 | 5 | 1.7 | 9.4 | 8.3 | 1.3 | 0.0 | 0.6 | 8.6 | 8.2 | 1.0 | 0.0 |
| | 100 | 25 | 0.0 | 7.5 | 4.2 | 0.0 | 0.0 | 0.0 | 8.8 | 5.8 | 0.0 | 0.0 |
| | 100 | 50 | 0.0 | 4.4 | 3.0 | 0.0 | 0.0 | 0.0 | 7.5 | 5.9 | 0.0 | 0.0 |
| Panel B: Contemporaneous, but no dynamic dependence | | | | | | | | | | | | |
| $A_1 = 0,$ | 25 | 5 | 0.0 | 5.4 | 5.0 | 0.0 | 0.0 | 0.2 | 2.5 | 3.6 | 0.0 | 0.0 |
| $B_1 = 0,$ | 25 | 25 | 0.0 | 0.8 | 1.4 | 0.0 | 0.0 | 0.0 | 0.4 | 0.5 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 0.0 | 0.2 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | 50 | 5 | 1.0 | 9.0 | 9.7 | 0.2 | 0.0 | 0.6 | 7.8 | 7.5 | 0.0 | 0.0 |
| | 50 | 25 | 0.0 | 4.2 | 5.0 | 0.0 | 0.0 | 0.0 | 3.8 | 5.8 | 0.0 | 0.0 |
| | 50 | 50 | 0.0 | 2.9 | 2.8 | 0.0 | 0.0 | 0.0 | 3.8 | 4.0 | 0.0 | 0.0 |
| | 100 | 5 | 0.9 | 9.3 | 7.8 | 0.8 | 0.0 | 0.5 | 8.9 | 7.2 | 0.6 | 0.0 |
| | 100 | 25 | 0.0 | 7.3 | 4.6 | 0.0 | 0.0 | 0.0 | 10.1 | 5.4 | 0.0 | 0.0 |
| | 100 | 50 | 0.0 | 3.4 | 3.4 | 0.0 | 0.0 | 0.0 | 5.7 | 5.0 | 0.0 | 0.0 |
| Panel C: No contemporaneous, but dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi,$ | 25 | 5 | 1.6 | 11.2 | 10.0 | 0.4 | 0.0 | 1.5 | 1.58 | 1.28 | 0.3 | 0.1 |
| $B_1 = \Omega,$ | 25 | 25 | 0.1 | 2.8 | 3.6 | 0.0 | 0.0 | 0.4 | 7.5 | 9.0 | 0.0 | 0.0 |
| $r = 1$ | 25 | 50 | 0.0 | 1.4 | 1.3 | 0.0 | 0.0 | 0.0 | 7.8 | 6.4 | 0.0 | 0.0 |
| | 50 | 5 | 1.8 | 0.8 | 12.6 | 2.4 | 0.0 | 2.5 | 19.0 | 19.4 | 3.3 | 0.0 |
| | 50 | 25 | 0.4 | 5.8 | 6.9 | 0.2 | 0.0 | 0.7 | 24.2 | 28.2 | 1.4 | 0.0 |
| | 50 | 50 | 0.1 | 3.9 | 5.0 | 0.0 | 0.0 | 0.6 | 28.4 | 31.4 | 0.8 | 0.0 |
| | 100 | 5 | 2.2 | 1.0 | 10.1 | 3.2 | 0.0 | 2.8 | 20.2 | 17.0 | 7.8 | 0.0 |
| | 100 | 25 | 0.4 | 9.0 | 7.5 | 1.0 | 0.0 | 1.2 | 33.4 | 29.9 | 6.6 | 0.0 |
| | 100 | 50 | 0.2 | 5.6 | 5.2 | 0.6 | 0.0 | 1.1 | 40.0 | 37.0 | 5.8 | 0.0 |
| Panel D: Contemporaneous and dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi,$ | 25 | 5 | 1.0 | 6.8 | 7.6 | 0.0 | 0.0 | 0.9 | 1.05 | 9.2 | 0.0 | 0.0 |
| $B_1 = \Omega,$ | 25 | 25 | 0.0 | 2.6 | 3.0 | 0.0 | 0.0 | 0.2 | 7.0 | 8.4 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 0.0 | 1.8 | 1.2 | 0.0 | 0.0 | 0.0 | 7.8 | 5.6 | 0.0 | 0.0 |
| | 50 | 5 | 2.3 | 11.6 | 11.4 | 1.4 | 0.0 | 1.5 | 16.4 | 16.8 | 1.6 | 0.0 |
| | 50 | 25 | 0.2 | 5.8 | 6.4 | 0.0 | 0.0 | 0.6 | 22.8 | 24.8 | 0.2 | 0.0 |
| | 50 | 50 | 0.0 | 4.2 | 4.4 | 0.0 | 0.0 | 0.4 | 29.2 | 30.2 | 0.0 | 0.0 |
| | 100 | 5 | 2.6 | 12.4 | 11.5 | 3.4 | 0.0 | 2.2 | 20.3 | 16.6 | 5.4 | 0.0 |
| | 100 | 25 | 0.6 | 7.2 | 6.8 | 1.0 | 0.0 | 1.0 | 32.0 | 27.4 | 4.7 | 0.0 |
| | 100 | 50 | 0.2 | 5.0 | 5.8 | 0.2 | 0.0 | 0.8 | 38.7 | 35.4 | 3.2 | 0.0 |

PSU (developed for DGP without trend) directly to our GDP (with trend), is denoted by MIS. The block length of the bootstrap test was taken as $b = 1.75T^{1/3}$, which amounts to blocks length 6, 7 and 9 for sample sizes 25, 50 and 100 respectively.

5.2. **Monte Carlo result.** The simulation results are shown in Table 1 to 5. It is clear that directly applying the test from PSU to data generated from the model with trend hardly rejects the unit root null hypothesis, whether the DGP is under the null or alternative hypothesis. This indicates that further generalization work is absolutely necessary.

Table 2. Size properties with common factors (setting II).

| $A_1, B_1, \Sigma$ | $T$ | $N$ | CON | OLS | $\tau_p^\tau$ GLS | REC | MIS | CON | OLS | $\tau_{gm}^\tau$ GLS | REC | MIS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: No short-run dependence | | | | | | | | | | | | |
| $A_1 = 0$, | 25 | 5 | 1.8 | 8.8 | 8.2 | 1.0 | 0.1 | 1.2 | 7.4 | 6.4 | 0.5 | 0.5 |
| $B_1 = 0$, | 25 | 25 | 1.4 | 6.8 | 6.4 | 0.4 | 0.0 | 0.5 | 4.3 | 4.3 | 0.2 | 0.1 |
| $r = 1$ | 25 | 50 | 1.1 | 6.2 | 6.7 | 0.5 | 0.0 | 0.3 | 4.0 | 5.3 | 0.1 | 0.0 |
| | 50 | 5 | 2.6 | 8.6 | 9.0 | 2.4 | 0.0 | 1.3 | 8.0 | 7.2 | 1.4 | 0.2 |
| | 50 | 25 | 1.4 | 9.8 | 8.6 | 1.4 | 0.0 | 0.8 | 8.2 | 7.4 | 1.0 | 0.0 |
| | 50 | 50 | 1.8 | 11.8 | 12.1 | 0.8 | 0.0 | 0.7 | 9.5 | 11.0 | 0.4 | 0.0 |
| | 100 | 5 | 2.0 | 8.7 | 8.0 | 3.4 | 0.0 | 1.7 | 8.1 | 8.0 | 2.7 | 0.0 |
| | 100 | 25 | 1.9 | 12.2 | 8.3 | 2.2 | 0.0 | 1.3 | 10.7 | 7.4 | 1.0 | 0.0 |
| | 100 | 50 | 1.4 | 9.7 | 8.7 | 1.8 | 0.0 | 1.0 | 10.9 | 8.6 | 1.0 | 0.0 |
| Panel B: Contemporaneous, but no dynamic dependence | | | | | | | | | | | | |
| $A_1 = 0$, | 25 | 5 | 1.5 | 5.6 | 6.5 | 1.0 | 0.1 | 1.2 | 4.4 | 5.4 | 0.2 | 0.2 |
| $B_1 = 0$, | 25 | 25 | 1.9 | 8.6 | 6.0 | 0.4 | 0.0 | 0.6 | 6.3 | 4.6 | 0.2 | 0.0 |
| $r = 0.1$ | 25 | 50 | 1.8 | 6.6 | 8.8 | 0.3 | 0.0 | 0.8 | 5.2 | 6.6 | 0.2 | 0.0 |
| | 50 | 5 | 2.0 | 9.2 | 9.1 | 2.2 | 0.0 | 1.6 | 10.0 | 8.1 | 1.5 | 0.1 |
| | 50 | 25 | 2.5 | 9.5 | 10.0 | 1.6 | 0.0 | 1.0 | 8.2 | 9.5 | 0.8 | 0.0 |
| | 50 | 50 | 2.2 | 10.0 | 9.2 | 0.8 | 0.0 | 1.1 | 9.9 | 9.8 | 0.2 | 0.0 |
| | 100 | 5 | 3.0 | 8.4 | 8.4 | 3.0 | 0.0 | 1.6 | 8.2 | 8.5 | 1.8 | 0.0 |
| | 100 | 25 | 2.6 | 10.1 | 7.8 | 3.8 | 0.0 | 1.4 | 9.8 | 7.6 | 2.2 | 0.0 |
| | 100 | 50 | 2.2 | 10.9 | 8.2 | 2.4 | 0.0 | 1.1 | 10.8 | 8.0 | 1.5 | 0.0 |
| Panel C: contemporaneous and dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi$, | 25 | 5 | 7.2 | 1.0 | 0.6 | 0.0 | 0.1 | 5.4 | 1.1 | 1.0 | 0.0 | 0.1 |
| $B_1 = \Omega$, | 25 | 25 | 6.0 | 0.7 | 0.8 | 0.0 | 0.0 | 5.0 | 0.6 | 0.5 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 6.6 | 0.6 | 0.9 | 0.0 | 0.0 | 4.7 | 0.4 | 0.6 | 0.0 | 0.0 |
| | 50 | 5 | 7.7 | 1.4 | 0.7 | 0.0 | 0.1 | 6.9 | 2.2 | 1.2 | 0.2 | 0.0 |
| | 50 | 25 | 7.2 | 1.1 | 0.6 | 0.0 | 0.0 | 6.7 | 1.0 | 0.7 | 0.1 | 0.0 |
| | 50 | 50 | 6.2 | 0.9 | 0.6 | 0.0 | 0.0 | 5.9 | 1.0 | 0.8 | 0.0 | 0.0 |
| | 100 | 5 | 6.2 | 2.6 | 1.3 | 0.4 | 0.0 | 6.6 | 3.2 | 3.1 | 0.8 | 0.0 |
| | 100 | 25 | 6.0 | 1.8 | 1.2 | 0.3 | 0.0 | 6.3 | 2.0 | 1.8 | 0.4 | 0.0 |
| | 100 | 50 | 6.2 | 1.5 | 0.9 | 0.2 | 0.0 | 5.8 | 2.0 | 1.3 | 0.1 | 0.0 |

Regarding the empirical size, the simulation results are collected in Table 1 to 3. For the setting without common factors, from Table 1 we can see that in panel A and panel B, OLS and GLS detrending methods for both $\tau_p$ and $\tau_{gm}$ have a good size. However, in the remaining parts (with dynamic dependence), the size of $\tau_p$ is still good, but the size of $\tau_{gm}$ increases with $T$ and the test is largely oversized for large $T$.[2] Both the conventional method and recursive detrending method give a lower size in most cases. Table 2 presents results for the setting with a nonstationary common factor and nonstationary idiosyncratic components. For all of the three types of cross-sectional dependence, OLS and GLS methods still have a satisfactory size. The recursive detrending method gives a lower size in general and tests are largely undersized when the contemporaneous and dynamic dependence are both presented. Furthermore, the conventional method gives a good size only for the group mean statistics and when the most complicated cross-sectional dependence is presented. At last, the results for the setting with cross-unit cointegration, i.e. with a nonstationary common factor and stationary idiosyncratic components are present in Table 3. Under the model without dynamic dependence, the OLS and GLS detrending methods give serious size distortions. This result is consistent with the corresponding result in PSU. As pointed out in PSU, the bootstrap tests do not perform very

[2]This result is very strange since the validity seems violated although we have proved the validity theoretically.

TABLE 3. Size properties with cross-unit cointegration (setting III).

| | | | | | $\tau_p^\tau$ | | | | | $\tau_{gm}^\tau$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1, B_1, \Sigma$ | $T$ | $N$ | CON | OLS | GLS | REC | MIS | CON | OLS | GLS | REC | MIS |
| Panel A: No short-run dependence | | | | | | | | | | | | |
| $A_1 = 0,$ | 25 | 5 | 2.0 | 7.3 | 8.0 | 1.2 | 0.2 | 1.8 | 6.6 | 7.9 | 0.6 | 0.2 |
| $B_1 = 0,$ | 25 | 25 | 1.7 | 7.6 | 10.8 | 1.0 | 0.0 | 0.9 | 7.5 | 8.8 | 0.4 | 0.2 |
| $r = 1$ | 25 | 50 | 1.6 | 10.6 | 9.5 | 1.0 | 0.0 | 0.8 | 9.0 | 7.4 | 0.4 | 0.1 |
| | 50 | 5 | 4.0 | 14.2 | 12.8 | 4.6 | 0.0 | 2.2 | 16.2 | 14.8 | 3.2 | 0.0 |
| | 50 | 25 | 4.2 | 17.7 | 17.9 | 4.6 | 0.0 | 3.2 | 21.2 | 18.7 | 3.0 | 0.0 |
| | 50 | 50 | 4.6 | 18.5 | 17.2 | 5.4 | 0.0 | 2.6 | 20.2 | 19.8 | 3.6 | 0.0 |
| | 100 | 5 | 5.2 | 18.0 | 16.8 | 8.3 | 0.0 | 6.9 | 24.6 | 24.7 | 9.9 | 0.0 |
| | 100 | 25 | 6.4 | 20.2 | 22.4 | 9.8 | 0.0 | 7.2 | 33.4 | 34.0 | 11.2 | 0.0 |
| | 100 | 50 | 6.7 | 22.8 | 23.2 | 10.9 | 0.0 | 7.4 | 35.4 | 34.4 | 13.7 | 0.0 |
| Panel B: Contemporaneous, but no dynamic dependence | | | | | | | | | | | | |
| $A_1 = 0,$ | 25 | 5 | 2.4 | 10.2 | 8.0 | 1.1 | 0.1 | 1.5 | 7.4 | 6.8 | 0.6 | 0.2 |
| $B_1 = 0,$ | 25 | 25 | 2.6 | 9.5 | 12.2 | 1.2 | 0.0 | 1.0 | 7.8 | 10.8 | 0.4 | 0.0 |
| $r = 0.1$ | 25 | 50 | 2.8 | 8.6 | 9.0 | 0.6 | 0.0 | 1.2 | 7.3 | 7.9 | 0.2 | 0.1 |
| | 50 | 5 | 2.8 | 11.4 | 10.3 | 4.4 | 0.1 | 2.0 | 12.4 | 13.2 | 3.0 | 0.1 |
| | 50 | 25 | 3.9 | 12.9 | 13.2 | 3.2 | 0.0 | 2.7 | 15.9 | 17.2 | 2.0 | 0.0 |
| | 50 | 50 | 4.8 | 12.8 | 16.2 | 3.0 | 0.0 | 2.0 | 15.5 | 18.2 | 1.8 | 0.0 |
| | 100 | 5 | 4.8 | 13.4 | 13.2 | 6.3 | 0.0 | 5.2 | 20.2 | 18.4 | 6.2 | 0.0 |
| | 100 | 25 | 4.6 | 16.0 | 16.2 | 7.2 | 0.0 | 5.5 | 23.4 | 25.7 | 7.3 | 0.0 |
| | 100 | 50 | 6.1 | 15.1 | 14.9 | 6.7 | 0.0 | 6.2 | 25.2 | 23.4 | 7.6 | 0.0 |
| Panel C: contemporaneous and dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi,$ | 25 | 5 | 7.9 | 0.8 | 0.6 | 0.0 | 0.2 | 5.0 | 1.2 | 1.2 | 0.0 | 0.0 |
| $B_1 = \Omega,$ | 25 | 25 | 8.5 | 0.8 | 0.4 | 0.0 | 0.0 | 6.4 | 0.6 | 0.5 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 9.4 | 0.4 | 0.3 | 0.0 | 0.0 | 7.5 | 0.8 | 0.5 | 0.0 | 0.0 |
| | 50 | 5 | 9.0 | 1.2 | 0.6 | 0.2 | 0.0 | 8.1 | 7.6 | 3.2 | 0.6 | 0.0 |
| | 50 | 25 | 10.9 | 1.2 | 0.4 | 0.2 | 0.0 | 9.9 | 9.6 | 1.4 | 0.2 | 0.0 |
| | 50 | 50 | 10.9 | 1.0 | 0.6 | 0.1 | 0.0 | 10.8 | 9.8 | 1.4 | 0.1 | 0.0 |
| | 100 | 5 | 9.1 | 2.2 | 1.8 | 0.6 | 0.0 | 11.3 | 6.0 | 7.2 | 2.8 | 0.0 |
| | 100 | 25 | 11.6 | 2.1 | 1.6 | 0.6 | 0.0 | 12.2 | 3.4 | 4.0 | 1.3 | 0.0 |
| | 100 | 50 | 11.7 | 2.1 | 1.4 | 0.8 | 0.0 | 11.6 | 3.8 | 3.3 | 1.4 | 0.0 |

well because for some units the loadings will be close to zero thereby making those units close to stationary. In contrast, the conventional method in this setting does not cause the oversize problem and the recursive detrending method gives undersized tests for small $T$ and the size increases with $T$.

Next, about the power properties, the simulation results are summarized in Table 4 and 5. For the model under the alternative without a common factor, from Table 4 we can see that the powers of the OLS and GLS based tests are satisfactory when $T = 50, 100$ and the power of the GLS based tests are larger than the power of the OLS based tests. However, when $T = 25$, the power is low and decreases with $N$, which is not as expected. The recursive detrending method and the conventional method all have acceptable power only for $T = 100$, and the former is always higher than the latter but less than those for the OLS and GLS methods. Table 5 gives results for the setting with a common factor. Compared with Table 4, all tests give low power and almost no power when both the contemporaneous and dynamic dependence are present in the model. It should be noticed that $\phi = 0.95$ is chosen for the autoregressive parameter in the

Table 4. Power properties without common factors (setting IV).

| | | | | | $\tau_p^\tau$ | | | | | $\tau_{gm}^\tau$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1, B_1, \Sigma$ | $T$ | $N$ | CON | OLS | GLS | REC | MIS | CON | OLS | GLS | REC | MIS |
| Panel A: No short-run dependence | | | | | | | | | | | | |
| $A_1 = 0$, | 25 | 5 | 0.6 | 10.4 | 11.6 | 0.0 | 0.0 | 0.4 | 10.4 | 11.0 | 0.0 | 0.0 |
| $B_1 = 0$, | 25 | 25 | 0.0 | 7.4 | 6.1 | 0.0 | 0.0 | 0.0 | 4.8 | 4.1 | 0.0 | 0.0 |
| $r = 1$ | 25 | 50 | 0.0 | 4.2 | 5.0 | 0.0 | 0.0 | 0.0 | 3.0 | 2.4 | 0.0 | 0.0 |
| | 50 | 5 | 4.0 | 33.0 | 39.0 | 5.8 | 0.0 | 1.8 | 32.8 | 35.4 | 2.9 | 0.0 |
| | 50 | 25 | 2.8 | 69.4 | 78.6 | 4.8 | 0.0 | 0.8 | 74.6 | 82.8 | 1.3 | 0.0 |
| | 50 | 50 | 0.4 | 80.4 | 93.4 | 1.6 | 0.0 | 0.1 | 86.6 | 96.6 | 0.4 | 0.0 |
| | 100 | 5 | 21.6 | 72.3 | 78.0 | 48.0 | 0.0 | 16.2 | 79.2 | 83.8 | 44.5 | 0.0 |
| | 100 | 25 | 49.5 | 99.6 | 99.8 | 93.0 | 0.0 | 49.4 | 100.0 | 100.0 | 96.7 | 0.0 |
| | 100 | 50 | 62.2 | 100.0 | 100.0 | 98.9 | 0.0 | 73.8 | 100.0 | 100.0 | 99.8 | 0.0 |
| Panel B: Contemporaneous, but no dynamic dependence | | | | | | | | | | | | |
| $A_1 = 0$, | 25 | 5 | 0.4 | 12.8 | 10.4 | 0.0 | 0.0 | 0.6 | 10.7 | 9.8 | 0.0 | 0.0 |
| $B_1 = 0$, | 25 | 25 | 0.0 | 7.4 | 7.9 | 0.0 | 0.0 | 0.0 | 6.6 | 5.6 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 0.0 | 3.5 | 4.4 | 0.0 | 0.0 | 0.0 | 2.6 | 2.5 | 0.0 | 0.0 |
| | 50 | 5 | 3.2 | 33.2 | 43.8 | 1.0 | 0.0 | 2.1 | 33.0 | 39.4 | 0.5 | 0.0 |
| | 50 | 25 | 0.8 | 67.9 | 76.3 | 0.2 | 0.0 | 0.8 | 72.6 | 79.0 | 0.0 | 0.0 |
| | 50 | 50 | 0.2 | 80.6 | 91.1 | 0.0 | 0.0 | 0.1 | 86.6 | 94.7 | 0.0 | 0.0 |
| | 100 | 5 | 23.0 | 71.3 | 78.0 | 35.4 | 0.0 | 19.1 | 77.8 | 84.4 | 27.7 | 0.0 |
| | 100 | 25 | 43.0 | 99.2 | 99.8 | 77.8 | 0.0 | 52.2 | 100.0 | 100.0 | 76.6 | 0.0 |
| | 100 | 50 | 61.2 | 100.0 | 100.0 | 92.6 | 0.0 | 65.8 | 100.0 | 100.0 | 94.2 | 0.0 |
| Panel C: contemporaneous and dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi$, | 25 | 5 | 2.9 | 12.8 | 11.6 | 0.8 | 0.0 | 2.2 | 18.0 | 14.0 | 0.4 | 0.0 |
| $B_1 = \Omega$, | 25 | 25 | 0.2 | 9.4 | 8.7 | 0.0 | 0.0 | 0.5 | 21.1 | 20.0 | 0.0 | 0.0 |
| $r = 1$ | 25 | 50 | 0.2 | 5.2 | 4.6 | 0.0 | 0.0 | 0.2 | 19.2 | 16.0 | 0.0 | 0.0 |
| | 50 | 5 | 5.2 | 27.8 | 31.4 | 10.8 | 0.0 | 6.3 | 42.0 | 46.7 | 13.9 | 0.0 |
| | 50 | 25 | 2.8 | 33.8 | 45.2 | 7.6 | 0.0 | 6.4 | 72.2 | 81.3 | 17.6 | 0.0 |
| | 50 | 50 | 2.0 | 32.8 | 46.4 | 7.2 | 0.0 | 5.4 | 80.8 | 89.9 | 22.5 | 0.0 |
| | 100 | 5 | 13.8 | 51.2 | 62.0 | 39.2 | 0.0 | 18.6 | 71.5 | 80.0 | 54.0 | 0.0 |
| | 100 | 25 | 20.0 | 81.4 | 87.8 | 66.0 | 0.0 | 36.0 | 99.2 | 99.7 | 92.7 | 0.0 |
| | 100 | 50 | 22.6 | 90.3 | 96.1 | 76.5 | 0.0 | 47.6 | 99.9 | 100.0 | 99.0 | 0.0 |
| Panel D: Contemporaneous and dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi$, | 25 | 5 | 2.0 | 14.7 | 13.0 | 0.1 | 0.0 | 2.1 | 20.2 | 17.4 | 0.0 | 0.0 |
| $B_1 = \Omega$, | 25 | 25 | 0.1 | 8.0 | 6.8 | 0.0 | 0.0 | 0.3 | 18.6 | 16.0 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 0.0 | 3.6 | 7.0 | 0.0 | 0.0 | 0.2 | 14.0 | 19.6 | 0.0 | 0.0 |
| | 50 | 5 | 5.2 | 27.5 | 28.6 | 5.5 | 0.0 | 6.6 | 41.4 | 41.5 | 6.3 | 0.0 |
| | 50 | 25 | 2.9 | 36.3 | 47.4 | 2.8 | 0.0 | 6.0 | 73.2 | 78.8 | 4.2 | 0.0 |
| | 50 | 50 | 1.8 | 36.4 | 47.3 | 1.8 | 0.0 | 4.6 | 83.5 | 88.7 | 3.0 | 0.0 |
| | 100 | 5 | 13.9 | 52.8 | 58.7 | 34.3 | 0.0 | 16.4 | 72.0 | 78.8 | 42.6 | 0.0 |
| | 100 | 25 | 20.0 | 80.6 | 88.4 | 56.8 | 0.0 | 31.6 | 99.2 | 99.8 | 82.2 | 0.0 |
| | 100 | 50 | 25.7 | 89.6 | 96.0 | 65.4 | 0.0 | 44.7 | 100.0 | 100.0 | 94.4 | 0.0 |

common factor and this makes the test reject the null hypothesis less frequently. [3] This result is in accordance with the result in PSU as well. In addition, the powers of GLS based tests are larger than the powers of OLS based tests.

In summary, comparing the two strategies of dealing with trend, the two stage method with OLS and GLS detrending is better than the conventional method in general. Expect for the low power in Table 5 and oversize in the cross-unit cointegration case, OLS and GLS based

[3] We conjecture that the power in this setting will increase with $T$. However, it is very time consuming if we continue increasing $T$. Instead, we did a simulation by setting $\phi = 0.7$ and see that the power increases to a large extent. We report the result for $\phi = 0.95$ because we want to make it comparable to the results from PSU.

Table 5. Power properties with common factors (setting V).

| | | | | | $\tau_p^\tau$ | | | | | $\tau_{gm}^\tau$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_1, B_1, \Sigma$ | $T$ | $N$ | CON | OLS | GLS | REC | MIS | CON | OLS | GLS | REC | MIS |
| Panel A: No short-run dependence | | | | | | | | | | | | |
| $A_1 = 0$, | 25 | 5 | 2.6 | 7.8 | 8.9 | 1.9 | 0.1 | 1.6 | 7.2 | 8.2 | 0.8 | 0.4 |
| $B_1 = 0$, | 25 | 25 | 1.8 | 8.6 | 12.4 | 1.5 | 0.0 | 0.8 | 7.5 | 9.0 | 0.5 | 0.1 |
| $r = 1$ | 25 | 50 | 2.4 | 10.6 | 10.2 | 1.2 | 0.0 | 0.7 | 9.0 | 8.8 | 0.5 | 0.0 |
| | 50 | 5 | 3.8 | 16.0 | 15.2 | 7.2 | 0.0 | 2.6 | 19.8 | 17.4 | 5.2 | 0.0 |
| | 50 | 25 | 4.1 | 20.4 | 19.6 | 5.7 | 0.0 | 2.4 | 23.2 | 21.8 | 4.4 | 0.0 |
| | 50 | 50 | 4.2 | 19.3 | 21.0 | 8.0 | 0.0 | 2.6 | 23.2 | 22.7 | 5.2 | 0.0 |
| | 100 | 5 | 8.1 | 30.6 | 31.6 | 19.0 | 0.0 | 9.3 | 38.8 | 43.0 | 19.0 | 0.0 |
| | 100 | 25 | 9.8 | 34.3 | 41.8 | 23.1 | 0.0 | 8.6 | 46.7 | 56.5 | 26.5 | 0.0 |
| | 100 | 50 | 11.4 | 36.1 | 42.9 | 25.7 | 0.0 | 10.6 | 52.6 | 58.9 | 31.3 | 0.0 |
| Panel B: Contemporaneous, but no dynamic dependence | | | | | | | | | | | | |
| $A_1 = 0$, | 25 | 5 | 2.6 | 10.5 | 8.0 | 1.6 | 0.2 | 1.2 | 7.9 | 5.6 | 0.8 | 0.4 |
| $B_1 = 0$, | 25 | 25 | 2.8 | 10.6 | 13.3 | 1.4 | 0.0 | 1.0 | 8.2 | 11.1 | 0.6 | 0.0 |
| $r = 0.1$ | 25 | 50 | 2.0 | 9.8 | 10.2 | 1.0 | 0.0 | 1.3 | 8.9 | 8.6 | 0.2 | 0.0 |
| | 50 | 5 | 3.5 | 12.5 | 15.2 | 5.4 | 0.0 | 2.0 | 13.6 | 14.8 | 3.2 | 0.0 |
| | 50 | 25 | 3.8 | 15.0 | 15.8 | 4.1 | 0.0 | 2.8 | 18.2 | 18.6 | 3.6 | 0.0 |
| | 50 | 50 | 3.7 | 15.2 | 18.0 | 4.4 | 0.0 | 2.4 | 18.4 | 19.8 | 3.0 | 0.0 |
| | 100 | 5 | 7.2 | 24.8 | 27.1 | 15.0 | 0.0 | 6.6 | 32.3 | 34.6 | 13.6 | 0.0 |
| | 100 | 25 | 6.8 | 28.0 | 28.2 | 15.2 | 0.0 | 8.2 | 41.2 | 44.1 | 17.0 | 0.0 |
| | 100 | 50 | 6.9 | 28.4 | 26.7 | 16.8 | 0.0 | 8.2 | 41.6 | 41.6 | 18.0 | 0.0 |
| Panel C: contemporaneous and dynamic dependence | | | | | | | | | | | | |
| $A_1 = \Xi$, | 25 | 5 | 0.6 | 0.8 | 0.8 | 0.0 | 0.2 | 0.5 | 1.6 | 1.4 | 0.0 | 0.0 |
| $B_1 = \Omega$, | 25 | 25 | 0.6 | 0.8 | 0.4 | 0.0 | 0.0 | 0.4 | 0.6 | 0.6 | 0.0 | 0.0 |
| $r = 0.1$ | 25 | 50 | 0.6 | 0.9 | 0.6 | 0.0 | 0.0 | 0.2 | 0.7 | 0.5 | 0.0 | 0.0 |
| | 50 | 5 | 1.0 | 1.2 | 0.7 | 0.2 | 0.2 | 1.6 | 3.6 | 4.1 | 0.8 | 0.1 |
| | 50 | 25 | 0.8 | 0.9 | 0.6 | 0.1 | 0.0 | 1.2 | 1.5 | 2.0 | 0.1 | 0.0 |
| | 50 | 50 | 1.0 | 1.2 | 0.6 | 0.1 | 0.0 | 0.8 | 1.6 | 1.7 | 0.1 | 0.0 |
| | 100 | 5 | 1.9 | 2.1 | 2.4 | 1.0 | 0.0 | 3.4 | 7.0 | 9.0 | 3.8 | 0.0 |
| | 100 | 25 | 1.4 | 2.0 | 1.9 | 0.8 | 0.0 | 2.2 | 4.2 | 5.5 | 1.8 | 0.0 |
| | 100 | 50 | 1.4 | 1.4 | 2.0 | 0.8 | 0.0 | 2.2 | 3.6 | 4.8 | 2.0 | 0.0 |

two stage methods can always provide satisfactory sizes and acceptable power. The GLS detrending method works slightly better than OLS. The conventional method usually has good size properties, and especially, it still can provide an acceptable size in the setting of cross-unit cointegration. Compared to the OLS and GLS based two stage method, however, the power of the conventional method is quite low. Furthermore, the recursive detrending method has very low rejection probabilities and low power in most cases.

It is difficult to find an overall winner of all the methods dealing with trend. In general, the test using $\tau_{gm}$ has larger power than the test using $\tau_p$ and the test using the GLS detrending method has the largest power. However, from the size perspective, the test using $\tau_{gm}$ and the GLS detrending method is largely oversized in Panel C and Panel D of setting 1. In contrast, the test using $\tau_p$ and the GLS method does not have this size problem but the power is less that for the test using $\tau_{gm}$.

## 6. Conclusion

In this paper, we have done a generalization for the block bootstrap panel unit root test which is proposed by PSU in the sense of considering the deterministic terms in the model. Two different methods of dealing with the deterministic terms are suggested. For both of them, we theoretically checked the asymptotic validity of bootstrap tests under the main null hypothesis. The consistency of the bootstrap test under the cross sectional cointegration null hypothesis and all the power properties have been investigated by a Monte Carlo simulation. The simulation results show that all tests have acceptable size properties.The tests which are based on detrending methods have better power than the tests which are based on the conventional methods. Furthermore the tests which are based on the GLS detrending method have the best power among all detrending methods. About the difference between pooled and group mean statistics, even though the group mean statistics have better power than pooled in general, we still recommend pooled statistics because of the unexpected results of size of group mean statistics in some cases.

## References

Bai, J. and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration, *Econometrica* **72**: 1127–177.

Breitung, J. and Das J. (2005). Panel unit root tests under cross-sectional dependence, *Statistica Neerlandica* **72**:414–433.

Chang Y. (2002). Nonliear IV unit root tests in panels with cross-sectional dependency *Journal of Econometrics* **110**:261–292.

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

Hamilton, J.D. (1994). Time Series Analysis. *Princeton, New Jersey: Princeton University Press.*

Im, K.S., M.H. Pesaran, and Y. Shin (2003). Testing for Unit Roots in Heterogenous Panels, *Journal of Econometrics* **115**: 53–74.

Levin, A., C. Lin, and C.J. Chu (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties, *Journal of Econometrics* **108**: 1–24.

Moon, R. and P. Perron (2004). Testing for Unit Root in Panels with Dynamic Factors, *Journal of Econometrics* **122**: 81–126.

O'Connell, P. (1998). The Overvaluation of Purchasing Power Parity, *Journal of International Economics* **44**: 1–19.

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

Paparoditis, E. and Politis, D.N. (2003). Residual-based block bootstrap for unit root testing, *Econometrica* **71**:813–855.

Pesaran, M.H. (2005). A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, *Journal of Applied Econometrics***22**: 265-312

Phillips, P.C.B. and H.R. Moon. (1999). Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica* **67**: 1057–111.

Smeekes, S. (2009). Detrending bootstrap unit root tests, *METEOR Research Memorandum 09/056. Maastricht University.*

Taylor, R. (2002). Regression-based unit root tests with recursive mean adjustment for seasonal and nonseasional time series *Journal of Business and Economic Statisitics* vol.20m issue2, pages 269-81

## Appendix

**Proof of Lemma 1.** About the limiting distribution of $\tau_p^\mu$ and $\tau_{gm}^\mu$: Under the null hypothesis setting A, we have $\Delta x_t = u_t$, where $u_t = \Delta y_t = \Gamma' \psi(L) \varepsilon_t$. Then $x_t = \beta_1 + \widetilde{u}_t$ where $\widetilde{u}_t = \sum_{j=1}^t u_j$. Then we have

$$\begin{aligned} & T^{-1} \sum_{t=1}^T (x_{i,t-1} - \overline{x}_{i,-1})(\Delta x_{it} - \Delta \overline{x}_i) \\ = \ & T^{-1} \sum_{t=1}^T \frac{\widetilde{u}_{i,t-1}}{\sqrt{T}} \sqrt{T} u_{it} - T^{-1} \sum_{t=1}^T \frac{\widetilde{u}_{i,t-1}}{\sqrt{T}} \sqrt{T} \frac{1}{T} \sum_{t=1}^T \Delta u_{it} \\ = \ & \int B_i(r)\, dB_i(r) - B_i(1) \int B_i(r)\, dr + \frac{1}{2}(\omega_{0i} + \omega_i) + o_p(1) \end{aligned}$$

$$\begin{aligned} & T^{-2} \sum_{t=1}^T (x_{i,t-1} - \overline{x}_{i,-1})^2 \\ = \ & T^{-1} \sum_{t=1}^T \left( \frac{\widetilde{u}_{i,t-1}}{\sqrt{T}} - \frac{\overline{\widetilde{u}}_{i,-1}}{\sqrt{T}} \right)^2 = \int B_i(r)^2\, dr - \left( \int B_i(r)\, dr \right)^2 + o_p(1) \end{aligned}$$

which follows from Lemma 1 and Lemma A.1. of PSU. Then the limiting distribution of $\tau_p^\mu$ and $\tau_{gm}^\mu$ can be derived. About the limiting distribution of $\tau_p^\tau$ and $\tau_{gm}^\tau$: under the main null hypothesis, $\Delta x_t = \beta_{2\cdot} + u_t$ and $x_t = \beta_{2\cdot} t + \widetilde{u}_t$. Then

$$x_{i,t-1}^\dagger \doteq \beta_{2,i} \left( t - \frac{T}{2} \right) + \left( \widetilde{u}_{i,t-1} - \overline{\widetilde{u}}_{i,-1} \right)$$

and

$$\Delta x_{it}^\dagger = u_{it} - \overline{u}_i$$

Then we have

$$\begin{aligned} & \sum_{t=1}^T \Delta x_{it}^\dagger x_{i,t-1}^\dagger \sum_{t=1}^T t^{\dagger 2} \\ \doteq \ & \frac{T^3}{12} \beta_{2,i} \sum_{t=1}^T \left( t - \frac{T}{2} \right)(u_{it} - \overline{u}_i) + \frac{T^3}{12} \sum_{t=1}^T \left( \widetilde{u}_{i,t-1} - \overline{\widetilde{u}}_{i,-1} \right)(u_{it} - \overline{u}_i) \end{aligned}$$

$$\begin{aligned} & \sum_{t=1}^T \Delta x_{it}^\dagger t^\dagger \sum_{t=1}^T x_{i,t-1}^\dagger t^\dagger \\ \doteq \ & \frac{T^3}{12} \beta_{2,i} \sum_{t=1}^T \left( t - \frac{T}{2} \right)(u_{it} - \overline{u}_i) + \sum_{t=1}^T (u_{it} - \overline{u}_i) \left( t - \frac{T}{2} \right) \sum_{t=1}^T \left( \widetilde{u}_{i,t-1} - \overline{\widetilde{u}}_{i,-1} \right) \left( t - \frac{T}{2} \right) \end{aligned}$$

Thus

$$\begin{aligned} & \sum_{t=1}^T \Delta x_{it}^\dagger x_{i,t-1}^\dagger \sum_{t=1}^T t^{\dagger 2} - \sum_{t=1}^T \Delta x_{it}^\dagger t^\dagger \sum_{t=1}^T x_{i,t-1}^\dagger t^\dagger \\ \doteq \ & \frac{T^3}{12} \sum_{t=1}^T \left( \widetilde{u}_{i,t-1} - \overline{\widetilde{u}}_{i,-1} \right)(u_{it} - \overline{u}_i) - \sum_{t=1}^T (u_{it} - \overline{u}_i) \left( t - \frac{T}{2} \right) \sum_{t=1}^T \left( \widetilde{u}_{i,t-1} - \overline{\widetilde{u}}_{i,-1} \right) \left( t - \frac{T}{2} \right) \end{aligned}$$

(Or we can say that the OLS estimation of the autoregressive coefficient is invariant with $\beta_{2,i}$, thus without loss generality, we set $\beta_{2,i} = 0$.) Then we have

$$\begin{aligned} & T^{-4} \sum_{t=1}^{T} \Delta x_{it}^{\dagger} x_{i,t-1}^{\dagger} \sum_{t=1}^{T} t^{\dagger 2} - \sum_{t=1}^{T} \Delta x_{it}^{\dagger} t^{\dagger} \sum_{t=1}^{T} x_{i,t-1}^{\dagger} t^{\dagger} \\ = \; & \frac{1}{12} \left( \int B_i dB_i - B_i(1) \int B_i + \frac{1}{2} (\omega_{0i} + \omega_i) \right) - \left( \frac{1}{2} B_i(1) - \int B_i \right) \left( \int r B_i - \frac{1}{2} \int B_i \right) + o_p(1) \end{aligned}$$

Similiary we have

$$\begin{aligned} & T^{-4} \left( \sum_{t=1}^{T} x_{i,t-1}^{\dagger 2} \sum_{t=1}^{T} t^{\dagger 2} - \left( \sum_{t=1}^{T} x_{i,t-1}^{\dagger} t^{\dagger} \right)^2 \right) \\ = \; & \frac{1}{12} \left( \int B_i(r)^2 dr - \left( \int B_i(r) dr \right)^2 \right) - \left( \int r B_i(r) dr - \frac{1}{2} \int B_i(r) dr \right)^2 + o_p(1) \end{aligned}$$

Then the limiting distribution of $\tau_p^{\tau}$ and $\tau_{gm}^{\tau}$ can be derived. □

**Proof of Lemma 2.** For model 0 and 1, the bootstrap invariance principle has been discussed in Palm et al. (2011). Here we investigate it for model 2. Under $H_0$ setting $(A)$, we can write

$$\Delta x_t = \beta_{2\cdot} + \Delta y_t = \beta_{2\cdot} + u_t$$

and

$$x_t = \beta_{2\cdot} t + \sum_{j=1}^{t} u_j = \beta_{2\cdot} t + \widetilde{u}_t$$

Note that

$$S_T^*(r) = T^{-1/2} \sum_{t=1}^{\lfloor Tr \rfloor} u_t^* = T^{-1/2} x_1 + T^{-1/2} \sum_{m=0}^{M_r - 1} \sum_{s=1}^{b} \widehat{u}_{i_m + s} + T^{-1/2} \sum_{s=1}^{N_r} \widehat{u}_{i_{M_r} + s}$$

where $M_r = \lfloor (\lfloor Tr \rfloor - 2)/b \rfloor$ and $N_r = \lfloor Tr \rfloor - M_r b - 1$. As $T^{-1/2} x_1 = O_p\left(T^{-1/2}\right)$, we write

$$S_T^*(r) = \underbrace{T^{-1/2} \sum_{m=0}^{M_r} \sum_{s=1}^{b} \widehat{u}_{i_m + s}}_{A_T} - \underbrace{T^{-1/2} \sum_{s=N_r+1}^{b} \widehat{u}_{i_{M_r} + s}}_{B_T} + O_p\left(T^{-1/2}\right) \tag{9}$$

Note that $B_T$ is a $N \times 1$ vector. For each $i = 1, ..., N$, the $i$th element of $B_T$ can be written as

$$\begin{aligned} & T^{-1/2} \sum_{s=N_r+1}^{b} \widehat{u}_{i, i_{M_r} + s} \\ = \; & T^{-1/2} \sum_{s=N_r+1}^{b} \left( \left( x_{i, i_{M_r} + s} - \widetilde{\rho}_i x_{i, i_{M_r} + s - 1} \right) - \frac{1}{T-1} \sum_{t=2}^{T} \left( x_{it} - \widetilde{\rho}_i x_{i,t-1} \right) \right) \\ = \; & \underbrace{T^{-1/2} \sum_{s=N_r+1}^{b} \left( u_{i, i_{M_r} + s} - \frac{1}{T-1} \sum_{t=2}^{T} u_{it} \right)}_{B_{1T}} - \underbrace{(\widetilde{\rho}_i - 1) T^{-1/2} \sum_{s=N_r+1}^{b} \left( x_{i, i_{M_r} + s - 1} - \frac{1}{T-1} \sum_{t=2}^{T} x_{i,t-1} \right)}_{B_{2T}} \end{aligned}$$

where

$$B_{1T} = T^{-1/2} \sum_{s=N_r+1}^{b} u_{i,i_{M_r}+s} - \frac{b-N_r}{(T-1)\,T^{1/2}} \sum_{t=2}^{T} u_{it}$$

Since the first term is in order $O_p\left(\sqrt{b/T}\right) = O_p\left(k^{-1/2}\right)$ and the second term is in order $O_p(b/T) = O_p\left(k^{-1}\right)$, $B_{1T} = O_p\left(k^{-1/2}\right)$. And

$$x_{it} = \beta_{2i} t + \widetilde{u}_{it}$$

$$\begin{aligned} & B_{2T} \\ = \; & (\widetilde{\rho}_i - 1)\, T^{-1/2} \sum_{s=N_r+1}^{b} \left( x_{i,i_{M_r}+s-1} - \frac{1}{T-1} \sum_{t=2}^{T} x_{i,t-1} \right) \\ = \; & (\widetilde{\rho}_i - 1) \left( \underbrace{\beta_{2i} T^{-1/2} \sum_{s=N_r+1}^{b} (i_{M_r} + s)}_{O_p\left(b^2 T^{-1/2}\right)} + \underbrace{T^{-1/2} \sum_{s=N_r+1}^{b} \widetilde{u}_{i,i_{M_r}+s-1}}_{O_p\left(b T^{-1/2}\right)} - \underbrace{\frac{b-N_r}{(T-1)\,T^{1/2}} \sum_{t=2}^{T} (\beta_{2i} t + \widetilde{u}_{i,t-1})}_{O_p\left(b T^{1/2}\right)} \right) \end{aligned}$$

In this case, $\widetilde{\rho}_i$ is the OLS estimate of the autoregressive coefficient of the regression of $x_{it}$ on $(1, x_{i,t-1})'$, thus $\widetilde{\rho}_i - 1 = O_p\left(T^{-3/2}\right)$ , see Hamilton (1994). Assuming $b/\sqrt{T} = o(1)$, we have that $B_{2T} = O_p(b/T) = O_p\left(k^{-1}\right)$. Thus we can get $B_T = O_p\left(k^{-1/2}\right)$. Now focus on $A_T$: for each $i$, the $i$th element of $A_T$ can be written as

$$\begin{aligned} & T^{-1/2} \sum_{m=0}^{M_r} \sum_{s=1}^{b} \widehat{u}_{i,i_m+s} \\ = \; & T^{-1/2} \sum_{m=0}^{M_r} \sum_{s=1}^{b} \left( u_{i,i_m+s} - \frac{1}{T-1} \sum_{t=2}^{T} u_{it} \right) - (\widetilde{\rho}_i - 1)\, T^{-1/2} \sum_{m=0}^{M_r} \sum_{s=1}^{b} \left( x_{i,i_m+s-1} - \frac{1}{T-1} \sum_{t=2}^{T} x_{i,t-1} \right) \end{aligned}$$

By the proof of PP, we have that the second term converge to 0 uniformly in $r$. At last, by the Lemma A.5 in PSU, this lemma can be proved. □

**Proof of Lemma 3.** Note that under the null hypothesis setting (A), we have

$$T^{-1/2} x^d_{\lfloor Tr \rfloor} = T^{-1/2} \sum_{j=1}^{\lfloor Tr \rfloor} u_j - T^{-1/2} \left( \widehat{\beta} - \beta \right) z_{\lfloor Tr \rfloor}$$

By Lemma 1 in PSU we have that

$$T^{-1/2} \sum_{j=1}^{\lfloor Tr \rfloor} u_j \xrightarrow{d} B(r) = \Gamma' \Psi(1)\, W(r) . \tag{10}$$

Now we focus on the second term, since $\widehat{\beta}$ is the OLS estimate of $\beta$, then we have

$$T^{-1/2} \left( \widehat{\beta} - \beta \right) z_{\lfloor Tr \rfloor} = T^{-1/2} \left( \sum_{t=1}^{T} y_t z_t' \right) \left( \sum_{t=1}^{T} z_t z_t' \right)^{-1} \begin{pmatrix} 1 \\ \lfloor Tr \rfloor \end{pmatrix}$$

where define

$$\Upsilon_{0T}^{-1} = \left( \sum_{t=1}^{T} z_t z_t' \right)^{-1} = \begin{pmatrix} T & T(T+1)/2 \\ T(T+1)/2 & T(T+1)(2T+1)/6 \end{pmatrix}^{-1}$$

Then

$$\begin{aligned} & T^{-1/2}\left(\widehat{\beta}-\beta\right) z_{\lfloor Tr \rfloor} \\ = & \left( T^{-1} \sum_{t=1}^{T} \frac{y_t}{\sqrt{T}}, T^{-1} \sum_{t=1}^{T} \frac{t}{T} \frac{y_t}{\sqrt{T}} \right) \left( T \Upsilon_{1T} \Upsilon_{0T}^{-1} \Upsilon_{1T} \right) \begin{pmatrix} 1 \\ \lfloor Tr \rfloor / T \end{pmatrix} \end{aligned}$$

where $\Upsilon_{1T} = diag\,(1, T)$, and

$$\begin{aligned} & T \Upsilon_{1T} \Upsilon_{0T}^{-1} \Upsilon_{1T} = \left( T^{-1} \Upsilon_{1T}^{-1} \Upsilon_{0T} \Upsilon_{1T}^{-1} \right)^{-1} \\ = & \begin{pmatrix} 1 & \frac{1}{2}\left(1 - T^{-1}\right) \\ \frac{1}{2}\left(1 - T^{-1}\right) & \left(\frac{T}{3} + \frac{1}{6}\right)(T+1) T^{-2} \end{pmatrix}^{-1} = \left( \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} + o(1) \right)^{-1} \end{aligned}$$

Then by Equation (10) we have

$$T^{-1/2}\left(\widehat{\beta}-\beta\right) z_{\lfloor Tr \rfloor} \xrightarrow{d} \left( \int_0^1 B(s)\, ds, \int_0^1 sB(s)\, ds \right) \begin{pmatrix} 4 - 6r \\ 12r - 6 \end{pmatrix} \tag{11}$$

At last, associate Equation (10) and (11)

$$T^{-1/2} x_{\lfloor Tr \rfloor}^{d} = B^{d}(r) + o_p(1)$$

where $B^d(r) = B(r) - \left( \int_0^1 B(s)\, ds, \int_0^1 sB(s)\, ds \right) \begin{pmatrix} 4 - 6r \\ 12r - 6 \end{pmatrix}$. □

**Proof of Lemma 4.** First, we check $\Omega^d$

$$\Omega^d = \lim_{T \to \infty} T^{-1} E \left( \sum_{t=1}^{T} u_t^d \right) \left( \sum_{t=1}^{T} u_t^d \right)' = \lim_{T \to \infty} E \left( T^{-1} \sum_{r,s=1}^{T} u_r^d u_s^{d\prime} \right)$$

where

$$\begin{aligned} T^{-1} \sum_{r,s=1}^{T} u_r^d u_s^{d\prime} = & \; T^{-1} \sum_{r,s=1}^{T} u_r u_s' - \sum_{r=1}^{T} u_r \left( \widehat{\beta}_{2\cdot} - \beta_{2\cdot} \right)' - \sum_{s=1}^{T} \left( \widehat{\beta}_{2\cdot} - \beta_{2\cdot} \right) u_s' \\ & + T \left( \widehat{\beta}_{2\cdot} - \beta_{2\cdot} \right) \left( \widehat{\beta}_{2\cdot} - \beta_{2\cdot} \right)' \end{aligned}$$

By Lemma A.1 in PSU we have that $\lim_{T\to\infty} E\left(T^{-1}\sum_{r,s=1}^{T} u_r u_s'\right) = \Omega$. About the limit value of the second term of $T^{-1}\sum_{r,s=1}^{T} u_r^d u_s^{d\prime}$, we have

$$\begin{aligned}
&\lim_{T\to\infty} E\left(\sum_{r=1}^{T} u_r\left(\widehat{\beta}_{2\cdot}-\beta_{2\cdot}\right)'\right)\\
=\;& 12\lim_{T\to\infty} E\left(\sum_{r=1}^{T} u_r T^{-3}\sum_{t=1}^{T}(t-T/2)\,y_t'\right)\\
=\;& 12\lim_{T\to\infty} T^{-3}E\left(\sum_{r,t=1}^{T}\left(t-\frac{T}{2}\right)\left(\Gamma'\sum_{i=0}^{\infty}\psi_i\varepsilon_{r-i}\right)\sum_{v=1}^{t}\left(\sum_{j=0}^{\infty}\varepsilon_{v-j}'\psi_j'\Gamma\right)\right)\\
=\;& 12\sum_{i,j=0}^{\infty}\Gamma'\psi_i\lim_{T\to\infty}T^{-3}\sum_{t=1}^{T}\sum_{r=1}^{T}\sum_{v=1}^{t}\left(t-\frac{T}{2}\right)E\varepsilon_{r-i}\varepsilon_{v-j}'\psi_j'\Gamma
\end{aligned}$$

Define $A_T = T^{-3}\sum_{t=1}^{T}\sum_{r=1}^{T}\sum_{v=1}^{t}\left(t-\frac{T}{2}\right)E\varepsilon_{r-i}\varepsilon_{v-j}'$. For fixed $i$ and $j$, suppose $i > j$, there exists a large $T_0$ s.t. $T > i-j$ when $T > T_0$. Now assume $T > T_0$, if $t > i-j$ then there are $t-(i-j)$ cases s.t. $r-i = v-j$ in all the combinations of $i, j, r$ and $v$, and so we have

$$A_T = T^{-3}\sum_{t=i-j+1}^{T}\left(t-\frac{T}{2}\right)(t-(i-j))\,\Sigma$$

Then we have

$$\begin{aligned}
&\lim_{T\to\infty} E\left(T^{-1}\sum_{r,s=1}^{T} u_r\left(\widehat{\beta}_{2\cdot}-\beta_{2\cdot}\right)'\right)\\
=\;& 12\left(\sum_{i,j=0}^{\infty}\Gamma'\psi_i\Sigma\psi_j'\Gamma\right)\left(\lim_{T\to\infty}T^{-3}\sum_{t=i-j+1}^{T}\left(t-\frac{T}{2}\right)(t-(i-j))\right)\\
=\;& \sum_{i,j=0}^{\infty}\Gamma'\psi_i\Sigma\psi_j'\Gamma = \Omega
\end{aligned}$$

About the last term of $T^{-1}\sum_{r,s=1}^{T} u_r^d u_s^{d\prime}$

$$\begin{aligned}
&\lim_{T\to\infty} E\left(T\left(\widehat{\beta}_{2\cdot}-\beta_{2\cdot}\right)\left(\widehat{\beta}_{2\cdot}-\beta_{2\cdot}\right)'\right)\\
=\;& 12^2\lim_{T\to\infty}E\left(T^{-5}\sum_{s,t=1}^{T}\left(s-\frac{T}{2}\right)\left(t-\frac{T}{2}\right)y_s y_t'\right)\\
=\;& 12^2\lim_{T\to\infty}E\left(T^{-5}\sum_{s,t=1}^{T}\left(s-\frac{T}{2}\right)\left(t-\frac{T}{2}\right)\sum_{k=1}^{s}\sum_{l=1}^{t}\left(\Gamma'\sum_{i=0}^{\infty}\psi_k\varepsilon_{i-k}\right)\left(\Gamma'\sum_{j=0}^{\infty}\psi_l\varepsilon_{j-l}\right)'\right)\\
=\;& 12^2\sum_{i,j=0}^{\infty}\Gamma'\psi_k\lim_{T\to\infty}T^{-5}\sum_{s,t=1}^{T}\left(s-\frac{T}{2}\right)\left(t-\frac{T}{2}\right)\sum_{k=1}^{s}\sum_{l=1}^{t}E\varepsilon_{i-k}\varepsilon_{j-l}'\psi_l'\Gamma
\end{aligned}$$

Define $B_T = \sum_{s,t=1}^{T} \left(s - \frac{T}{2}\right)\left(t - \frac{T}{2}\right) \sum_{k=1}^{s} \sum_{l=1}^{t} E\varepsilon_{i-k}\varepsilon'_{j-l}$. Again, For fixed $i$ and $j$, suppose $i > j$, there should exist a large $T_0$ s.t. $T > i - j$ when $T > T_0$. Now assume $T > T_0$, if $t > i - j$ and $s > i - j$ then there should be $\min(s,t) - (i-j)$ cases s.t. $i - k = j - l$ in all the combinations of $i, j, k$ and $l$, then we have

$$B_T = \Sigma \sum_{s,t=i-j+1}^{T} \left(s - \frac{T}{2}\right)\left(t - \frac{T}{2}\right)(\min(s,t) - (i-j))$$

Then we have

$$\begin{aligned}
&\lim_{T\to\infty} E\left(T^{-1}\sum_{r,s=1}^{T}\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)'\right) \\
= \;& 12^2\left(\sum_{k,l=0}^{\infty}\Gamma'\psi_k\Sigma\psi_l'\Gamma\right)\left(\lim_{T\to\infty} T^{-5}\sum_{s,t=i-j+1}^{T}\left(s - \frac{T}{2}\right)\left(t - \frac{T}{2}\right)(\min(s,t) - (i-j))\right) \\
= \;& \frac{6}{5}\sum_{k,l=0}^{\infty}\Gamma'\psi_k\Sigma\psi_l'\Gamma = \frac{6}{5}\Omega
\end{aligned}$$

Then we have

$$\Omega^d = \Omega/5$$

Next, check $\Omega_0^d$

$$\Omega_0^d = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E\left(u_t^d u_t^{d\prime}\right)$$

Since

$$\begin{aligned}
u_t^d u_t^{d\prime} &= \left(u_t - \left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)\right)\left(u_t' - \left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)'\right) \\
&= u_t u_t' - u_t\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)' - \left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)u_t' + \left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)'
\end{aligned}$$

By Lemma A.1 in PSU we have that $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} Eu_t u_t' = \Omega_0$. About the second term

$$\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E\left(u_t\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)'\right) = 12\lim_{T\to\infty} E\left(T^{-4}\sum_{t=1}^{T} u_t \sum_{t=1}^{T}(t - T/2)\, y_t'\right)$$

It is easy to see that $\sum_{t=1}^{T} u_t \sum_{t=1}^{T}(t - T/2)\, y_t' = O_p\left(T^{-3}\right)$, i.e. $T^{-1}\sum_{t=1}^{T} u_t\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)'$ converges to 0 in probability. Then we can apply Lebesgue dominated convergence theorem to change the order of limit and expectation, and have that

$$\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E\left(u_t\left(\widehat{\beta}_{2\cdot} - \beta_{2\cdot}\right)'\right) = 0$$

By the similar arguments, we have that the last two terms both converge to 0, and thus we have $\Omega_0^d = \Omega_0$ □

**Proof of Theorem 1.** By the results of Lemma 3 and 4, and the continuous mapping theorem we have

$$T^{-1}\sum_{t=1}^{T} x_{i,t-1}^{d}\Delta x_{i,t}^{d} \xrightarrow{d} \int_0^1 B_i^d(r)\, dB_i^d(r) + \frac{1}{2}\left(\omega_i^d - \omega_{0,i}^d\right)$$

and

$$T^{-1}\sum_{t=1}^{T} x_{i,t-1}^{d2} \xrightarrow{d} \int_0^1 B_i^d(r)^2\, dr.$$

Then applying the Cramér-Wold device, the convergence also holds jointly. Thus Theorem 1 can be proved immediately. □

**Proof of Lemma 5**. First, under the null hypothesis setting (A) we have

$$\begin{aligned}
T^{-1/2}x_t^{*d} &= T^{-1/2}y_t^{*} - T^{-1/2}\left(\widehat{\beta}^{*} - \beta^{*}\right)z_t \\
&= \underbrace{T^{-1/2}\left(x_1^d + \sum_{j=2}^{t} u_j^{d*}\right)}_{A_t} - \underbrace{T^{-1/2}\left(\widehat{\beta}^{*} - \beta^{*}\right)z_t}_{B_t}
\end{aligned}$$

About $A_t$, for some $r \in (0,1)$, it can be written as

$$\begin{aligned}
A_{\lfloor Tr \rfloor} &= T^{-1/2}x_1^d + T^{-1/2}\sum_{m=0}^{M_r-1}\sum_{s=1}^{b}\widehat{u}_{i_m+s}^{d} + T^{-1/2}\sum_{s=1}^{N_r}\widehat{u}_{i_{M_r}+s}^{d} \\
&= \underbrace{T^{-1/2}\sum_{m=0}^{M_r}\sum_{s=1}^{b}\widehat{u}_{i_m+s}^{d}}_{A_{1t}} - \underbrace{T^{-1/2}\sum_{s=N_r+1}^{b}\widehat{u}_{i_{M_r}+s}^{d}}_{A_{2t}} + O_p\left(T^{-1/2}\right)
\end{aligned}$$

where $M_r = \lfloor(\lfloor Tr \rfloor - 2)/b\rfloor$ and $N_r = \lfloor Tr \rfloor - M_r b - 1$. Note that $A_t$ and $B_t$ are both $N \times 1$ vectors. For $i = 1, ..., N$, each element of $A_{2t}$ can be expressed as

$$
\begin{aligned}
& T^{-1/2}\sum_{s=N_r+1}^{b}\left(\left(x^d_{i,i_{M_r}+s}-\widehat{\rho}_i x^d_{i,i_{M_r}+s-1}\right)-\frac{1}{T-1}\sum_{t=2}^{T}\left(x^d_{i,t}-\widehat{\rho}_i x^d_{i,t}\right)\right)\\
= & T^{-1/2}\sum_{s=N_r+1}^{b}\left(\Delta x^d_{i,i_{M_r}+s}-\frac{1}{T-1}\sum_{t=2}^{T}\Delta x^d_{i,t}\right)\\
& -\left(\widehat{\rho}_i-1\right)T^{-1/2}\sum_{s=N_r+1}^{b}\left(x^d_{i,i_{M_r}+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}x^d_{i,t-1}\right)\\
= & T^{-1/2}\sum_{s=N_r+1}^{b}\left(u_{i,i_{M_r}+s}-\frac{1}{T-1}\sum_{t=2}^{T}u_{i,t}\right)\\
& -\left(\widehat{\rho}_i-1\right)T^{-1/2}\sum_{s=N_r+1}^{b}\left(y_{i,i_{M_r}+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}y_{i,t-1}\right)\\
& -\left(\widehat{\rho}_i-1\right)\left(\widetilde{\beta}_i-\beta_i\right)'T^{-1/2}\sum_{s=N_r+1}^{b}\left(z_{i_{M_r}+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}z_{t-1}\right)
\end{aligned} \tag{12}
$$

About the first term of Equation (12)

$$
\begin{aligned}
& T^{-1/2}\sum_{s=N_r+1}^{b}\left(u_{i,i_{M_r}+s}-\frac{1}{T-1}\sum_{t=2}^{T}u_{i,t}\right)\\
= & \underbrace{T^{-1/2}\sum_{s=N_r+1}^{b}u_{i,i_{M_r}+s}}_{O_p\left(\sqrt{b/T}\right)=O_p\left(k^{-1/2}\right)}-\underbrace{\frac{b-N_r}{(T-1)\,T^{1/2}}\sum_{t=2}^{T}u_{i,t}}_{O_p(b/T)=O_p(k^{-1})}
\end{aligned}
$$

Thus the first term of $A_{2it}$ is $O_p\left(k^{-1/2}\right)$. About the second term of $A_{2it}$, since Theorem 1 indicates that $\widehat{\rho}_i-1=O_p\left(T^{-1}\right)$, then

$$
\begin{aligned}
& \left(\widehat{\rho}_i-1\right)T^{-1/2}\sum_{s=N_r+1}^{b}\left(y_{i,i_{M_r}+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}y_{i,t-1}\right)\\
= & \underbrace{\left(\widehat{\rho}_i-1\right)}_{O_p(T^{-1})}\left(\underbrace{T^{-1/2}\sum_{s=N_r+1}^{b}y_{i,i_{M_r}+s-1}}_{O_p(b)}-\underbrace{\frac{b-N_r}{(T-1)\,T^{1/2}}\sum_{t=2}^{T}y_{i,t-1}}_{O_p(b)}\right)
\end{aligned}
$$

where $b\pm N_r\doteq b$. Then the second term of $A_{2it}$ is in order $O_p\left(k^{-1}\right)$. Furthermore, as discussed in Remark 3, $\widetilde{\beta}-\beta=\left(O_p\left(T^{1/2}\right),O_p\left(T^{-1/2}\right)\right)$ and

$$\begin{aligned}
&\sum_{s=N_r+1}^{b}\left(z_{i_{M_r}+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}z_{t-1}\right)\\
=\ &\sum_{s=N_r+1}^{b}\left(\left(\begin{array}{c}1\\ i_{M_r}+s-1\end{array}\right)-\left(\begin{array}{c}1\\ T/2\end{array}\right)\right)=\left(\begin{array}{c}0\\ i_{M_r}+\left(N_r+b-1\right)/2-T/2\end{array}\right)\left(b-N_r\right)\\
=\ &\left(\begin{array}{c}0\\ O_p\left(T\right)+O_p\left(b\right)-O_p\left(T\right)\end{array}\right)O_p\left(b\right)=\left(\begin{array}{c}0\\ O_p\left(bT\right)\end{array}\right)
\end{aligned}$$

where since $i_{M_r}$ is uniformly drawn from $\{1,2,...,T-b\}$, then $i_{M_r}=O_p\left(T\right)$. Then the third term of Equation (12) is $O_p\left(T^{-1}\right)O_p\left(T^{-1}\right)O_p\left(bT\right)=O_p\left(bT^{-1}\right)=O_p\left(k^{-1}\right)$. Thus we have $A_{2t}$ is $O_p\left(k^{-1/2}\right)$. Next, focusing on $A_{1t}$. For $i=1,...,N$, each element of $A_{1t}$ can be expressed as

$$T^{-1/2}\left(\sum_{m=0}^{M_r}\sum_{s=1}^{b}\left(x_{i,i_m+s}^{d}-\widehat{\rho}_i x_{i,,i_m+s-1}^{d}\right)-\frac{1}{T-1}\sum_{t=2}^{T}\left(x_{i,t}^{d}-\widehat{\rho}_i x_{i,t-1}^{d}\right)\right)\tag{13}$$

$$\begin{aligned}
=\ &T^{-1/2}\sum_{m=0}^{M_r}\sum_{s=1}^{b}\left(u_{i,i_m+s}-\frac{1}{T-1}\sum_{t=2}^{T}u_{i,t}\right)\\
&-\left(\widehat{\rho}_i-1\right)T^{-1/2}\sum_{m=0}^{M_r}\sum_{s=1}^{b}\left(y_{i,i_m+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}y_{i,t-1}\right)\\
&-\left(\widehat{\rho}_i-1\right)\left(\widetilde{\beta}_i-\beta_i\right)'T^{-1/2}\sum_{m=0}^{M_r}\sum_{s=1}^{b}\left(z_{i_m+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}z_{t-1}\right)
\end{aligned}$$

About the first term of Equation (13), by Lemma A.5. of PSU, we know that it is asymptoticly equivalent to

$$T^{-1/2}\sum_{m=0}^{\lfloor (k-1)r\rfloor}\sum_{s=1}^{b}\left(u_{i_m+s}-E^{*}u_{i_m+s}\right)$$

By the same proof as in PP, the second term of Equation (13) will converge to 0 uniformly in $r$. About the third term of Equation (13), since

$$\begin{aligned}
&\sum_{m=0}^{M_r}\sum_{s=1}^{b}\left(z_{i_m+s-1}-\frac{1}{T-1}\sum_{t=2}^{T}z_{t-1}\right)\\
=\ &\sum_{m=0}^{M_r}\sum_{s=1}^{b}\left(\left(\begin{array}{c}1\\ i_m+s-1\end{array}\right)-\left(\begin{array}{c}1\\ T/2\end{array}\right)\right)\\
=\ &\left(\begin{array}{c}0\\ b\sum_{m=0}^{M_r}i_m+\left(M_r+1\right)\left(b-1\right)b/2-\left(M_r+1\right)Tb/2\end{array}\right)
\end{aligned}$$

then we rewrite the last term of Equation (13) as

$$T\left(\widehat{\rho}_i-1\right)\left(\widetilde{\beta}_i-\beta_i\right)'\left(\begin{array}{cc}T^{-1/2}&0\\0&T^{1/2}\end{array}\right)\left(\begin{array}{c}0\\ bT^{-2}\sum_{m=0}^{M_r}i_m+M_r b^2T^{-2}/2-M_rT^{-1}b/2\end{array}\right)$$

Since

$$b\sum_{m=0}^{M_r} i_m = bM_r \underbrace{M_r^{-1}\sum_{m=0}^{M_r} i_m}_{\stackrel{P}{\to}E(i_m)} = O_p\left(T\right)$$

thus $T^{-2}b\sum_{m=0}^{M_r} i_m = O_p^*\left(T^{-1}\right)$. Additionally we have $M_r b^2 T^{-2}/2 = O_p\left(k^{-1}\right)$ and $M_r T^{-1}b/2 = r/2 + o_p\left(1\right)$. Then applying the following notations

$$T\left(\widehat{\rho}_i - 1\right) = \Theta_{1i} + o_p\left(1\right)$$
$$\left(\widetilde{\beta}_i - \beta_i\right)' diag(T^{-1/2}, T^{1/2}) = \begin{pmatrix} \Theta_{2i}^{\mu} \\ \Theta_{2i}^{\tau} \end{pmatrix} + o_p\left(1\right)$$

the last term of Equation (13) can be expressed as

$$-\frac{1}{2}\Theta_i r + o_p\left(1\right)$$

where $\Theta_i = \Theta_{1i}\Theta_{2i}^{\tau}$. At last, summarizing all the results about Equation (13), we have

$$\begin{aligned} A_{\lfloor Tr\rfloor} &= T^{-1/2}\sum_{m=0}^{\lfloor (k-1)r\rfloor}\sum_{s=1}^{b}\left(u_{i_m+s} - E^* u_{i_m+s}\right) + \Theta r + o_p^*\left(1\right) \\ &= B\left(r\right) + \Theta r + o_p^*\left(1\right) \end{aligned} \tag{14}$$

where $\Theta = \left(\Theta_1, .., \Theta_N\right)'/2$. Now focus on $B_t$

$$\begin{aligned} B_t &= T^{-1/2}\left(\widehat{\beta}^* - \beta^*\right) z_t = T^{-1/2}\left(\sum_{t=1}^{T} y_t^* z_t'\right)\left(\sum_{t=1}^{T} z_t z_t'\right)^{-1} z_t \\ &= T^{-1/2}\left(\sum_{t=1}^{T}\left(\sum_{j=1}^{t} u_j^{d*}\right) z_t'\right)\left(\sum_{t=1}^{T} z_t z_t'\right)^{-1} z_t \end{aligned}$$

Then

$$B_{\lfloor Tr\rfloor} = \left(\sum_{t=1}^{T}\left(T^{-1/2}\sum_{j=1}^{\lfloor Tr\rfloor} u_j^{d*}\right) z_t'\right)\left(\sum_{t=1}^{T} z_t z_t'\right)^{-1} z_t$$

Applying the previous results of $A_t$, we have

$$\begin{aligned} B_{\lfloor Tr\rfloor} &= \left(\int_0^1 \left(B\left(r\right) + \Theta r\right) dr, \int_0^1 r\left(B\left(r\right) + \Theta r\right) dr\right)\begin{pmatrix} 4 - 6r \\ 12r - 6 \end{pmatrix} + o_p^*\left(1\right) \\ &= \left(\int_0^1 B\left(r\right) dr, \int_0^1 rB\left(r\right) dr\right)\begin{pmatrix} 4 - 6r \\ 12r - 6 \end{pmatrix} + \Theta\left(\frac{1}{2}, \frac{1}{3}\right)\begin{pmatrix} 4 - 6r \\ 12r - 6 \end{pmatrix} + o_p^*\left(1\right) \\ &= \left(\int_0^1 B\left(r\right) dr, \int_0^1 rB\left(r\right) dr\right)\begin{pmatrix} 4 - 6r \\ 12r - 6 \end{pmatrix} + \Theta r + o_p^*\left(1\right) \end{aligned} \tag{15}$$

Then by Equation (14) and (15), we have

$$T^{-1/2}x_{\lfloor Tr\rfloor}^{*d} = B^d\left(r\right) + o_p^*\left(1\right).$$

□

**Proof of Lemma 6**. About $\Omega^{d*}$, first we check $E^*\left(T^{-1}\sum_{t=1}^{T} u_t^{*d}\right)$

$$E^*\left(T^{-1/2}\sum_{t=1}^{T}u_t^{*d}\right)=E^*T^{-1/2}\sum_{t=1}^{T}u_t^{d*}-E^*T^{1/2}\left(\widehat{\beta}_{2\cdot}^*-\beta_{2\cdot}^*\right)$$

For the first term of $E^*\left(T^{-1/2}\sum_{t=1}^{T}u_t^{*d}\right)$, by Equation (14) we have that

$$T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}u_t^{d*}=T^{-1/2}\sum_{m=0}^{\lfloor (k-1)r\rfloor}\sum_{s=1}^{b}\left(u_{i_m+s}-E^*u_{i_m+s}\right)+\Theta r+o_p^*\left(1\right) \tag{16}$$

By setting $r=1$, we have

$$T^{-1/2}\sum_{t=1}^{T}u_t^{d*}=T^{-1/2}\sum_{m=0}^{k-1}\sum_{s=1}^{b}\left(u_{i_m+s}-E^*u_{i_m+s}\right)+\Theta+o_p^*\left(1\right) \tag{17}$$

Then

$$E^*\left(T^{-1/2}\sum_{t=1}^{T}u_t^{d*}\right)=E^*\Theta+o_p\left(1\right) \tag{18}$$

For the second term of $E^*\left(T^{-1/2}\sum_{t=1}^{T}u_t^{*d}\right)$,

$$T^{1/2}\left(\widehat{\beta}_{2\cdot}^*-\beta_{2\cdot}^*\right)=12T^{-2}\sum_{t=1}^{T}\left(t-\frac{T}{2}\right)T^{-1/2}y_t^*$$

and $T^{-1/2}y_t^*$ can be expressed as

$$\begin{aligned}
T^{-1/2}y_t^* &= T^{-1/2}\sum_{j=1}^{t}u_j^{d*}=T^{-1/2}\sum_{m=0}^{\lfloor (t-2)/b\rfloor}\sum_{s=1}^{b}\left(u_{i_m+s}-E^*u_{i_m+s}\right)+\frac{t}{T}\Theta+o_p^*\left(1\right)\\
&= T^{-1/2}\Gamma'\psi\left(1\right)\sum_{m=0}^{\lfloor (t-2)/b\rfloor}\sum_{s=1}^{b}\left(\varepsilon_{i_m+s}-E^*\varepsilon_{i_m+s}\right)+\frac{t}{T}\Theta+o_p^*\left(1\right)\\
&= T^{-1/2}\Gamma'\psi\left(1\right)\sum_{m=0}^{\lfloor (t-2)/b\rfloor}b^{1/2}H_m^*+\frac{t}{T}\Theta+o_p^*\left(1\right)
\end{aligned}$$

where we use the same notation in PSU and define $H_m^*=b^{-1/2}\sum_{s=1}^{b}\left(\varepsilon_{i_m+s}-E^*\varepsilon_{i_m+s}\right)$. And by Lemma A.3. (i) in PSU, $E^*H_m^*=0$, and so we have

$$E^*\left(T^{1/2}\left(\widehat{\beta}_{2\cdot}^*-\beta_{2\cdot}^*\right)\right)=E^*\Theta 12T^{-2}\sum_{t=1}^{T}\left(t-\frac{T}{2}\right)\frac{t}{T}+o_p\left(1\right)=E^*\Theta+o_p\left(1\right) \tag{19}$$

Then by Equation (18) and (19) we have

$$E^*\left(\sum_{t=1}^{T}u_t^{*d}\right)=o_p\left(1\right)$$

Now focus on the first part of $\Omega^{d*}$

$$\begin{aligned} & E^*\left(T^{-1/2}\sum_{t=1}^T u_t^{*d}\right)\left(T^{-1/2}\sum_{t=1}^T u_t^{*d}\right)' \\ = \; & E^*\left(T^{-1/2}\sum_{t=1}^T u_t^{d*}\right)\left(T^{-1/2}\sum_{t=1}^T u_t^{d*}\right)' - E^*\left(\sum_{t=1}^T u_t^{d*}\right)\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' \\ & -E^*\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)\left(\sum_{t=1}^T u_t^{d*}\right)' + TE^*\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' \end{aligned} \tag{20}$$

About the first term of Equation (20), by Equation (17) we have

$$\begin{aligned} & E^*T^{-1}\left(\sum_{t=1}^T u_t^{d*}\right)\left(\sum_{t=1}^T u_t^{d*}\right)' \\ = \; & E^*\left(T^{-1/2}\Gamma'\psi(1)\sum_{m=0}^{k-1} b^{1/2}H_m^* + \Theta\right)\left(T^{-1/2}\Gamma'\psi(1)\sum_{m=0}^{k-1} b^{1/2}H_m^{*\prime} + \Theta'\right) + o_p(1) \end{aligned}$$

by Lemma A.6. in PSU we have that

$$\begin{aligned} & E^*T\left(\sum_{t=1}^T u_t^{d*}\right)\left(\sum_{t=1}^T u_t^{d*}\right)' \\ = \; & \Omega + E^*\left\{\left(T^{-1/2}\Gamma'\psi(1)\sum_{m=0}^{k-1} b^{1/2}H_m^*\right)\Theta'\right\} + E^*\left\{\Theta\left(T^{-1/2}\Gamma'\psi(1)\sum_{m=0}^{k-1} b^{1/2}H_m^{*\prime}\right)\right\} + E^*\Theta\Theta' + o_p(1) \end{aligned} \tag{21}$$

About the second term of Equation (20), by Equation (17) and (16) we have

$$\begin{aligned} & E^*\left(\sum_{t=1}^T u_t^{d*}\right)\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' \\ = \; & 12T^{-3}E^*\left(\sum_{t=1}^T u_t^{d*}\right)\left(\sum_{t=1}^T\left(t-\frac{T}{2}\right)\sum_{j=1}^t u_j^{d*\prime}\right) = 12T^{-2}\sum_{t=1}^T\left(t-\frac{T}{2}\right)E^*T^{-1}\left(\sum_{t=1}^T u_t^{d*}\right)\left(\sum_{j=1}^t u_j^{d*\prime}\right) \\ = \; & 12T^{-2}\sum_{t=1}^T\left(t-\frac{T}{2}\right)E^*\left\{\left(T^{-1/2}\Gamma'\psi(1)\sum_{m_1=0}^{k-1} b^{1/2}H_m^* + \Theta\right)\right. \\ & \left.\left(T^{-1/2}\sum_{m_2=0}^{\lfloor (t-2)/b\rfloor} b^{1/2}H_m^{*\prime}\psi(1)'\Gamma + \frac{t}{T}\Theta'\right)\right\} + o_p(1) \end{aligned}$$

About the product of $T^{-1/2}\Gamma'\psi(1)\sum_{m_1=0}^{k-1} b^{1/2}H_m^*$ and $T^{-1/2}\sum_{m_2=0}^{\lfloor (t-2)/b\rfloor} b^{1/2}H_m^{*\prime}\psi(1)'\Gamma$, for fixed $T$, there exist $t_0$ s.t. $\lfloor (t-2)/b\rfloor \geq 1$ when $t > t_0$. Assume $t > t_0$, then there should be $\lfloor (t-2)/b\rfloor$ cases s.t. $m_1 = m_2$. Applying the independence of blocks and Lemma A.3 of PSU,

$E^* H_m^* H_m^{*\prime} = \Sigma + o_p(1)$, then we have

$$\begin{aligned}
& 12T^{-3} \sum_{t=1}^{T} \left(t - \frac{T}{2}\right) \Gamma' \psi(1) \sum_{m_1=0}^{k-1} \sum_{m_2=0}^{\lfloor (t-2)/b \rfloor} b E^* H_m^* H_m^{*\prime} \psi(1)' \Gamma \\
= \; & 12T^{-3} \Gamma' \psi(1) \left\{ \sum_{t=1}^{T} \left(t - \frac{T}{2}\right) \lfloor (t-2)/b \rfloor b E^* H_{m_1}^* H_{m_1}^{*\prime} \right\} \psi(1)' \Gamma + o_p(1) \\
= \; & 12\Gamma' \psi(1) \left\{ T^{-3} \sum_{t=\lfloor (t-2)/b \rfloor + 1}^{T} \left(t - \frac{T}{2}\right) \lfloor (t-2)/b \rfloor b \right\} \Sigma \psi(1)' \Gamma + o_p(1) \\
= \; & \Gamma' \psi(1) \Sigma \psi(1)' \Gamma + o_p(1) = \Omega + o_p(1)
\end{aligned}$$

Then by tedious calculations we have

$$E^* \left( \sum_{t=1}^{T} u_t^{d*} \right) \left( \widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^* \right)' = \Omega + E^* \left( T^{-1/2} \Gamma' \psi(1) \sum_{m_1=0}^{k-1} b^{1/2} H_m^* \Theta' \right) + E^* \Theta \Theta' + o_p(1) \tag{22}$$

Similarly, for the third term of Equation (20) we have

$$E^* \left( \widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^* \right) \left( \sum_{t=1}^{T} u_t^{d*} \right)' = \Omega + E^* \Theta \left( T^{-1/2} \Gamma' \psi(1) \sum_{m=0}^{k-1} b^{1/2} H_m^{*\prime} \right) + E^* \Theta \Theta' + o_p(1) \tag{23}$$

At last, about the last term of Equation (20),

$$\begin{aligned}
& T E^* \left( \widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^* \right) \left( \widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^* \right)' \\
= \; & 12^2 T^{-5} E^* \left( \sum_{t=1}^{T} \left(t - \frac{T}{2}\right) \sum_{j=1}^{t} u_j^{d*} \right) \left( \sum_{t=1}^{T} \left(t - \frac{T}{2}\right) \sum_{j=1}^{t} u_j^{d*\prime} \right) \\
= \; & 12^2 T^{-4} E^* \left( \sum_{t=1}^{T} \left(t - \frac{T}{2}\right) \left( T^{-1/2} \sum_{m=0}^{\lfloor (t-2)/b \rfloor} \sum_{s=1}^{b} (u_{i_m+s} - E^* u_{i_m+s}) + t T^{-1} \Theta \right) \right) \\
& \left( \sum_{s=1}^{T} \left(t - \frac{T}{2}\right) \left( T^{-1/2} \sum_{m=0}^{\lfloor (s-2)/b \rfloor} \sum_{s=1}^{b} (u_{i_m+s} - E^* u_{i_m+s}) + t T^{-1} \Theta \right)' \right) + o_p^*(1)
\end{aligned}$$

By the same argument, we have

$$\begin{aligned}
& T E^* \left( \widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^* \right) \left( \widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^* \right)' \qquad (24) \\
= \; & 12^2 T^{-4} \sum_{t,s=1}^{T} \left(t - \frac{T}{2}\right) \left(s - \frac{T}{2}\right) \Gamma' \psi(1) T^{-1} b \sum_{m_1=0}^{\lfloor (t-2)/b \rfloor} \sum_{m_2=0}^{\lfloor (s-2)/b \rfloor} E H_m^* H_m^{*'} + E^* \Theta \Theta' + o_p(1) \\
= \; & \frac{6}{5} \Omega + E^* \Theta \Theta' + o_p(1)
\end{aligned}$$

Then by associating equations from (21) to (24), we have

$$\Omega^{d*} = \Omega^d + o_p(1)$$

Next, about

$$\Omega_0^{d*} = T^{-1}\sum_{t=1}^{T}\left(E^* u_t^{*d}u_t^{*d\prime} - E^* u_t^{*d}E^* u_t^{*d\prime}\right) = \Omega_0^d + o_p(1)$$

$$T^{-1}\sum_{t=1}^{T} E^* u_t^{*d}u_t^{*d\prime} = E^* T^{-1}\sum_{t=1}^{T} u_t^{*d}u_t^{*d\prime}$$

where

$$\begin{aligned} u_t^{*d}u_t^{*d\prime} &= \left(u_t^{d*} - \left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)\right)\left(u_t^{d*\prime} - \left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)'\right) \\ &= u_t^{d*}u_t^{d*\prime} - u_t^{d*}\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' - \left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)u_t^{d*\prime} + \left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' \end{aligned}$$

Apply the previous results, we have that

$$T^{-1}\sum_{t=1}^{T} u_t^{d*}\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' = o_p^*(1)$$

and

$$T^{-1}\sum_{t=1}^{T}\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)\left(\widehat{\beta}_{2\cdot}^* - \beta_{2\cdot}^*\right)' = o_p^*(1)$$

Then we have

$$\Omega_0^{d*} = \Omega_0^d + o_p^*(1)$$

□

**Proof of Theorem 2**. By the results of Lemma 5 and 6, and the continuous mapping theorem we have

$$T^{-1}\sum_{t=1}^{T} x_{i,t-1}^{*d}\Delta x_{i,t}^{*d} \xrightarrow{d} \int_0^1 B_i^d(r)\, dB_i^d(r) + \frac{1}{2}\left(\omega_i^d - \omega_{0,i}^d\right)$$

and

$$T^{-1}\sum_{t=1}^{T} x_{i,t-1}^{*d2} \xrightarrow{d} \int_0^1 B_i^d(r)^2\, dr.$$

Then applying the Cramér-Wold device, the convergence also holds jointly. Thus Theorem 2 can be proved immediately. □