Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependency[1]

Yoosoon Chang
Department of Economics
Rice University

Abstract

We apply bootstrap methodology to unit root tests for dependent panels with N cross-sectional units and T time series observations. More specifically, we let each panel be driven by a general linear process which may be different across cross-sectional units, and approximate it by a finite order autoregressive integrated process of order increasing with T. As we allow the dependency among the innovations generating the individual panels, we construct our unit root tests from the estimation of the system of the entire N panels. The limit distributions of the tests are derived by passing T to infinity, with N fixed. We then apply the bootstrap method to the approximated autoregressions to obtain the critical values for the panel unit root tests, and establish the asymptotic validity of such bootstrap panel unit root tests under general conditions. The proposed bootstrap tests are indeed quite general covering a wide class of panel models. They in particular allow for very general dynamic structures which may vary across individual units, and more importantly for the presence of arbitrary cross-sectional dependency. The finite sample performance of the bootstrap tests is examined via simulations, and compared to that of the t-bar statistics by Im, Pesaran and Shin (1997), which is one of the commonly used unit root tests for panel data. We find that our bootstrap panel unit root tests perform well relative to the t-bar statistics, especially when N is small.

This version: January, 2000

JEL Classification: C12, C15, C32, C33.
Key words and phrases: Panels with cross-sectional dependency, unit root tests, sieve bootstrap, AR approximation.

[1]The paper was written while I was visiting the Cowles Foundation for Research in Economics at Yale University during the fall of 1999. I am grateful to Don Andrews, Bill Brown, Joon Park and Peter Phillips for helpful discussions and comments. My thanks also go to the seminar participants at Yale and the Texas Econometrics Camp. This research is supported in part by the CSIV fund from Rice University.
Correspondence address to: Yoosoon Chang, Department of Economics - MS 22, Rice University, 6100 Main Street, Houston, TX 77005-1892, Tel: 713-348-2796, Fax: 713-348-5278, Email: yoosoon@rice.edu.

1. Introduction

Recently, nonstationary panels have drawn much attention in both theoretical and empirical research, as a number of panel data sets covering relatively long time periods have become available. Various statistics for testing the unit roots and cointegration for panel models have been proposed, and frequently used for testing growth convergence theories, purchasing power parity hypothesis and for estimating long-run relationships among many macroeconomic and international financial series including exchange rates and spot and future interest rates. Such panel data based tests appeared attractive to many empirical researchers, since they offer alternatives to the tests based only on individual time series observations that are known to have low discriminatory power. A number of unit roots and cointegration tests have been developed for panel models by many authors. See Levin and Lin (1992,1993), Quah (1994), Im, Pesaran and Shin (1997) and Maddala and Wu (1996) for some of the panel unit root tests, and Pedroni (1996,1997) and McCoskey and Kao (1998) for the panel cointegration tests available in the current literature. Banerjee (1999) gives a good survey on the recent developments in the econometric analysis of panel data whose time series component is nonstationary.[2]

Since the work by Levin and Lin (1992), a number of unit root tests for panel data have been proposed. Levin and Lin (1992,1993) consider unit root tests for homogeneous panels, which are simply the usual t-statistics constructed from the pooled estimator with some appropriate modifications. Such unit root tests for homogeneous panels can therefore be represented as a simple sum taken over $i = 1, \ldots, N$ and $t = 1, \ldots, T$. They show under cross-sectional independency that the sequential limit of the standard t-statistics for testing the unit root is the standard normal distribution.[3] For heterogeneous panels, the unit root test can no longer be represented as a simple sum since the pooled estimator is inconsistent for such heterogeneous panels as shown in Pesaran and Smith (1995). Consequently the second stage N-asymptotics in the above sequential asymptotics does not work here. Im, Pesaran and Shin (1997) looks at the heterogeneous panels and proposes unit root tests which are based on the average of the independent individual unit root tests, t-statistics and LM statistics, computed from each individual panel. They show that their tests also converge to the standard normal distribution upon taking sequential limits. Though they allow for the heterogeneity, their limit theory is still based on the cross-sectional independecy, which can be quite a restrictive assumption for most of the panel data we encounter.

The tests suggested by Levin and Lin (1993) and Im, Pesaran and Shin (1997) are not valid in the presence of cross-correlations among the cross-sectional units. The limit

[2]Stationary panels have a much longer history and have been intensely investigated by many researchers. The readers are referred to the books by Hsiao (1986), Matyas and Sevestre (1996) and Baltagi (1995) for the literature on the econometric analysis of panel data.

[3]The sequential limit is taken by first passing T to infinity with N fixed and subsequently let N tend to infinity. Regression limit theory for nonstationary panel data is developed rigorously by Phillips and Moon (1999). They show that the limit of the double indexed processes may depend on the way N and T tend to infinity. They formally develop the asymptotics of sequential limit, diagonal path limit (N and T tend to infinity at a controlled rate of the type $T = T(N)$) and joint limits (N and T tend to infinity simultaneously without any restrictions imposed on the divergence rate). Their limit thoery, however, assumes cross-sectional independence.

limit distributions of their tests are no longer valid and unknown when the independency assumption is violated. Indeed, Maddala and Wu (1996) show through simulations that their tests have substantial size distortions when used for cross-sectionally dependent panels. As a way to deal with such inferential difficulty in panels with cross-correlations, they suggest to bootstrap the panel unit root tests, such as those proposed by Levin and Lin (1993), Im, Pesaran and Shin (1997) and Fisher (1933), for cross-sectinally dependent panels. They show through simulations that the bootstrap version of such tests perform much better, but do not provide the validity of using bootstrap methodology.

In this paper, we apply bootstrap methodology to unit root tests for cross-sectionally dependent panels. More specifically, we let each panel be driven by a general linear process which may be different across cross-sectional units, and approximate it by a finite order autoregressive integrated process of order increasing with T. As we allow the dependency among the innovations generating the individual panels, we construct our unit root tests from the estimation of the system of the entire N panels. The limit distributions of the tests are derived by passing T to infinity, with N fixed. We then apply the bootstrap method to the approximated autoregressions to obtain the critical values for the panel unit root tests based on the original sample, and establish the asymptotic validity of such bootstrap panel unit root tests under general conditions.

The rest of the paper is organized as follows. Section 2 introduces the panel unit root tests for cross-sectionally dependent panels based on the original sample and derives the limit theory for the sample tests. Section 3 applies the sieve bootstrap methodology to the sample panel unit root tests considered in Section 2 and establishes asymptotic validity of the sieve bootstrap unit root tests. Also discussed in Section 3 are the practical issues arising from the implementation of the sieve bootstrap methodology. In Section 4, we conduct simulations to investigate finite sample performance of the bootstrap unit root tests. Section 5 concludes, and mathematical proofs are provided in an Appendix.

2. Unit Root Tests for Dependent Panels

We consider a panel model generated as the following first order autoregressive regression:

$$\triangle y_{it} = \alpha_i y_{i,t-1} + u_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T. \tag{1}$$

As usual, the index i denotes individual cross-sectional units, such as individuals, households, industries or countries, and the index t denotes time periods. We are interested in testing the unit root null hypothesis, $\alpha_i = 0$ for all y_{it} given as in (1), against the alternative, $\alpha_i < 0$ for some y_{it}, $i = 1, \ldots, N$. Thus, the null implies that all y_{it}'s have unit roots, and is rejected if any one of y_{it}'s is stationary with $\alpha_i < 0$. The rejection of the null therefore does not imply that the entire panel is stationary. The initial values $(y_{10}, \ldots, y_{N0})$ of $(y_{1t}, \ldots, y_{Nt})$ do not affect our subsequent asymptotic analysis as long as they are stochastically bounded, and therefore we set them at zero for expositional brevity.

It is assumed that the error term (u_{it}) in the model (1) is given by a general linear process specified as

$$u_{it} = \pi_i(L)\varepsilon_{it} \tag{2}$$

where L is the usual lag operator and

$$\pi_i(z) = \sum_{k=0}^{\infty} \pi_{i,k} z^k$$

for $i = 1, \ldots, N$. Note that we let π_i vary across i, thereby allowing heterogeneity in individual serial correlation structures. We also allow for the cross-sectional dependency through the cross-correlation of the innovations ε_{it}, $i = 1, \ldots, N$ that generate the error u_{it}'s. To define the cross-sectional dependecy more explicitly, we define the time series innovation $(\varepsilon_t)_{t=1}^T$ by

$$\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})' \tag{3}$$

and denote by $|\cdot|$ the Euclidean norm: for a vector $x = (x_i)$, $|x|^2 = \sum_i x_i^2$, and for a matrix $A = (a_{ij})$, $|A| = \sum_{i,j} a_{ij}^2$. We assume the following:

Assumption A1 Let $(\varepsilon_t, \mathcal{F}_t)$ be a martingale difference sequence, with some filtration $(\mathcal{F}_t)$, such that $\mathbf{E}(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}) = \Sigma$ a.s., and $\mathbf{E}|\varepsilon_t|^r < \infty$ for some $r \geq 4$.

Assumption A2 Let $\pi_i(z) \neq 0$ for all $|z| \leq 1$, and $\sum_{k=0}^{\infty} |k|^s|\pi_{i,k}| < \infty$ for some $s \geq 1$, for all $i = 1, \ldots, N$.

The conditions in Assumptions A1 and A2 are routinely imposed on the linear processes given by (2). It is well known that an invariance principle holds for a partial sum process of (ε_t) defined in (3) under Assumption A1. That is,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\cdot]} \varepsilon_t \;\to_d\; B = \begin{pmatrix} B_1 \\ \vdots \\ B_N \end{pmatrix} = BM(0, \Sigma) \tag{4}$$

as $T \to \infty$, where $[x]$ denotes the maximum integer which does not exceed x.

We may write (u_{it}) as

$$u_{it} = \pi_i(1)\varepsilon_{it} + (\bar{u}_{i,t-1} - \bar{u}_{it}) \tag{5}$$

where

$$\bar{u}_{it} = \sum_{k=0}^{\infty} \bar{\pi}_{i,k}\varepsilon_{i,t-k}, \quad \bar{\pi}_{i,k} = \sum_{j=k+1}^{\infty} \pi_{i,j}$$

Under our condition in Assumption A2, we have $\sum_{k=0}^{\infty} |\bar{\pi}_{i,k}| < \infty$ [see Phillips and Solo (1992)] and therefore $(\bar{u}_{it})$ is well defined both in a.s. and L^r sense [see Brockwell and Davis (1991, Proposition 3.1.1)].

Under the unit root hypothesis $\alpha_1 = \cdots = \alpha_N = 0$, we may now write

$$y_{it} = \pi_i(1)w_{it} + (\bar{u}_{i0} - \bar{u}_{it}) \tag{6}$$

where $w_{it} = \sum_{k=1}^{t} \varepsilon_{ik}$. Consequently, (y_{it}) behaves asymptotically as the constant $\pi_i(1)$ multiple of (w_{it}). Note that $(\bar{u}_{it})$ is stochastically of smaller order of magnitude than (w_{it}), and therefore will not contribute to our limit theory.

Under Assumptions A1 and A2, we may write the linear process given in (2) as an infinite order autoregressive (AR) process

$$\alpha_i(L)u_{it} = \varepsilon_{it}$$

with

$$\alpha_i(z) = 1 - \sum_{k=1}^{\infty} \alpha_{i,k} z^k$$

and approximate (u_{it}) by a finite order AR process

$$u_{it} = \alpha_{i,1}u_{i,t-1} + \cdots + \alpha_{i,p_i}u_{i,t-p_i} + \varepsilon_{it}^{p_i} \tag{7}$$

with

$$\varepsilon_{it}^{p_i} = \varepsilon_{it} + \sum_{k=p_i+1}^{\infty} \alpha_{i,k}u_{i,t-k}$$

Under Assumptions A1 and A2, we have for each $i = 1, \ldots, N$

$$\mathbf{E}|\varepsilon_{it}^{p_i} - \varepsilon_{it}|^r \leq \mathbf{E}|u_{it}|^r \left(\sum_{k=p_i+1}^{\infty} |\alpha_{i,k}| \right)^r = o(p_i^{-rs})$$

Note that we have under Assumptions A1 and A2

$$\mathbf{E}|u_{it}|^r \leq c \left(\sum_{k=0}^{\infty} \pi_{i,k}^2 \right)^{r/2} \mathbf{E}|\varepsilon_{it}|^r < \infty$$

for some constant c, due to the Marcinkiewicz-Zygmund inequality [see, e.g., Stout (1974, Theorem 3.3.6)]. The error in approximating (u_{it}) by a finite order AR process thus becomes small as p_i gets large.

Using the AR approximation of (u_{it}) given in (7), we write the model (1) as

$$\triangle y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k}u_{i,t-k} + \varepsilon_{it}^{p_i}$$

which, since $\triangle y_{it} = u_{it}$ under the null hypothesis, can be seen as an autoregression of $\triangle y_{it}$ augmented by $y_{i,t-1}$, viz.

$$\triangle y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k} \triangle y_{i,t-k} + \varepsilon_{it}^{p_i} \tag{8}$$

Our unit root tests will be based on the above approximated autoregression.

For the order p_i in the regression (8), we assume

Assumption A3 $p_i \to \infty$ and $p_i = o(T^{1/2})$ as $T \to \infty$, for all $i = 1, \ldots, N$.

The AR order p_i should, in particular, be increasing with T.[4] We may choose p_i's using the usual order selection criteria such as Schwartz information criterion (BIC) or Akaike information criterion (AIC).[5] The order selection can be based either on the regression (8) with no restriction on α_i's, or on the approximated AR regression in (7) where α_i's are restricted to be zero. This will not affect our subsequent limit theory.

2.1 Unit Root Tests for Heterogeneous Panels

The augmented autoregression (8) can be written in the following matrix form by taking the individual units, with all their T observations, one after the other, viz.

$$\begin{pmatrix} \triangle y_1 \\ \vdots \\ \triangle y_N \end{pmatrix} = \begin{pmatrix} y_{\ell,1} & & 0 \\ & \ddots & \\ 0 & & y_{\ell,N} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} + \begin{pmatrix} X_1^{p_1} & & 0 \\ & \ddots & \\ 0 & & X_N^{p_N} \end{pmatrix} \begin{pmatrix} \beta_1^{p_1} \\ \vdots \\ \beta_N^{p_N} \end{pmatrix} + \begin{pmatrix} \varepsilon_1^{p_1} \\ \vdots \\ \varepsilon_N^{p_N} \end{pmatrix}$$

or more compactly

$$\triangle y = Y_\ell \alpha + X_p \beta_p + \varepsilon_p \tag{9}$$

where we use the following notation

$$y_{\ell,i} = \begin{pmatrix} y_{i,0} \\ \vdots \\ y_{i,T-1} \end{pmatrix}, \quad X_i^{p_i} = \begin{pmatrix} x_{i1}^{p_i\prime} \\ \vdots \\ x_{iT}^{p_i\prime} \end{pmatrix} \quad \text{and} \quad \beta_i^{p_i} = \begin{pmatrix} \alpha_{i,1} \\ \vdots \\ \alpha_{i,p_i} \end{pmatrix}$$

with $x_{it}^{p_i\prime} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-p_i})$, for all $i = 1, \ldots, N$.

We construct the tests for the null hypothesis of the unit roots in $y_t = (y_{1t}, \ldots, y_{Nt})'$ generated by (1) and (2) based on the system GLS and OLS estimation of the augmented AR (9). The feasible GLS estimator of α in (9) is given by

$$\hat{\alpha}_{GT} = B_{GT}^{-1} A_{GT}$$

where A_{GT} and B_{GT} are defined below. For the test of the null $\alpha = 0$, we consider the following F-type test based on the feasible GLS estimator $\hat{\alpha}_{GT}$:

$$F_{GT} = \hat{\alpha}_{GT}' (\text{var}(\hat{\alpha}_{GT}))^{-1} \hat{\alpha}_{GT} = A_{GT}' B_{GT}^{-1} A_{GT} \tag{10}$$

where

$$\begin{aligned} A_{GT} &= Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p - Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)X_p \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T)X_p\right)^{-1} X_p'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p \\ B_{GT} &= Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)Y_\ell - Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)X_p \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T)X_p\right)^{-1} X_p'(\tilde{\Sigma}^{-1} \otimes I_T)Y_\ell \end{aligned}$$

[4] Our regression (8) here may be viewed as the extension of the unit root regression considered in Said and Dickey (1984) to the panel models. However, our assumption on the AR order p_i is substantially weaker than that used by Said and Dickey (1984), due to the result in Chang and Park (1999).

[5] As for the choice among the selection criteria, BIC might be preferred if (u_{it}) is believed to be generated by a finite autoregression, since it yields a consistent estimator for p_i. See, e.g., An, Chen and Hannan (1982). If not, AIC may be a better choice, since it leads to an asymptotically efficient choice for the optimal order of some projected infinite order autoregressive process as shown by Shibata (1980). See Choi (1992) for more discussions on the model selection issue for ARMA models.

and $\tilde{\Sigma}$ is a consistent estimator of the covariance matrix Σ. The limit distribution for the test F_{GT} is easily drived from the asymptotic behaviors of A_{GT} and B_{GT}, and is given in Theorem 2.1 below.

On the other hand, the OLS estimator of α in (9) is given by

$$\hat{\alpha}_{OT} = B_{OT}^{-1} A_{OT}$$

and use the following OLS-based F-type test for testing $\alpha = 0$

$$F_{OT} = \hat{\alpha}_{OT}' (\mathrm{var}(\hat{\alpha}_{OT}))^{-1} \hat{\alpha}_{OT} = A_{OT}' M_{FOT}^{-1} A_{OT} \tag{11}$$

where

$$\begin{aligned} A_{OT} &= Y_\ell' \varepsilon_p - Y_\ell' X_p (X_p' X_p)^{-1} X_p' \varepsilon_p \\ B_{OT} &= Y_\ell' Y_\ell - Y_\ell' X_p (X_p' X_p)^{-1} X_p' Y_\ell \\ M_{FOT} &= Y_\ell' (\tilde{\Sigma} \otimes I_T) Y_\ell - Y_\ell' X_p (X_p' X_p)^{-1} X_p' (\tilde{\Sigma} \otimes I_T) Y_\ell - Y_\ell' (\tilde{\Sigma} \otimes I_T) X_p (X_p' X_p)^{-1} X_p' Y_\ell \\ &\quad + Y_\ell' X_p (X_p' X_p)^{-1} X_p' (\tilde{\Sigma} \otimes I_T) X_p (X_p' X_p)^{-1} X_p' Y_\ell \end{aligned}$$

The OLS estimator $\hat{\alpha}_{OT}$ is less efficient that the GLS estimator $\hat{\alpha}_{GT}$ in our context. The OLS-based test F_{OT} in (11) is thus expected to be less powerful than the GLS-based test F_{GT} in (10). However, we observe in our simulations that F_{OT} often performs better than F_{GT} in finite samples, especially when N is large.

To construct a consistent estimator for the covariance matrix Σ, we may estimate the regression

$$u_{it} = \tilde{\alpha}_{i,1}^{p_i} u_{i,t-1} + \cdots + \tilde{\alpha}_{i,p_i}^{p_i} u_{i,t-p_i} + \tilde{\varepsilon}_{it}^{p_i} \tag{12}$$

by single-equation OLS for $i = 1, \ldots, N$, with the unit root restriction $\alpha_i = 0$ imposed. The fitted residuals $(\tilde{\varepsilon}_{it}^{p_i})$ are consistent for (ε_{it}), since $\tilde{\alpha}_{i,k}^{p_i}$ are consistent for $\alpha_{i,k}$ for $1 \leq k \leq p_i$, and the autoregressive coefficients $(\alpha_{i,k})$ for $k > p_i$ become negligible in the limit as long as we let $p_i \to \infty$. This is shown in Park (1999, Lemma 3.1). Of course, one may obtain consistent fitted residuals by estimating the unrestricted regession (8). This again will not affect our limit theory. From $(\tilde{\varepsilon}_{it}^{p_i})$, form the time series residual vectors

$$\tilde{\varepsilon}_t^p = (\tilde{\varepsilon}_{1t}^{p_1}, \ldots, \tilde{\varepsilon}_{Nt}^{p_N})' \tag{13}$$

for $t = 1, \ldots, T$. We then estimate Σ by

$$\tilde{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \tilde{\varepsilon}_t^p \tilde{\varepsilon}_t^{p\prime}$$

Notice that

$$\tilde{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^p \varepsilon_t^{p\prime} + o_p(1) = \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t \varepsilon_t' + o_p(1) = \mathbf{E} \varepsilon_t \varepsilon_t' + o_p(1)$$

where the second equality follows from Lemma A1 (c) in Appendix. We use $(\tilde{\Sigma} \otimes I_T)$ as an estimator for the variance of the regression error in (9).

Let σ_{ij} and σ^{ij} denote, respectively, the (i,j)-elements of the covariance matrix Σ and its inverse Σ^{-1}. The limit theories for the tests F_{GT} and F_{OT} are given in

Theorem 2.1 Under Assumptions A1, A2 and A3, we have
(a) $F_{GT} \;\to_d\; Q'_{A_G} Q^{-1}_{B_G} Q_{A_G}$
(b) $F_{OT} \;\to_d\; Q'_{A_O} Q^{-1}_{M_{FO}} Q_{A_O}$
as $T \to \infty$, where

$$Q_{A_G} = \begin{pmatrix} \pi_1(1)\sum_{j=1}^{N}\sigma^{1j}\int_0^1 B_1 dB_j \\ \vdots \\ \pi_N(1)\sum_{j=1}^{N}\sigma^{Nj}\int_0^1 B_N dB_j \end{pmatrix}, \quad Q_{A_O} = \begin{pmatrix} \pi_1(1)\int_0^1 B_1 dB_1 \\ \vdots \\ \pi_N(1)\int_0^1 B_N dB_N \end{pmatrix}$$

$$Q_{B_G} = \begin{pmatrix} \sigma^{11}\pi_1(1)^2\int_0^1 B_1^2 & \dots & \sigma^{1N}\pi_1(1)\pi_N(1)\int_0^1 B_1 B_N \\ \vdots & \vdots & \vdots \\ \sigma^{N1}\pi_N(1)\pi_1(1)\int_0^1 B_N B_1 & \dots & \sigma^{NN}\pi_N(1)^2\int_0^1 B_N^2 \end{pmatrix}$$

and

$$Q_{M_{FO}} = \begin{pmatrix} \sigma_{11}\pi_1(1)^2\int_0^1 B_1^2 & \dots & \sigma_{1N}\pi_1(1)\pi_N(1)\int_0^1 B_1 B_N \\ \vdots & \vdots & \vdots \\ \sigma_{N1}\pi_N(1)\pi_1(1)\int_0^1 B_N B_1 & \dots & \sigma_{NN}\pi_N(1)^2\int_0^1 B_N^2 \end{pmatrix}$$

Remarks
(a) The limit distributions of the F_{GT} and F_{OT} are nonstandard and depend heavily on the nuisance parameters that define the cross-sectional dependency and the heterogeneous serial dependence. Therefore, it is impossible to base inference on the tests F_{GT} and F_{OT}. In the next section, we propose bootstrap version of these tests to deal with the nuisance parameter dependency problem and to overcome the inferential difficulty.
(b) The F-type tests F_{GT} and F_{OT} considered here are two-tailed tests which reject the null $\alpha_i = 0$ for all i when $\alpha_i \neq 0$ for some i. Hence, they reject the null of the unit roots not only against the stationarity $\alpha_i < 0$ but also against the explosive cases with $\alpha_i > 0$ for some i. This will have a negative effect on the powers of the tests.

The framework within which we may effectively deal with the problem in Remark (b) above has been recently developed by Andrews (1999).[6] To deal with the problem, we may

[6] Here we consider testing $\alpha_i = 0$ against $\alpha_i < 0$, and the parameter set is given by $\alpha_i \leq 0$ for each cross-sectional unit $i = 1, \dots, N$. The value of α_i under the null hypothesis is therefore on the boundary of the parameter set.

replace zeros for the members of $\hat{\alpha}_{GT}$ and $\hat{\alpha}_{OT}$ which have positive values. This can be easily carried out by multiplying element by element the estimators $\hat{\alpha}_{GT} = (\hat{\alpha}_{GT,1}, \ldots, \hat{\alpha}_{GT,N})'$ and $\hat{\alpha}_{OT} = (\hat{\alpha}_{OT,1}, \ldots, \hat{\alpha}_{OT,N})'$ respectively to the N-dimensional indicator functions $1\{\hat{\alpha}_{GT} \leq 0\}$ and $1\{\hat{\alpha}_{OT} \leq 0\}$. Denote by $.*$ the element by element multiplication, and use this to modify the estimators $\hat{\alpha}_{GT}$ and $\hat{\alpha}_{OT}$ as follows

$$\hat{\alpha}_{GT} .* 1\{\hat{\alpha}_{GT} \leq 0\} = \begin{pmatrix} \hat{\alpha}_{GT,1} 1\{\hat{\alpha}_{GT,1} \leq 0\} \\ \vdots \\ \hat{\alpha}_{GT,N} 1\{\hat{\alpha}_{GT,N} \leq 0\} \end{pmatrix} \tag{14}$$

$$\hat{\alpha}_{OT} .* 1\{\hat{\alpha}_{OT} \leq 0\} = \begin{pmatrix} \hat{\alpha}_{OT,1} 1\{\hat{\alpha}_{OT,1} \leq 0\} \\ \vdots \\ \hat{\alpha}_{OT,N} 1\{\hat{\alpha}_{OT,N} \leq 0\} \end{pmatrix}$$

We now define new statistics, which we call K-statistics. From the modified GLS estimator above, we define the GLS-based K-statistics K_{GT} as follows

$$\begin{aligned} K_{GT} &= (\hat{\alpha}_{GT} .* 1\{\hat{\alpha}_{GT} \leq 0\})' (\mathrm{var}(\hat{\alpha}_{GT}))^{-1} (\hat{\alpha}_{GT} .* 1\{\hat{\alpha}_{GT} \leq 0\}) \\ &= (A_{GT} .* 1\{\hat{\alpha}_{GT} \leq 0\})' B_{GT}^{-1} (A_{GT} .* 1\{\hat{\alpha}_{GT} \leq 0\}) \end{aligned} \tag{15}$$

and the OLS-based K-statistics K_{OT} as

$$\begin{aligned} K_{OT} &= (\hat{\alpha}_{OT} .* 1\{\hat{\alpha}_{OT} \leq 0\})' (\mathrm{var}(\hat{\alpha}_{OT}))^{-1} (\hat{\alpha}_{OT} .* 1\{\hat{\alpha}_{OT} \leq 0\}) \\ &= (A_{OT} .* 1\{\hat{\alpha}_{OT} \leq 0\})' M_{FOT}^{-1} (A_{OT} .* 1\{\hat{\alpha}_{OT} \leq 0\}) \end{aligned} \tag{16}$$

The K-statistics constructed as above are essentially one-sided tests, since they effectively elliminate the probability of rejecting the null against the explosive alternative. Therefore they are expected to improve the power properties of the corresponding two-tailed F-type tests for testing of the unit root null against the one-way stationary alternative.

The limit distributions of the K-statistics can be easily obtained in a manner similar to that used to derive the limit theories for the F-type tests, and are given in

Corollary 2.1 Under Assumptions A1, A2 and A3, we have
(a) $K_{GT} \to_d (Q_{A_G} .* 1\{Q_{B_G}^{-1} Q_{A_G} \leq 0\})' Q_{B_G}^{-1} (Q_{A_G} .* 1\{Q_{B_G}^{-1} Q_{A_G} \leq 0\})$
(b) $K_{OT} \to_d (Q_{A_O} .* 1\{Q_{B_O}^{-1} Q_{A_O} \leq 0\})' Q_{M_{FO}}^{-1} (Q_{A_O} .* 1\{Q_{B_O}^{-1} Q_{A_O} \leq 0\})$
as $T \to \infty$, where

$$Q_{B_O} = \begin{pmatrix} \pi_1(1)^2 \int_0^1 B_1^2 & \ldots & \pi_1(1)\pi_N(1) \int_0^1 B_1 B_N \\ \vdots & \vdots & \vdots \\ \pi_N(1)\pi_1(1) \int_0^1 B_N B_1 & \ldots & \pi_N(1)^2 \int_0^1 B_N^2 \end{pmatrix}$$

and the terms Q_{A_G}, Q_{B_G}, Q_{A_O} and $Q_{M_{FO}}$ are defined in Theroem 2.1.

As can be seen clearly from the above Corollary, the limit distributions of the K-tests are also nonstandard and depend heavily on the nuisance parameters. In the next section, we will also consider bootstrapping the K-type tests.

2.2 Unit Root Tests for Homogeneous Panels

For the test of the unit root, we are testing $\alpha_i = 0$ for all i. Therefore, we are essentially looking at a homogeneous panel, as far as testing of the null hypothesis is concerned. If AR coefficients α_i's in our original model (1) are homogeneous, i.e., $\alpha_1 = \cdots = \alpha_N = \alpha$, then the corresponding augmented AR in matrix form is given by

$$\triangle y = y_\ell \alpha + X_p \beta_p + \varepsilon_p \tag{17}$$

which is the same as the augmented AR in matrix form for the original heterogeneous model (9), except that here we have an $(NT \times 1)$-vector $y_\ell = (y'_{\ell,1}, \ldots, y'_{\ell,N})'$ in the place of the $(NT \times N)$-matrix Y_ℓ and the parameter α is now a scalar instead of an $(N \times 1)$-vector.

It is natural to consider the t-statistics for testing the null hypothesis of the unit roots in the homogeneous model (17), since the parameter α to be tested is now a scalar. Here we do not allow for the heterogeneity of the AR coefficient, as in Levin and Lin (1992,1993). Obviously, the unit root test based on the homogeneous panel (17) is valid, since the model is correctly specified under the null hypothesis of the unit roots. The homogeneous panel, however, may not provide appropriate modellings under the alternative hypothesis, and this may have an adverse effect on the power of the tests. However, we may use the one-sided t-type tests, if based on the homogeneous panels, which have a clear a clear advantage over the two-tailed F-type tests constructed from the heterogeneous panels.

The OLS and GLS based t-statistics are constructed from the GLS and OLS estimators of the scalar parameter α in the homogeneous augmented AR (17) and are given by

$$t_{GT} = a_{GT} b_{GT}^{-1/2} \quad \text{and} \quad t_{OT} = a_{OT} M_{tOT}^{-1/2} \tag{18}$$

where

$$\begin{aligned}
a_{GT} &= y'_\ell(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p - y'_\ell(\tilde{\Sigma}^{-1} \otimes I_T)X_p(X'_p(\tilde{\Sigma}^{-1} \otimes I_T)X_p)^{-1}X'_p(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p \\
b_{GT} &= y'_\ell(\tilde{\Sigma}^{-1} \otimes I_T)y_\ell - y'_\ell(\tilde{\Sigma}^{-1} \otimes I_T)X_p(X'_p(\tilde{\Sigma}^{-1} \otimes I_T)X_p)^{-1}X'_p(\tilde{\Sigma}^{-1} \otimes I_T)y_\ell \\
a_{OT} &= y'_\ell \varepsilon_p - y'_\ell X_p(X'_p X_p)^{-1}X'_p \varepsilon_p \\
M_{tOT} &= y'_\ell(\tilde{\Sigma} \otimes I_T)y_\ell - 2y'_\ell X_p(X'_p X_p)^{-1}X'_p(\tilde{\Sigma} \otimes I_T)y_\ell \\
&\quad + y'_\ell X_p(X'_p X_p)^{-1}X'_p(\tilde{\Sigma} \otimes I_T)X_p(X'_p X_p)^{-1}X'_p y_\ell
\end{aligned}$$

In the following theorem we present the limit theories for the t_{GT} and T_{OT} tests.

Theorem 2.2 Under Assumptions A1, A2 and A3, we have

(a) $t_{GT} \;\to_d\; Q_{a_G} Q_{b_G}^{-1/2}$

(b) $t_{OT} \;\to_d\; Q_{a_O} Q_{M_{tO}}^{-1/2}$

as $T \to \infty$, where

$$Q_{a_G} = \sum_{i=1}^{N}\sum_{j=1}^{N} \sigma^{ij} \int_0^1 B_i dB_j, \quad Q_{b_G} = \sum_{i=1}^{N}\sum_{j=1}^{N} \sigma^{ij} \int_0^1 B_i B_j$$

and

$$Q_{a_O} = \sum_{i=1}^{N} \pi_i \int_0^1 B_i dB_i, \quad Q_{M_{tO}} = \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij} \pi_i \pi_j \int_0^1 B_i B_j$$

The limit processes $Q_{M_{tO}}$ appearing in the limit distributions of t_{GT} and t_{OT} are the sums of the individual elements in the corresponding limit processes Q_{A_G}, Q_{B_G}, Q_{A_O} and $Q_{M_{FO}}$ defined in Theorem 2.1, which constitute the statistics K_{GT} and K_{OT} developed for the heterogenous panels.[7] The limit distributions of the t-statistics t_{GT} and t_{OT} are also non-standard and suffer from nuisance parameter dependency, as in the cases with the F-tests and K-statistics. Hence it is not possible to use these statistics for inference as they stand. In the next section, we consider boostrapping the panel unit root tests proposed in this section to resolve the nuisance prameter dependency problem and to provide a valid basis for inference in nonstationary panels with cross-sectional dependency.

3. Bootstrap Unit Root Tests for Dependent Panels

In this section, we consider the sieve bootstraps for the various panel unit root tests, F_{GT}, F_{OT}, K_{GT}, K_{OT}, t_{GT} and t_{OT} considered in the previous section. In particular, we establish the asymptotic validity of the bootstrapped tests by showing bootstrap consistency of the tests. We use the conventional notation $*$ to signify the bootstrap samples, and use $\mathbf{P}^*$ and $\mathbf{E}^*$ to denote, respectively, the probability and expectation conditional upon the realization of the original sample. While developing the asymptotic theories for the bootstrapped tests, we also discuss various issues and problems arising in practical implementation of the sieve bootstrap methodology in this section.

To construct the bootstrapped tests, we first generate the bootstrap samples for (ε_{it}^*), (u_{it}^*) and (y_{it}^*). For the generation of (ε_{it}^*), we need to make sure that the dependence structure among cross-sectional units, $i = 1, \ldots, N$, is preserved. To do so, we generate the N-dimensional vector $(\varepsilon_t^*) = (\varepsilon_{1t}^*, \ldots, \varepsilon_{Nt}^*)'$ by resampling from the centered residual vectors $(\tilde{\varepsilon}_t^p)$ defined in (13) from the regression (12). That is, obtain (ε_t^*) from the empirical distribution of

$$\left(\tilde{\varepsilon}_t^p - \frac{1}{T} \sum_{t=1}^{T} \tilde{\varepsilon}_t^p \right)_{t=1}^{T}$$

The bootstrap samples (ε_t^*) constructed as such will, in particular, satisfy $\mathbf{E}^* \varepsilon_t^* = 0$ and $\mathbf{E}^* \varepsilon_t^* \varepsilon_t^* = \tilde{\Sigma}$.[8]

[7] Levin and Lin (1992,1993) considers t-statistics for homogeneous panels under cross-sectional independency. Consequently, they can apply N-asymptotics after the limit as T tends to infinity is taken, and derive the limit distribution that is the standard normal. Their theory, however, does not extend to our statistics, since we allow for dependency across cross-sectional units.

[8] Of course, we may resample ε_{it}^*'s individually from the $\tilde{\varepsilon}_{it}^{p_i}$'s for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. In this case, preserving the original correlation structure among the cross-sectional units needs more care. We basically need to pre-whiten $\tilde{\varepsilon}_{it}^{p_i}$'s before resampling, and then re-color the resamples to recover the correlation structure. More specifically, we first pre-whiten $\tilde{\varepsilon}_{it}^{p_i}$'s by pre-multiplying $\tilde{\Sigma}^{-1/2}$ to $\tilde{\varepsilon}_t^p = (\tilde{\varepsilon}_{1t}^{p_1}, \ldots, \tilde{\varepsilon}_{Nt}^{p_N})'$, for $t = 1, \ldots, T$. Next, generate ε_{it}^*'s by resampling from the pre-whitened $\tilde{\varepsilon}_{it}^{p_i}$'s, and then re-color them by pre-multiplying $\tilde{\Sigma}^{1/2}$ to $\varepsilon_t^* = (\varepsilon_{1t}^*, \ldots, \varepsilon_{Nt}^*)'$ to restore the original dependence structure.

Next, we generate (u_{it}^*) recursively from (ε_{it}^*) as

$$u_{it}^* = \tilde{\alpha}_{i,1}^{p_i} u_{i,t-1}^* + \cdots + \tilde{\alpha}_{i,p_i}^{p_i} u_{i,t-p_i}^* + \varepsilon_{it}^* \tag{19}$$

where $(\tilde{\alpha}_{i,1}^{p_i}, \ldots, \tilde{\alpha}_{i,p_i}^{p_i})$ are the coefficient estimates from the fitted regression (12). Initialization of (u_{it}^*) is unimportant for our subsequent theoretical development, though it may play an important role in finite samples.[9] The coefficient estimates $(\tilde{\alpha}_{i,1}^{p_i}, \ldots, \tilde{\alpha}_{i,p_i}^{p_i})$ used in (19) may be obtained from estimating (12) by the Yule-Walker method instead of the OLS. The two methods are asymptotically equivalent. However, in small samples the Yule-Walker method may be preferred to the OLS, since it always yields an invertible autoregression, thereby ensuring the stationarity of the process (u_{it}^*). See Brockwell and Davis (1991, Sections 8.1 and 8.2). However, the probability of having the noninvertibility problem in the OLS estimation becomes negligible as the sample size increases.

Finally, obtain (y_{it}^*) by taking partial sums of (u_{it}^*), viz.

$$y_{it}^* = y_{i0}^* + \sum_{k=1}^{t} u_{ik}^*$$

with some initial initial value y_{i0}^*. Notice that the bootstrap samples (y_{it}^*) are generated with the unit root imposed. The samples generated according to the unrestricted regression (1) will not necessarily have the unit root property, and this will make the subsequent bootstrap procedure inconsistent as shown in Basawa *et al* (1991). The choice of the initial value y_{i0}^* does not affect the asymptotics as long as it is stochastically bounded. Therefore, we simply set it equal to zero for the subsequent analysis in this section.

We may obtain the Beveridge-Nelson representations for the bootstrapped series (u_{it}^*) and (y_{it}^*) similar to those for (u_{it}) and (y_{it}) given in (5) and (6) in the previous section. Let $\tilde{\alpha}_i(1) = 1 - \sum_{k=1}^{p_i} \tilde{\alpha}_{i,k}^{p_i}$. Then it is indeed easy to get

$$\begin{aligned} u_{it}^* &= \frac{1}{\tilde{\alpha}_i(1)} \varepsilon_{it}^* + \sum_{k=1}^{p_i} \frac{\sum_{j=k}^{p_i} \tilde{\alpha}_{i,j}^{p_i}}{\tilde{\alpha}_i(1)} (u_{i,t-k}^* - u_{i,t-k+1}^*) \\ &= \tilde{\pi}_i(1)\varepsilon_{it}^* + (\bar{u}_{i,t-1}^* - \bar{u}_{it}^*) \end{aligned}$$

where $\tilde{\pi}_i(1) = 1/\tilde{\alpha}_i(1)$ and $\bar{u}_t^* = \tilde{\pi}_i(1) \sum_{k=1}^{p_i} (\sum_{j=k}^{p_i} \tilde{\alpha}_{i,j}^{p_i}) u_{i,t-k+1}^*$, and therefore,

$$y_{it}^* = \sum_{k=1}^{t} u_{ik}^* = \tilde{\pi}_i(1) w_{it}^* + (\bar{u}_{i0}^* - \bar{u}_{it}^*)$$

where $w_{it}^* = \sum_{k=1}^{t} \varepsilon_{ik}^*$.

For the development of the limit theories for the bootstrapped test statistics, we assume

[9] We may use the first p_i-values of (u_{it}) as the initial values of (u_{it}^*). The bootstrap samples (u_{it}^*) generated as such may not be stationary processes. Alternatively, we may generate a larger number, say $T + M$, of (u_{it}^*) and discard first M-values of (u_{it}^*). This will ensure that (u_{it}^*) become more stationary. In this case the initialization becomes unimportant, and we may therefore simply choose zeros for the initial values.

Assumption B1 Let (ε_t) be a sequence of iid random variables such that $\mathbf{E}\varepsilon_t = 0$, $\mathbf{E}\varepsilon_t\varepsilon_t' = \Sigma$ and $\mathbf{E}|\varepsilon_t|^r < \infty$ for some $r \geq 4$.

Assumption B2 Let $\pi_i(z) \neq 0$ for all $|z| \leq 1$, and $\sum_{k=0}^{\infty} |k|^s |\pi_{i,k}| < \infty$ for some $s \geq 1$, for all $i = 1, \ldots, N$.

Assumption B3a Let $p_i \to \infty$ and $p_i = o(T^\kappa)$ with $\kappa < 1/2$ as $T \to \infty$, for all $i = 1, \ldots, N$.

Assumption B3b Let $p_i = cn^\kappa$ for some constant c and $1/rs < \kappa < 1/2$, for all $i = 1, \ldots, N$.

The iid assumption in Assumption B1, instead of the martingale difference condition in Assumption A1, is made to make the usual bootstrap procedure meaningful. Assumption B2 is identical to Assumption A2. In the place of Assumption A3 for the expansion rate of AR order p_i's, we impose either Assumption B3a or B3b. Both Assumptions B3a and B3b are stronger than Assumption A3. We will impose the condition in Assumption B3a to prove the consistency of the bootstrap tests in the weak form, i.e., the convergence of conditional bootstrap distributions in probability. To establish the strong consistency or the a.s. convergence of conditional bootstrap distributions, we need a stronger condition in Assumption B3b. Notice that we only require $0 < \kappa < 1/2$, for the Gaussian model with $r = \infty$ or the finite order ARMA model with $s = \infty$. The condition is therefore not very stringent.

Conventions
(a) Assumptions B1, B2 and B3a together will be refered to as Assumption (W), with 'W' standing for *weak*, and the set of Assumptions B1, B2 and B3b will be called as Assumption (S), with 'S' for *strong*.
(b) We will use the symbol $o_p^*(1)$ to signify the bootstrap convergence in probability. For a sequence of bootstrapped random variables Z_n^*, for instance, $Z_n^* = o_p^*(1)$ a.s. and in $\mathbf{P}$ imply respectively that

$$\mathbf{P}^*\{|Z_n^*| > \delta\} \to 0 \ \text{ a.s. or in } \mathbf{P}$$

for any $\delta > 0$. Similarly, we will use the symbol $O_p^*(1)$ to denote the bootstrap version of the boundedness in probability. Needless to say, the definitions of $o_p^*(1)$ and $O_p^*(1)$ naturally extend to $o_p^*(c_n)$ and $O_p^*(c_n)$ for some nonconstant numerical sequence (c_n).

We need following lemmas for the derivation of the limit distributions for the sieve bootstrap panel unit root tests.

Lemma 3.1 Under Assumptions (W), we have

(a) $$\frac{1}{T}\sum_{t=1}^{T} y_{i,t-1}^* \varepsilon_{jt}^* = \tilde{\pi}_i(1)\,\frac{1}{T}\sum_{t=1}^{T} w_{i,t-1}^* \varepsilon_{jt}^* + o_p^*(1)$$

(b) $$\frac{1}{T^2}\sum_{t=1}^{T} y_{i,t-1}^* y_{j,t-1}^* = \tilde{\pi}_i(1)\tilde{\pi}_j(1)\,\frac{1}{T^2}\sum_{t=1}^{T} w_{i,t-1}^* w_{j,t-1}^* + o_p^*(1)$$

In the following lemma, we use an operator norm for matrices: if $C = (c_{ij})$ is a matrix, then we let $\|C\| = \max_x |Cx|/|x|$.

Lemma 3.2 Let $x_{it}^{*p_i} = (\triangle y_{i,t-1}^*, \ldots, \triangle y_{i,t-p_i}^*)'$. Then we have

(a) $\mathbf{E}^* \left\| \left(\frac{1}{T} \sum_{t=1}^T x_{it}^{*p_i} x_{it}^{*p_i\prime} \right)^{-1} \right\| = O_p(1)$ or $O(1)$ a.s. under Assumptions (W) and (S), respectively, for all $i = 1, \ldots, N$.

(b) $\mathbf{E}^* \left| \sum_{t=1}^T x_{it}^{*p_i} y_{j,t-1}^* \right| = O(Tp_i^{1/2})$ a.s. under Assumption (W), for all $i, j = 1, \ldots, N$.

(c) $\mathbf{E}^* \left| \sum_{t=1}^T x_{it}^{*p_i} \varepsilon_{jt}^* \right| = O(T^{1/2} p_i^{1/2})$ a.s. under Assumptions (W), for all $i, j = 1, \ldots, N$.

3.1 Bootstrap Unit Root Tests for Heterogeneous Panels

To construct the bootstrapped tests, we consider the following bootstrap version of the augmented autoregression (8) which was used to construct the sample test statistics

$$\triangle y_{it}^* = \alpha_i y_{i,t-1}^* + \sum_{k=1}^{p_i} \alpha_{i,k} \triangle y_{i,t-k}^* + \varepsilon_{it}^* \tag{20}$$

and write this in matrix form

$$\triangle y^* = Y_\ell^* \alpha + X_p^* \beta_p + \varepsilon^* \tag{21}$$

where the variables are defined in the same manner as in the regression (9) with

$$y_{\ell,i}^* = \begin{pmatrix} y_{i,0}^* \\ \vdots \\ y_{i,T-1}^* \end{pmatrix}, \quad X_i^{*p_i} = \begin{pmatrix} x_{i1}^{*p_i\prime} \\ \vdots \\ x_{iT}^{*p_i\prime} \end{pmatrix} \quad \text{and} \quad \varepsilon_i^* = \begin{pmatrix} \varepsilon_{i,1}^* \\ \vdots \\ \varepsilon_{i,T}^* \end{pmatrix}$$

for $i = 1, \ldots, N$.

We test for the unit root hypothesis $\alpha = 0$ in (21), using the bootstrap versions of F-type tests that are defined analogously as the sample F-type tests considered in the previous section. The bootstrap F-tests are constructed from the bootstrap GLS and OLS estimators of α in the bootstrap augmented AR regression (21). More explicitly, we define the boostrap GLS-based F-test as

$$F_{GT}^* = A_{GT}^{*\prime} B_{GT}^{*-1} A_{GT}^* \tag{22}$$

analogously as the sample GLS-based F-test F_{GT} given in (10), where

$$\begin{aligned}
A_{GT}^* &= Y_\ell^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)\varepsilon^* - Y_\ell^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)X_p^* \left(X_p^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)X_p^* \right)^{-1} X_p^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)\varepsilon^* \\
B_{GT}^* &= Y_\ell^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)Y_\ell^* - Y_\ell^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)X_p^* \left(X_p^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)X_p^* \right)^{-1} X_p^{*\prime}(\tilde\Sigma^{-1} \otimes I_T)Y_\ell^*
\end{aligned}$$

The bootstrap OLS-based F-test is also defined analogously as the sample OLS-based F-test F_{OT} defined in (11), viz.

$$F^*_{OT} = A^{*\prime}_{OT} M^{*-1}_{FOT} A^*_{OT} \tag{23}$$

where

$$\begin{aligned} A^*_{OT} &= Y^{*\prime}_\ell \varepsilon^* - Y^{*\prime}_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p \varepsilon^*_p \\ M^*_{FOT} &= Y^{*\prime}_\ell (\tilde\Sigma \otimes I_T) Y^*_\ell - Y^{*\prime}_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p (\tilde\Sigma \otimes I_T) Y^*_\ell \\ &\quad - Y^{*\prime}_\ell (\tilde\Sigma \otimes I_T) X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p Y^*_\ell \\ &\quad + Y^{*\prime}_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p (\tilde\Sigma \otimes I_T) X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p Y^*_\ell \end{aligned}$$

The bootstrap F-statistics F^*_{GT} and F^*_{OT} given in (22) and (23) involve the covariance matrix estimator $\tilde\Sigma$ defined below (13). The estimate $\tilde\Sigma$ is the population parameter for the bootstrap samples (ε^*_t), corresponding to Σ for the original samples (ε_t). We may of course use the bootstrap estimate $\tilde\Sigma^*$, say, for the construction of the statistics F^*_{GT} and F^*_{OT} for each bootstrap iteration. The two versions of the bootstrap tests are asymptotically equivalent at least for the first order asymptotics, and we use $\tilde\Sigma$ in the construction of the bootsrap tests for convenience.[10]

We now present the limit theory for the bootstrap F-type tests F^*_{GT} and F^*_{OT} in

Theorem 3.1 We have as $T \to \infty$,
(a) $F^*_{GT} \;\to_{d^*}\; Q'_{A_G} Q^{-1}_{B_G} Q_{A_G}$ in $\mathbf{P}$ or a.s.
(b) $F^*_{OT} \;\to_{d^*}\; Q'_{A_O} Q^{-1}_{M_{FO}} Q_{A_O}$ in $\mathbf{P}$ or a.s.
respectively under Assumption (W) or (S), where Q_{A_G}, Q_{B_G}, Q_{A_O} and $Q_{M_{FO}}$ are defined in Theorem 2.1.

The results in Part (a) and (b) above show that the bootstrap F-statistics F^*_{GT} and F^*_{OT} have the same limit distributions as the corresponding sample F-statistics F_{GT} and F_{OT} given in Theorem 2.1. This establishes the asymptotic validity of the boostrap tests F^*_{GT} and F^*_{OT}.

The bootstrap K-statistics are constructed from the bootstrap samples in the analogous manner in which the sample K-statistics are defined in (15) and (16).

$$\begin{aligned} K^*_{GT} &= (A^*_{GT} .\!* 1\{\hat\alpha^*_{GT} \le 0\})' \, B^{*-1}_{GT} \, (A^*_{GT} .\!* 1\{\hat\alpha^*_{GT} \le 0\}) \\ K^*_{OT} &= (A^*_{OT} .\!* 1\{\hat\alpha^*_{OT} \le 0\})' \, M^{*-1}_{FOT} \, (A^*_{OT} .\!* 1\{\hat\alpha^*_{OT} \le 0\}) \end{aligned} \tag{24}$$

and their limit theories are given in

Corollary 3.1 We have as $T \to \infty$,
(a) $K^*_{GT} \;\to_{d^*}\; (Q_{A_G} .\!* 1\{Q^{-1}_{B_G} Q_{A_G} \le 0\})' Q^{-1}_{B_G} (Q_{A_G} .\!* 1\{Q^{-1}_{B_G} Q_{A_G} \le 0\})$ in $\mathbf{P}$ or a.s.

[10] The bootstrap tests based on the bootstrap estimate $\tilde\Sigma^*$ may be better for higher order asymptotics, since they more closely mimic the sample statistics than the bootstrap tests based on the population parameter $\tilde\Sigma$. The statistics considered in the paper are, however, non-pivotal and therefore the higher order asymptotics are irrelevant here.

(b) $K^*_{OT} \;\to_{d^*}\; (Q_{A_O} .* 1\{Q^{-1}_{B_O} Q_{A_O} \leq 0\})' Q^{-1}_{M_{FO}} (Q_{A_O} .* 1\{Q^{-1}_{B_O} Q_{A_O} \leq 0\})$ in $\mathbf{P}$ or a.s.
respectively under Assumption (W) or (S), where Q_{A_G}, Q_{B_G}, Q_{A_O}, $Q_{M_{FO}}$ and Q_{B_O} are defined in Theorem 2.1 and Corollary 2.1.

Corollary 3.1 shows that the bootstrap K-statistics K^*_{GT} and K^*_{OT} have the same limiting distributions as the corresponding sample K-statistics K_{GT} and K_{GT} given in Corollary 2.1, thereby proving the asymptotic validity of the bootstrap K-statistics.

3.2 Bootstrap Unit Root Tests for Homogeneous Panels

The bootstrap t-statistics are also constructed in an analogous manner as we constructed the sample t-statistics, t_{GT} and t_{OT}, in Section 2.2. Thus, we consider the homogeneous panel of the bootstrap samples, with $\alpha_1 = \cdots = \alpha_N = \alpha$ imposed, and compute the t-statistics from the corresponding augemented AR, which is written in matrix form as

$$\triangle y^* = y^*_\ell \alpha + X^*_p \beta_p + \varepsilon^* \tag{25}$$

The variables appearing in the above regression are defined in the same way as in the augmented AR in matrix form for the bootstrap heterogeneous model (21), except that here we have an $(NT \times 1)$-vector $y^*_\ell = (y^{*\prime}_{\ell,1}, \ldots, y^{*\prime}_{\ell,N})'$ in the place of the $(NT \times N)$-matrix Y^*_ℓ and the parameter α is now a scalar instead of an $(N \times 1)$-vector.

The bootstrapped GLS and OLS based t-statistics are based on the GLS and OLS estimator of α in the homogeneous augmented AR (25), and are given by

$$t^*_{GT} = a^*_{GT} b^{*-1/2}_{GT} \quad \text{and} \quad t^*_{OT} = a^*_{OT} M^{*-1/2}_{tOT} \tag{26}$$

where

$$\begin{aligned}
a^*_{GT} &= y^{*\prime}_\ell (\tilde\Sigma^{-1} \otimes I_T)\varepsilon^* - y^{*\prime}_\ell (\tilde\Sigma^{-1} \otimes I_T) X^*_p (X^{*\prime}_p (\tilde\Sigma^{-1} \otimes I_T) X^*_p)^{-1} X^{*\prime}_p (\tilde\Sigma^{-1} \otimes I_T)\varepsilon^* \\
b^*_{GT} &= y^{*\prime}_\ell (\tilde\Sigma^{-1} \otimes I_T) y^*_\ell - y^{*\prime}_\ell (\tilde\Sigma^{-1} \otimes I_T) X^*_p (X^{*\prime}_p (\tilde\Sigma^{-1} \otimes I_T) X^*_p)^{-1} X^{*\prime}_p (\tilde\Sigma^{-1} \otimes I_T) y^*_\ell \\
a^*_{OT} &= y^{*\prime}_\ell \varepsilon^* - y^*_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p \varepsilon^* \\
M^*_{tOT} &= y^{*\prime}_\ell (\tilde\Sigma \otimes I_T) y^*_\ell - 2 y^{*\prime}_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p (\tilde\Sigma \otimes I_T) y^*_\ell \\
&\quad + y^{*\prime}_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p (\tilde\Sigma \otimes I_T) X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p y^*_\ell
\end{aligned}$$

The limit distributions of t^*_{GT} and t^*_{OT} are given in

Theorem 3.2 We have as $T \to \infty$,

(a) $t^*_{GT} \;\to_{d^*}\; Q_{a_G} Q^{-1/2}_{b_G}$ in $\mathbf{P}$ or a.s.

(b) $t^*_{OT} \;\to_{d^*}\; Q_{a_O} Q^{-1/2}_{M_{tO}}$ in $\mathbf{P}$ or a.s.

respectively under Assumption (W) or (S), where Q_{a_G}, Q_{b_G}, Q_{a_O} and $Q_{M_{tO}}$ are defined in Theorem 2.2.

The results in Theorem 3.2 show that the bootstrap t-statistics t^*_{GT} and t^*_{OT} have the limit distributions that are equivalent to those of the sample t-statistics t_{GT} and t_{OT} given in Theorem 2.2, thereby establishing the asymptotically validity of the bootstrap t-statistics.

4. Simulations

We conduct a set of simulations to investigate the finite sample performance of the bootstrap panel unit root tests, F^*_{GT}, F^*_{OT}, K^*_{GT}, K^*_{OT}, t^*_{GT} and t^*_{OT}, proposed in the paper. For the simulation, we consider the (y_t) given by the model (1) with (u_t) generated as either AR(1) or MA(1) processes, viz.,

$$\text{(A)} \quad u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it}$$
$$\text{(B)} \quad u_{it} = \varepsilon_{it} + \theta_i \varepsilon_{i,t-1}$$

The innovations $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ that generate $u_t = (u_{1t}, \ldots, u_{Nt})'$ are drawn from an N-dimensional multivariate normal distribution with mean zero and covariance matrix Σ.[11] The AR and MA coefficients, ρ_i's and θ_i's, used in the generation of the errors (u_{it}) are drawn randomly from the uniform distribution. More specifically, $\rho_i \sim$ Uniform[0.2,0.4] and $\theta_i \sim$ Uniform$[-0.4, 0.4]$.[12]

The parameter values for the $(N \times N)$ covariance matrix $\Sigma = (\sigma_{ij})$ are also randomly drawn, but with particular attention. To ensure that Σ is a symmetric positive definite matrix and to avoid the near singularity problem, we generate Σ via following steps:

(1) Generate an $(N \times N)$ matrix U from Uniform[0,1].
(2) Construct from U an orthogonal matrix $H = U(U'U)^{-1/2}$.
(3) Generate a set of N eigen values, $\lambda_1, \ldots, \lambda_N$. Let $\lambda_1 = r > 0$ and $\lambda_N = 1$ and draw $\lambda_2, \ldots, \lambda_{N-1}$ from Uniform[r,1].
(4) Form a diagonal matrix Λ with $(\lambda_1, \ldots, \lambda_N)$ on the diagonal.
(5) Construct the covariance matrix Σ as a spectral representation $\Sigma = H\Lambda H'$.

The covariance matrix constructed this way will surely be symmetric and nonsingular with eigenvalues taking values from r to 1. We set the maximum eigenvalue at 1 since the scale does not matter. The ratio of the minimum eigenvalue to the maximum is therefore determined by the same parameter r. The covariance matrix becomes singular as r tends to zero, and becomes spherical as r approaches to 1. For the simulations, we set r at $r = 0.1$.[13]

For the test of the unit root hypothesis, we set $\alpha_i = 0$ for all $i = 1, \ldots, N$, and investigate the finite sample sizes in relation to the corresponding nominal test sizes. To examine the rejection probabilities of the tests under the alternative of stationarity, we generate α_i's randomly from Uniform$[-0.8, 0]$. The model is thus heterogenous under the alternative. The finite sample performance of the bootstrap tests are compared with that of the t-bar statistics by Im, Pesaran and Shin (1997), which is based on the average of the individual t-statistics computed from the sample ADF regressions (8) with mean and variance

[11] The simulation model for the case (B) is generated from an MA(1) process (u_{it}), which can be represented as an infinite order AR process. Using the lag order p_i selected by AIC, we approximate (u_{it}) by an AR(p_i) process as in (12). The approximated autoregression is then estimated by the Yule-Walker method.

[12] Maddala and Wu (1996) and Im, Pesaran and Shin (1997) also generate parameters for their simulation models radomly from uniform distributions.

[13] Our bootstrap tests do not seem to depend on the the value of r, but the t-bar statistics does. Though we do not report the details, we observe from a set of simulations that the t-bar tends to have higher rejection probabilities when r is close to 0, and that it seems to have substantial size distortions even when Σ is nearly spherical with $r = 0.99$.

modifications. More explicitly, the t-bar statistics is defined as

$$t\text{-bar} = \frac{\sqrt{N}(\bar{t}_N - N^{-1}\sum_{i=1}^{N}\mathbf{E}(t_i))}{\sqrt{N^{-1}\sum_{i=1}^{N}\text{var}(t_i)}}$$

where t_i is the t-statistics for testing $\alpha_i = 0$ for the i-th sample ADF regression (8), and $\bar{t}_N = N^{-1}\sum_{i=1}^{N} t_i$. The values of the expectation and variance, $\mathbf{E}(t_i)$ and $\text{var}(t_i)$, for each individual t_i depend on T and the lag order p_i, and computed via simulations from independent nomal samples. Table 2 in Im, Pesaran and Shin (1997) tabulates the values of $\mathbf{E}(t_i)$ and $\text{var}(t_i)$ for $T = 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 100$ and for $p_i = 1, \ldots, 8$.

The panels with the cross-sectional dimensions $N = 5, 20$ and the time series dimensions $T = 50, 100$ are considered for the 1%, 5% and 10% size tests. Since we are using random parameter values, we simulate 20 times and report the ranges of the finite sample performances of the bootstrap tests. Each simulation run is carried out with 1,000 simulation iterations, each of which uses bootstrap critical values computed from 500 bootstrap repetitions. The simulation results for the t-bar statistics and our bootstrap tests F^*_{GT}, F^*_{OT}, K^*_{GT}, K^*_{OT}, t^*_{GT} and t^*_{OT} are reported in Tables A1-B2. Tables A1 and A2 reports, respectively, the finite sample sizes and powers of the tests for Case A with AR errors, and Tables B1 and B2 reports those for Case B with MA errors. For each statistics, we report the minimum, mean, median and maximum of the rejection probabilities under the null and under the alternative hypothesis.

As can be seen from Tables A1 and B1, the t-bar test suffers from serious size distortions. The direction of the size distortions is, however, not in one way. For the 1% tests, the t-bar statistics suffers from upward size distortions except for the MA case with $N = 5$, where the t-bar is slightly downward biased. The degree of the upward distortions seems to be higher for the AR case and increases as N gets large. For the 5% and 10% tests, the t-bar test is mostly downward biased except for the 5% test with $N = 20$, where the test is upward biased.[14] The downward distortion is more serious for the MA case with smaller $N = 5$. On the other hand, the finite sample sizes of the bootstrap tests are quite close to the nominal test sizes for both AR and MA cases and for all $N = 5{,}20$ and $T = 50{,}100$.

The bootstrap tests are more powerful than the t-bar statistics for most cases with the smaller $N = 5$, as can be seen from Tables A2 and B2. Indeed, for the 5% and 10% tests all of our bootstrap tests have higher rejection probabilities than the t-bar for both AR and MA cases. For 1% tests, only the GLS based bootstrap tests F^*_{GT} and K^*_{GT} perform better than the t-bar. As the number of the cross-sectional units increases to $N = 20$, the performance of the t-bar statistics improves. With the smaller number of observations over time $T = 50$, it actually performs better than the bootstrap tests except the OLS based t-statistics t^*_{OT}, but the difference becomes negligible as T increases.

Among the bootstrap tests, the GLS based tests, F^*_{GT} and K^*_{GT}, are more powerful than the OLS based tests, F^*_{OT} and K^*_{OT}, for the smaller $N = 5$, but for the larger $N = 20$, the advantage from the GLS efficiency vanishes. This is perhaps due to the error involved in

[14]The downward size distortions of the t-bar statistics have been well noted in several simulation works. Maddala and Wu (1996), for example, report that the t-bar statistics suffers from substantial downward size distortions in the presence of cross-correlations among the cross-sectional units.

estimating large dimensional covariance matrix. For t-type tests, the OLS based t-statistics t^*_{OT} is indeed noticeably more powerful than its GLS couterpart t^*_{GT} when the larger $N = 20$ is used. They are also more powerful than the F-type tests and K-statistics in this case. The advantage of the one-tail tests based on the homogeneous panels appears to be quite important in finite samples.

The K-statistics was proposed as alternatives to the two-sided F-type tests to come up with more powerful tests for the unit roots against the one-way alternative of the stationarity. The simulation results in Tables A2 and B2, however, show that the improvement the K-statistics make over the F-type tests are not noticeable. One possible reason is that the finite sample distributions of the $\hat{\alpha}_{GT}$ and $\hat{\alpha}_{OT}$, upon which the modifications for the K-statistics are made, are skew to the left so much that the modification does not have actual effect. Thus, one may correct for the biases in the distributions of $\hat{\alpha}_{GT}$ and $\hat{\alpha}_{OT}$ before applying the modifications in (14). This can be done by carrying out a nested bootstrap. We do not pursue this in this paper due to the computation time, but will report in a future work.

All bootstrap tests are more powerful for the case with the smaller $N = 5$ and the larger $T = 100$ than the cases with the larger $N = 20$ and the smaller $T = 50$. This is because our bootstrap tests are T-asymptotic tests, which will work better for a large T. The t-bar test is, however, noticeably more powerful for the cases with $N = 20$ and $T = 50$ than for the cases with $N = 5$ and $T = 100$. This indicates that the t-bar test works much better for panels with larger number of N, which is expected since the test is based on the average of individual tests.

4. Conclusion

There has been much recent empirical and theoretical econometric work on models with nonstationary panel data. In particular, much attention has been paid to the development and implementation of the panel unit root tests which have been used frequently to test for various covergence theories, such as growth covergence theories and purchasing power parity hypothesis. A variety of tests have been proposed, including the tests proposed by Levin and Lin (1993) and Im, Pesaran and Shin (1997) that appear to be most commonly used. All the existing tests, however, assume the independence across cross-sectional units, which is quite restrictive. Cross-sectional dependency seems indeed quite apparent for most of interesting panel data.

In the paper, we investigate various unit root tests for panel models which explicitly allow for the cross-correlation across cross-sectional units as well as heterogeneous serial dependence. The limit theories for the panel unit root tests are derived by passing the number of time series observations T to infinity with the number of cross-sectional units N fixed. As expected the limit distributions of the tests are nonstandard and depend heavily on the nuisance parameters, rendering the standard inferential procedure invalid. To overcome the inferential difficulty of the panel unit root tests in the presence of cross-sectional dependency, we propose to use the bootstrap method. Limit theories for the bootstrap tests are developed, and in particular their asymptotic validity is established by proving the consistency of the boostrap tests. The simulations show that the bootstrap panel unit root tests perform well in finite samples relative to the t-bar statistics by Im, Pesaran and Shin (1997).

5. Appendix: Mathematical Proofs

The following lemmas provide asymptotic results for the sample moments appearing in the sample test statistics F_{GT}, F_{OT}, K_{GT}, K_{OT}, t_{GT} and t_{OT} defined in (10), (11), (15), (16) and (18).

Lemma A1 Under Assumptions A1, A2 and A3, we have

(a) $\frac{1}{T}\sum_{t=1}^{N} y_{i,t-1}\varepsilon_{jt}^{p_j} = \pi_i(1)\,\frac{1}{T}\sum_{t=1}^{T} w_{i,t-1}\varepsilon_{jt} + o_p(1)$, for all $i,j=1,\ldots,N$

(b) $\frac{1}{T^2}\sum_{t=1}^{T} y_{i,t-1}y_{j,t-1} = \pi_i(1)\pi_j(1)\,\frac{1}{T^2}\sum_{t=1}^{T} w_{i,t-1}w_{j,t-1} + o_p(1)$, for all $i,j=1,\ldots,N$

(c) $\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t^p\varepsilon_t^{p\prime} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t' + o_p(1)$

Proof of Lemma A1

Part (a) The stated results follow immediately if we apply the results in Lemma A1 (a) of Chang and Park (1999) to each (i,j) pair, for $i,j=1,\ldots,N$.

Part (b) The stated result follows directly from Phillips and Solo (1992).

Part (c) Let

$$Q_T = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^p\varepsilon_t^{p\prime} - \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t'$$

Then for each (i,j)-element of Q, the following holds

$$\begin{aligned} Q_{T,ij} &= \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^{p_i}\varepsilon_{jt}^{p_j} - \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}\varepsilon_{jt} \\ &= \frac{1}{T}\sum_{t=1}^{T}(\varepsilon_{it}^{p_i}-\varepsilon_{it})\varepsilon_{jt}^{p_j} + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}(\varepsilon_{jt}^{p_j}-\varepsilon_{j,t}) \\ &= o_p(p_i^{-s}) + o_p(p_j^{-s}) \end{aligned}$$

for all $i,j=1,\ldots,N$, due to Lemma A1 (c) in Chang and Park (1999). Now the stated result is immediate.

Lemma A2 Under Assumptions A1, A2 and A3, we have

(a) $\left\| \left(\frac{1}{T}\sum_{t=1}^{T} x_{it}^{p_i}x_{it}^{p_i\prime}\right)^{-1} \right\| = O_p(1)$, for all p_i and $i=1,\ldots,N$

(b) $\left|\sum_{t=1}^{T} x_{it}^{p_i}y_{j,t-1}\right| = O_p(Tp_i^{1/2})$, for all $i,j=1,\ldots,N$

(c) $\left|\sum_{t=1}^{T} x_{it}^{p_i}\varepsilon_{jt}^{p_j}\right| = O_p(T^{1/2}p_i^{1/2}) + o_p(Tp_i^{1/2}p_j^{-s})$, for all $i,j=1,\ldots,N$.

Proof of Lemma A2 The stated result in Part (a) follows directly from the application of the result in Lemma A2 (a) for each $i = 1, \ldots, N$, and those in Parts (b) and (c) are easily obtained using the results in Lemma A2 (b) and (c) of Chang and Park (1999) for each (i, j) pair for $i, j = 1, \ldots, N$, with some obvious modification with respect to the orders p_i's of the AR approximations involved.

Proof of Theorem 2.1

Part (a) We begin by writing out explicitly the component sample moments appearing in A_{GT} and B_{GT} defined below (11).

$$
\begin{aligned}
Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)Y_\ell &= \begin{pmatrix} y_{\ell,1}' & & 0 \\ & \ddots & \\ 0 & & y_{\ell,N}' \end{pmatrix} \begin{pmatrix} \tilde{\sigma}^{11} I_T & \cdots & \tilde{\sigma}^{1N} I_T \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{11} I_T & \cdots & \tilde{\sigma}^{1N} I_T \end{pmatrix} \begin{pmatrix} y_{\ell,1} & & 0 \\ & \ddots & \\ 0 & & y_{\ell,N} \end{pmatrix} \\
&= \begin{pmatrix} \tilde{\sigma}^{11} \sum_{t=1}^{T} y_{1,t-1}^2 & \cdots & \tilde{\sigma}^{1N} \sum_{t=1}^{T} y_{1,t-1} y_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{N1} \sum_{t=1}^{T} y_{N,t-1} y_{1,t-1} & \cdots & \tilde{\sigma}^{NN} \sum_{t=1}^{T} y_{N,t-1}^2 \end{pmatrix}
\end{aligned} \tag{27}
$$

and

$$
\begin{aligned}
X_p'(\tilde{\Sigma}^{-1} \otimes I_T)Y_\ell &= \begin{pmatrix} X_1^{p_1\prime} & & 0 \\ & \ddots & \\ 0 & & X_N^{p_N\prime} \end{pmatrix} \begin{pmatrix} \tilde{\sigma}^{11} I_T & \cdots & \tilde{\sigma}^{1N} I_T \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{11} I_T & \cdots & \tilde{\sigma}^{1N} I_T \end{pmatrix} \begin{pmatrix} y_{\ell,1} & & 0 \\ & \ddots & \\ 0 & & y_{\ell,N} \end{pmatrix} \\
&= \begin{pmatrix} \tilde{\sigma}^{11} \sum_{t=1}^{T} x_{1t}^{p_1} y_{1,t-1} & \cdots & \tilde{\sigma}^{1N} \sum_{t=1}^{T} x_{1t}^{p_1} y_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{N1} \sum_{t=1}^{T} x_{Nt}^{p_N} y_{1,t-1} & \cdots & \tilde{\sigma}^{NN} \sum_{t=1}^{T} x_{Nt}^{p_N} y_{N,t-1} \end{pmatrix}
\end{aligned} \tag{28}
$$

where $\tilde{\sigma}^{ij}$ denotes (i, j)-element of the inverse covariance matrix estimate $\tilde{\Sigma}^{-1}$. Similarly, we have

$$
X_p'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p = \begin{pmatrix} \tilde{\sigma}^{11} \sum_{t=1}^{T} x_{1t}^{p_1} \varepsilon_{1t}^{p_1} + & \cdots & + \tilde{\sigma}^{1N} \sum_{t=1}^{T} x_{1t}^{p_1} \varepsilon_{Nt}^{p_N} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{N1} \sum_{t=1}^{T} x_{Nt}^{p_N} \varepsilon_{1t}^{p_1} + & \cdots & + \tilde{\sigma}^{NN} \sum_{t=1}^{T} x_{Nt}^{p_N} \varepsilon_{Nt}^{p_N} \end{pmatrix}
$$

$$= \begin{pmatrix} \sum_{j=1}^{N} \tilde{\sigma}^{1j} \sum_{t=1}^{T} x_{1t}^{p_1} \varepsilon_{jt}^{p_j} \\ \vdots \\ \sum_{j=1}^{N} \tilde{\sigma}^{Nj} \sum_{t=1}^{T} x_{Nt}^{p_N} \varepsilon_{jt}^{p_j} \end{pmatrix} \tag{29}$$

$$Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p = \begin{pmatrix} \tilde{\sigma}^{11} \sum_{t=1}^{T} y_{1,t-1}\varepsilon_{1t}^{p_1} + & \cdots & + \tilde{\sigma}^{1N} \sum_{t=1}^{T} y_{1,t-1}\varepsilon_{Nt}^{p_N} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{N1} \sum_{t=1}^{T} y_{N,t-1}\varepsilon_{1t}^{p_1} + & \cdots & + \tilde{\sigma}^{NN} \sum_{t=1}^{T} y_{N,t-1}\varepsilon_{Nt}^{p_N} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{N} \tilde{\sigma}^{1j} \sum_{t=1}^{T} y_{1,t-1} \varepsilon_{jt}^{p_j} \\ \vdots \\ \sum_{j=1}^{N} \tilde{\sigma}^{Nj} \sum_{t=1}^{T} y_{N,t-1} \varepsilon_{jt}^{p_j} \end{pmatrix}$$

We now examine the stochastic orders of the components included in A_{GT} and B_{GT}. Let $\lambda(\cdot)$ denote eigenvalues of a matrix. We have

$$\lambda_{\min}(\tilde{\Sigma}^{-1} \otimes I_T) X_p' X_p \leq X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p$$

Notice that $\lambda_{\min}(\tilde{\Sigma}^{-1} \otimes I_T) = \lambda_{\min}(\tilde{\Sigma}^{-1})$ and $\lambda_{\min}(\tilde{\Sigma}^{-1}) = 1/\lambda_{\max}(\tilde{\Sigma})$. Then we have

$$\left(\frac{X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p}{T} \right)^{-1} \leq \lambda_{\max}(\tilde{\Sigma}) \left(\frac{X_p' X_p}{T} \right)^{-1} = O_p(1) \tag{30}$$

since $\lambda_{\max}(\tilde{\Sigma}) \to_p \lambda_{\max}(\Sigma) < \infty$, and

$$\left(\frac{X_p' X_p}{T} \right)^{-1} = \begin{pmatrix} \left(\frac{1}{T} \sum_{t=1}^{T} x_{1t}^{p_1} x_{1t}^{p_1\prime} \right)^{-1} & & 0 \\ & \ddots & \\ 0 & & \left(\frac{1}{T} \sum_{t=1}^{T} x_{Nt}^{p_N} x_{Nt}^{p_N\prime} \right)^{-1} \end{pmatrix} = O_p(1) \tag{31}$$

due to Lemma A2 (a). Moreover it follows from Lemma A2 (b) and (28) that

$$X_p'(\tilde{\Sigma}^{-1} \otimes I_T) Y_\ell = O_p(T\bar{p}^{1/2}) \tag{32}$$

where $\bar{p} = \max_{1 \leq i \leq N} p_i$, and from Lemma A2 (c) and (29) that

$$X_p'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p = O_p(T^{1/2}\bar{p}^{1/2}) + o_p(T\bar{p}^{1/2}\underline{p}^{-s}) \tag{33}$$

where $\underline{p} = \min_{1 \le i \le N} p_i$, as defined earlier. Notice that $\bar{p} = \underline{p} = o(T^{1/2})$ as $T \to \infty$ under Assumption 3.

It follows from (30), (32) and (33) that

$$\begin{aligned} & \left| Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right)^{-1} X_p'(\tilde{\Sigma}^{-1} \otimes I_T) \varepsilon_p \right| \\ & \le \left| Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right| \left\| \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right)^{-1} \right\| \left| X_p'(\tilde{\Sigma}^{-1} \otimes I_T) \varepsilon_p \right| \\ & = o_p(T \bar{p} \underline{p}^{-s}) + O_p(T^{1/2} \bar{p}) \end{aligned}$$

which implies

$$\frac{A_{GT}}{T} = \frac{Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) \varepsilon_p}{T} + o_p(1) = Q_{A_{GT}} + o_p(1) \tag{34}$$

due to Lemma A1 (a), where

$$Q_{A_{GT}} = \begin{pmatrix} \sum_{j=1}^{N} \tilde{\sigma}^{1j} \pi_1(1) \frac{1}{T} \sum_{t=1}^{T} w_{1,t-1} \varepsilon_{jt} \\ \vdots \\ \sum_{j=1}^{N} \tilde{\sigma}^{Nj} \pi_N(1) \frac{1}{T} \sum_{t=1}^{T} w_{N,t-1} \varepsilon_{jt} \end{pmatrix} + o_p(1)$$

Moreover, we have from (30) and (32) that

$$\begin{aligned} & \left| Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right)^{-1} X_p'(\tilde{\Sigma}^{-1} \otimes I_T) Y_\ell \right| \\ & \le \left| Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right| \left\| \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right)^{-1} \right\| \left| X_p'(\tilde{\Sigma}^{-1} \otimes I_T) Y_\ell \right| \\ & = O_p(T \bar{p}) \end{aligned}$$

which, together with Lemma A1 (b) and (27), gives

$$\frac{B_{GT}}{T^2} = \frac{Y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) Y_\ell}{T^2} + o_p(1) = Q_{B_{GT}} + o_p(1) \tag{35}$$

where

$$Q_{B_{GT}} = \begin{pmatrix} \tilde{\sigma}^{11} \pi_1(1)^2 \frac{1}{T^2} \sum_{t=1}^{T} w_{1,t-1}^2 & \cdots & \tilde{\sigma}^{1N} \pi_1(1) \pi_N(1) \frac{1}{T^2} \sum_{t=1}^{T} w_{1,t-1} w_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}^{N1} \pi_N(1) \pi_1(1) \frac{1}{T^2} \sum_{t=1}^{T} w_{N,t-1} w_{1,t-1} & \cdots & \tilde{\sigma}^{NN} \pi_N(1)^2 \frac{1}{T^2} \sum_{t=1}^{T} w_{N,t-1}^2 \end{pmatrix}$$

Using the asymptotic results in (34) and (35, we write

$$F_{GT} = \left(\frac{A_{GT}}{T} \right)' \left(\frac{B_{GT}}{T^2} \right)^{-1} \left(\frac{A_{GT}}{T} \right) = Q_{A_{GT}}' Q_{B_{GT}}^{-1} Q_{A_{GT}} + o_p(1)$$

Then the limit distribution of F_{GT} follows immediately from the invariance principle given in (4).

Part (b) We have from Lemma A2 (b) and (c) that

$$X_p'Y_\ell = \begin{pmatrix} \sum_{t=1}^T x_{1t}^{p_1} y_{1,t-1} & & 0 \\ & \ddots & \\ 0 & & \sum_{t=1}^T x_{Nt}^{p_N} y_{Nt,t-1} \end{pmatrix} = O_p(T\bar{p}^{1/2}) \tag{36}$$

$$X_p'\varepsilon_p = \begin{pmatrix} \sum_{t=1}^T x_{1t}^{p_1} \varepsilon_{1t}^{p_1} \\ \vdots \\ \sum_{t=1}^T x_{Nt}^{p_N} \varepsilon_{Ntt}^{p_N} \end{pmatrix} = O_p(T^{1/2}\bar{p}^{1/2}) + o_p(T\bar{p}^{1/2}\underline{p}^{-s}) \tag{37}$$

These together with (31) give

$$\left| Y_\ell' X_p (X_p'X_p)^{-1} X_p'\varepsilon_p \right| \le |Y_\ell' X_p| \left\| (X_p'X_p)^{-1} \right\| \left| X_p'\varepsilon_p \right| = o_p(T\bar{p}\underline{p}^{-s}) + O_p(T^{1/2}\bar{p})$$

which in turn gives

$$\frac{A_{OT}}{T} = \frac{Y_\ell'\varepsilon_p}{T} + o_p(1) = Q_{A_{OT}} + o_p(1) \tag{38}$$

due to Lemma A1 (a), where

$$Q_{A_{OT}} = \begin{pmatrix} \pi_1(1)\frac{1}{T}\sum_{t=1}^T w_{1,t-1}\varepsilon_{1t} \\ \vdots \\ \pi_N(1)\frac{1}{T}\sum_{t=1}^T w_{N,t-1}\varepsilon_{Nt} \end{pmatrix}$$

We have from (30) that

$$X_p'(\tilde{\Sigma} \otimes I_T)X_p \le \lambda_{\max(\tilde{\Sigma})}(X_p'X_p) = O_p(T) \tag{39}$$

We also have from Lemma A2 (b) that

$$X_p'(\tilde{\Sigma} \otimes I_T)Y_\ell = \begin{pmatrix} \tilde{\sigma}_{11}\sum_{t=1}^T x_{1t}^{p_1} y_{1,t-1} & \cdots & \tilde{\sigma}_{1N}\sum_{t=1}^T x_{1t}^{p_1} y_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}_{N1}\sum_{t=1}^T x_{Nt}^{p_N} y_{1,t-1} & \cdots & \tilde{\sigma}_{NN}\sum_{t=1}^T x_{Nt}^{p_N} y_{N,t-1} \end{pmatrix} = O_p(T\bar{p}^{1/2}) \tag{40}$$

where $\tilde{\sigma}_{ij}$ denotes (i,j)-element of the covariance matrix estimate $\tilde{\Sigma}$. Then we have

$$\left|Y_\ell' X_p (X_p' X_p)^{-1} X_p' (\tilde{\Sigma} \otimes I_T) Y_\ell\right| = O_p(T\bar{p})$$

and

$$\left|Y_\ell' X_p (X_p' X_p)^{-1} X_p' (\tilde{\Sigma} \otimes I_T) X_p (X_p' X_p)^{-1} X_p' Y_\ell\right| = O_p(T\bar{p})$$

which then give

$$\frac{M_{FOT}}{T^2} = \frac{Y_\ell' (\tilde{\Sigma} \otimes I_T) Y_\ell}{T^2} + o_p(1) = Q_{M_{FOT}} + o_p(1) \tag{41}$$

due to Lemma A1 (b), where

$$Q_{M_{FOT}} = \begin{pmatrix} \tilde{\sigma}_{11}\pi_1(1)^2 \frac{1}{T^2}\sum_{t=1}^T w_{1,t-1}^2 & \cdots & \tilde{\sigma}_{1N}\pi_1(1)\pi_N(1)\frac{1}{T^2}\sum_{t=1}^T w_{1,t-1}w_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}_{N1}\pi_N(1)\pi_1(1)\frac{1}{T^2}\sum_{t=1}^T w_{N,t-1}w_{1,t-1} & \cdots & \tilde{\sigma}_{NN}\pi_N(1)^2\frac{1}{T^2}\sum_{t=1}^T w_{N,t-1}^2 \end{pmatrix}$$

We now have from the results in (38) and (41) that

$$F_{OT} = \left(\frac{A_{OT}}{T}\right)' \left(\frac{M_{FOT}}{T^2}\right)^{-1} \left(\frac{A_{OT}}{T}\right) = Q_{A_{OT}}' Q_{M_{FOT}}^{-1} Q_{A_{OT}} + o_p(1)$$

from which the stated result follows immediately.

Proof of Corollary 2.1

Part(a) It follows from (34) and (35) that

$$T\hat{\alpha}_{GT} = \left(\frac{B_{GT}}{T^2}\right)^{-1} \left(\frac{A_{GT}}{T}\right) = Q_{B_{GT}}^{-1} Q_{A_{GT}} + o_p(1)$$

which implies

$$\begin{aligned} \frac{1}{T}\Big(A_{GT} .\!* 1\{\hat{\alpha}_{GT} \le 0\}\Big) &= \left(\frac{A_{GT}}{T} .\!* 1\left\{\frac{\hat{\alpha}_{GT}}{T} \le 0\right\}\right) \\ &= \left(\frac{A_{GT}}{T} .\!* 1\left\{T\hat{\alpha}_{GT} \le 0\right\}\right) \\ &= \left(Q_{A_{GT}} .\!* 1\left\{Q_{B_{GT}}^{-1} Q_{A_{GT}} \le 0\right\}\right) + o_p(1) \end{aligned}$$

Due to the above result and (35), we may write the K_{GT} statistics given in (15) as

$$\begin{aligned} K_{GT} &= \left(\frac{1}{T}\Big(A_{GT} .\!* 1\{\hat{\alpha}_{GT} \le 0\}\Big)\right)' \left(\frac{B_{GT}}{T^2}\right)^{-1} \left(\frac{1}{T}\Big(A_{GT} .\!* 1\{\hat{\alpha}_{GT} \le 0\}\Big)\right) \\ &= \left(Q_{A_{GT}} .\!* 1\left\{Q_{B_{GT}}^{-1} Q_{A_{GT}} \le 0\right\}\right)' Q_{B_{GT}}^{-1} \left(Q_{A_{GT}} .\!* 1\left\{Q_{B_{GT}}^{-1} Q_{A_{GT}} \le 0\right\}\right) + o_p(1) \end{aligned}$$

Now the stated result follows immediately from (4).

Part (b) From (31) and (36), we have

$$\left| Y_\ell' X_p (X_p' X_p)^{-1} X_p' Y_\ell \right| = O_p(T\bar{p})$$

which together with Lemma A1 (b) gives

$$\frac{B_{OT}}{T^2} = \frac{Y_\ell' Y_\ell}{T^2} + o_p(1) = Q_{B_{OT}} + o_p(1)$$

where

$$Q_{B_{OT}} = \begin{pmatrix} \pi_1(1)^2 \frac{1}{T^2}\sum_{t=1}^T w_{1,t-1}^2 & \cdots & \pi_1(1)\pi_N(1)\frac{1}{T^2}\sum_{t=1}^T w_{1,t-1}w_{N,t-1} \\ \vdots & \vdots & \vdots \\ \pi_N(1)\pi_1(1)\frac{1}{T^2}\sum_{t=1}^T w_{N,t-1}w_{1,t-1} & \cdots & \pi_N(1)^2 \frac{1}{T^2}\sum_{t=1}^T w_{N,t-1}^2 \end{pmatrix}$$

It follows from (38) and the above result that

$$T\hat{\alpha}_{OT} = \left(\frac{B_{OT}}{T^2}\right)^{-1}\left(\frac{A_{OT}}{T}\right) = Q_{B_{OT}}^{-1} Q_{A_{OT}} + o_p(1)$$

and

$$\frac{1}{T}\Big(A_{OT} .\!* 1\{\hat{\alpha}_{OT} \le 0\}\Big) = \Big(Q_{A_{OT}} .\!* 1\left\{Q_{B_{OT}}^{-1} Q_{A_{OT}} \le 0\right\}\Big) + o_p(1)$$

From this and the result in (41), we may express the statistics K_{OT} given in (16) as

$$\begin{aligned} K_{OT} &= \left(\frac{1}{T}\Big(A_{OT} .\!* 1\{\hat{\alpha}_{OT} \le 0\}\Big)\right)' \left(\frac{M_{FOT}}{T^2}\right)^{-1} \left(\frac{1}{T}\Big(A_{OT} .\!* 1\{\hat{\alpha}_{OT} \le 0\}\Big)\right) \\ &= \Big(Q_{A_{OT}} .\!* 1\left\{Q_{B_{OT}}^{-1} Q_{A_{OT}} \le 0\right\}\Big)' Q_{M_{FOT}}^{-1} \Big(Q_{A_{OT}} .\!* 1\left\{Q_{B_{OT}}^{-1} Q_{A_{OT}} \le 0\right\}\Big) + o_p(1) \end{aligned}$$

which is required for the stated result.

Proof of Theorem 2.2 The limit theories for the GLS and OLS based t-statistics t_{GT} and t_{OT} defined in (18) can be derived in the similar manner as we did for the F-type tests F_{GT} and F_{OT} in the proof of Theorem 2.1. We just have to take into account that the lagged level variables come in a $(NT \times 1)$-vector y_ℓ instead of the $(NT \times N)$-matrix Y_ℓ.

Part (a) Since

$$X_p'(\tilde{\Sigma}^{-1} \otimes I_T) y_\ell = \begin{pmatrix} \sum_{j=1}^N \tilde{\sigma}^{1j} \sum_{t=1}^T x_{1t}^{p_1} \varepsilon_{jt}^{p_j} \\ \vdots \\ \sum_{j=1}^N \tilde{\sigma}^{Nj} \sum_{t=1}^T x_{Nt}^{p_N} \varepsilon_{jt}^{p_j} \end{pmatrix} = O_p(T\bar{p}^{1/2})$$

due to Lemma A2 (b), it follows from (30) and (33) that

$$\left| y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right)^{-1} X_p'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p \right| = o_p(T\bar{p}\underline{p}^{-s}) + O_p(T^{1/2}\bar{p})$$

and

$$\left| y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \left(X_p'(\tilde{\Sigma}^{-1} \otimes I_T) X_p \right)^{-1} X_p'(\tilde{\Sigma}^{-1} \otimes I_T) y_\ell \right| = O_p(T\bar{p})$$

Next, we write out the following sample moments appearing in a_{GT} and b_{GT}, defined below (18):

$$\begin{aligned} y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) y_\ell &= \sum_{i=1}^N \sum_{j=1}^N \tilde{\sigma}^{ij} \sum_{t=1}^T y_{i,t-1} y_{j,t-1} \\ y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p &= \sum_{i=1}^N \sum_{j=1}^N \tilde{\sigma}^{ij} \sum_{t=1}^T y_{i,t-1} \varepsilon_{jt}^{p_j} \end{aligned}$$

Then from the above results and Lemma A1 (a) and (b), it follows that

$$\begin{aligned} \frac{a_{GT}}{T} &= \frac{y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon_p}{T} + o_p(1) = \sum_{i=1}^N \sum_{j=1}^N \tilde{\sigma}^{ij} \frac{1}{T} \sum_{t=1}^T y_{i,t-1} \varepsilon_{jt}^{p_j} + o_p(1) = Q_{a_{GT}} + o_p(1) \\ \frac{b_{GT}}{T^2} &= \frac{y_\ell'(\tilde{\Sigma}^{-1} \otimes I_T) y_\ell}{T^2} + o_p(1) = \sum_{i=1}^N \sum_{j=1}^N \tilde{\sigma}^{ij} \frac{1}{T^2} \sum_{t=1}^T y_{i,t-1} y_{j,t-1} + o_p(1) = Q_{b_{GT}} + o_p(1) \end{aligned}$$

where

$$\begin{aligned} Q_{a_{GT}} &= \sum_{i=1}^N \sum_{j=1}^N \tilde{\sigma}^{ij} \pi_i(1) \frac{1}{T} \sum_{t=1}^T w_{i,t-1} \varepsilon_{jt} \\ Q_{b_{GT}} &= \sum_{i=1}^N \sum_{j=1}^N \tilde{\sigma}^{ij} \pi_i(1)\pi_j(1) \frac{1}{T^2} \sum_{t=1}^T w_{i,t-1} w_{j,t-1} \end{aligned}$$

We may now write t_{GT} defined in (18) as follows

$$t_{GT} = \frac{a_{GT}}{T} \left(\frac{b_{GT}}{T^2} \right)^{-1/2} = Q_{a_{GT}} Q_{b_{GT}}^{-1/2} + o_p(1)$$

and the limit theory for t_{GT} is directly obtained from applying the invariance principle in (4) to $Q_{a_{GT}}$ and $Q_{b_{GT}}$.

Part (b) Again, we first analyze the components a_{OT} and M_{tOT}, defined below (18), that constitute the OLS based t-statistics t_{OT} given in (18). Since

$$X_p' y_\ell = \begin{pmatrix} \sum_{t=1}^T x_{1t}^{p_1} y_{1,t-1} \\ \vdots \\ \sum_{t=1}^T x_{Nt}^{p_N} y_{N,t-1} \end{pmatrix} = O_p(T\bar{p}^{1/2})$$

$$X_p'(\tilde{\Sigma}\otimes I_T)y_\ell = \begin{pmatrix} \sum_{j=1}^N \tilde{\sigma}_{1j}\sum_{t=1}^T x_{1t}^{p_1}y_{j,t-1} \\ \vdots \\ \sum_{j=1}^N \tilde{\sigma}_{Nj}\sum_{t=1}^T x_{Nt}^{p_N}y_{j,t-1} \end{pmatrix} = O_p(T\bar{p}^{1/2})$$

by Lemma A2 (b), we have from (39) that

$$\begin{aligned}
\left|Y_\ell' X_p(X_p'X_p)^{-1}X_p'\varepsilon_p\right| &= o_p(T\bar{p}\underline{p}^{-s}) + O_p(T^{1/2}\bar{p}) \\
\left|Y_\ell' X_p(X_p'X_p)^{-1}X_p'(\tilde{\Sigma}\otimes I_T)Y_\ell\right| &= O_p(T\bar{p}) \\
\left|Y_\ell' X_p(X_p'X_p)^{-1}X_p'(\tilde{\Sigma}\otimes I_T)X_p(X_p'X_p)^{-1}X_p'Y_\ell\right| &= O_p(T\bar{p})
\end{aligned}$$

We now deduce from Lemma A1 (a) and (b) that

$$\begin{aligned}
\frac{a_{OT}}{T} &= \frac{y_\ell'\varepsilon_p}{T} + o_p(1) = \sum_{i=1}^N \frac{1}{T}\sum_{t=1}^T y_{i,t-1}\varepsilon_{it}^{p_i} + o_p(1) = Q_{a_{OT}} + o_p(1) \\
\frac{M_{tOT}}{T^2} &= \frac{y_\ell'(\tilde{\Sigma}\otimes I_T)y_\ell}{T^2} + o_p(1) = \sum_{i=1}^N\sum_{j=1}^N \tilde{\sigma}_{ij}\frac{1}{T^2}\sum_{t=1}^T y_{i,t-1}y_{j,t-1} + o_p(1) = Q_{M_{tOT}} + o_p(1)
\end{aligned}$$

where

$$\begin{aligned}
Q_{a_{OT}} &= \sum_{i=1}^N \pi_i(1)\frac{1}{T}\sum_{t=1}^T w_{i,t-1}\varepsilon_{it} \\
Q_{M_{tOT}} &= \sum_{i=1}^N\sum_{j=1}^N \tilde{\sigma}_{ij}\pi_i(1)\pi_j(1)\frac{1}{T^2}\sum_{t=1}^T w_{i,t-1}w_{j,t-1}
\end{aligned}$$

Then we have

$$t_{OT} = \frac{a_{OT}}{T}\left(\frac{M_{tOT}}{T^2}\right)^{-1/2} = Q_{a_{OT}}Q_{M_{tOT}}^{-1/2} + o_p(1)$$

from which the stated result follows immediately.

Proofs for the Bootstrap Asymptotics

Proof of Lemma 3.1 The stated results in parts (a)–(c) follow from Lemma 1 of Chang and Park (199).

Proof of Lemma 3.2 See Proof of Lemma 2 in Chang and Park (1999).

Proof of Theorem 3.1

Part (a) From

$$\left(\frac{X_p^{*\prime}(\tilde{\Sigma}^{-1}\otimes I_T)X_p^*}{T}\right)^{-1} \leq \lambda_{\max}(\tilde{\Sigma})\left(\frac{X_p^{*\prime}X_p^*}{T}\right)^{-1} = O_p^*(1) \tag{42}$$

and the results in Lemma 2 (a)–(c), we have

$$\begin{aligned}
&\left|Y_\ell^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\left(X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\right)^{-1}X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)\varepsilon^*\right| \\
&\quad\le \left|Y_\ell^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\right|\left\|\left(X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\right)^{-1}\right\|\left|X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)\varepsilon^*\right| \\
&\quad= O_p^*(T^{1/2}\bar p)
\end{aligned}$$

This together with Lemma 1(b) implies that

$$\frac{A_{GT}^*}{T} = Y_\ell^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)\varepsilon^* + o_p^*(1) = Q_{A_{GT}^*} + o_p^*(1) \tag{43}$$

in **P** or a.s. under Assumption (W) or (S), where

$$Q_{A_{GT}^*} = \begin{pmatrix} \sum_{j=1}^N \tilde\sigma^{1j}\tilde\pi_1(1)\frac{1}{T}\sum_{t=1}^T w_{1,t-1}^*\varepsilon_{jt}^* \\ \vdots \\ \sum_{j=1}^N \tilde\sigma^{Nj}\tilde\pi_N(1)\frac{1}{T}\sum_{t=1}^T w_{N,t-1}^*\varepsilon_{jt}^* \end{pmatrix}$$

Similarly, we have from (42), Lemma 2 (a) and (b) that

$$\begin{aligned}
&\left|Y_\ell^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\left(X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\right)^{-1}X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)Y_\ell^*\right| \\
&\quad\le \left|Y_\ell^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\right|\left\|\left(X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)X_p^*\right)^{-1}\right\|\left|X_p^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)Y_\ell^*\right| \\
&\quad= O_p^*(T\bar p)
\end{aligned}$$

and this along with Lemma 1 (a) gives

$$\frac{B_{GT}^*}{T^2} = Y_\ell^{*\prime}(\tilde\Sigma^{-1}\otimes I_T)Y_\ell^* + o_p^*(1) = Q_{B_{GT}^*} + o_p^*(1) \tag{44}$$

in **P** or a.s. under Assumption (W) or (S), where

$$Q_{B_{GT}^*} = \begin{pmatrix} \tilde\sigma^{11}\tilde\pi_1(1)^2\frac{1}{T^2}\sum_{t=1}^T w_{1,t-1}^{*2} & \cdots & \tilde\sigma^{1N}\tilde\pi_1(1)\tilde\pi_N(1)\frac{1}{T^2}\sum_{t=1}^T w_{1,t-1}^* w_{N,t-1}^* \\ \vdots & \vdots & \vdots \\ \tilde\sigma^{N1}\tilde\pi_N(1)\tilde\pi_1(1)\frac{1}{T^2}\sum_{t=1}^T w_{N,t-1}^* w_{1,t-1}^* & \cdots & \tilde\sigma^{NN}\tilde\pi_N(1)^2\frac{1}{T^2}\sum_{t=1}^T w_{N,t-1}^{*2} \end{pmatrix}$$

in **P** or a.s. under Assumption (W) or (S), analogously as before.

We now write the bootstrapped statistic F_{GT}^* as

$$F_{GT}^* = \left(\frac{A_{GT}^*}{T}\right)'\left(\frac{B_{GT}^*}{T^2}\right)^{-1}\left(\frac{A_{GT}^*}{T}\right) = Q_{A_{GT}^*}' Q_{B_{GT}^*}^{-1} Q_{A_{GT}^*} + o_p^*(1)$$

due to (43) and (44). It is shown in Park (1999) that

$$\tilde{\pi}_i(1) \to_{a.s.} \pi_i(1) \tag{45}$$

and, using the multivariate bootstrap invariance principle developed in Chang (2000), we have

$$\frac{1}{T}\sum_{t=1}^{T} w_{t-1}^{*}\varepsilon_t^{*\prime} \to_{d^*} \int_0^1 BdB' \quad \text{a.s.} \quad \text{and} \quad \frac{1}{T^2}\sum_{t=1}^{T} w_{t-1}^{*}w_{t-1}^{*\prime} \to_{d^*} \int_0^1 BB' \quad \text{a.s.} \tag{46}$$

under Assumption (W). Now, the limiting distribution of the F_{GT}^{*} follows immediately.

Part (b) It follows from Parts (b) and (c) of Lemma 2 that

$$X_p^{*\prime}Y_\ell^{*} = O_p^{*}(T\bar{p}^{1/2}), \quad X_p^{*\prime}\varepsilon^{*} = O_p^{*}(T^{1/2}\bar{p}^{1/2}) \tag{47}$$

which together with (42) gives

$$\left|Y_\ell^{*\prime}X_p^{*}(X_p^{*\prime}X_p^{*})^{-1}X_p^{*\prime}\varepsilon^{*}\right| \leq \left|Y_\ell^{*\prime}X_p^{*}\right| \left\|(X_p^{*\prime}X_p)^{-1}\right\| \left|X_p^{*\prime}\varepsilon^{*}\right| = O_p^{*}(T^{1/2}\bar{p})$$

Then we have from Lemma 1(a) that

$$\frac{A_{OT}^{*}}{T} = \frac{Y_\ell^{*\prime}\varepsilon^{*}}{T} + o_p^{*}(1) = Q_{A_{OT}^{*}} + o_p^{*}(1) \tag{48}$$

where

$$Q_{A_{OT}^{*}} = \begin{pmatrix} \tilde{\pi}_1(1)\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} w_{1,t-1}^{*}\varepsilon_{1t}^{*} \\ \vdots \\ \tilde{\pi}_N(1)\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} w_{N,t-1}^{*}\varepsilon_{Nt}^{*} \end{pmatrix}$$

Next, we deduce from (42) and Lemma 2(b) that

$$X_p^{*\prime}(\tilde{\Sigma}\otimes I_T)X_p^{*} = O_p^{*}(T^{-1}), \quad X_p^{*\prime}(\tilde{\Sigma}\otimes I_T)Y_\ell^{*} = O_p^{*}(T\bar{p}^{1/2}) \tag{49}$$

and this together with (47) gives

$$\left|Y_\ell^{*\prime}X_p^{*}(X_p^{*\prime}X_p^{*})^{-1}X_p^{*\prime}(\tilde{\Sigma}\otimes I_T)Y_\ell^{*}\right| = O_p^{*}(T\bar{p})$$

and

$$\left|Y_\ell^{*\prime}X_p^{*}(X_p^{*\prime}X_p^{*})^{-1}X_p^{*\prime}(\tilde{\Sigma}\otimes I_T)X_p^{*}(X_p^{*\prime}X_p^{*})^{-1}X_p^{*\prime}Y_\ell^{*}\right| = O_p^{*}(T\bar{p})$$

Then we have

$$\frac{M_{FOT}^{*}}{T^2} = \frac{Y_\ell^{*\prime}(\tilde{\Sigma}\otimes I_T)Y_\ell^{*}}{T^2} + o_p^{*}(1) = Q_{M_{FOT}^{*}} + o_p^{*}(1) \tag{50}$$

due to Lemma 1(b), where

$$Q_{M^*_{FOT}} = \begin{pmatrix} \tilde{\sigma}_{11}\tilde{\pi}_1(1)^2 \frac{1}{T^2}\sum_{t=1}^{T} w^{*2}_{1,t-1} & \cdots & \tilde{\sigma}_{1N}\tilde{\pi}_1(1)\tilde{\pi}_N(1)\frac{1}{T^2}\sum_{t=1}^{T} w^*_{1,t-1}w^*_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\sigma}_{N1}\tilde{\pi}_N(1)\tilde{\pi}_1(1)\frac{1}{T^2}\sum_{t=1}^{T} w^*_{N,t-1}w^*_{1,t-1} & \cdots & \tilde{\sigma}_{NN}\tilde{\pi}_N(1)^2\frac{1}{T^2}\sum_{t=1}^{T} w^{*2}_{N,t-1} \end{pmatrix}$$

Finally, we have from the results in (49) and (50)

$$F^*_{OT} = \left(\frac{A^*_{OT}}{T}\right)' \left(\frac{M^*_{FOT}}{T^2}\right)^{-1} \left(\frac{A^*_{OT}}{T}\right) = Q'_{A^*_{OT}} Q^{-1}_{M^*_{FOT}} Q_{A^*_{OT}} + o^*_p(1)$$

and the stated result now follows immediately from (45) and (46).

Proof of Corollary 3.1 The proof is analogous to the proof of Corollary 2.1.

Part(a) It follows from (43) and (44) that

$$T\hat{\alpha}^*_{GT} = \left(\frac{B^*_{GT}}{T^2}\right)^{-1} \left(\frac{A^*_{GT}}{T}\right) = Q^{-1}_{B^*_{GT}} Q_{A^*_{GT}} + o^*_p(1)$$

giving

$$\begin{aligned} \frac{1}{T}\Big(A^*_{GT} .* 1\{\hat{\alpha}^*_{GT} \le 0\}\Big) &= \left(\frac{A^*_{GT}}{T} .* 1\left\{T\hat{\alpha}^*_{GT} \le 0\right\}\right) \\ &= \left(Q_{A^*_{GT}} .* 1\left\{Q^{-1}_{B^*_{GT}} Q_{A^*_{GT}} \le 0\right\}\right) + o^*_p(1) \end{aligned}$$

From the above result and (44), we may write the K^*_{GT} statistics given in (24) as

$$\begin{aligned} K^*_{GT} &= \left(\frac{1}{T}\Big(A^*_{GT} .* 1\{\hat{\alpha}^*_{GT} \le 0\}\Big)\right)' \left(\frac{B^*_{GT}}{T^2}\right)^{-1} \left(\frac{1}{T}\Big(A^*_{GT} .* 1\{\hat{\alpha}^*_{GT} \le 0\}\Big)\right) \\ &= \left(Q_{A^*_{GT}} .* 1\left\{Q^{-1}_{B^*_{GT}} Q_{A^*_{GT}} \le 0\right\}\right)' Q^{-1}_{B^*_{GT}} \left(Q_{A^*_{GT}} .* 1\left\{Q^{-1}_{B^*_{GT}} Q_{A^*_{GT}} \le 0\right\}\right) + o^*_p(1) \end{aligned}$$

Now the stated result follows immediately from (45) and (46).

Part (b) It follows from (42) and (47) that

$$\left|Y^{*\prime}_\ell X^*_p (X^{*\prime}_p X^*_p)^{-1} X^{*\prime}_p Y^*_\ell\right| \le \left|Y^{*\prime}_\ell X^*_p\right| \left\|(X^{*\prime}_p X_p)^{-1}\right\| \left|X^{*\prime}_p Y^*_\ell\right| = O^*_p(T\bar{p})$$

which together with Lemma 3.1 (b) gives

$$\frac{B^*_{OT}}{T^2} = \frac{Y^{*\prime}_\ell Y^*_\ell}{T^2} + o_p(1) = Q_{B^*_{OT}} + o_p(1)$$

where

$$
Q_{B^*_{OT}} = \begin{pmatrix} \tilde{\pi}_1(1)^2 \frac{1}{T^2} \sum_{t=1}^{T} w^{*2}_{1,t-1} & \cdots & \tilde{\pi}_1(1)\tilde{\pi}_N(1) \frac{1}{T^2} \sum_{t=1}^{T} w^*_{1,t-1} w^*_{N,t-1} \\ \vdots & \vdots & \vdots \\ \tilde{\pi}_N(1)\tilde{\pi}_1(1) \frac{1}{T^2} \sum_{t=1}^{T} w^*_{N,t-1} w^*_{1,t-1} & \cdots & \tilde{\pi}_N(1)^2 \frac{1}{T^2} \sum_{t=1}^{T} w^{*2}_{N,t-1} \end{pmatrix}
$$

It follows from (48) and the above result that

$$
T\hat{\alpha}^*_{OT} = \left(\frac{B^*_{OT}}{T^2}\right)^{-1} \left(\frac{A^*_{OT}}{T}\right) = Q^{-1}_{B^*_{OT}} Q_{A^*_{OT}} + o^*_p(1)
$$

and

$$
\frac{1}{T}\Big(A^*_{OT} .\!* 1\{\hat{\alpha}^*_{OT} \leq 0\}\Big) = \Big(Q_{A^*_{OT}} .\!* 1\left\{Q^{-1}_{B^*_{OT}} Q_{A^*_{OT}} \leq 0\right\}\Big) + o^*_p(1)
$$

From this and the result in (50), we may express the test K^*_{OT} defined in (24) as

$$
\begin{aligned}
K^*_{OT} &= \left(\frac{1}{T}\Big(A^*_{OT} .\!* 1\{\hat{\alpha}^*_{OT} \leq 0\}\Big)\right)' \left(\frac{M^*_{FOT}}{T^2}\right)^{-1} \left(\frac{1}{T}\Big(A^*_{OT} .\!* 1\{\hat{\alpha}^*_{OT} \leq 0\}\Big)\right) \\
&= \Big(Q_{A^*_{OT}} .\!* 1\left\{Q^{-1}_{B^*_{OT}} Q_{A^*_{OT}} \leq 0\right\}\Big)' Q^{-1}_{M^*_{FOT}} \Big(Q_{A^*_{OT}} .\!* 1\left\{Q^{-1}_{B^*_{OT}} Q_{A^*_{OT}} \leq 0\right\}\Big) + o^*_p(1)
\end{aligned}
$$

which together with (45) and (46) gives the stated result.

Proof of Theorem 3.2 The limit distributions of the bootstrap GLS and OLS based t-statistics, t^*_{GT} and t^*_{OT}, defined in (26) are derived analogously as we did for the sample t-statistics t_{GT} and t_{OT} in the proof of Theorem 2.2.

Part (a) It follows from Parts (b) and (c) of Lemma 2 that

$$
X^{*\prime}_p(\tilde{\Sigma}^{-1} \otimes I_T)y^*_\ell = O^*_p(T\bar{p}^{1/2}), \quad X^{*\prime}_p(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon^* = O^*_p(T\bar{p}^{1/2})
$$

which along with (42) gives

$$
\left| y^{*\prime}_\ell(\tilde{\Sigma}^{-1} \otimes I_T)X^*_p \left(X^{*\prime}_p(\tilde{\Sigma}^{-1} \otimes I_T)X^*_p\right)^{-1} X^{*\prime}_p(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon^* \right| = O^*_p(T^{1/2}\bar{p})
$$

and

$$
\left| y^{*\prime}_\ell(\tilde{\Sigma}^{-1} \otimes I_T)X^*_p \left(X^{*\prime}_p(\tilde{\Sigma}^{-1} \otimes I_T)X^*_p\right)^{-1} X^{*\prime}_p(\tilde{\Sigma}^{-1} \otimes I_T)y^*_\ell \right| = O^*_p(T\bar{p})
$$

Then we have due to the results in Parts (a) and (b) of Lemma 1 that

$$
\begin{aligned}
\frac{a^*_{GT}}{T} &= \frac{y^{*\prime}_\ell(\tilde{\Sigma}^{-1} \otimes I_T)\varepsilon^*}{T} + o^*_p(1) \;=\; Q_{a^*_{GT}} + o^*_p(1) \\
\frac{b^*_{GT}}{T^2} &= \frac{y^{*\prime}_\ell(\tilde{\Sigma}^{-1} \otimes I_T)y^*_\ell}{T^2} + o^*_p(1) \;=\; Q_{b^*_{GT}} + o^*_p(1)
\end{aligned}
$$

where

$$
\begin{aligned}
Q_{a^*_{GT}} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{\sigma}^{ij}\tilde{\pi}_i(1)\frac{1}{T}\sum_{t=1}^{T}w^*_{i,t-1}\varepsilon^*_{jt} \\
Q_{b^*_{GT}} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{\sigma}^{ij}\tilde{\pi}_i(1)\tilde{\pi}_j(1)\frac{1}{T^2}\sum_{t=1}^{T}w^*_{i,t-1}w^*_{j,t-1}
\end{aligned}
$$

We may now write t^*_{GT} as

$$
t^*_{GT} = \frac{a^*_{GT}}{T}\left(\frac{b^*_{GT}}{T^2}\right)^{-1/2} = Q_{a^*_{GT}}Q^{-1/2}_{b^*_{GT}} + o^*_p(1)
$$

and the limit theory for t^*_{GT} is directly obtained from (45) and (46).

Part (b) Since,

$$
X^{*\prime}_p y^*_\ell = O^*_p(T\bar{p}^{1/2}), \quad X^{*\prime}_p(\tilde{\Sigma}\otimes I_T)y^*_\ell = O^*_p(T\bar{p}^{1/2})
$$

by Lemma A2 (b), we have from (49) that

$$
\begin{aligned}
\left|Y^{*\prime}_\ell X^*_p(X^{*\prime}_p X^*_p)^{-1}X^{*\prime}_p\varepsilon^*\right| &= O^*_p(T^{1/2}\bar{p}) \\
\left|Y^{*\prime}_\ell X^*_p(X^{*\prime}_p X^*_p)^{-1}X^{*\prime}_p(\tilde{\Sigma}\otimes I_T)Y^*_\ell\right| &= O^*_p(T\bar{p}) \\
\left|Y^{*\prime}_\ell X^*_p(X^{*\prime}_p X^*_p)^{-1}X^{*\prime}_p(\tilde{\Sigma}\otimes I_T)X^*_p(X^{*\prime}_p X^*_p)^{-1}X^{*\prime}_p Y^*_\ell\right| &= O^*_p(T\bar{p})
\end{aligned}
$$

We now deduce from Lemma 1 that

$$
\begin{aligned}
\frac{a^*_{OT}}{T} &= \frac{y^{*\prime}_\ell\varepsilon^*}{T} + o^*_p(1) = Q_{a^*_{OT}} + o^*_p(1) \\
\frac{M^*_{tOT}}{T^2} &= \frac{y^{*\prime}_\ell(\tilde{\Sigma}\otimes I_T)y^*_\ell}{T^2} + o^*_p(1) = Q_{M^*_{tOT}} + o^*_p(1)
\end{aligned}
$$

where

$$
\begin{aligned}
Q_{a^*_{OT}} &= \sum_{i=1}^{N}\tilde{\pi}_i(1)\frac{1}{T}\sum_{t=1}^{T}w^*_{i,t-1}\varepsilon^*_{it} \\
Q_{M^*_{tOT}} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{\sigma}_{ij}\tilde{\pi}_i(1)\tilde{\pi}_j(1)\frac{1}{T^2}\sum_{t=1}^{T}w^*_{i,t-1}w^*_{j,t-1}
\end{aligned}
$$

Then we have

$$
t^*_{OT} = \frac{a^*_{OT}}{T}\left(\frac{M^*_{tOT}}{T^2}\right)^{-1/2} = Q_{a^*_{OT}}Q^{-1/2}_{M^*_{tOT}} + o^*_p(1)
$$

from which the stated result follows immediately from (45) and (46).

7. References

An, H.-Z., Z.-G. Chen and E.J. Hannan (1982). "Autocorrelation, autoregression and autoregressive approximation," *Annals of Statistics*, 10, 926-936 (Corr: 11, p.1018).

Andrews, D.W.K. (1999). "Estimation when a parameter is on a boundary," *Econometrica*, 67, 1341-1383.

Baltagi, B.H. (1995). *Econometric Analysis of Panel Data*. Wiley: Chichester.

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

Benerjee, A. (1999). "Panel data unit roots and cointegration: An Overview". Oxford University, mimeographed.

Brillinger, D.R. (1975). *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston: New York.

Brockwell, P.J. and R.A. Davis (1991). *Time Series: Theory and Methods*. Springer-Verlag: New York.

Chang, Y. (2000). "Bootstrapping unit root tests with covariates," Rice University, mimeographed.

Chang, Y. and J.Y. Park (1999). "A sieve bootstrap for the test of a unit root," Rice University, mimeographed.

Choi, B. (1992). *ARMA Model Identification*. Springer-Verlag: New York.

Fisher, R.A. (1933). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, 4th Edition.

Hsiao, C. (1986). *Analysis of Panel Data*. Cambridge University Press: Cambridge.

Im, K.S., M.H. Pesaran and Y. Shin (1997). "Testing for unit roots in heterogeneous panels," mimeographed.

Levin, A. and C.F. Lin (1992). "Unit root tests in panel data: Asymptotic and finite sample properties," University of California, San Diego, mimeographed.

Levin, A. and C.F. Lin (1993). "Unit root tests in panel data: New results," University of California, San Diego, mimeographed.

McCoskey, S. and C. Kao (1998). "A residual based test for the null of cointegration in panel data," *Econometric Reviews*, 17, 57-84.

Maddala, G.S. and S. Wu (1996). "A comparative study of unit root tests with panel data and a new simple test: Evidence from siulations and bootstrap". Ohio State University, mimeographed.

Matyas, L. and P. Sevestre (eds.) (1996). *The Econometrics of Panel Data*, Kluwer Academic Publishers: Boston.

Park, J.Y. (1999). "An invariance principle for sieve bootstrap in time series". Seoul National University, mimeographed.

Pedroni, P. (1996). "Fully modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity," Indiana University, mimeographed.

Pedroni, P. (1997). "Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis, New results," University of Indiana, mimeographed.

Phillips, P.C.B. and H.R. Moon (1999). "Linear regression limit theory for nonstationary panel data," *Econometrica*, 67, 1057-1111.

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

Quah, D. (1994). "Exploiting cross-section variations for unit root inference in dynamic data," *Economics Letters*, 7, 175-189.

Said, S.E. and D.A. Dickey (1984). "Testing for unit roots in autoregressive-moving average models of unknown order," *Biometrika*, 71, 599-608.

Shao, Q.-M. (1995). "Strong approximation theorems for independent random variables and their applications," *Journal of Multivariate Analysis*, 52, 107-130.

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

Stout, W.F. (1974). *Almost Sure Convergence*. Academic Press: New York.

Table A1: Finite Sample Sizes for AR Errors

				1%	tests			5%	tests			10%	tests	
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	50	t-bar	0.011	0.016	0.015	0.023	0.022	0.030	0.030	0.039	0.032	0.040	0.040	0.048
		F^*_{GT}	0.001	0.009	0.009	0.014	0.035	0.047	0.048	0.061	0.084	0.098	0.098	0.114
		F^*_{OT}	0.007	0.012	0.012	0.016	0.038	0.053	0.052	0.064	0.080	0.107	0.111	0.121
		K^*_{GT}	0.001	0.009	0.009	0.014	0.034	0.047	0.047	0.059	0.084	0.098	0.097	0.114
		K^*_{OT}	0.007	0.011	0.012	0.016	0.038	0.052	0.052	0.065	0.079	0.107	0.111	0.122
		t^*_{GT}	0.005	0.009	0.009	0.015	0.035	0.049	0.049	0.067	0.085	0.103	0.102	0.120
		t^*_{OT}	0.006	0.010	0.010	0.015	0.044	0.052	0.050	0.061	0.075	0.105	0.103	0.121
5	100	t-bar	0.009	0.013	0.014	0.016	0.018	0.025	0.026	0.028	0.023	0.034	0.034	0.041
		F^*_{GT}	0.005	0.011	0.010	0.017	0.039	0.051	0.049	0.068	0.088	0.103	0.102	0.125
		F^*_{OT}	0.006	0.011	0.011	0.018	0.041	0.052	0.051	0.062	0.085	0.103	0.105	0.119
		K^*_{GT}	0.005	0.011	0.011	0.018	0.039	0.051	0.049	0.068	0.088	0.103	0.102	0.126
		K^*_{OT}	0.006	0.011	0.012	0.018	0.040	0.052	0.051	0.063	0.086	0.103	0.104	0.122
		t^*_{GT}	0.004	0.009	0.008	0.021	0.038	0.049	0.050	0.064	0.082	0.106	0.107	0.126
		t^*_{OT}	0.004	0.008	0.007	0.011	0.042	0.050	0.048	0.061	0.087	0.102	0.101	0.125
20	50	t-bar	0.032	0.050	0.049	0.072	0.043	0.063	0.063	0.081	0.054	0.072	0.074	0.089
		F^*_{GT}	0.004	0.006	0.005	0.009	0.025	0.036	0.037	0.043	0.068	0.083	0.085	0.096
		F^*_{OT}	0.005	0.011	0.010	0.017	0.041	0.055	0.055	0.068	0.090	0.112	0.116	0.125
		K^*_{GT}	0.003	0.005	0.005	0.009	0.025	0.037	0.037	0.042	0.068	0.083	0.085	0.096
		K^*_{OT}	0.005	0.010	0.011	0.016	0.036	0.054	0.054	0.066	0.092	0.111	0.114	0.123
		t^*_{GT}	0.003	0.006	0.006	0.010	0.024	0.040	0.040	0.050	0.073	0.090	0.092	0.103
		t^*_{OT}	0.005	0.008	0.007	0.013	0.032	0.044	0.045	0.058	0.079	0.094	0.098	0.105
20	100	t-bar	0.029	0.039	0.039	0.049	0.040	0.052	0.052	0.066	0.045	0.060	0.061	0.073
		F^*_{GT}	0.004	0.009	0.009	0.016	0.039	0.045	0.046	0.054	0.077	0.095	0.095	0.110
		F^*_{OT}	0.007	0.011	0.010	0.015	0.036	0.051	0.052	0.064	0.097	0.109	0.109	0.124
		K^*_{GT}	0.004	0.009	0.009	0.016	0.036	0.045	0.045	0.053	0.074	0.094	0.095	0.111
		K^*_{OT}	0.006	0.011	0.010	0.015	0.039	0.051	0.052	0.062	0.094	0.107	0.107	0.123
		t^*_{GT}	0.005	0.008	0.008	0.015	0.036	0.046	0.047	0.061	0.084	0.095	0.094	0.108
		t^*_{OT}	0.005	0.009	0.009	0.017	0.035	0.046	0.045	0.063	0.073	0.095	0.095	0.126

Table A2: Finite Sample Powers for AR Errors

				1%	tests			5%	tests			10%	tests	
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	50	t-bar	0.069	0.166	0.155	0.271	0.113	0.243	0.231	0.373	0.148	0.290	0.279	0.439
		F^*_{GT}	0.038	0.120	0.121	0.199	0.178	0.347	0.343	0.492	0.302	0.509	0.506	0.660
		F^*_{OT}	0.037	0.081	0.075	0.128	0.140	0.258	0.247	0.354	0.249	0.407	0.399	0.532
		K^*_{GT}	0.039	0.120	0.119	0.200	0.178	0.347	0.346	0.492	0.302	0.510	0.509	0.658
		K^*_{OT}	0.038	0.082	0.076	0.128	0.141	0.260	0.247	0.356	0.252	0.409	0.401	0.532
		t^*_{GT}	0.033	0.104	0.100	0.257	0.138	0.307	0.304	0.551	0.227	0.456	0.453	0.721
		t^*_{OT}	0.027	0.097	0.088	0.199	0.129	0.309	0.293	0.476	0.250	0.467	0.449	0.643
5	100	t-bar	0.208	0.598	0.631	0.902	0.302	0.691	0.730	0.948	0.361	0.738	0.785	0.965
		F^*_{GT}	0.228	0.646	0.674	0.912	0.515	0.864	0.911	0.988	0.692	0.930	0.964	0.998
		F^*_{OT}	0.117	0.412	0.406	0.670	0.342	0.700	0.720	0.906	0.497	0.820	0.854	0.965
		K^*_{GT}	0.228	0.647	0.675	0.910	0.519	0.865	0.913	0.987	0.695	0.930	0.964	0.998
		K^*_{OT}	0.118	0.414	0.407	0.672	0.342	0.702	0.720	0.909	0.500	0.822	0.855	0.967
		t^*_{GT}	0.079	0.411	0.398	0.893	0.240	0.649	0.693	0.984	0.356	0.752	0.813	0.996
		t^*_{OT}	0.069	0.403	0.376	0.746	0.265	0.690	0.712	0.941	0.430	0.807	0.831	0.977
20	50	t-bar	0.766	0.867	0.863	0.961	0.805	0.895	0.891	0.976	0.828	0.910	0.905	0.982
		F^*_{GT}	0.268	0.363	0.347	0.527	0.555	0.656	0.644	0.811	0.706	0.793	0.790	0.908
		F^*_{OT}	0.286	0.381	0.356	0.551	0.561	0.676	0.658	0.833	0.738	0.811	0.811	0.914
		K^*_{GT}	0.270	0.365	0.348	0.532	0.560	0.659	0.646	0.811	0.706	0.794	0.792	0.907
		K^*_{OT}	0.291	0.388	0.364	0.562	0.571	0.684	0.664	0.839	0.743	0.818	0.819	0.919
		t^*_{GT}	0.133	0.286	0.301	0.472	0.354	0.557	0.577	0.749	0.495	0.699	0.719	0.855
		t^*_{OT}	0.363	0.513	0.506	0.698	0.665	0.801	0.820	0.919	0.806	0.898	0.919	0.969
20	100	t-bar	0.998	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999	1.000	1.000	1.000
		F^*_{GT}	0.978	0.994	0.997	1.000	0.996	0.999	1.000	1.000	0.999	1.000	1.000	1.000
		F^*_{OT}	0.959	0.985	0.986	1.000	0.993	0.999	0.999	1.000	0.998	1.000	1.000	1.000
		K^*_{GT}	0.978	0.994	0.997	1.000	0.996	0.999	1.000	1.000	0.999	1.000	1.000	1.000
		K^*_{OT}	0.961	0.986	0.988	1.000	0.992	0.999	0.999	1.000	0.998	1.000	1.000	1.000
		t^*_{GT}	0.539	0.842	0.880	0.988	0.769	0.938	0.964	0.999	0.849	0.963	0.984	1.000
		t^*_{OT}	0.828	0.943	0.964	0.999	0.946	0.987	0.994	1.000	0.976	0.994	0.998	1.000

Table B1: Finite Sample Sizes for MA Errors

				1%	tests			5%	tests			10%	tests	
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	50	t-bar	0.002	0.006	0.006	0.008	0.005	0.012	0.013	0.017	0.010	0.018	0.018	0.026
		F_{GT}^{*}	0.003	0.007	0.007	0.013	0.032	0.043	0.042	0.054	0.080	0.094	0.094	0.109
		F_{OT}^{*}	0.002	0.006	0.006	0.012	0.030	0.040	0.040	0.051	0.076	0.094	0.095	0.107
		K_{GT}^{*}	0.003	0.007	0.007	0.014	0.035	0.044	0.042	0.055	0.080	0.095	0.096	0.113
		K_{OT}^{*}	0.003	0.006	0.006	0.011	0.031	0.041	0.039	0.052	0.077	0.094	0.094	0.108
		t_{GT}^{*}	0.005	0.009	0.009	0.014	0.040	0.053	0.053	0.063	0.089	0.109	0.107	0.127
		t_{OT}^{*}	0.004	0.008	0.008	0.013	0.036	0.050	0.051	0.066	0.092	0.106	0.108	0.120
5	100	t-bar	0.003	0.007	0.006	0.011	0.009	0.015	0.014	0.021	0.013	0.020	0.019	0.032
		F_{GT}^{*}	0.003	0.009	0.009	0.017	0.043	0.052	0.051	0.063	0.081	0.105	0.105	0.124
		F_{OT}^{*}	0.004	0.009	0.008	0.018	0.036	0.044	0.047	0.053	0.078	0.098	0.094	0.117
		K_{GT}^{*}	0.003	0.009	0.009	0.017	0.044	0.052	0.052	0.064	0.080	0.105	0.106	0.121
		K_{OT}^{*}	0.005	0.009	0.008	0.018	0.037	0.045	0.046	0.054	0.078	0.098	0.095	0.117
		t_{GT}^{*}	0.002	0.009	0.009	0.013	0.035	0.048	0.050	0.059	0.086	0.103	0.102	0.115
		t_{OT}^{*}	0.006	0.009	0.009	0.015	0.038	0.048	0.045	0.064	0.074	0.102	0.102	0.118
20	50	t-bar	0.013	0.023	0.024	0.031	0.023	0.032	0.032	0.040	0.031	0.038	0.037	0.047
		F_{GT}^{*}	0.003	0.008	0.007	0.014	0.024	0.041	0.041	0.056	0.070	0.090	0.089	0.109
		F_{OT}^{*}	0.003	0.008	0.009	0.013	0.033	0.046	0.047	0.055	0.091	0.103	0.103	0.113
		K_{GT}^{*}	0.004	0.008	0.007	0.014	0.026	0.042	0.042	0.058	0.067	0.090	0.089	0.110
		K_{OT}^{*}	0.003	0.009	0.009	0.015	0.032	0.047	0.047	0.054	0.092	0.103	0.102	0.115
		t_{GT}^{*}	0.003	0.008	0.008	0.013	0.037	0.050	0.051	0.060	0.094	0.114	0.115	0.133
		t_{OT}^{*}	0.005	0.009	0.009	0.013	0.044	0.055	0.056	0.074	0.101	0.116	0.114	0.139
20	100	t-bar	0.018	0.026	0.026	0.038	0.031	0.035	0.035	0.048	0.036	0.042	0.042	0.052
		F_{GT}^{*}	0.005	0.010	0.009	0.013	0.040	0.051	0.050	0.064	0.094	0.104	0.103	0.113
		F_{OT}^{*}	0.005	0.009	0.009	0.013	0.039	0.048	0.049	0.056	0.095	0.104	0.105	0.118
		K_{GT}^{*}	0.006	0.010	0.010	0.014	0.041	0.051	0.050	0.063	0.096	0.104	0.105	0.112
		K_{OT}^{*}	0.005	0.008	0.009	0.013	0.039	0.049	0.049	0.057	0.095	0.105	0.106	0.119
		t_{GT}^{*}	0.005	0.010	0.010	0.018	0.049	0.057	0.056	0.070	0.099	0.115	0.117	0.132
		t_{OT}^{*}	0.004	0.011	0.010	0.018	0.042	0.057	0.057	0.068	0.092	0.117	0.118	0.138

Table B2: Finite Sample Powers for MA Errors

				1%	tests			5%	tests			10%	tests	
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	50	t-bar	0.030	0.075	0.063	0.152	0.062	0.134	0.117	0.258	0.084	0.172	0.153	0.318
		F_{GT}^{*}	0.036	0.112	0.100	0.210	0.159	0.334	0.324	0.509	0.309	0.496	0.502	0.673
		F_{OT}^{*}	0.029	0.062	0.052	0.129	0.126	0.238	0.230	0.374	0.254	0.396	0.395	0.565
		K_{GT}^{*}	0.036	0.113	0.100	0.210	0.158	0.336	0.323	0.513	0.310	0.498	0.504	0.673
		K_{OT}^{*}	0.029	0.063	0.052	0.129	0.128	0.240	0.232	0.375	0.255	0.399	0.400	0.567
		t_{GT}^{*}	0.042	0.091	0.070	0.216	0.157	0.287	0.255	0.559	0.268	0.432	0.400	0.701
		t_{OT}^{*}	0.040	0.089	0.073	0.190	0.189	0.303	0.268	0.499	0.328	0.468	0.435	0.671
5	100	t-bar	0.120	0.406	0.338	0.763	0.212	0.516	0.456	0.853	0.268	0.579	0.533	0.896
		F_{GT}^{*}	0.186	0.551	0.532	0.836	0.495	0.800	0.820	0.969	0.674	0.894	0.909	0.993
		F_{OT}^{*}	0.081	0.333	0.281	0.649	0.280	0.628	0.601	0.907	0.451	0.769	0.774	0.956
		K_{GT}^{*}	0.186	0.552	0.534	0.837	0.499	0.802	0.821	0.970	0.673	0.895	0.909	0.993
		K_{OT}^{*}	0.084	0.335	0.283	0.650	0.283	0.630	0.602	0.908	0.454	0.771	0.776	0.957
		t_{GT}^{*}	0.088	0.300	0.222	0.794	0.235	0.546	0.493	0.958	0.334	0.667	0.637	0.983
		t_{OT}^{*}	0.119	0.321	0.224	0.723	0.359	0.607	0.532	0.939	0.538	0.742	0.681	0.978
20	50	t-bar	0.578	0.710	0.685	0.862	0.648	0.761	0.744	0.893	0.683	0.787	0.772	0.906
		F_{GT}^{*}	0.258	0.348	0.316	0.497	0.540	0.639	0.615	0.776	0.704	0.780	0.754	0.892
		F_{OT}^{*}	0.230	0.312	0.283	0.478	0.525	0.617	0.597	0.751	0.699	0.775	0.754	0.871
		K_{GT}^{*}	0.267	0.354	0.322	0.504	0.541	0.645	0.621	0.781	0.711	0.785	0.758	0.896
		K_{OT}^{*}	0.234	0.323	0.294	0.491	0.545	0.630	0.611	0.770	0.709	0.786	0.764	0.884
		t_{GT}^{*}	0.148	0.284	0.276	0.511	0.383	0.555	0.555	0.792	0.542	0.692	0.698	0.901
		t_{OT}^{*}	0.378	0.516	0.518	0.665	0.712	0.809	0.825	0.902	0.840	0.905	0.911	0.963
20	100	t-bar	0.980	0.998	0.999	1.000	0.989	0.999	1.000	1.000	0.991	0.999	1.000	1.000
		F_{GT}^{*}	0.947	0.988	0.994	1.000	0.992	0.999	1.000	1.000	0.998	1.000	1.000	1.000
		F_{OT}^{*}	0.835	0.947	0.962	0.991	0.964	0.992	0.995	1.000	0.983	0.997	0.999	1.000
		K_{GT}^{*}	0.949	0.988	0.994	1.000	0.992	0.999	1.000	1.000	0.998	1.000	1.000	1.000
		K_{OT}^{*}	0.840	0.950	0.962	0.992	0.966	0.992	0.996	1.000	0.984	0.998	0.999	1.000
		t_{GT}^{*}	0.556	0.786	0.765	0.983	0.749	0.913	0.915	0.998	0.835	0.950	0.954	1.000
		t_{OT}^{*}	0.779	0.903	0.915	0.985	0.913	0.974	0.983	1.000	0.965	0.989	0.994	1.000