

Cross-Sectional Effects of Common and Heterogeneous Regressors on Asymptotic Properties of Panel Autoregressive Unit Root Tests

Katsuto Tanaka
Faculty of Economics, Gakushuin University, Tokyo, Japan

The present paper deals with nonstationary panel autoregressive (AR) models, and examines cross-sectional effects of regressors on the asymptotic properties of panel unit root tests for the AR (1) coefficient. We consider various types of common and heterogeneous regressors and compute limiting local powers of tests as $T \rightarrow \infty$ for each N , where T and N are the time and cross section dimensions, respectively. Dealing with tests based on the ordinary least squares estimator (OLSE) and the generalized LSE (GLSE), we examine how common and heterogeneous regressors affect the tests as N becomes large. It is shown that the existence of common regressors does not affect the tests asymptotically as $N \rightarrow \infty$. This means that the power of the tests remains the same even if the model contains common regressors. We further derive the limiting power envelopes of the most powerful invariant (MPI) tests, which yields the conclusion that the GLSE-based tests are asymptotically efficient, unlike the time series case.

Keywords *Asymptotically efficient test, Common regressor, Cross-sectional effect, Heterogeneous regressor, Moment generating function, Numerical integration, Panel unit root tests.*

Address correspondence to Katsuto Tanaka:

Faculty of Economics
Gakushuin University
Mejiro, Toshima-ku, Tokyo 171-8588
Japan
e-mail: katsuto.tanaka@gakushuin.ac.jp

1. INTRODUCTION

Nonstationary panel AR models were extensively discussed in Moon and Perron (2008) and Moon, Perron, and Phillips (2007), where the former deals with the case of heterogeneous intercepts, whereas the latter discusses the case of heterogeneous trends. In these papers, the limiting local powers of various panel AR unit root tests are computed as T and N jointly tend to ∞ under the local alternative that shrinks to the null at the rate of $1/(TN^k)$, where T is the time series dimension and N is the cross section

dimension with $0 < \kappa < 1$.

Unlike the above works, the present paper examines the effect of the cross section dimension N on the unit root tests as $T \rightarrow \infty$. This may be useful when T is bigger than N and it is desirable to see the intermediate situation rather than the final situation as both T and N go to ∞ . We consider four types of regressors: (1) a common intercept and trend, (2) heterogeneous intercepts and a common trend, (3) a common intercept and heterogeneous trends, (4) heterogeneous intercepts and trends. For these models we conduct panel unit root tests based on the OLSE and GLSE of the AR coefficient, and some other tests based on these residuals. To see the cross-sectional effect, we compute limiting local powers of these unit root tests as $T \rightarrow \infty$ for each intermediate N under the AR coefficient close to unity in the order of $1/T$. It is theoretically and graphically shown that, as N becomes large, the existence of common regressors does not affect the asymptotic properties of these tests, although that of heterogeneous regressors does affect. This fact was also partly observed in panel AR models discussed in Breitung (2000) and Moon et al. (2007). We give more detailed analysis of this fact for each intermediate value of N . We also derive the limiting powers of these tests and envelopes of the most powerful invariant (MPI) tests as $N \rightarrow \infty$, utilizing the joint moment generating functions (m.g.f.s) associated with the test statistics obtained in Nabeya and Tanaka (1990), Tanaka (1996, Chap. 7), and Tanaka (2017, Chap. 10).

The outline of the paper is as follows. In Section 2 we present panel AR models to be dealt with in this paper. In Section 3 we compute limiting local powers of various unit root tests. In Section 3.1 we deal with OLSE-based tests, followed by GLSE-based tests in Section 3.2. The limiting power envelopes are derived in Section 3.3, and it is found that the GLSE-based tests are asymptotically efficient, unlike the time series case. The effect of temporal or cross-sectional dependence of the error term on the tests is discussed in Section 3.4. Section 4 concludes the paper. Proofs of theorems are provided in the Appendix.

2. PANEL AR MODELS

The panel AR models to be discussed in this paper are the following types:

$$\text{Model A: } y_{it} = \alpha + \beta t + \eta_{it}, \quad (1)$$

$$\text{Model B: } y_{it} = \alpha_i + \beta t + \eta_{it}, \quad (2)$$

$$\text{Model C: } y_{it} = \alpha + \beta_i t + \eta_{it}, \quad (3)$$

$$\text{Model D: } y_{it} = \alpha_i + \beta_i t + \eta_{it}, \quad (4)$$

where i refers to cross section, whereas t refers to time series. The process $\{\eta_{it}\}$ is defined for all models by

$$\eta_{it} = \rho_i \eta_{i,t-1} + \varepsilon_{it}, \quad (i = 1, \dots, N; t = 1, \dots, T), \quad (5)$$

where it is assumed that $\{\eta_{it}\}$ starts from $\eta_{i0} = 0$ for each i , and is driven by $\{\varepsilon_{it}\}$. We initially assume $\{\varepsilon_{it}\} \sim \text{i.i.d. } (0, \sigma^2)$ for simplicity of presentation. The case of temporal or cross-sectional dependence will be discussed in Section 3.4.

Model A is the most restricted model with common intercept and trend. Model B has heterogeneous intercepts, whereas Model C has heterogeneous trends. Model D is the most unrestricted model with heterogeneous intercepts and trends. Note that these four models coincide with each other when $N = 1$.

For the above models we consider the panel AR unit root test

$$H_0 : \rho_i = 1 \text{ versus } H_1 : \rho_i < 1 \text{ for some } i, \quad (6)$$

where we assume that, under H_1 , ρ_i takes the following form:

$$\rho_i = 1 - \frac{c_N}{T}, \quad c_N = \frac{c}{N^\kappa}, \quad (7)$$

with $c > 0$ and $0 < \kappa < 1$. This is a simple extension of the time series unit root test. A more general alternative allows the true value of ρ_i to be different among cross sections. Moon and Perron (2008) and Moon, Perron, and Phillips (2007) assume such an alternative, but we maintain (7) to simplify subsequent discussions.

Under the above setting we shall explore asymptotic properties of various unit root tests. For this purpose we define the Ornstein-Uhlenbeck (O-U) process by

$$dY_i(r) = -c_N Y_i(r) dr + dW_i(r), \quad Y_i(0) = 0, \quad (i = 1, \dots, N), \quad (8)$$

where $r \in [0, 1]$ and $\{W_i(r)\}$ is the standard Brownian motion independent of $\{W_k(r)\}$ ($i \neq k$) so that $Y_1(r), \dots, Y_N(r)$ are i.i.d. for any $r \in [0, 1]$.

3. LIMITING POWERS AND POWER ENVELOPES

We first compute the limiting local powers of various unit root tests for Models A through D. In Section 3.1 we deal with OLSE-based tests, followed by GLSE-based tests in Section 3.2. The limiting power envelopes of the MPI tests are derived in Section 3.3. The effect of temporal or cross-sectional dependence of the error term is discussed in Section 3.4.

3.1. OLSE-Based Tests

The present test was earlier considered in Moon et al. (2007), and Moon and Perron (2008). The limiting local power was also computed in these works as both T and N go to ∞ under a more general setting. Here we examine the cross-sectional effect of regressors as $T \rightarrow \infty$ for each N .

Let $\hat{\eta}_i^{(M)}$ be the OLS residual obtained from Model M ($M = A, B, C, D$). Then we compute the estimator $\hat{\rho}^{(M)}$ of $\rho_i = \rho$ under H_0 by

$$\hat{\rho}^{(M)} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{\eta}_{i,t-1}^{(M)} \hat{\eta}_{it}^{(M)}}{\sum_{i=1}^N \sum_{t=2}^T \left\{ \hat{\eta}_{i,t-1}^{(M)} \right\}^2} = 1 + \frac{1}{T} \frac{\sum_{i=1}^N U_{iT}^{(M)}}{\sum_{i=1}^N V_{iT}^{(M)}}, \quad (9)$$

where

$$\begin{aligned} U_{iT}^{(M)} &= \frac{1}{T\sigma^2} \sum_{t=2}^T \hat{\eta}_{i,t-1}^{(M)} \left(\hat{\eta}_{it}^{(M)} - \hat{\eta}_{i,t-1}^{(M)} \right) \\ &= \frac{1}{2T\sigma^2} \left[\left\{ \hat{\eta}_{iT}^{(M)} \right\}^2 - \left\{ \hat{\eta}_{i1}^{(M)} \right\}^2 - \sum_{t=2}^T \left(\hat{\eta}_{it}^{(M)} - \hat{\eta}_{i,t-1}^{(M)} \right)^2 \right], \end{aligned} \quad (10)$$

$$V_{iT}^{(M)} = \frac{1}{T^2\sigma^2} \sum_{t=2}^T \left\{ \hat{\eta}_{i,t-1}^{(M)} \right\}^2. \quad (11)$$

The following theorem describes the asymptotic distribution of $\hat{\rho}^{(M)}$ as $T \rightarrow \infty$ for each N , the proof of which is given in the Appendix.

Theorem 1. As $T \rightarrow \infty$ with N fixed under $\rho_i = 1 - c_N/T$, the asymptotic distribution of $\hat{\rho}^{(M)}$ in Model M ($M = A, B, C, D$) follows

$$T(\hat{\rho}^{(M)} - 1) = \frac{\sum_{i=1}^N U_{iT}^{(M)}}{\sum_{i=1}^N V_{iT}^{(M)}} \Rightarrow Q_N^{(M)} = \frac{U_1^{(D)} + \sum_{i=2}^N U_i^{(M)}}{V_1^{(D)} + \sum_{i=2}^N V_i^{(M)}}, \quad (12)$$

where

$$\begin{aligned} U_i^{(A)} &= \int_0^1 Y_i(r) dY_i(r), \\ V_i^{(A)} &= \int_0^1 Y_i^2(r) dr, \\ U_i^{(B)} &= \int_0^1 \left(Y_i(r) - \int_0^1 Y_i(s) ds \right) dY_i(r), \\ V_i^{(B)} &= \int_0^1 \left(Y_i(r) - \int_0^1 Y_i(s) ds \right)^2 dr, \\ U_i^{(C)} &= \int_0^1 \left(Y_i(r) - 3r \int_0^1 s Y_i(s) ds \right) \left(dY_i(r) - 3 \int_0^1 s Y_i(s) ds dr \right), \\ V_i^{(C)} &= \int_0^1 \left(Y_i(r) - 3r \int_0^1 s Y_i(s) ds \right)^2 dr, \\ U_i^{(D)} &= \int_0^1 \left(Y_i(r) - (4-6r) \int_0^1 Y_i(s) ds - (12r-6) \int_0^1 s Y_i(s) ds \right) dY_i(r), \\ V_i^{(D)} &= \int_0^1 \left(Y_i(r) - (4-6r) \int_0^1 Y_i(s) ds - (12r-6) \int_0^1 s Y_i(s) ds \right)^2 dr. \end{aligned}$$

Some remarks follow.

- (a) When $N = 1$, that is, in the time series case, the distribution of $Q_N^{(M)}$ in (12) reduces to $Q_i^{(D)} = U_1^{(D)}/V_1^{(D)}$ for all M. Note also that $U_1^{(A)}/V_1^{(A)}$ corresponds to the popular near-unit root distribution associated with the time series model $x_t = \rho x_{t-1} + \varepsilon_t$ with $\rho = 1 - c/T$.

- (b) As N becomes large, it holds that

$$Q_N^{(M)} = \frac{U_1^{(D)} + \sum_{i=2}^N U_i^{(M)}}{V_1^{(D)} + \sum_{i=2}^N V_i^{(M)}} = \frac{\sum_{i=2}^N U_i^{(M)} + O_p(1)}{\sum_{i=2}^N V_i^{(M)} + O_p(1)} \approx \frac{\sum_{i=1}^N U_i^{(M)}}{\sum_{i=1}^N V_i^{(M)}},$$

where the distribution of this last quantity is obtained from Model M without common regressors, which means that the effect of common regressors fades away as N becomes large.

- (c) We can deal with some other variations of the above models, for which we can also consider the statistics $T(\hat{\rho}^{(c)} - 1)$. For example, we can show that, as $T \rightarrow \infty$ with N fixed under $\rho_i = 1 - c_N/T$,

$$\text{Model A1: } y_{it} = \eta_{it}, \quad T(\hat{\rho}^{(A1)} - 1) \Rightarrow \frac{\sum_{i=1}^N U_i^{(A)}}{\sum_{i=1}^N V_i^{(A)}},$$

$$\text{Model A2: } y_{it} = \alpha + \eta_{it}, \quad T(\hat{\rho}^{(A2)} - 1) \Rightarrow \frac{U_1^{(B)} + \sum_{i=2}^N U_i^{(A)}}{V_1^{(B)} + \sum_{i=2}^N V_i^{(A)}},$$

$$\text{Model A3: } y_{it} = \beta t + \eta_{it}, \quad T(\hat{\rho}^{(A3)} - 1) \Rightarrow \frac{U_1^{(C)} + \sum_{i=2}^N U_i^{(A)}}{V_M^{(C)} + \sum_{i=2}^N V_i^{(A)}},$$

$$\text{Model B1: } y_{it} = \alpha_i + \eta_{it}, \quad T(\hat{\rho}^{(B1)} - 1) \Rightarrow \frac{\sum_{i=1}^N U_i^{(B)}}{\sum_{i=1}^N V_i^{(B)}},$$

$$\text{Model C1: } y_{it} = \beta_i t + \eta_{it}, \quad T(\hat{\rho}^{(C1)} - 1) \Rightarrow \frac{\sum_{i=1}^N U_i^{(C)}}{\sum_{i=1}^N V_i^{(C)}}.$$

Thus we also conclude that, for these models, the existence of common regressors does not affect the asymptotic behavior of the OLSE-based tests as $N \rightarrow \infty$, which was also described in (b).

- (d) $U_i^{(C)}$ and $V_i^{(C)}$ behave differently from the other quantities, which may be because $U_i^{(C)}/V_i^{(C)}$ results from the restricted regression without intercept $y_{it} = \beta_i t + \eta_{it}$. It can also be shown that $U_i^{(M)}$ and $V_i^{(M)}$ are uncorrelated under $\rho = 1$ for M = B, D, but are correlated for M = A, C. In fact, it holds that $\text{Cov}(U_i^{(A)}, V_i^{(A)}) = 1/3$ and $\text{Cov}(U_i^{(C)}, V_i^{(C)}) = 1/175$ when $c_N = 0$. These can be computed easily from the joint moment generating function (m.g.f.) described below.

To compute the distribution of $Q_N^{(M)}$ in (12) for each N , we use the joint m.g.f. $m^{(M)}(x, y)$ of $U_i^{(M)}$ and $V_i^{(M)}$ defined by

$$m^{(M)}(x, y) = \mathbb{E} \left[\exp \left\{ x U_i^{(M)} + y V_i^{(M)} \right\} \right] = e^{(c_N - x)/2} \left[H^{(M)}(x, y) \right]^{-1/2}, \quad (13)$$

where, by putting $\mu = \sqrt{c_N^2 - 2y}$, we have [Tanaka (2017, Chap. 10)]

$$\begin{aligned} H^{(A)}(x, y) &= \cosh \mu + (c_N - x) \frac{\sinh \mu}{\mu}, \\ H^{(B)}(x, y) &= -\frac{x^2 + c_N^2 x - c_N^3 + 2y}{\mu^2} \frac{\sinh \mu}{\mu} + \frac{c_N^2}{\mu^2} \cosh \mu \\ &\quad + \frac{2(x^2 + c_N^2 x - 2c_N y)}{\mu^4} (\cosh \mu - 1), \\ H^{(C)}(x, y) &= -\frac{x(c_N^2 + 3c_N + 3) - c_N^3}{\mu^2} \frac{\sinh \mu}{\mu} + \frac{c_N^2}{\mu^2} \cosh \mu \\ &\quad - \frac{3x(c_N^2 + 3c_N + 3) - 6y(c_N + 1)}{\mu^4} \left(\frac{\sinh \mu}{\mu} - \cosh \mu \right), \\ H^{(D)}(x, y) &= \frac{c_N^5 - c_N^4 x - 4x^2(c_N^2 + 3c_N + 27) - 8y(c_N^2 - 3c_N - 3)}{\mu^4} \frac{\sinh \mu}{\mu} \\ &\quad - \frac{24(c_N^4 x + 8x^2 y + 4(c_N + 1)(3x^2 - y^2))}{\mu^6} \left(\frac{\sinh \mu}{\mu} - \frac{\cosh \mu}{\mu^2} + \frac{1}{\mu^2} \right) \\ &\quad + \left(\frac{c_N^4}{\mu^4} + \frac{8(c_N^3(c_N x - 2y) + 4x^2(c_N^2 + 3c_N + 6))}{\mu^6} \right) \cosh \mu \\ &\quad + \frac{4(c_N^4 x + 4x^2(c_N^2 + 3c_N - 3) - 2c_N^2 y(c_N + 3))}{\mu^6}. \end{aligned}$$

Then the distribution of $Q_N^{(M)}$ can be computed by using Imhof's formula [Imhof (1961)]

$$\begin{aligned} P(Q_N^{(M)} \leq z) &= P \left(z(V_1^{(D)} + \sum_{i=2}^N V_i^{(M)}) - U_1^{(D)} - \sum_{i=2}^N U_i^{(M)} \geq 0 \right) \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{1}{\theta} \operatorname{Im} \left[m^{(D)}(-i\theta, i\theta z) \left\{ m^{(M)}(-i\theta, i\theta z) \right\}^{N-1} \right] d\theta. \end{aligned} \quad (14)$$

Numerical computation like Simpson's formula can be used to compute (14) by taking care of the computation of square roots of complex-valued quantities [Tanaka (1996)].

Figure 1 draws the probability densities of $Q_N^{(B)}$ and $Q_N^{(D)}$ for various values of N under H_0 ($c_N = 0$) to examine the cross-sectional effect of N . As Theorem 2 below indicates, these distributions converge to -3 and $-15/2$, respectively, as N becomes large. Note that $Q_1^{(B)} = Q_1^{(D)}$. The distributions $Q_N^{(B)}$ for $N > 1$ are shifted from $Q_1^{(D)}$, whereas $Q_N^{(D)}$ for $N > 1$ are just the convolution since $P(Q_N^{(D)} \leq z) = P(\sum_{i=1}^N (zV_i^{(D)} - U_i^{(D)}) \geq 0)$, as is seen from (14). The general feature of $Q_N^{(A)}$ and $Q_N^{(C)}$ are the same as $Q_N^{(B)}$, although those densities are not presented here.

We next compute limiting powers of the tests based on $Q_N^{(M)}$ as $N \rightarrow \infty$ under $\rho = 1 - c_N$ with $c_N =$

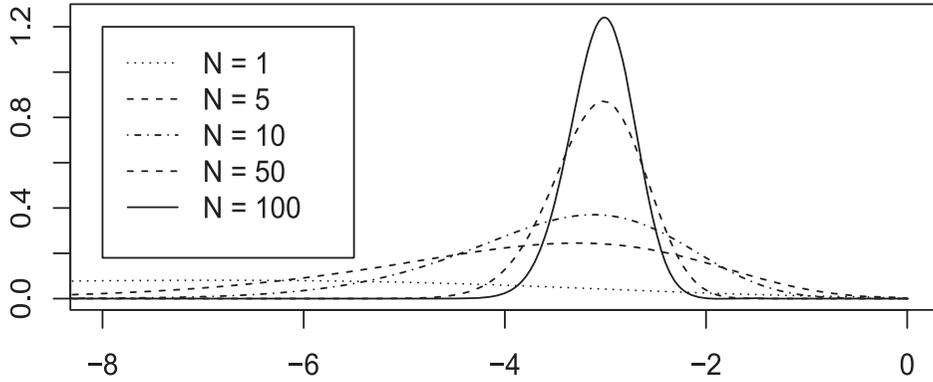
c/N^κ . We need to find the limiting distribution of normalized $Q_N^{(M)}$ by suitably choosing κ . For this purpose, let us put

$$E(U_i^{(M)}) = a_0 + a_1 c_N + a_2 c_N^2 + O(c_N^3), \quad E(V_i^{(M)}) = b_0 + b_1 c_N + b_2 c_N^2 + O(c_N^3).$$

The joint m.g.f. $m^{(M)}(x, y)$ of $U_i^{(M)}$ and $V_i^{(M)}$ shown above can be used to compute these moments using the Taylor expansion, as is shown in the Appendix. We have, by the weak law of large numbers (WLLN) and the central limit theorem (CLT),

$$\sqrt{N} \left(Q_N^{(M)} - \frac{a_0}{b_0} \right) = \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N (U_i^{(M)} - \frac{a_0}{b_0} V_i^{(M)})}{\frac{1}{N} \sum_{i=1}^N V_i^{(M)}} + o_p(1) \Rightarrow N(\mu_0, \sigma_0^2),$$

Model B



Model D

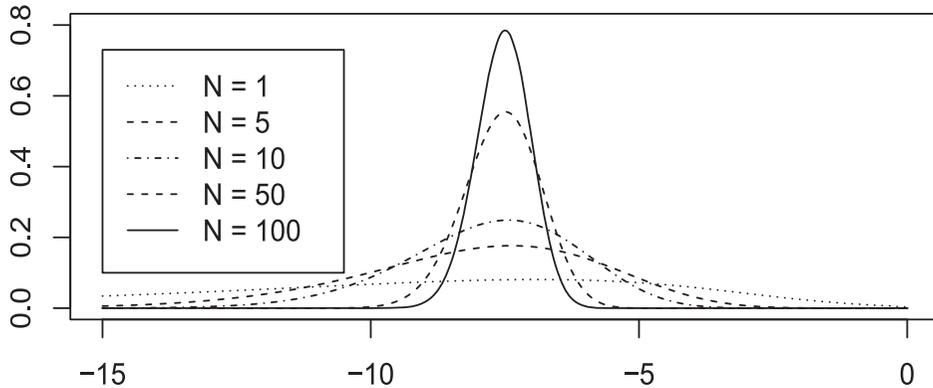


FIGURE 1 Densities of $Q_N^{(B)}$ under H_0 are drawn at the top, whereas those of $Q_N^{(D)}$ at the bottom. Densities are computed for $N=1, 5, 10, 50, 100$ for both graphs.

where

$$\mu_0 = \lim_{N \rightarrow \infty} \sqrt{N} \left(\frac{a_1 b_0 - a_0 b_1}{b_0} c_N + \frac{a_2 b_0 - a_0 b_2}{b_0} c_N^2 \right),$$

$$\sigma_0^2 = \lim_{N \rightarrow \infty} \text{Var} \left(U_i^{(M)} - \frac{a_0}{b_0} V_i^{(M)} \right).$$

Then it is recognized that, for the asymptotic distribution of normalized $Q_N^{(M)}$ to be nondegenerate, $c_N = O(1/\sqrt{N})$ when $a_1 b_0 - a_0 b_1 \neq 0$, and $c_N = O(1/N^{1/4})$ when $a_1 b_0 - a_0 b_1 = 0$ and $a_2 b_0 - a_0 b_2 \neq 0$. It is shown in the Appendix that $c_N = O(1/\sqrt{N})$ for $Q_N^{(A)}$ and $Q_N^{(B)}$, whereas $c_N = O(1/N^{1/4})$ for $Q_N^{(C)}$ and $Q_N^{(D)}$, and we have the the following theorem.

Theorem 2. The limiting powers of the tests based on $Q_N^{(M)}$ ($M = A, B, C, D$) under $\rho = 1 - c/(N^\kappa T)$ at the 100 γ % level are given as follows:

$$P \left(\sqrt{\frac{N}{2}} Q_N^{(A)} < z_\gamma \right) \rightarrow \Phi(z_\gamma + 0.707c),$$

$$P \left(\sqrt{\frac{5N}{51}} (Q_N^{(B)} + 3) < z_\gamma \right) \rightarrow \Phi(z_\gamma + 0.470c),$$

$$P \left(\sqrt{\frac{7N}{110}} (Q_N^{(C)} + 4) < z_\gamma \right) \rightarrow \Phi(z_\gamma + 0.0721c^2),$$

$$P \left(\sqrt{\frac{112N}{2895}} \left(Q_N^{(D)} + \frac{15}{2} \right) < z_\gamma \right) \rightarrow \Phi(z_\gamma + 0.0527c^2),$$

where $\Phi(\cdot)$ is the distribution function of $N(0, 1)$, and z_γ is the 100 γ % point of $N(0, 1)$, whereas $\kappa = 1/2$ for Models A and B, and $\kappa = 1/4$ for Models C and D.

It follows that the OLSE-based unit root tests in Models A and B have nontrivial powers in a $N^{-1/2}T^{-1}$ neighborhood of unity, whereas the powers for Models C and D are nontrivial in a $N^{-1/4}T^{-1}$ neighborhood of unity. It is also seen that the limiting power decreases as the model complexity increases.

Figure 2 shows powers of the $Q_N^{(B)}$ - and $Q_N^{(D)}$ -tests against $c = c_N N^\kappa \in [0, 20]$ at the 5% level for $N = 1, 10, 100, \infty$, where the powers for $N < \infty$ are obtained from (14) by putting z at the 5% point of the null distribution of $Q_N^{(M)}$, whereas those for $N = \infty$ are obtained from Theorem 2. It seems that the powers for $N = 100$ are still not well approximated by the limiting powers. This is particularly true of Model D. The powers of the $Q_N^{(B)}$ -test are higher than those of the $Q_N^{(D)}$ -test, which is also evident from Theorem 2. This means that the existence of heterogeneous trends decreases the power.

In the next subsection we consider the GLSE-based tests, which will be shown to be better than the

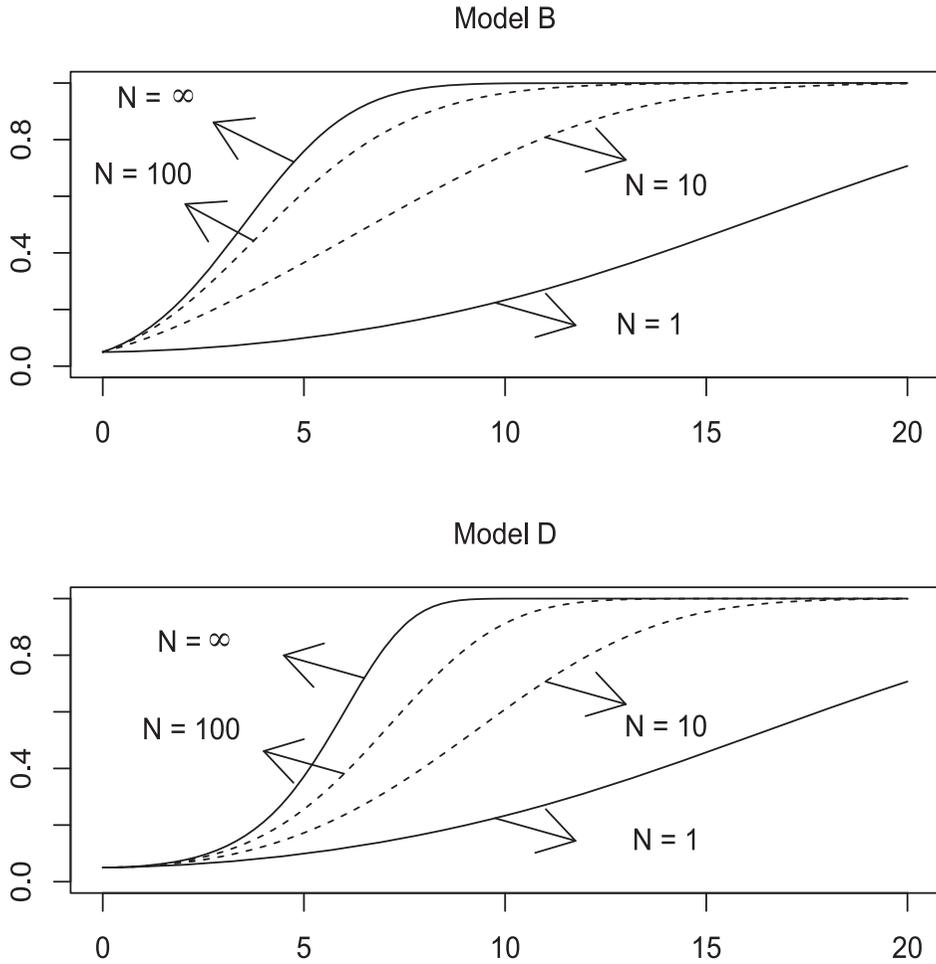


FIGURE 2 Powers of the $Q_N^{(B)}$ -test are shown at the top, whereas those of the $Q_N^{(D)}$ -test are at the bottom. Powers are computed for $N = 1, 10, 100, \infty$ for both graphs.

OLSE-based tests.

3.2. GLSE-Based Tests

Let us express Model M ($M = A, B, C, D$) as $\mathbf{y} = X^{(M)}\boldsymbol{\gamma}^{(M)} + \boldsymbol{\eta}$, where $X^{(M)}$ and $\boldsymbol{\gamma}^{(M)}$ are the regression matrix and parameter vector in Model M, respectively, whereas

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}, \quad \mathbf{y}_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix}, \quad \boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_N \end{pmatrix}, \quad \boldsymbol{\eta}_i = \begin{pmatrix} \eta_{i1} \\ \vdots \\ \eta_{iT} \end{pmatrix}.$$

Then we define the GLS residual by $\tilde{\boldsymbol{\eta}}^{(M)} = \mathbf{y} - X^{(M)} \tilde{\boldsymbol{\gamma}}^{(M)}$, where

$$\tilde{\boldsymbol{\gamma}}^{(M)} = \left((X^{(M)})' (I_N \otimes CC')^{-1} X^{(M)} \right)^{-1} (X^{(M)})' (I_N \otimes CC')^{-1} \mathbf{y}. \quad (15)$$

Here \otimes is the Kronecker product and C is the $T \times T$ lower triangular matrix with (s, t) -th element being 1 for $s \geq t$ and 0 otherwise. The GLSE $\tilde{\rho}^{(M)}$ of ρ can be computed following (9) with $\hat{\boldsymbol{\eta}}^{(M)}$ replaced by $\tilde{\boldsymbol{\eta}}^{(M)}$.

The following theorem describes the asymptotic distribution of $\tilde{\rho}^{(M)}$ as $T \rightarrow \infty$ for each N , the proof of which is given in the Appendix.

Theorem 3. As $T \rightarrow \infty$ with N fixed under $\rho = 1 - c_N/T$, the asymptotic distribution of $\tilde{\rho}^{(M)}$ for Model M (M = A, B, C, D) follows

$$T(\tilde{\rho}^{(M)} - 1) \Rightarrow R_N^{(M)} = \frac{W_1^{(D)} + \sum_{i=2}^N W_i^{(M)}}{X_1^{(D)} + \sum_{i=2}^N X_i^{(M)}}, \quad (16)$$

where

$$W_i^{(A)} = W_i^{(B)} = \int_0^1 Y_i(r) dY_i(r), \quad X_i^{(A)} = X_i^{(B)} = \int_0^1 Y_i^2(r) dr,$$

$$W_i^{(C)} = W_i^{(D)} = -\frac{1}{2}, \quad X_i^{(C)} = X_i^{(D)} = \int_0^1 (Y_i(r) - rY_i(1))^2 dr.$$

It is noticed that the distributional structure of the GLSE-based statistics $R_N^{(M)}$ remains the same as that of the OLSE-based statistics $Q_N^{(M)}$. It is also seen that $R_N^{(A)}$ coincides with $R_N^{(B)}$. The same is true of $R_N^{(C)}$ and $R_N^{(D)}$, and these properties are also shared in the time series case [Tanaka (1996)]. The densities of $R_N^{(A)}$ ($= R_N^{(B)}$) under H_0 are drawn at the top of Figure 3 for $N=1, 10, 30$, whereas those of $R_N^{(C)}$ ($= R_N^{(D)}$) at the bottom for $N=1, 10, 50$. The former densities are seen to be shifted from the latter as N becomes large. Both $R_N^{(A)}$ and $R_N^{(B)}$ converge to 0, whereas both $R_N^{(C)}$ and $R_N^{(D)}$ converge to -3 , as is described in Theorem 4 below.

We next consider limiting powers of the tests based on $R_N^{(M)}$ as $N \rightarrow \infty$, which is described in the following theorem, the proof of which is given in the Appendix.

Theorem 4. The limiting powers of the tests based on $R_N^{(M)}$ (M = A, B, C, D) as $N \rightarrow \infty$ under $\rho = 1 - c/(N^k T)$ at the 100 γ % level are given as follows:

$$P \left(\sqrt{\frac{N}{2}} R_N^{(M)} < z_\gamma \right) \rightarrow \Phi \left(z_\gamma + 0.707c \right), \quad (\text{M} = \text{A, B}),$$

$$P \left(\sqrt{\frac{5N}{6}} (R_N^{(M)} + 3) < z_\gamma \right) \rightarrow \Phi \left(z_\gamma + 0.0745c^2 \right), \quad (\text{M} = \text{C, D}),$$

where $\kappa = 1/2$ for Models A and B, and $\kappa = 1/4$ for Models C and D.

It follows from Theorem 4 that $R_N^{(M)}$ converges to 0 for $M = A, B$, whereas it converges to -3 for $M = C, D$, as was mentioned before. It is also noticed from Theorems 4 and 2 that the GLSE-based tests are better than the OLSE-based tests in Models B, C, D, although those are the same in Model A. The top of Figure 4 shows powers of $R_N^{(A)} (= R_N^{(B)})$ -tests at the 5% level for $N = 1, 10, 50, \infty$, whereas the bottom of Figure 4 those of the $R_N^{(C)} (= R_N^{(D)})$ -tests. It is seen that, for Models A and B, the powers for $N = 50$ are reasonably well approximated by the limiting powers, whereas, for Models C and D, the approximation is still not good enough for $N = 50$. It is seen that the former powers are higher than the latter, as is anticipated from Theorem 4. This means that the existence of heterogeneous trends decreases the power, as in the OLSE-based tests.

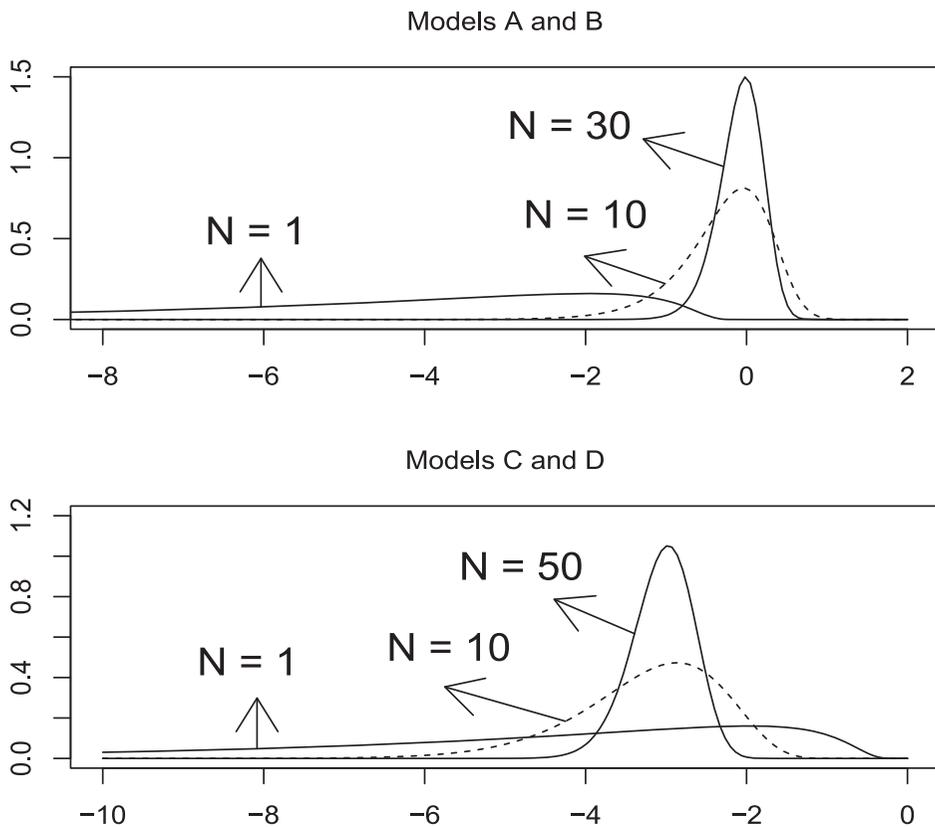


FIGURE 3 Densities of $R_N^{(A)} (= R_N^{(B)})$ under H_0 are drawn at the top, whereas those of $R_N^{(C)} (= R_N^{(D)})$ at the bottom. Densities at the top are computed for $N = 1, 10, 30$, and those at the bottom for $N = 1, 10, 50$.

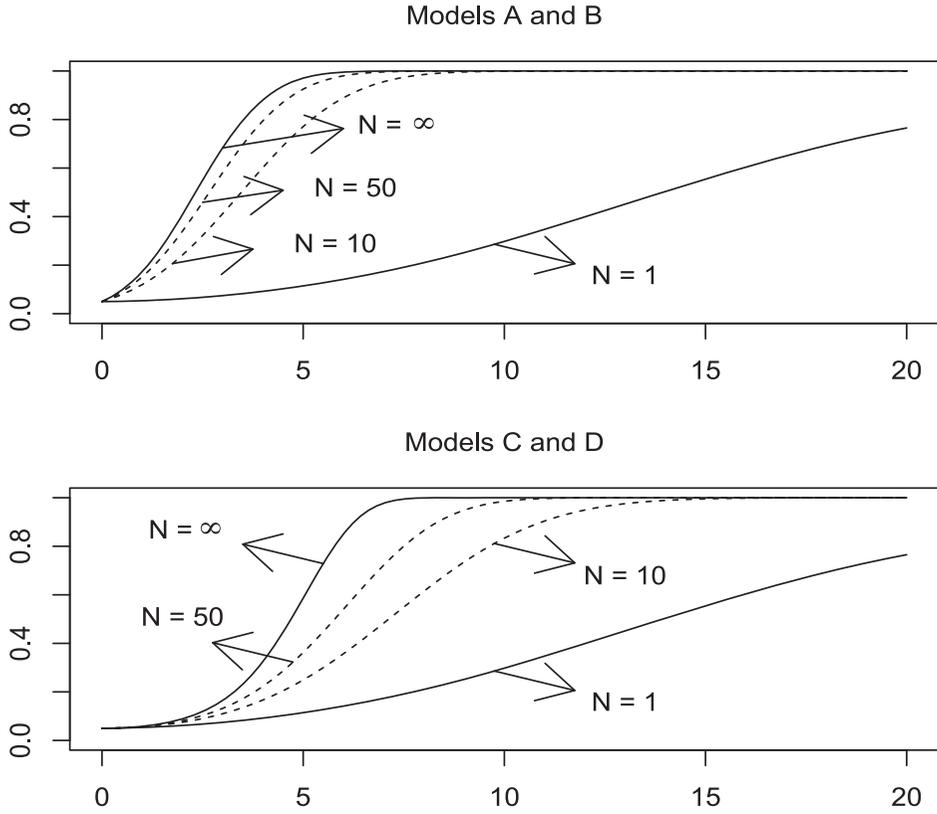


FIGURE 4 Powers of the $R_N^{(A)}$ ($= R_N^{(B)}$)-test are shown at the top, whereas those of the $R_N^{(C)}$ ($= R_N^{(D)}$)-test at the bottom. Powers are computed for $N = 1, 10, 50, \infty$ for both graphs.

3.3. Limiting Power Envelopes

In previous subsections, dealing with Models A through D, we considered panel unit root tests based on OLSE and GLSE, for which the limiting local powers were computed and power comparisons were made among those tests, examining the cross-sectional effects. In this subsection we derive the power envelopes, from which the performance of these tests can be evaluated. The idea was earlier developed in the time series context by Elliott et al. (1996), and was extended to the nonstationary panel data by Moon et al. (2007). Here we derive the power envelopes for Models A through D, paying attention to the cross-sectional effects.

Let us consider the testing problem

$$H_0 : \rho_i = 1 \quad \text{versus} \quad H_1 : \rho_i = 1 - \frac{\theta_N}{T} = \rho(\theta), \quad (17)$$

where $\theta_N = \theta/N^k$ with θ being a known positive constant. We assume that the true value of ρ under H_1 is

given by $\rho(c) = 1 - c_N/T$ with $c_N = c/N^\kappa$. Assuming $\{\varepsilon_{it}\} \sim \text{NID}(0, \sigma^2)$, the Neyman-Pearson lemma tells us that the test which rejects H_0 for small values of

$$S_{NT}^{(M)}(\theta) = \frac{\sum_{i=1}^N \sum_{t=1}^T \left[\left(\tilde{\eta}_{it}^{(M)}(1) - \rho(\theta) \tilde{\eta}_{i,t-1}^{(M)}(1) \right)^2 - \left(\tilde{\eta}_{it}^{(M)}(0) - \tilde{\eta}_{i,t-1}^{(M)}(0) \right)^2 \right]}{\frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T \left(\tilde{\eta}_{it}^{(M)}(0) - \tilde{\eta}_{i,t-1}^{(M)}(0) \right)^2} \quad (18)$$

is MPI, where $\tilde{\eta}_{it}^{(M)}(0)$ and $\tilde{\eta}_{it}^{(M)}(1)$ are the GLS residuals obtained from Model M under H_0 and H_1 , respectively. The residual $\tilde{\eta}_{it}^{(M)}(0)$ is the same as the GLS residual dealt with in the last subsection, that is, $\tilde{\eta}_{it}^{(M)}(0) = \tilde{\eta}_{it}^{(M)}$, whereas $\tilde{\eta}_{it}^{(M)}(1) = \mathbf{y} - X^{(M)}\gamma^{(M)}(1)$, where

$$\gamma^{(M)}(1) = \left((X^{(M)})' \Omega^{-1}(\theta) X^{(M)} \right)^{-1} (X^{(M)})' \Omega^{-1}(\theta) \mathbf{y},$$

with $\Omega(\theta) = I_N \otimes C(\rho(\theta)) C'(\rho(\theta))$. Here $C(\rho(\theta))$ is the $T \times T$ lower triangular matrix with (s, t) -th element being $\rho^{|s-t|}(\theta)$ for $s \geq t$ and 0 otherwise. The test based on $S_{NT}^{(M)}(\theta)$ with fixed θ is called the point optimal invariant (POI) test [King (1987)].

The following theorem gives the weak convergence of $S_{NT}^{(M)}(\theta)$ as $T \rightarrow \infty$ for each N , the proof of which is given in the Appendix.

Theorem 5. As $T \rightarrow \infty$ under $\rho = 1 - c_N/T$ for each N , the MPI test statistic $S_{NT}^{(M)}(\theta)$ in (18) follows

$$S_{NT}^{(M)}(\theta) \Rightarrow S_N^{(M)}(\theta) = \frac{1}{N} \left[Z_1^{(D)}(\theta) + \sum_{i=2}^N Z_i^{(M)}(\theta) \right], \quad (\mathbf{M} = \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}), \quad (19)$$

where

$$\begin{aligned} Z_i^{(A)}(\theta) &= Z_i^{(B)}(\theta) = \theta_N^2 \int_0^1 Y_i^2(r) dr + 2\theta_N \int_0^1 Y_i(r) dY_i(r), \\ Z_i^{(C)}(\theta) &= Z_i^{(D)}(\theta) = \theta_N^2 \left[\int_0^1 Y_i^2(r) dr - \frac{2(\theta_N + 1)}{\delta_N} Y_i(1) \int_0^1 r Y_i(r) dr \right. \\ &\quad \left. + \frac{\theta_N + 1}{3\delta_N} Y_i^2(1) - \frac{\theta_N^2}{\delta_N} \left(\int_0^1 r Y_i(r) dr \right)^2 \right] - \theta_N, \end{aligned}$$

with $\delta_N = 1 + \theta_N + \frac{1}{3}\theta_N^2$.

It is seen that the expression for $S_N^{(M)}(\theta)$ in (19) is of a similar nature to $Q_N^{(M)}$ in (12) and $R_N^{(M)}$ in (16). It is also noticed that the distribution of $S_N^{(M)}(\theta)$ depends on θ that is the value under H_1 . Thus the MPI test based on $S_N^{(M)}(\theta)$ is not uniformly best, but we can modify $S_N^{(M)}(\theta)$ so that the distribution of the modified statistic does not depend on θ as $N \rightarrow \infty$. Then we can compute the limiting power of the test based on a modified statistic which yields the limiting power envelope of all the invariant tests for Model M. The following theorem gives such statistics and the power envelopes.

Theorem 6. The limiting powers of the tests based on the MPI statistics $S_N^{(M)}(\theta)$ in (19) at the 100 γ % level as $N \rightarrow \infty$ under $\theta_N = \theta/N^\kappa$ and $c_N = c/N^\kappa$ are given by

$$P\left(\sqrt{\frac{N}{2}}\left(\frac{S_N^{(M)}(\theta)}{\theta_N} - \frac{1}{2}\theta_N\right) < z_\gamma\right) \rightarrow \Phi(z_\gamma + 0.707c), \quad (\mathbf{M} = \mathbf{A}, \mathbf{B}), \quad (20)$$

$$P\left(3\sqrt{5N}\left(\frac{S_N^{(M)}(\theta) + \theta_N}{\theta_N^2} - \frac{1}{6} + \frac{1}{45}\theta_N^2\right) < z_\gamma\right) \rightarrow \Phi(z_\gamma + 0.0745c^2), \quad (\mathbf{M} = \mathbf{C}, \mathbf{D}), \quad (21)$$

where $\kappa = 1/2$ for Models A and B, and $\kappa = 1/4$ for Models C and D.

The limiting powers of the modified tests give the power envelope of all the invariant tests. Comparing Theorem 6 with Theorem 4 it is seen that the power functions of the GLSE-based tests coincide with the power envelopes. Thus the GLSE-based tests are asymptotically efficient, unlike the time series case. This is a merit of panel tests as $N \rightarrow \infty$.

There are some other tests that are asymptotically efficient. Here we take up two such tests. Define

$$K_{NT}^{(M)} = \sum_{i=1}^N \left\{ \tilde{\eta}_{it}^{(M)}(0) \right\}^2 / \sum_{i=1}^N \sum_{t=1}^T \left(\tilde{\eta}_{it}^{(M)}(0) - \tilde{\eta}_{it-1}^{(M)}(0) \right)^2, \quad (22)$$

$$L_{NT}^{(M)} = \frac{1}{T} \sum_{i=1}^N \sum_{t=1}^T \left\{ \tilde{\eta}_{it}^{(M)}(0) \right\}^2 / \sum_{i=1}^N \sum_{t=1}^T \left(\tilde{\eta}_{it}^{(M)}(0) - \tilde{\eta}_{it-1}^{(M)}(0) \right)^2. \quad (23)$$

The test that rejects H_0 for $K_{NT}^{(M)}$ small is locally best invariant (LBI), although the test is inapplicable to Models C and D because $K_{NT}^{(M)} \equiv 0$ for $\mathbf{M} = \mathbf{C}, \mathbf{D}$, whereas the test that rejects H_0 for $L_{NT}^{(M)}$ small is LBI and unbiased (LBIU) for $\mathbf{M} = \mathbf{C}, \mathbf{D}$ [Tanaka (2017, Chap. 10)]. We have, as $T \rightarrow \infty$ for each N ,

$$K_{NT}^{(M)} \Rightarrow K_N^{(M)} = \begin{cases} \frac{1}{N} \sum_{i=1}^N Y_i^2(1), & (\mathbf{M} = \mathbf{A}, \mathbf{B}), \\ 0, & (\mathbf{M} = \mathbf{C}, \mathbf{D}), \end{cases}$$

$$L_{NT}^{(M)} \Rightarrow L_N^{(M)} = \begin{cases} \frac{1}{N} \sum_{i=1}^N \int_0^1 Y_i^2(r) dr, & (\mathbf{M} = \mathbf{A}, \mathbf{B}), \\ \frac{1}{N} \sum_{i=1}^N \int_0^1 \left(Y_i(r) - rY_i(1) \right)^2 dr, & (\mathbf{M} = \mathbf{C}, \mathbf{D}). \end{cases}$$

Since it can be shown that, for $\mathbf{M} = \mathbf{A}, \mathbf{B}$,

$$\mathbb{E}(K_N^{(M)}) = 1 - c_N + O(c_N^2), \quad \text{Var}(K_N^{(M)}) = \frac{1}{N} \left(2 - 4c_N + O(c_N^2) \right),$$

we have $\sqrt{N}(K_N^{(M)} - 1) \Rightarrow N(-c, 2)$ by putting $c_N = c/\sqrt{N}$, which implies that the $K_N^{(M)}$ -tests for $\mathbf{M} = \mathbf{A}, \mathbf{B}$ are asymptotically efficient. It is evident that the $L_N^{(M)}$ -tests are asymptotically efficient for $\mathbf{M} = \mathbf{C}, \mathbf{D}$.

Note that the LBI and LBIU tests in the time series case ($N=1$) are asymptotically inefficient [Tanaka (1996, Chap. 9)]. We also note in passing that, if the GLS residual in (23) is replaced by the OLS residual, the resulting statistic is essentially the Durbin-Watson statistic and the corresponding test is asymptotically inefficient.

3.4. Effect of Temporal or Cross-Sectional Dependence

Here we consider the situation where there exists temporal or cross-sectional dependence of the error term $\{\varepsilon_{it}\}$ in (5) and examine the effect of such dependence on the test statistics obtained in previous subsections.

Let us first consider temporal dependence. For this purpose we assume

$$\varepsilon_{it} = \sum_{k=0}^{\infty} \varphi_{ik} \zeta_{i,t-k} = \varphi_i(L) \zeta_{it}, \quad \sum_{k=1}^{\infty} k |\varphi_{ik}| < \infty, \quad \{\zeta_{it}\} \sim \text{i.i.d.}(0, \sigma^2), \quad (24)$$

where $\phi_i(L) = 1 + \phi_{i1}L + \phi_{i2}L^2 + \dots$ with L being the lag-operator. The distributional properties of the statistics $T(\hat{\rho}^{(M)} - 1)$ in (12) and $T(\tilde{\rho}^{(M)} - 1)$ in (16) are affected by this relaxation. In fact, it can be shown that, as $T \rightarrow \infty$ for each N ,

$$T(\hat{\rho}^{(M)} - 1) \Rightarrow \frac{\sum_{i=1}^N \varphi_i^2(1) \left[U_i^{(M)} + \frac{1}{2}(1 - \lambda_i) \right] + O_p(1)}{\sum_{i=1}^N \varphi_i^2(1) V_i^{(M)} + O_p(1)},$$

$$T(\tilde{\rho}^{(M)} - 1) \Rightarrow \begin{cases} \frac{\sum_{i=1}^N \varphi_i^2(1) \left[W_i^{(M)} + \frac{1}{2}(1 - \lambda_i) \right] + O_p(1)}{\sum_{i=1}^N \varphi_i^2(1) X_i^{(M)} + O_p(1)}, & (\text{M} = \text{A, B}), \\ \frac{\sum_{i=1}^N \sum_{k=0}^{\infty} \varphi_{ik}^2 W_i^{(M)} + O_p(1)}{\sum_{i=1}^N \varphi_i^2(1) X_i^{(M)} + O_p(1)}, & (\text{M} = \text{C, D}), \end{cases}$$

where $U_i^{(M)}$ and $V_i^{(M)}$ are defined in (12), and $W_i^{(M)}$ and $X_i^{(M)}$ are defined in (16), whereas λ_i is the ratio of the short-run to long-run variances of $\{\varepsilon_{it}\}$ given by $\lambda_i = \sum_{k=0}^{\infty} \varphi_{ik}^2 / \varphi_i^2(1)$. The above statistics depend on the short-run and long-run variances of the error term that characterize temporal dependence.

We next consider cross-sectional dependence, for which we assume that

$$\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{Nt} \end{pmatrix} : N \times 1 \sim \text{i.i.d.}(\mathbf{0}, \Sigma), \quad \Sigma = [\sigma_{ik}] : N \times N.$$

It then follows that

$$T(\hat{\rho}^{(M)} - 1) \Rightarrow \frac{\sum_{i=1}^N \sigma_{ii} \tilde{U}_i^{(M)} + O_p(1)}{\sum_{i=1}^N \sigma_{ii} \tilde{V}_i^{(M)} + O_p(1)}, \quad T(\tilde{\rho}^{(M)} - 1) \Rightarrow \frac{\sum_{i=1}^N \sigma_{ii} \tilde{W}_i^{(M)} + O_p(1)}{\sum_{i=1}^N \sigma_{ii} \tilde{X}_i^{(M)} + O_p(1)}.$$

Here $\tilde{U}_i^{(M)}$, $\tilde{V}_i^{(M)}$, $\tilde{W}_i^{(M)}$, and $\tilde{X}_i^{(M)}$ replace $U_i^{(M)}$, $V_i^{(M)}$, $W_i^{(M)}$, and $X_i^{(M)}$, respectively, with $Y_i(r)$ replaced by $\tilde{Y}_i(r)$, where

$$d\tilde{Y}_i(r) = -c_N \tilde{Y}_i(r) dr + d\tilde{W}_i(r),$$

and $\{\tilde{W}_i(r)\}$ is the standard Brownian motion with

$$\text{Cov}(\tilde{W}_i(r), \tilde{W}_k(s)) = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii} \sigma_{kk}}} \min(r, s).$$

The test statistics depend on the covariances σ_{ik} of the error term that characterize cross-sectional dependence.

It is recognized from the above observations that, to use the asymptotic results obtained in previous subsections, we need to modify the statistics to make them independent of nuisance parameters. This remains to be done.

4. CONCLUDING REMARKS

Under a simple setting, we have presented a unified approach to deriving the limiting local powers of panel AR unit root tests, paying attention to the cross-sectional effect of N . For this purpose it is necessary to compute moments up to the second order of the limiting statistic in the time series direction. We found it easier to use its m.g.f., unlike in the literature. It happened that the tests that were not powerful in the time series case become more powerful in the panel case. It was also found that the existence of a common intercept and/or a common trend does not affect the asymptotic behavior of the tests. This holds for not only the tests based on OLS and GLS residuals, but also power envelopes.

The present approach can be applied to unit root tests for other types of panel models such as panel moving average models or panel error components models. Some simple extensions are found in Tanaka (2017, Chap. 10). For these models the panel LBI or LBIU tests can be used and the corresponding statistics have a distributional structure similar to the panel AR unit root tests discussed in this paper. Details are reported in Tanaka (2018).

5. APPENDIX: PROOFS OF THEOREMS

Proof of Theorem 1: We first deal with Model D. Given the OLSEs $\hat{\alpha}_i$ and $\hat{\beta}_i$ of α_i and β_i , respectively, the OLS residual is

$$\begin{aligned}
 \hat{\eta}_{it} &= y_{it} - \hat{\alpha}_i - \hat{\beta}_i t \\
 &= \eta_{it} - \frac{\left(\sum_{s=1}^T s^2 - t \sum_{s=1}^T s\right) \sum_{s=1}^T \eta_{is} + \left(tT - \sum_{s=1}^T s\right) \sum_{s=1}^T s \eta_{is}}{T \sum_{s=1}^T s^2 - \left(\sum_{s=1}^T s\right)^2} \\
 &= \eta_{it} - \left(\frac{4}{T} - \frac{6t}{T^2}\right) \sum_{s=1}^T \eta_{is} - \left(\frac{12t}{T^3} - \frac{6}{T^2}\right) \sum_{s=1}^T s \eta_{is} + O_p(T^{-1}).
 \end{aligned}$$

The continuous mapping theorem (CMT) yields, as $T \rightarrow \infty$ with N fixed,

$$\begin{aligned}
 U_{iT} &= \frac{1}{T\sigma^2} \sum_{t=2}^T \hat{\eta}_{i,t-1} (\hat{\eta}_{it} - \hat{\eta}_{i,t-1}) = \frac{1}{2T\sigma^2} \left[\hat{\eta}_{iT}^2 - \hat{\eta}_{i1}^2 - \sum_{t=2}^T (\hat{\eta}_{it} - \hat{\eta}_{i,t-1})^2 \right] \\
 &= \frac{1}{2T\sigma^2} \left[\left(\eta_{iT} + \frac{2}{T} \sum_{t=1}^T \eta_{it} - \frac{6}{T^2} \sum_{t=1}^T t \eta_{it} \right)^2 - \left(\frac{4}{T} \sum_{t=1}^T \eta_{it} - \frac{6}{T^2} \sum_{t=1}^T t \eta_{it} \right)^2 \right. \\
 &\quad \left. - \sum_{t=2}^T (\eta_{it} - \eta_{i,t-1})^2 \right] + o_p(1) \\
 \Rightarrow U_i &= \frac{1}{2} \left[\left(Y_i(1) + 2 \int_0^1 Y_i(r) dr - 6 \int_0^1 r Y_i(r) dr \right)^2 \right. \\
 &\quad \left. - \left(4 \int_0^1 Y_i(r) dr - 6 \int_0^1 r Y_i(r) dr \right)^2 - 1 \right] \\
 &= \int_0^1 \left(Y_i(r) - (4-6r) \int_0^1 Y_i(s) ds \right. \\
 &\quad \left. - (12r-6) \int_0^1 s Y_i(s) ds \right) dY_i(r),
 \end{aligned}$$

$$\begin{aligned}
 V_{iT} &= \frac{1}{T^2\sigma^2} \sum_{t=2}^T \hat{\eta}_{i,t-1}^2 \\
 &= \frac{1}{T^2\sigma^2} \sum_{t=2}^T \left(\eta_{i,t-1} - \left(\frac{4}{T} - \frac{6(t-1)}{T^2} \right) \sum_{s=1}^T \eta_{is} - \left(\frac{12(t-1)}{T^3} - \frac{6}{T^2} \right) \sum_{s=1}^T s \eta_{is} \right)^2 \\
 \Rightarrow V_i &= \int_0^1 \left(Y_i(r) - (4-6r) \int_0^1 Y_i(s) ds - (12r-6) \int_0^1 s Y_i(s) ds \right)^2 dr.
 \end{aligned}$$

Thus the relation (12) is proved for Model D by the CMT.

We next deal with Model B, for which the OLSes $\hat{\alpha}$ and $\hat{\beta}$ of α and β are given by

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \left[\begin{pmatrix} I_N & \otimes \mathbf{i}'_T \\ \mathbf{i}'_N & \otimes \mathbf{d}'_T \end{pmatrix} \begin{pmatrix} I_N & \otimes \mathbf{i}_T & \mathbf{i}_N & \otimes \mathbf{d}_T \end{pmatrix} \right]^{-1} \begin{pmatrix} I_N & \otimes \mathbf{i}'_T \\ \mathbf{i}'_N & \otimes \mathbf{d}'_T \end{pmatrix} \mathbf{y},$$

where I_N is the identity matrix of order N , \otimes is the Kronecker product, and

$$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_N \end{pmatrix}, \quad \mathbf{i}_T = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \mathbf{d}_T = \begin{pmatrix} 1 \\ \vdots \\ T \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{pmatrix}, \quad \mathbf{y}_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix},$$

We then have

$$\begin{aligned}
\hat{\alpha} - \alpha &= \frac{1}{T} \sum_{t=1}^T \boldsymbol{\eta}_t + \frac{3}{NT} \mathbf{i}_N \sum_{i=1}^N \sum_{t=1}^T \left(1 - \frac{2t}{T}\right) \eta_{it} + o_p(\sqrt{T}), \\
\hat{\beta} - \beta &= -\frac{6}{NT^2} \sum_{i=1}^N \sum_{t=1}^T \left(1 - \frac{2t}{T}\right) \eta_{it} + o_p(T^{-1/2}), \\
\hat{\eta}_{it} &= y_{it} - \hat{\alpha}_i - \hat{\beta} t = \eta_{it} - (\hat{\alpha}_i - \alpha) - (\hat{\beta} - \beta) t \\
&= \eta_{it} - \frac{1}{T} \sum_{t=1}^T \eta_{it} - \frac{3}{NT} \left(1 - \frac{2t}{T}\right) \sum_{i=1}^N \sum_{s=1}^T \eta_{is} \\
&\quad + \frac{6}{NT^2} \left(1 - \frac{2t}{T}\right) \sum_{i=1}^N \sum_{s=1}^T s \eta_{is} + o_p(\sqrt{T}),
\end{aligned}$$

where $\boldsymbol{\eta}_t = (\eta_{1t}, \dots, \eta_{Nt})'$. It follows that $T(\hat{\rho}^{(B)} - 1) = \sum_{i=1}^N U_{iT}^{(B)} / \sum_{i=1}^N V_{iT}^{(B)}$, where it holds that, as $T \rightarrow \infty$ with N fixed,

$$\begin{aligned}
\sum_{i=1}^N U_{iT}^{(B)} &= \frac{1}{2T\sigma^2} \sum_{i=1}^N \left[\hat{\eta}_{iT}^2 - \hat{\eta}_{i1}^2 - \sum_{t=2}^T (\hat{\eta}_{it} - \hat{\eta}_{i,t-1})^2 \right] \\
&\Rightarrow \frac{1}{2} \sum_{i=1}^N \left[Y_i(1) - \int_0^1 Y_i(r) dr + \frac{3}{N} \sum_{k=1}^N \int_0^1 (1-2s) Y_k(s) ds \right]^2 \\
&\quad - \frac{1}{2} \sum_{i=1}^N \left[-\int_0^1 Y_i(r) dr - \frac{3}{N} \sum_{k=1}^N \int_0^1 (1-2s) Y_k(s) ds \right]^2 - \frac{N}{2} \\
&= \sum_{i=1}^N \int_0^1 (Y_i(r) - \int_0^1 Y_i(s) ds) dY_i(r) \\
&\quad - \frac{3}{N} \sum_{i=1}^N \int_0^1 (2Y_i(r) - Y_i(1)) dr \sum_{k=1}^N \int_0^1 (Y_k(s) - 2sY_k(s)) ds \\
&= \sum_{i=1}^N U_i^{(B)}, \\
\sum_{i=1}^N V_{iT}^{(B)} &= \frac{1}{T^2\sigma^2} \sum_{i=1}^N \sum_{t=2}^T \hat{\eta}_{i,t-1}^2 \\
&\Rightarrow \sum_{i=1}^N \int_0^1 \left[Y_i(r) - \int_0^1 Y_i(s) ds + (2r-1) \frac{3}{N} \sum_{k=1}^N \int_0^1 (Y_k(s) - 2sY_k(s)) ds \right]^2 dr \\
&= \sum_{i=1}^N V_i^{(B)}.
\end{aligned}$$

Let us put $\mathbf{Y}(r) = (Y_1(r), \dots, Y_N(r))'$. Following Nabeya (2000), consider $\mathbf{Z}(r) = H\mathbf{Y}(r)$, where H is the $N \times N$ orthogonal matrix with the first row being \mathbf{i}'_N / \sqrt{N} so that $\{\mathbf{Z}(r)\} \stackrel{D}{=} \{\mathbf{Y}(r)\}$ and $\mathbf{Y}'(r) \mathbf{i}_N = \mathbf{Y}'(r) H' H \mathbf{i}_N = \sqrt{N} Z_1(r)$. Then it holds that

$$\begin{aligned}
\sum_{i=1}^N U_i^{(B)} &= \int_0^1 (\mathbf{Y}'(r) - \int_0^1 \mathbf{Y}'(s) ds) H' H d\mathbf{Y}(r) \\
&\quad - \frac{3}{N} \int_0^1 (2\mathbf{Y}'(r) - \mathbf{Y}'(1)) H' H \mathbf{i}_N dr \int_0^1 (\mathbf{Y}'(s) - 2s\mathbf{Y}'(s)) H' H \mathbf{i}_N ds
\end{aligned}$$

パネル自己回帰単位根検定の漸近的性質に与える説明変数のクロスセクション効果 (田中)

$$\begin{aligned}
&= \sum_{i=1}^N \int_0^1 \left(Z_i(r) - \int_0^1 Z_i(s) ds \right) dZ_i(r) \\
&\quad - 3 \int_0^1 (2Z_1(r) - Z_1(1)) dr \int_0^1 (Z_1(s) - 2sZ_1(s)) ds \\
&= \sum_{i=2}^N \int_0^1 \left(Z_i(r) - \int_0^1 Z_i(s) ds \right) dZ_i(r) + \int_0^1 Z_1(r) dZ_1(r) + 2Z_1(1) \int_0^1 Z_1(r) dr \\
&\quad - 6 \left(\int_0^1 Z_1(r) dr \right)^2 - 6Z_1(1) \int_0^1 rZ_1(r) dr + 12 \int_0^1 Z_1(r) dr \int_0^1 sZ_1(s) ds, \\
&\stackrel{\text{D}}{=} U_1^{(D)} + \sum_{i=2}^N U_i^{(B)}, \\
\sum_{i=1}^N V_i^{(B)} &= \sum_{i=1}^N \int_0^1 \left(Z_i(r) - \int_0^1 Z_i(s) ds \right)^2 dr - 3 \left(\int_0^1 Z_1(r) dr \right)^2 \\
&\quad - 12 \left(\int_0^1 rZ_1(r) dr \right)^2 + 12 \int_0^1 Z_1(r) dr \int_0^1 sZ_1(s) ds \\
&\stackrel{\text{D}}{=} V_1^{(D)} + \sum_{i=2}^N V_i^{(B)},
\end{aligned}$$

which shows that (12) holds for $M = B$ by the CMT. We can prove (12) for $M = A$ and C similarly.

Proof of Theorem 2: Let us consider Model B. The joint m.g.f of $U_i^{(B)}$ and $V_i^{(B)}$ is given in the text, which can be used to compute first and second moments of these quantities as $c_N \rightarrow 0$ by employing computerized algebra and the Taylor expansion. Here we show how to compute $E(U_i^{(B)})$ and $\text{Var}(U_i^{(B)})$. It follows from (13) that

$$m(x) = E\left(e^{xU_i^{(B)}}\right) = e^{(c_N^{-1}x)^2} g(x)^{-1/2},$$

where

$$g(x) = -\frac{x^2 + c_N^2 x - c_N^3}{c_N^2} a_1 + a_2 + \frac{2(x^2 + c_N^2 x)}{c_N^4} (a_2 - 1),$$

with $g(0) = e^{c_N}$, $a_1 = (\sinh c_N)/c_N$ and $a_2 = \cosh c_N$. Then we have

$$\begin{aligned}
m'(0) &= E(U_i^{(B)}) = -\frac{1}{2} - \frac{1}{2} e^{-c_N} g'(0), \\
m''(0) &= E\left((U_i^{(B)})^2\right) = \frac{1}{4} + \frac{1}{2} e^{-c_N} \left(g'(0) - g''(0)\right) + \frac{3}{4} e^{-2c_N} \{g'(0)\}^2.
\end{aligned}$$

Here it holds that

$$g'(0) = -a_1 + \frac{2}{c_N^2} (a_2 - 1) \quad g''(0) = -\frac{2}{c_N^2} a_1 + \frac{4}{c_N^4} (a_2 - 1).$$

Then, using the expansions for e^{-c_N} , e^{-2c_N} , a_1 , and a_2 as $c_N \rightarrow 0$,

$$e^{-c_N} = 1 - c_N + \frac{c_N^2}{2} - \frac{c_N^3}{6} + \frac{c_N^4}{24} + O(c_N^5),$$

$$\begin{aligned}
e^{-2c_N} &= 1 - 2c_N + 2c_N^2 - \frac{4c_N^3}{3} + \frac{2c_N^4}{3} + O(c_N^5), \\
a_1 &= 1 + \frac{c_N^2}{6} + \frac{c_N^4}{120} + \frac{c_N^6}{5040} + O(c_N^8), \\
a_2 &= 1 + \frac{c_N^2}{2} + \frac{c_N^4}{24} + \frac{c_N^6}{720} + O(c_N^8),
\end{aligned}$$

we can obtain moments of $U_i^{(B)}$. We now have

$$\begin{aligned}
\mathbb{E}(U_i^{(B)}) &= -\frac{1}{2} + \frac{1}{24}c_N^2 + O(c_N^3), \quad \text{Var}(U_i^{(B)}) = \frac{1}{12} - \frac{1}{12}c_N + O(c_N^2), \\
\mathbb{E}(V_i^{(B)}) &= \frac{1}{6} - \frac{1}{12}c_N + O(c_N^2), \quad \text{Var}(V_i^{(B)}) = \frac{1}{45} - \frac{1}{30}c_N + O(c_N^2).
\end{aligned}$$

It follows from the WLLN that $Q_N^{(B)} \rightarrow -3$ in probability. Thus we consider

$$\sqrt{N}(Q_N^{(B)} + 3) \Rightarrow \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N (U_i^{(B)} + 3V_i^{(B)})}{\frac{1}{N} \sum_{i=1}^N V_i} \Rightarrow \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N (U_i^{(B)} + 3V_i^{(B)})}{1/6}.$$

The m.g.f. of $U_i^{(B)} + 3V_i^{(B)}$ can be obtained from $m^{(B)}(x, y)$ by putting $y = 3x$, which yields

$$\mathbb{E}(U_i^{(B)} + 3V_i^{(B)}) = -\frac{1}{4}c_N + O(c_N^2), \quad \text{Var}(U_i + 3V_i) = \frac{17}{60} - \frac{23}{60}c_N + O(c_N^2).$$

Putting $c_N = c/\sqrt{N}$, we now have

$$\sqrt{N}(Q_N^{(B)} + 3) \Rightarrow \frac{N\left(-\frac{c}{4}, \frac{17}{60}\right)}{1/6} = N\left(-\frac{3}{2}c, \frac{51}{5}\right),$$

which proves the second relation in the theorem. We note in passing that

$$\begin{aligned}
\text{Var}(U_i + 3V_i) &= \text{Var}(U_i) + 9\text{Var}(V_i) - \frac{1}{20}c_N^2 + O(c_N^3), \\
\text{Cov}(U_i, V_i) &= -\frac{1}{120}c_N^2 + O(c_N^3),
\end{aligned}$$

so that it holds that

$$\int_0^1 \left(W_i(t) - \int_0^1 W_i(s) ds \right) dW_i(t) \quad \text{and} \quad \int_0^1 \left(W_i(t) - \int_0^1 W_i(s) ds \right)^2 dt$$

are uncorrelated.

For the other models the first two moments of $U_i^{(M)}$ and $V_i^{(M)}$ as $c_N \rightarrow 0$ can be obtained from their joint m.g.f.s given in Tanaka (2017, Chap. 10). More specifically we have

$$\mathbb{E}(U_i^{(A)}) = -\frac{1}{2}c_N + O(c_N^2), \quad \text{Var}(U_i^{(A)}) = \frac{1}{2} - c_N + O(c_N^2),$$

パネル自己回帰単位根検定の漸近的性質に与える説明変数のクロスセクション効果 (田中)

$$E(V_i^{(A)}) = \frac{1}{2} - \frac{1}{3} c_N + O(c_N^2), \quad \text{Var}(V_i^{(A)}) = \frac{1}{3} - \frac{8}{15} c_N + O(c_N^2).$$

$$E(U_i^{(C)}) = -\frac{2}{5} - \frac{1}{105} c_N^2 + O(c_N^4), \quad \text{Var}(U_i^{(C)}) = \frac{1}{50} - \frac{2}{525} c_N^2 + O(c_N^4),$$

$$E(V_i^{(C)}) = \frac{1}{10} - \frac{1}{210} c_N^2 + O(c_N^4), \quad \text{Var}(V_i^{(C)}) = \frac{1}{175} - \frac{4}{4725} c_N^2 + O(c_N^4).$$

$$E(U_i^{(C)} + 4V_i^{(C)}) = -\frac{1}{35} c_N^2 + O(c_N^4), \quad \text{Var}(U_i^{(C)} + 4V_i^{(C)}) = \frac{11}{70} - \frac{26}{945} c_N^2 + O(c_N^4),$$

$$E(U_i^{(D)}) = -\frac{1}{2} + \frac{1}{1440} c_N^4 + O(c_N^5), \quad \text{Var}(U_i^{(D)}) = \frac{1}{60} - \frac{1}{504} c_N^2 + O(c_N^4),$$

$$E(V_i^{(D)}) = \frac{1}{15} - \frac{1}{420} c_N^2 + O(c_N^4), \quad \text{Var}(V_i^{(D)}) = \frac{11}{6300} - \frac{1}{4200} c_N^2 + O(c_N^4).$$

$$E\left(U_i^{(D)} + \frac{15}{2} V_i^{(D)}\right) = \frac{1}{56} c_N^2 + O(c_N^4), \quad \text{Var}\left(U_i^{(D)} + \frac{15}{2} V_i^{(D)}\right) = \frac{193}{1680} - \frac{31}{2016} c_N^2 + O(c_N^4).$$

These moments yield the limiting powers for M=A, C, D shown in the theorem, which establishes Theorem 2. We note in passing that, when $c_N=0$,

$$\begin{aligned} \text{Cov}(U_i^{(C)}, 4V_i^{(C)}) &= \frac{1}{2} \left(\text{Var}(U_i^{(C)} + 4V_i^{(C)}) - \text{Var}(U_i^{(C)}) - \text{Var}(4V_i^{(C)}) \right) \\ &= \frac{4}{175}, \\ \text{Cov}\left(U_i^{(D)}, \frac{15}{2} V_i^{(D)}\right) &= \frac{1}{2} \left(\text{Var}(U_i^{(D)} + \frac{15}{2} V_i^{(D)}) - \text{Var}(U_i^{(D)}) - \frac{225}{4} \text{Var}(V_i^{(D)}) \right) \\ &= 0 \end{aligned}$$

so that $U_i^{(C)}$ and $V_i^{(C)}$ are correlated, whereas $U_i^{(D)}$ and $V_i^{(D)}$ are uncorrelated.

Proof of Theorem 3: We first deal with Model D. Defining the $T \times T$ lower triangular matrix C with $C(s, t) = 1$ for $s \geq t$ and $C(s, t) = 0$ for $s < t$, the GLSEs of α_i and β_i are given by

$$\begin{aligned} \begin{pmatrix} \tilde{\alpha}_i \\ \tilde{\beta}_i \end{pmatrix} &= \left(\begin{pmatrix} \mathbf{i}'_T \\ \mathbf{d}'_T \end{pmatrix} (CC')^{-1} \begin{pmatrix} \mathbf{i}_T \\ \mathbf{d}_T \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{i}'_T \\ \mathbf{d}'_T \end{pmatrix} (CC')^{-1} \mathbf{y}_i \\ &= \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & T \end{pmatrix}^{-1} \begin{pmatrix} \eta_{i1} \\ \eta_{iT} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} + \frac{1}{T-1} \begin{pmatrix} T\eta_{i1} - \eta_{iT} \\ -\eta_{i1} + \eta_{iT} \end{pmatrix}, \end{aligned}$$

so that we have

$$\begin{aligned} \tilde{\eta}_{it} &= y_{it} - \tilde{\alpha}_i - \tilde{\beta}_i t = \eta_{it} - (\tilde{\alpha}_i - \alpha_i) - (\tilde{\beta}_i - \beta_i) t \\ &= \eta_{it} - \frac{t}{T} \eta_{iT} + o_p(\sqrt{T}). \end{aligned}$$

Consider $T(\tilde{\rho} - 1) = \sum_{i=1}^N U_{iT}^{(D)} / \sum_{i=1}^N V_{iT}^{(D)}$. It holds that, as $T \rightarrow \infty$ with N fixed,

$$\begin{aligned}
U_{iT}^{(D)} &= \frac{1}{T\sigma^2} \sum_{t=2}^T \tilde{\eta}_{i,t-1} (\tilde{\eta}_{it} - \tilde{\eta}_{i,t-1}) = \frac{1}{2T\sigma^2} \left[\tilde{\eta}_{iT}^2 - \tilde{\eta}_{i1}^2 - \sum_{t=2}^T (\tilde{\eta}_{it} - \tilde{\eta}_{i,t-1})^2 \right] \\
&= -\frac{1}{2T\sigma^2} \left[\sum_{t=2}^T (\eta_{it} - \eta_{i,t-1})^2 \right] + o_p(1) \Rightarrow -\frac{1}{2},
\end{aligned}$$

$$V_{iT}^{(D)} = \frac{1}{T^2\sigma^2} \sum_{t=2}^T \tilde{\eta}_{i,t-1}^2 \Rightarrow \int_0^1 (Y_i(r) - rY_i(1))^2 dr,$$

which proves (16) for M=D.

We next consider Model B. The GLSEs of $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)'$ and β are given by

$$\begin{aligned}
\begin{pmatrix} \tilde{\boldsymbol{\alpha}} \\ \tilde{\beta} \end{pmatrix} &= \left(\begin{pmatrix} I_N \otimes \mathbf{i}'_T \\ \mathbf{i}'_N \otimes \mathbf{d}'_T \end{pmatrix} (I_N \otimes CC')^{-1} (I_N \otimes \mathbf{i}_T \mathbf{i}_N \otimes \mathbf{d}_T) \right)^{-1} \\
&\quad \begin{pmatrix} I_N \otimes \mathbf{i}'_T \\ \mathbf{i}'_N \otimes \mathbf{d}'_T \end{pmatrix} (I_N \otimes CC')^{-1} \mathbf{y} \\
&= \begin{pmatrix} \boldsymbol{\alpha} \\ \beta \end{pmatrix} + \begin{pmatrix} O_p(1) \\ \frac{1}{NT} \sum_{i=1}^N \eta_{iT} + o_p(T^{-1/2}) \end{pmatrix},
\end{aligned}$$

which yields

$$\tilde{\eta}_{it} = \eta_{it} - (\tilde{\alpha}_i - \alpha_i) - (\tilde{\beta} - \beta)t = \eta_{it} - \frac{t}{T} \frac{1}{N} \sum_{i=1}^N \eta_{iT} + o_p(\sqrt{T}).$$

Then, applying the orthogonal transformation $\mathbf{Z}(r) = H\mathbf{Y}(r)$ used in the proof of Theorem 1, it holds that, as $T \rightarrow \infty$ with N fixed,

$$\begin{aligned}
\sum_{i=1}^N U_{iT}^{(B)} &\Rightarrow \frac{1}{2} \sum_{i=1}^N \left[\left(Y_i(1) - \frac{1}{N} \sum_{k=1}^N Y_k(1) \right)^2 - 1 \right] \\
&= \frac{1}{2} \sum_{i=1}^N (Y_i^2(1) - 1) - \frac{1}{2N} \left(\sum_{k=1}^N Y_k(1) \right)^2 \\
&\stackrel{D}{=} \frac{1}{2} \sum_{i=1}^N (Z_i^2(1) - 1) - \frac{1}{2} Z_1^2(1) = -\frac{1}{2} + \frac{1}{2} \sum_{i=2}^N (Z_i^2(1) - 1), \\
\sum_{i=1}^N V_{iT}^{(B)} &\Rightarrow \sum_{i=1}^N \int_0^1 \left(Y_i(r) - \frac{r}{N} \sum_{k=1}^N Y_k(1) \right)^2 dr \\
&= \sum_{i=1}^N \int_0^1 Y_i^2(r) dr - \frac{2}{N} \sum_{i=1}^N \int_0^1 r Y_i(r) dr \sum_{k=1}^N Y_k(1) + \frac{1}{3N} \left(\sum_{k=1}^N Y_k(1) \right)^2 \\
&\stackrel{D}{=} \sum_{i=1}^N \int_0^1 Z_i^2(r) dr - 2Z_1(1) \int_0^1 r Z_1(r) dr + \frac{1}{3} Z_1^2(1) \\
&= \int_0^1 (Z_1(r) - rZ_1(1))^2 dr + \sum_{i=2}^N \int_0^1 Z_i^2(r) dr.
\end{aligned}$$

Noting that $\{\mathbf{Z}(r)\} \stackrel{D}{=} \{\mathbf{Y}(r)\}$, the relation (16) is proved for M=B. We can prove (16) for M=A and C similarly, which establishes Theorem 3.

Proof of Theorem 4: It follows from Theorem 3 that, as N becomes large,

$$R_N^{(M)} \Rightarrow \sum_{i=1}^N W_i^{(M)} \Big/ \sum_{i=1}^N X_i^{(M)},$$

where $W_i^{(M)}$ and $X_i^{(M)}$ are defined in (16). The limiting distribution of $R_N^{(M)}$ for $M = A, B$ is the same as that of $Q_N^{(A)}$ in (12), which proves the first relation. For Models C and D, $W_i^{(M)} = -1/2$, whereas the m.g.f. of $X_i^{(M)}$ is given in Tanaka (2017, Chap. 10) and

$$E(X_i^{(M)}) = \frac{1}{6} - \frac{1}{90} c_N^2 + O(c_N^4), \quad \text{Var}(X_i^{(M)}) = \frac{1}{45} - \frac{4}{945} c_N^2 + O(c_N^4),$$

which yields $R_N^{(M)} \rightarrow -3$ in probability. Then we consider

$$\sqrt{N} \left(R_N^{(M)} + 3 \right) \Rightarrow \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \left(-\frac{1}{2} + 3X_i^{(M)} \right)}{\frac{1}{N} \sum_{i=1}^N X_i^{(M)}},$$

where

$$E \left(-\frac{1}{2} + 3X_i^{(M)} \right) = -\frac{1}{30} c_N^2 + O(c_N^4), \quad \text{Var} \left(-\frac{1}{2} + 3X_i^{(M)} \right) = \frac{1}{5} + O(c_N^2).$$

Putting $c_N = c/N^{1/4}$, we now have

$$\sqrt{N} \left(R_N^{(M)} + 3 \right) \Rightarrow \frac{N \left(-\frac{c^2}{30}, \frac{1}{5} \right)}{1/6} = N \left(-\frac{c^2}{5}, \frac{36}{5} \right),$$

which proves the second relation, and Theorem 4 has been established.

Proof of Theorem 5: The relation (19) for Model D was proved in Tanaka (2017, Chap. 10). Let us consider Model B, for which it holds that

$$\tilde{\eta}_{it}^{(B)}(0) = \eta_{it} - \frac{t}{T} \frac{1}{N} \sum_{i=1}^N \eta_{iT} + o_p(\sqrt{T}), \quad \tilde{\eta}_{it}^{(B)}(1) = \eta_{it} - \frac{t}{T} \frac{1}{N} \sum_{i=1}^N A_{iT} + o_p(\sqrt{T}),$$

where

$$\begin{aligned} A_{iT} &= \frac{1}{\delta_N} \left[\sum_{t=2}^T (\eta_{it} - \rho(\theta) \eta_{i,t-1}) + \frac{\theta_N}{T} \sum_{t=2}^T t (\eta_{it} - \rho(\theta) \eta_{i,t-1}) \right] \\ &= \frac{1}{\delta_N} \left[(\theta_N + 1) \eta_{iT} + \frac{\theta_N^2}{T^2} \sum_{t=2}^T t \eta_{i,t-1} \right]. \end{aligned}$$

It holds that

$$\begin{aligned}
\sum_{t=2}^T \left(\tilde{\eta}_{it}^{(B)}(1) - \tilde{\eta}_{i,t-1}^{(B)}(1) \right)^2 &= \sum_{t=2}^T \left(\eta_{it} - \eta_{i,t-1} \right)^2 - \frac{2}{NT} \eta_{iT} \sum_{k=1}^N A_{kT} + \frac{1}{T} \left(\frac{1}{N} \sum_{k=1}^N A_{kT} \right)^2 \\
&\quad - \frac{\theta_N}{T} \left\{ \sum_{t=2}^T \left(\eta_{it} - \eta_{i,t-1} \right)^2 - \left(\eta_{iT} - \frac{1}{N} \sum_{k=1}^N A_{kT} \right)^2 \right\} \\
&\quad + \frac{\theta_N^2}{T^2} \sum_{t=2}^T \left(\eta_{it} - \frac{t}{NT} \sum_{k=1}^N A_{kT} \right)^2 + o_p(1), \\
\sum_{t=2}^T \left(\tilde{\eta}_{it}^{(B)}(0) - \tilde{\eta}_{i,t-1}^{(B)}(0) \right)^2 &= \sum_{t=2}^T \left(\eta_{it} - \eta_{i,t-1} \right)^2 - \frac{2}{NT} \eta_{iT} \sum_{k=1}^N \eta_{kT} \\
&\quad + \frac{1}{T} \left(\frac{1}{N} \sum_{k=1}^N \eta_{kT} \right)^2 + o_p(1),
\end{aligned}$$

which gives

$$\begin{aligned}
D_{iT}^{(B)} &= \sum_{t=2}^T \left(\tilde{\eta}_{it}^{(B)}(1) - \tilde{\eta}_{i,t-1}^{(B)}(1) \right)^2 - \left(\tilde{\eta}_{it}^{(B)}(0) - \tilde{\eta}_{i,t-1}^{(B)}(0) \right)^2 \\
&= -\frac{2}{NT} \eta_{iT} \sum_{k=1}^N A_{kT} + \frac{1}{T} \left(\frac{1}{N} \sum_{k=1}^N A_{kT} \right)^2 \\
&\quad - \frac{\theta_N}{T} \left\{ \sum_{t=2}^T \left(\eta_{it} - \eta_{i,t-1} \right)^2 - \left(\eta_{iT} - \frac{1}{N} \sum_{k=1}^N A_{kT} \right)^2 \right\} + \frac{2}{NT} \eta_{iT} \sum_{k=1}^N \eta_{kT} \\
&\quad + \frac{\theta_N^2}{T^2} \sum_{t=2}^T \left(\eta_{it} - \frac{t}{NT} \sum_{k=1}^N A_{kT} \right)^2 - \frac{1}{T} \left(\frac{1}{N} \sum_{k=1}^N \eta_{kT} \right)^2 + o_p(1).
\end{aligned}$$

Thus it holds that, as $T \rightarrow \infty$ with N fixed,

$$\begin{aligned}
\frac{1}{\sigma^2} D_{iT}^{(B)} &\Rightarrow -\frac{2}{N} Y_i(1) \sum_{k=1}^N X_k + \left(\frac{1}{N} \sum_{k=1}^N X_k \right)^2 - \theta_N \left\{ 1 - \left(Y_i(1) - \frac{1}{N} \sum_{k=1}^N X_k \right)^2 \right\} \\
&\quad + \frac{2}{N} Y_i(1) \sum_{k=1}^N Y_k(1) + \theta_N^2 \int_0^1 \left(Y_i(r) - \frac{r}{N} \sum_{k=1}^N X_k \right)^2 dr - \left(\frac{1}{N} \sum_{k=1}^N Y_k(1) \right)^2,
\end{aligned}$$

where

$$\frac{1}{\sqrt{T}\sigma} A_{iT} \Rightarrow X_i = \frac{1}{\delta_N} \left\{ (\theta_N + 1) Y_i(1) + \theta_N^2 \int_0^1 r Y_i(r) dr \right\}.$$

Applying the orthogonal transformation $\mathbf{Z}(r) = H\mathbf{Y}(r)$, we obtain

$$\begin{aligned}
\frac{1}{\sigma^2} \sum_{i=1}^N D_{iT}^{(B)} &\Rightarrow \theta_N^2 \sum_{i=1}^N \int_0^1 Z_i^2(r) dr + 2\theta_N \sum_{i=1}^N \int_0^1 Z_i(r) dZ_i(r) + \left(Z_1(1) - \tilde{X}_1 \right)^2 \\
&\quad - \theta_N \left(2Z_1(1) \tilde{X}_1 - \tilde{X}_1^2 \right) - \theta_N^2 \left\{ 2\tilde{X}_1 \int_0^1 r Z_1(r) dr - \frac{1}{3} \tilde{X}_1^2 \right\} \\
&= (\theta_N + 1) \left(Z_1(1) - \tilde{X}_1 \right)^2 + \theta_N^2 \int_0^1 \left(Z_1(r) - r \tilde{X}_1 \right)^2 dr \\
&\quad + \theta_N^2 \sum_{i=2}^N \int_0^1 Z_i^2(r) dr + 2\theta_N \sum_{i=2}^N \int_0^1 Z_i(r) dZ_i(r) - \theta_N,
\end{aligned}$$

where

$$\tilde{X}_1 = \frac{1}{\delta_N} \left\{ (\theta_N + 1) Z_1(1) + \theta_N^2 \int_0^1 r Z_1(r) dr \right\}.$$

Noting that $\{Z(r)\} \stackrel{D}{=} \{Y(r)\}$ and $\sum_{t=2}^T (\tilde{\eta}_t^{(B)}(0) - \tilde{\eta}_{t-1}^{(B)}(0))^2 / T \rightarrow \sigma^2$ in probability, we prove (19) for $M=B$ after some manipulations. We can prove (19) for $M=A$ and C similarly, which establishes Theorem 5.

Proof of Theorem 6: For Models A and B we have $S_N^{(A)}(\theta) = \sum_{i=1}^N Z_i^{(A)}(\theta) / N + o_p(1)$, where $Z_i^{(A)}(\theta)$ is defined in (19) and the m.g.f. of $(Z_i^{(A)}(\theta) + \theta_N) / \theta_N^2$ is given in Tanaka (2017, Chap. 10) as

$$\begin{aligned} m^{(A)}(x) &= \mathbb{E} \left[\exp \left\{ x (Z_i^{(A)}(\theta) + \theta_N) / \theta_N^2 \right\} \right] \\ &= e^{c_N / 2} \left[\cosh \mu + \left(c_N - \frac{2x}{\theta_N} \right) \frac{\sinh \mu}{\mu} \right]^{-1/2}, \end{aligned}$$

where $\mu = \sqrt{c_N^2 - 2x}$. Differentiation of $m^{(A)}(x)$ gives us

$$\begin{aligned} \mathbb{E} \left((Z_i^{(A)}(\theta) + \theta_N) / \theta_N^2 \right) &= \frac{1}{\theta_N} \left(1 - c_N + O(c_N^2) \right) + \frac{1}{2} + O(c_N), \\ \text{Var} \left((Z_i^{(A)}(\theta) + \theta_N) / \theta_N^2 \right) &= \frac{1}{\theta_N^2} \left(2 + O(c_N) \right) + O(1). \end{aligned}$$

It follows from the CLT that, as $N \rightarrow \infty$ under $\theta_N = \theta / N^{1/2}$ and $c_N = c / N^{1/2}$,

$$\sqrt{N} \left(\frac{S_N^{(A)}(\theta)}{\theta_N} - \frac{1}{2} \theta_N \right) \Rightarrow N(-c, 2),$$

which yields (20).

For Models C and D we have $S_N^{(C)}(\theta) = \sum_{i=1}^N Z_i^{(C)}(\theta) / N + o_p(1)$, where $Z_i^{(C)}(\theta)$ is defined in (19) and the m.g.f. of $(Z_i^{(C)}(\theta) + \theta_N) / \theta_N^2$ is given in Tanaka (2017, Chap. 10) as

$$\begin{aligned} m^{(C)}(x) &= \mathbb{E} \left[\exp \left\{ x (Z_i^{(C)}(\theta) + \theta_N) / \theta_N^2 \right\} \right] \\ &= e^{c_N / 2} \left[\left\{ \frac{c_N^2}{\mu^2} + \frac{2x}{\delta_N \mu^4} (c_N - \theta_N) (c_N + c_N \theta_N + \theta_N) \right\} \cosh \mu \right. \\ &\quad \left. + \left\{ c_N - \frac{2x}{3\delta_N \mu_N^2} \left((c_N^2 + 6) (\theta_N + 1) - c_N \theta_N^2 \right) \right. \right. \\ &\quad \left. \left. + \frac{2x}{\delta_N \mu^4} \left((c_N + 1) \theta_N^2 - 2x(\theta_N + 1) \right) \right\} \frac{\sinh \mu}{\mu} \right]^{-1/2}, \end{aligned}$$

where $\mu = \sqrt{c_N^2 - 2x}$. Differentiation of $m^{(C)}(x)$ gives us

$$\begin{aligned} \mathbb{E} \left((Z_i^{(C)}(\theta) + \theta_N) / \theta_N^2 \right) &= \frac{1}{6} - \frac{1}{45} \theta_N^2 - \frac{1}{90} c_N^2 + o(\theta_N^2) + o(c_N^2), \\ \text{Var} \left((Z_i^{(C)}(\theta) + \theta_N) / \theta_N^2 \right) &= \frac{1}{45} + o(1). \end{aligned}$$

It follows from the CLT that, as $N \rightarrow \infty$ under $\theta_N = \theta / N^{1/4}$ and $c_N = c / N^{1/4}$,

$$\sqrt{N} \left(\frac{S_N^{(C)}(\theta) + \theta_N}{\theta_N^2} - \frac{1}{6} + \frac{1}{45} \theta_N^2 \right) \Rightarrow N \left(-\frac{c^2}{90}, \frac{1}{45} \right),$$

which proves (21), and Theorem 6 has been established.

REFERENCES

- Breitung, J. (2000). The local power of some unit root tests for panel data. *Advances in Econometrics* 15:161-178.
- Elliott, G., T. Rothenberg, and J. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica* 64:813-836.
- Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* 48:419-426.
- King, M.L. (1987). Towards a theory of point optimal testing. *Econometric Reviews* 6:169-218.
- Moon, H.R. and B. Perron (2008). Asymptotic local power of pooled t-ratio tests for unit roots in panels with fixed effects. *Econometrics Journal* 11:80-104.
- Moon, H.R., B. Perron, and P.C.B. Phillips (2007). Incidental trends and the power of panel unit root tests. *Journal of Econometrics* 141:416-459.
- Nabeya, S. (2000). Asymptotic distributions for unit root test statistics in nearly integrated seasonal autoregressive models. *Econometric Theory* 16:200-230.
- Nabeya, S. and K. Tanaka (1990). A general approach to the limiting distribution for estimators in time series regression with nonstable autoregressive errors. *Econometrica* 58:145-163.
- Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. New York: John Wiley & Sons.
- Tanaka, K. (2017). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. 2nd ed. New York: John Wiley & Sons.
- Tanaka, K. (2018). Computing limiting local powers and power envelopes of panel MA unit root tests and stationarity tests. to appear in *Econometric Theory*.