POINT-OPTIMAL PANEL UNIT ROOT TESTS WITH SERIALLY CORRELATED ERRORS

by

Hyungsik R. Moon, Benoit Perron, and Peter C. B. Phillips

COWLES FOUNDATION PAPER NO. 1448

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

2015

http://cowles.econ.yale.edu/

The Econometrics Journal

Econometrics Journal (2014), volume **17**, pp. 338–372.
doi: 10.1111/ectj.12030

Point-optimal panel unit root tests with serially correlated errors

HYUNGSIK ROGER MOON[†,‡], BENOIT PERRON[§]
AND PETER C. B. PHILLIPS[‖,#]

[†]*Department of Economics, University of Southern California, Los Angeles, CA 90089, USA.*
E-mail: moonr@usc.edu

[‡]*Department of Economics, Yonsei University, Seoul, Korea.*

[§]*Universite de Montreal, CIREQ, CIRANO, C.P. 6128, Succ. centre-ville, Montreal, QC H3C 3J7, Canada.*
E-mail: benoit.perron@umontreal.ca

[‖]*Cowles Foundation, Yale University, PO Box 208281, New Haven, CT 06520-8281, USA.*
E-mail: peter.phillips@yale.edu

[#]*Department of Economics, University of Auckland, Owen G. Glenn Building, Private Bag 92019, Auckland 1142, New Zealand.*

First version received: February 2012; final version accepted: March 2014

Summary Generalizations of the point-optimal panel unit root tests of Moon, Perron and Phillips (MPP) are developed to cover cases of serially correlated errors. The resulting statistics involve two modifications relative to those of MPP: (a) the error variance is replaced by the long-run variance; (b) centring of the statistic is adjusted to correct for second-order bias effects induced by the correlation between the error and lagged dependent variable.

Keywords: *Bias correction, Incidental trends, Long-run variance, Point-optimal test, Serial dependence.*

1. INTRODUCTION

There has been much recent interest in testing for the presence of stochastic trends in large panels; see, e.g., Breitung and Pesaran (2008) and Breitung and Westerlund (2013). A prototypical model consists of a deterministic trend component d_{it} and an (unobserved) stochastic component y_{it} for some observable panel observations z_{it} for individual $i = 1, \ldots, n$ in period $t = 1, \ldots, T$ satisfying

$$\begin{aligned} z_{it} &= d_{it} + y_{it}, \\ y_{it} &= \rho_i y_{it-1} + u_{it}, \end{aligned} \tag{1.1}$$

where u_{it} is an error term that has zero mean and is stationary over time, and $y_{i0} = 0$ for simplicity. Dynamic panel models with incidental trend components of this type arise in many applications in microeconometrics, multicountry growth studies and international finance.

Empirical interest often centres on the individual dynamics and on whether there is commonality and persistence across individuals (i.e. that the autoregressive parameters ρ_i are all unity) or whether such commonality occurs for certain subgroups of individuals.

Moon et al. (2007, hereafter MPP) developed tests that are point optimal against a specific alternative hypothesis. MPP adopted a local-alternative set-up, specifying the autoregressive parameter as lying in a local vicinity of unity whose width narrows as the sample size increases, according to the form

$$\rho_i = 1 - \frac{\theta_i}{n^\kappa T} \quad \text{for some constant } \kappa > 0, \tag{1.2}$$

where θ_i is a sequence of i.i.d. random variables and κ is a parameter defining the width of the vicinity as $n \to \infty$. The null hypothesis of interest is then

$$\mathbb{H}_0 : \theta_i = 0 \text{ a.s. (i.e. } \rho_i = 1) \quad \text{for all } i, \tag{1.3}$$

with the alternative

$$\mathbb{H}_1 : \theta_i \neq 0 \text{ (i.e. } \rho_i \neq 1) \quad \text{for some } i. \tag{1.4}$$

The MPP tests are point optimal in the sense of giving highest power against a specific set of θ_i. These tests were derived under the assumption that the error term u_{it} is independent across individual units and over time. They represent a panel extension of the work of Elliott et al. (1996) in the time-series case where the autoregressive parameter converges to unity at a rate of $1/T$, regardless of the deterministic component in the model.

Independence assumptions are not realistic in many empirical applications and in this work we extend the MPP tests by allowing for serially correlated errors u_{it}. In their Section 6.4 (p. 436), MPP briefly mentioned this extension. Here, we provide explicit test statistics that have optimal asymptotic properties. The modified tests replace estimated variances of the errors in MPP with estimated long-run variances, and adjust centring terms. Our main purpose is to provide the form of the modified tests and to give their asymptotic properties so that they can be used in empirical work.

The paper is organized as follows. In Section 2, we show how to construct the tests, give results for cases with no fixed effects, fixed effects and incidental trends, and discuss implementation. In Section 3, we report some simulation findings. We conclude in Section 4, and in the Appendix we provide technical derivations and supporting lemmata.

2. TESTS UNDER SERIAL CORRELATION

Following MPP, in the following analysis, we consider three deterministic trend cases: (a) no individual effects (i.e. $d_{it} = 0$ and $z_{it} = y_{it}$); (b) fixed effects (i.e. $d_{it} = b_{0i}$); (c) heterogeneous or incidental linear trends (i.e. $d_{it} = b_{i0} + b_{i1}(t-1)$). In each case, we proceed in three steps. First, we define the likelihood ratio (LR) statistic under Gaussianity, which is known to be optimal by the Neyman–Pearson lemma when the null and alternative hypotheses are simple. Then, we show that this statistic can be approximated by a simpler version with parameters that are consistently estimable. Finally, we derive the asymptotic distribution of this approximation (with appropriate recentring). In all three cases, this asymptotic distribution coincides with the one in MPP.

Our notation is similar to that of MPP. Denote by Z, D, Y, Y_{-1} and U the $(n \times T)$ observation matrices whose (i, t)th elements are z_{it}, d_{it}, y_{it}, y_{it-1} and u_{it}, respectively. Define the $T-$ vectors $G_0 = (1, \ldots, 1)'$, $G_1 = (0, 1, \ldots, T-1)'$ and set $G = (G_0, G_1) = (g_1, \ldots, g_T)'$, where $g_t = (1, t-1)'$. Define $\beta_0 = (b_{01}, \ldots, b_{0n})'$, $\beta_1 = (b_{11}, \ldots, b_{1n})'$ and $\beta = (\beta_0, \beta_1) = (b_1, \ldots, b_n)'$, where $b_i = (b_{0i}, b_{1i})'$. Let $\underline{Z}_i$, $\underline{Y}_i$, $\underline{Y}_{-1,i}$ and $\underline{U}_i$ denote the transpose of the ith row of Z, Y, Y_{-1} and U, respectively. With this notation, the model has the matrix form

$$Z = D + Y, \qquad Y = \rho Y_{-1} + U,$$

where $\rho = \mathrm{diag}(\rho_1, \ldots, \rho_n)$.

Define σ_i^2, ω_i^2 and λ_i as the variance of u_{it}, the long-run variance of u_{it} and the one-sided long-run variance of u_{it}, respectively, so that $\omega_i^2 = \sigma_i^2 + 2\lambda_i$. Let Σ, Ω and Λ be the diagonal matrices with elements σ_i^2, ω_i^2 and λ_i, respectively. Define $\Omega_{u,i} = E[\underline{U}_i \underline{U}_i']$ as the $(T \times T)$ covariance matrix of $\underline{U}_i$ and $\Omega_u = \mathrm{diag}(\Omega_{u,1}, \ldots, \Omega_{u,n})$ as the $(nT \times nT)$ covariance matrix of $\mathrm{vec}(U')$. As in MPP, we assume that the errors u_{it} are cross-section independent over i.

We assume that the localizing coefficient θ_i in the local alternative (1.2) is a sequence of i.i.d. random variables with bounded support.[1] Let $\mu_{\theta,k} = E[\theta_i^k]$, $\rho_{c_i} = 1 - (c_i/(n^\kappa T))$ and define the quasi-difference operator

$$\underset{(T\times T)}{\Delta_{c_i}} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ -\rho_{c_i} & 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & \ddots & 0 & 0 \\ \vdots & & -\rho_{c_i} & 1 & 0 \\ 0 & \ldots & 0 & -\rho_{c_i} & 1 \end{bmatrix}.$$

Set $\mathbb{C} = \mathrm{diag}(c_1, \ldots, c_n)$ and $\Delta_{\mathbb{C}} = \mathrm{diag}(\Delta_{c_1}, \ldots, \Delta_{c_n})$.

The quasi-log-likelihood function of the panel Z, which we use in defining the likelihood ratio test statistic, has the form

$$L_{nT}(\mathbb{C}, D, B) = -\frac{1}{2}(\mathrm{vec}(Z' - D'))' \Delta_{\mathbb{C}}' B \Delta_{\mathbb{C}} (\mathrm{vec}(Z' - D')),$$

for some weight matrix B.

Throughout the paper, we assume panel linear process errors with conditions similar to those in the literature (e.g. Phillips and Moon, 1999). Let $c_j = \sup_i |c_{ij}|$, let $f_i(\lambda)$ be the spectral density of u_{it}, and let $\gamma_j(k) = \int_{-\pi}^{\pi} \exp(\mathrm{i}k) f_j(\lambda) \mathrm{d}\lambda$.

ASSUMPTION 2.1. *(a) $u_{it} = \sum_{j=0}^{\infty} c_{ij} v_{it-j}$, where $v_{it} \sim iid$ with $E[v_{it}] = 0$ and $E|v_{it}|^{8+\epsilon} < \infty$ for some $\epsilon > 0$; (b) $\sum_{j=0}^{\infty} j^m c_j < \infty$ for some $m > 1$, $\gamma(k) = \sup_i |\gamma_i(k)|$, $\phi_j(k) = \int_{-\pi}^{\pi} \exp(\mathrm{i}k)(4\pi^2 f_j(\lambda))^{-1} \mathrm{d}\lambda$, $\phi(k) = \sup_i |\phi_i(k)|$; (c) $\gamma(k), \phi(k) \le Mk^{-s}$ and for $s > 2$ and some constant M.*

These conditions extend the serial dependence restrictions of Elliott et al. (1996) (e.g. their Condition A) to heterogeneous panels. Assumption 2.1(a) assumes that the error term u_{it} follows a linear process that is heterogeneous across i. The higher moments are needed to ensure the

[1] As mentioned in MPP, the assumption of a bounded support for θ_i is made for convenience, and could be relaxed at the cost of stronger moment conditions. It is also convenient to assume that θ_i are identically distributed, and this assumption could be relaxed as long as cross-sectional averages of the moments θ_i have well-defined limits.

large N, T asymptotics of panel data that are heterogeneous across i and serially correlated over t. Under cross-sectional homoscedasticity, these moment conditions could be weakened. Assumptions 2.1(b) and 2.1(c) restrict the temporal dependence of the error term u_{it} to be 'weak' uniformly across i. These restrictions exclude long memory type strong dependence. The conditions in Assumption 2.1 are quite weak and are satisfied by many parametric weak dependent processes, such as stationary and invertible ARMA processes.

While Assumption 2.1 is quite general in terms of the serial correlation that is allowed, it is restrictive in that it assumes that all cross-sectional units are independent. This assumption is not reasonable for many interesting empirical data sets, such as cross-country studies where business-cycle effects are likely to induce correlation across countries. As in MPP, we conjecture that the procedures proposed below are valid after appropriate orthogonalization is applied, for example, after the removal of common factors as in Moon and Perron (2004), Bai and Ng (2004) or Phillips and Sul (2003). Moreover, the development of optimal procedures under cross-sectional dependence is beyond the scope of the current paper.

2.1. *No fixed effect:* $d_{it} = 0$

When $d_{it} = 0$, the model becomes

$$Z = Y, \qquad Y = \rho Y_{-1} + U.$$

Following MPP, in this case we consider local neighbourhoods of unity that shrink at the rate of $1/(n^{1/2}T)$, so that the rate coefficient $\kappa = 1/2$, and one-sided alternatives in which the support of θ_i is a bounded interval $[0, M_\theta]$ for some $M_\theta \geq 0$ so that $\rho_i \leq 1$ under this alternative. In terms of the first moment of θ_i, the hypotheses about ρ_i are

$$\mathbb{H}_0 : \mu_{\theta,1} = 0 \tag{2.1}$$

and

$$\mathbb{H}_1 : \mu_{\theta,1} > 0. \tag{2.2}$$

Suppose that u_{it} are Gaussian so that vec $(U') \sim N(0, \Omega_u)$ with known Ω_u and the initial conditions y_{i0} are all zeros.[2] By the Neyman–Pearson lemma, rejecting a small value of the log-likelihood ratio test statistic

$$-2L_{nT}\left(\mathbb{C}, 0, \Omega_u^{-1}\right) + 2L_{nT}\left(0, 0, \Omega_u^{-1}\right) \tag{2.3}$$

would be the uniformly most powerful test for the null $\rho_i = 1$ for $i = 1, \ldots, n$ against the simple alternative $\rho_i = 1 - (c_i/(n^{1/2}T))$ for $i = 1, \ldots, n$. When the alternative is (1.4) with (2.2), this becomes a point-optimal test.

In order to implement the optimal test statistic (2.3), we need an estimate of the entire $(nT \times nT)$ covariance matrix Ω_u. This is a huge high-dimensional covariance estimation problem in a non-parametric set-up. The following theorem provides an approximation of the likelihood ratio test statistic in (2.3) with a statistic where the unknown nuisance parameters are consistently estimable.

[2] This assumption of a zero initial value is strong. The treatment of initial values in panel unit root tests is still an open problem. In MPP, we showed that the assumption that the initial observation is drawn from the unconditional distribution cannot be easily extended to the panel case because the resulting test statistic diverges to infinity with probability 1.

THEOREM 2.1. *Let Assumption 2.1 hold with* $E[\mathrm{vec}(U')\mathrm{vec}(U')'] = \Omega_u$. *Assume that* $(n/T) \to 0$ *as* $n, T \to \infty$. *Then, for* $\rho_i = 1 - (\theta_i/(n^{1/2}T))$, *we have*

$$-2L_{nT}\big(\mathbb{C}, 0, \Omega_u^{-1}\big) + 2L_{nT}\big(0, 0, \Omega_u^{-1}\big) = -2L_{nT}\big(\mathbb{C}, 0, \Omega^{-1} \otimes I_T\big) + 2L_{nT}\big(0, 0, \Omega^{-1} \otimes I_T\big) - \frac{2}{n^{1/2}} l_n' \mathbb{C} \Omega^{-1} \Lambda l_n + o_p(1).$$

Note that the approximate likelihood ratio statistic

$$-2L_{nT}\big(\mathbb{C}, 0, \Omega^{-1} \otimes I_T\big) + 2L_{nT}\big(0, 0, \Omega^{-1} \otimes I_T\big) - \frac{2}{n^{1/2}} l_n' \mathbb{C} \Omega^{-1} \Lambda l_n \tag{2.4}$$

in Theorem 2.1 employs the Gaussian log-likelihood based on the long-run variance $\Omega \otimes I_T$ with an adjustment of the one-sided long run variance $(2/n^{1/2}) l_n' \mathbb{C} \Omega^{-1} \Lambda l_n$. The one-sided long-run drift correction appears to be a result of the correlation between the stationary error u_{it} and the lagged dependent variable $z_{it-1} = y_{it-1}$. The main advantage of this formulation is that it involves quantities (Ω and Λ) that can be estimated consistently.

The test statistic we propose is to use the approximated log-likelihood ratio (2.4) with appropriate centring. Define

$$V_{nT}(\mathbb{C}) = -2L_{nT}\big(\mathbb{C}, 0, \Omega^{-1} \otimes I_T\big) + 2L_{nT}\big(0, 0, \Omega^{-1} \otimes I_T\big) - \frac{1}{2}\mu_{c,2} - \frac{2}{\sqrt{n}} l_n' \mathbb{C} \Omega^{-1} \Lambda l_n,$$

where $l_n = (1, \ldots, 1)$ is the sum vector and $\mu_{c,2} = E[c_i^2]$.

THEOREM 2.2. *Let Assumption 2.1 hold and* $(n/T) \to 0$ *as* $n, T \to \infty$. *Then, under the local alternative* $\rho_i = 1 - (\theta_i/(n^{1/2}T))$, *we have*

$$V_{nT}(\mathbb{C}) \Rightarrow N(-E[c_i\theta_i], 2\mu_{c,2}),$$

where $\mu_{c,2} = E[c_i^2]$.

REMARK 2.1. We can interpret the test statistic $V_{nT}(\mathbb{C})$ as an asymptotic version of the point-optimal test for panel unit roots with possible serial correlation of unknown form in the error term.

REMARK 2.2. Compared to the corresponding statistic in MPP, which makes no allowance for serial correlation, there are two differences in $V_{nT}(\mathbb{C})$. First, as discussed in MPP, we use the long-run covariance matrix $\Omega \otimes I_T$ instead of the variance matrix $\Sigma \otimes I_T$ as the weight matrix. In addition, we recentre the statistic by subtracting the term $(2/\sqrt{n}) l_n' \mathbb{C} \Omega^{-1} \Lambda l_n$, which corrects for the correlation between the stationary error u_{it} and the lagged dependent variable $z_{it-1} = y_{it-1}$. This term is not required for the test under temporal independence.

REMARK 2.3. The limit distribution of $V_{nT}(\mathbb{C})$ is the same limit as in MPP (see their Theorem 6).

2.2. *Time-invariant fixed effects: $d_{it} = b_{i0}$*

In this section, we consider the case where the incidental trends $d_{it} = b_{0i}$ are fixed over time. This corresponds to the standard fixed effects model. In this case, the model has matrix form

$$Z = \beta_0 G_0' + Y, \qquad Y = \rho Y_{-1} + U.$$

As before, suppose that $\mathrm{vec}(U') \sim N(0, \Omega_u)$ with known Ω_u and the initial conditions y_{i0} are all zeros. Then, rejecting a small value of the test statistic,

$$-2\big(\min_{\beta_0} L_{nT}\big(\mathbb{C}, \beta_0 G_0', \Omega_u^{-1}\big) - \min_{\beta_0} L_{nT}\big(0, \beta_0 G_0', \Omega_u^{-1}\big)\big), \tag{2.5}$$

for the null $\rho_i = 1$ for $i = 1, \ldots, n$ and the alternative $\rho_i = 1 - (c_i/(n^{1/2}T))$ for $i = 1, \ldots, n$, is known as the uniformly most powerful invariant test that is invariant with respect to the transformation $Z \to Z + \beta_0^* G_0'$ for arbitrary β_0^*. Against the alternative in (1.4), this becomes a point-optimal invariant test (e.g. Dufour and King, 1991).

As mentioned in the previous section, this statistic is difficult to implement because of the presence of Ω_u, the full $(nT \times nT)$ covariance matrix of the error. This again motivates the use of an approximation.

THEOREM 2.3. *Let Assumption 2.1 hold with $E[\mathrm{vec}\,(U')\mathrm{vec}\,(U')'] = \Omega_u$ and let $(n/T^{1/2}) \to 0$ as $n, T \to \infty$. Then, for $\rho_i = 1 - (\theta_i/(n^{1/2}T))$, we have*

$$\begin{aligned}
&-2\big(\min_{\beta_0} L_{nT}\big(\mathbb{C}, \beta_0 G_0', \Omega_u^{-1}\big) - \min_{\beta_0} L_{nT}\big(0, \beta_0 G_0', \Omega_u^{-1}\big)\big)\\
&\quad = -2\big(\min_{\beta_0} L_{nT}\big(\mathbb{C}, \beta_0 G_0', \Omega^{-1} \otimes I_T\big) - \min_{\beta_0} L_{nT}\big(0, \beta_0 G_0', \Omega^{-1} \otimes I_T\big)\big)\\
&\qquad - \frac{2}{n^{1/2}} l_n' \mathbb{C} \Omega^{-1} \Lambda l_n + o_p(1).
\end{aligned}$$

REMARK 2.4. This approximation is derived under the stronger rate condition $(n/T^{1/2}) \to 0$ as $n, T \to \infty$ in place of the condition $(n/T) \to 0$ as $n, T \to \infty$ that is used without fixed effects.

REMARK 2.5. The approximation involves the same correction for second-order bias as in the case without fixed effects.

Again, the test statistic we propose is the approximate log-likelihood ratio (2.4) with appropriate centring. Define

$$\begin{aligned}
V_{nT,fe1}(\mathbb{C}) &= -2\big(\min_{\beta_0} L_{nT}\big(\mathbb{C}, \beta_0 G_0', \Omega^{-1} \otimes I_T\big) - \min_{\beta_0} L_{nT}\big(0, \beta_0 G_0', \Omega^{-1} \otimes I_T\big)\big)\\
&\quad - \frac{1}{2}\mu_{c,2} - \frac{2}{\sqrt{n}} l_n' \mathbb{C} \Omega^{-1} \Lambda l_n.
\end{aligned}$$

THEOREM 2.4. *Let Assumption 2.1 hold and let $(n/T) \to 0$ as $n, T \to \infty$. Then, under the local alternative $\rho_i = 1 - (\theta_i/(n^{1/2}T))$, we have*

$$V_{nT,fe1}(\mathbb{C}) \Rightarrow N\big(-E[c_i\theta_i], 2\mu_{c,2}\big),$$

where $\mu_{c,2} = E[c_i^2]$.

This asymptotic distribution is the same as without fixed effects and as in MPP (see their Theorem 9).

2.3. *Incidental trends: $d_{it} = b_{i0} + b_{i1}t$*

Under heterogeneous linear trends, we follow MPP and use local neighbourhoods of unity that shrink at the slower rate of $1/(n^{1/4}T)$, so that the rate coefficient is $\kappa = 1/4$. The alternative might be two-sided; that is, $\theta_i \sim iid$ with mean μ_θ and variance σ_θ^2, with a support that is a subset of a bounded interval $[-M_{l\theta}, M_{u\theta}]$, where $M_{l\theta}, M_{u\theta} \geq 0$. The slower rate of shrinkage in the local neighbourhoods of unity is the result of the presence of heterogeneous trend effects in the panel. The presence of these incidental trends reduces discriminatory power in testing for the presence of common stochastic trends, so wider localizing intervals are needed to attain non-trivial power functions.

Under these conditions, hypotheses (1.3) and (1.4) can be re-expressed as

$$\mathbb{H}_0 : \mu_{\theta,2} = 0 \tag{2.6}$$

and

$$\mathbb{H}_1 : \mu_{\theta,2} > 0. \tag{2.7}$$

Again, suppose that $\text{vec}(U') \sim N(0, \Omega_u)$ with known Ω_u and the initial conditions y_{i0} are all zeros. Then, similar to the case of time-invariant fixed effects, rejecting a small value of the test statistic,

$$-2\big(\min_\beta L_{nT}\big(\mathbb{C}, \beta G', \Omega_u^{-1}\big) - \min_\beta L_{nT}(0, \beta G', \Omega_u^{-1})\big),$$

for the null $\rho_i = 1$ for $i = 1, \ldots, n$ and the alternative $\rho_i = 1 - (c_i/(n^{1/4}T))$ for $i = 1, \ldots, n$, is known as the uniformly most powerful invariant test (with respect to the linear transformation $Z \to Z + \beta^* G'$ for arbitrary β^*), and against the alternative in (1.4), it becomes a point-optimal invariant test. As before, we start by proving the validity of an approximation to this log-likelihood ratio.

THEOREM 2.5. *Let Assumption 2.1 hold with $E[\text{vec}(U')\text{vec}(U')'] = \Omega_u$ and let $(n/T^{1/4}) \to 0$ as $n, T \to \infty$. Then, for $\rho_i = 1 - (\theta_i/(n^{1/4}T))$, we have*

$$-2\big(\min_\beta L_{nT}\big(\mathbb{C}, \beta G', \Omega_u^{-1}\big) - \min_\beta L_{nT}\big(0, \beta G', \Omega_u^{-1}\big)\big) = -2\big(\min_\beta L_{nT}\big(\mathbb{C}, \beta G', \Omega^{-1} \otimes I_T\big)$$
$$- \min_\beta L_{nT}\big(0, \beta G', \Omega^{-1} \otimes I_T\big)\big) - \frac{2}{n^{1/4}} l_n' \mathbb{C} \Omega^{-1} \Lambda l_n + o_p(1).$$

REMARK 2.6. This approximation is derived under the condition $(n/T^{1/4}) \to 0$ as $n, T \to \infty$, which is a stronger rate condition than that used for the intercepts case.

REMARK 2.7. As before, the correction is a result of the presence of a second-order bias term arising from the correlation between the lagged dependent variables and the error term.

Again, we propose to use the approximate log-likelihood ratio with appropriate centring as a test statistic. Define

$$V_{nT,fe2}(\mathbb{C}) = -2\big(\min_{\beta} L_{nT}\big(\mathbb{C}, \beta G', \Omega^{-1} \otimes I_T\big) - \min_{\beta} L_{nT}\big(0, \beta G', \Omega^{-1} \otimes I_T\big)\big)$$
$$+\frac{1}{n^{1/4}} l_n' \mathbb{C}\Omega^{-1}\Sigma l_n + \frac{1}{n^{1/2}}\big(l_n'\mathbb{C}^2 l_n\big)\omega_{p2T} + \frac{1}{n}\big(l_n'\mathbb{C}^4 l_n\big)\omega_{p4T},$$

where

$$\omega_{p2T} = -\frac{1}{T}\sum_{t=1}^{T}\frac{t}{T} + \frac{2}{T}\sum_{t=1}^{T}\left(\frac{t}{T}\right)^2 - \frac{1}{3},$$
$$\omega_{p4T} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{t}{T}\frac{s}{T}\min\left(\frac{t}{T}, \frac{s}{T}\right) - \frac{2}{3}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{t}{T}\right)^2 + \frac{1}{9}.$$

THEOREM 2.6. *Let Assumption 2.1 hold and let $(n/T) \to 0$ as $n, T \to \infty$. Then, under the local alternative $\rho_i = 1 - (\theta_i/(n^{1/4}T))$, we have*

$$V_{nT,fe2}(\mathbb{C}) \Rightarrow N\Big(-\frac{1}{90}E[c_i^2\theta_i^2], \frac{1}{45}E[c_i^4]\Big).$$

As before, $V_{fe2,nT}(\mathbb{C})$ reduces to the statistic from MPP when there is no serial correlation, and it has the same asymptotic distribution as in Theorem 13 of MPP.

2.4. *Implementation of the tests*

The test statistics $V_{nT}(\mathbb{C})$, $V_{nT,fe1}(\mathbb{C})$ and $V_{nT,fe2}(\mathbb{C})$ depend on unknown parameters $\{\sigma_i^2\}$, $\{\omega_i^2\}$ and $\{\lambda_i\}$. Let $\hat{\sigma}_i^2$, $\hat{\omega}_i^2$ and $\hat{\lambda}_i$ be consistent estimators of σ_i^2, ω_i^2 and λ_i, respectively. Similarly define the diagonal matrices of these elements as $\hat{\Sigma}$, $\hat{\Omega}$ and $\hat{\Lambda}$. To implement these tests, we can replace Σ, Ω and Λ in $V_{nT}(\mathbb{C})$, $V_{nT,fe1}(\mathbb{C})$ and $V_{nT,fe2}(\mathbb{C})$ with $\hat{\Sigma}$, $\hat{\Omega}$ and $\hat{\Lambda}$ and we denote the test statistics as $\hat{V}_{nT}(\mathbb{C})$, $\hat{V}_{nT,fe1}(\mathbb{C})$ and $\hat{V}_{nT,fe2}(\mathbb{C})$. We assume the following regarding these estimators.

ASSUMPTION 2.2. *Under the local alternative,* $\sup_i E[\hat{\sigma}_i^2 - \sigma_i^2]^2 = o(1/n)$, $\sup_i E[\hat{\omega}_i^2 - \omega_i^2]^2 = o(1/n)$ *and* $\sup_i E[\hat{\lambda}_i^2 - \lambda_i^2]^2 = o(1/n)$.

REMARK 2.8. An example of $\hat{\sigma}_i^2$ that satisfies Assumption 2.2 is the time-series sample variance of Δz_{it}:

$$\hat{\sigma}_i^2 = \frac{1}{T-1}\sum_{t=2}^{T}\left(\Delta z_{it} - \left(\frac{1}{T-1}\sum_{t=2}^{T}\Delta z_{it}\right)\right)^2.$$

REMARK 2.9. When kernel spectral density estimation is used for $\hat{\omega}_i^2$ and $\hat{\lambda}_i^2$ with bandwidth h, Assumption 2.2 is satisfied if: (a) the kernel function $K(\cdot) : \mathbb{R} \to [0, 1]$ is continuous at zero and all but a finite number of other points, satisfying $K(0) = 1$, $K(x) = K(-x)$, $\int_{-\infty}^{\infty} K(x)^2 \mathrm{d}x < M$, and $K_q = \lim_{x\to 0}[1 - K(x)/|x|^q] < \infty$ for some $0 < q \leq m$, where parameter m is defined in Assumption 2.1(b); (b) the bandwidth h satisfies

$$\frac{nh}{T} + \frac{n}{h^{2q}} = o(1) \tag{2.8}$$

as $(n/T) \to 0$ and $h \to \infty$*;* see, e.g., Moon and Perron (2004). If $(n/T) = o(T^{-a})$ for some $0 < a < 1$ and $q \geq (1/2)((1-a)/a)$*,* then the bandwidth condition (2.8) is satisfied if

$$T^{(1/2q)(1-a)} \lesssim h \lesssim T^a;$$

that is,

$$\frac{h}{T^a}, \frac{T^{(1/2q)(1-a)}}{h} = O(1).$$

THEOREM 2.7. *Under Assumptions 2.1 and 2.2, as* $n, T \to \infty$ *with* $(n/T) \to \infty$, *we have* $\hat{V}_{nT}(\mathbb{C}) = V_{nT}(\mathbb{C}) + o_p(1)$, $\hat{V}_{nT,fe1}(\mathbb{C}) = V_{nT,fe1}(\mathbb{C}) + o_p(1)$ *and* $\hat{V}_{nT,fe2}(\mathbb{C}) = V_{nT,fe2}(\mathbb{C}) + o_p(1)$ *under the local alternative.*

Implementation of the tests also requires the choice of an alternative to define the likelihood ratio, $\mathbb{C}$. MPP have shown that the optimal choice of c_i is θ_i. With this choice, the likelihood ratio statistics above attain the power envelope. However, this choice is infeasible because θ_i are the parameters under test. MPP proposed the assumption of a common $c_i = c$ for all cross-sectional units, and they called this test a common point-optimal (CPO) test. With this choice, $\mathbb{C} = cI_n$, and we can deduce from Theorems 2.2, 2.4 and 2.6 that under the null hypothesis,

$$\sqrt{\frac{1}{2c^2}} V_{nT}(cI_n) \Rightarrow N(0, 1),$$

$$\sqrt{\frac{1}{2c^2}} V_{nT,fe1}(cI_n) \Rightarrow N(0, 1),$$

$$\sqrt{\frac{45}{c^4}} V_{nT,fe2}(cI_n) \Rightarrow N(0, 1),$$

while under the alternative hypothesis,

$$\sqrt{\frac{1}{2c^2}} V_{nT}(cI_n) \Rightarrow N\Big(-\frac{1}{\sqrt{2}} E[\theta_i], 1\Big),$$

$$\sqrt{\frac{1}{2c^2}} V_{nT,fe1}(cI_n) \Rightarrow N\Big(-\frac{1}{\sqrt{2}} E[\theta_i], 1\Big),$$

$$\sqrt{\frac{45}{c^4}} V_{nT,fe2}(cI_n) \Rightarrow N\Big(-\frac{\sqrt{5}}{30} E[\theta_i^2], 1\Big).$$

The surprising result here is that neither distribution depends on the choice of c used to construct the test. This feature implies that the power is the same for all choices of c asymptotically, although that choice matters in finite samples. Based on the simulation evidence provided in MPP, we set $c = 1$ in the simulation below. Of course, this choice of $\mathbb{C}$ is not optimal unless the alternative hypothesis is homogeneous ($\theta_i = \theta \neq 0$ for all i) and results in a power loss relative to the power envelope.

3. MONTE CARLO SIMULATIONS

In this section, we report the results of a Monte Carlo experiment designed to assess the finite-sample properties of the tests presented above and compare them with other existing results. For this purpose, we use the same DGP as MPP but employ either an AR(1) or MA(1) process for the innovations. Thus, the generating model has the following form,

$$z_{it} = b_{0i} + b_{1i}t + y_{it},$$
$$y_{it} = \rho_i y_{it-1} + u_{it},$$

where the innovations follow either an AR(1) process

$$u_{it} = \gamma_\iota u_{it,t-1} + \varepsilon_{it}$$
$$\varepsilon_{it} \sim iid\ N\big(0, \sigma_i^2\big(1 - \gamma_i^2\big)\big),$$

or an MA(1) process

$$u_{it} = \varphi_i \varepsilon_{it-1} + \varepsilon_{it}$$
$$\varepsilon_{it} \sim iid\ N\left(0, \sigma_i^2\left(\frac{1}{1 + \varphi_i^2}\right)\right).$$

We have also looked at some ARMA(1,1) cases but we do not report those results in order to ease presentation (these results are available from the authors upon request). We consider five specifications for serial correlation: white noise ($\gamma_i = \varphi_i = 0$), positive AR ($\varphi_i = 0$ and $\gamma_i \sim U[0, 0.4]$), negative AR ($\varphi_i = 0$ and $\gamma_i \sim U[-0.4, 0]$), positive MA ($\gamma_i = 0$ and $\varphi_i \sim U[0, 0.4]$) and negative MA ($\gamma_i = 0$ and $\varphi_i \sim U[-0.4, 0]$).

In all cases, we allow for heterogeneity and draw the idiosyncratic variance σ_i^2 from a uniform distribution, $\sigma_i^2 \sim U[0.5, 1.5]$. This variance is scaled such that the scale of u_{it} is the same for all cases. In both the incidental intercepts case ($b_{1i} = 0$) and the incidental trends case ($b_{1i} \neq 0$), the parameters are drawn from $iid\ N(0, 1)$.

We focus the study on the size and size-adjusted power of the CPO test with $c_i = 1$ for all i, because MPP advocated that choice. For size calculations, we set $\rho_i = 1$ for all i, which corresponds to $\theta_i = 0$ for all i in our local-to-unity framework. For power, we draw θ_i from a uniform distribution between 0 and 8, as in one of the experiments of MPP. This specification ensures that power should be roughly constant as N and T increase.

We draw comparisons with two existing tests, those of Levin et al. (2002), hereafter LLC, and Im et al. (2003), hereafter IPS. We take three values for n (10, 25 and 100) and two values of T (100 and 250). All tests are conducted at the 5% significance level, and the number of replications is set at 2,000.

Estimation of the long-run variance and one-sided long-run variance is critical to the performance of the CPO test. We estimate these quantities in two ways based on demeaned first differences, as in Remark 2.8. The first method is a non-parametric estimator with quadratic

Table 1. Size of tests: incidental intercepts.

		N = 10			N = 25			N = 100		
		T =			T =			T =		
		100	250	500	100	250	500	100	250	500
White noise	This paper, no PW	1.6	2.2	2.6	2.8	3.2	3.6	4.2	3.7	5.2
	This paper, PW	2.4	2.4	2.5	3.0	3.8	4.5	0.6	2.5	4.4
$\gamma_i = 0$	MPP (2007)	2.9	2.4	2.5	4.8	3.7	4.4	4.3	3.6	4.7
$\varphi_i = 0$	IPS	6.4	3.7	4.9	7.0	5.2	5.2	8.4	6.9	5.1
	LLC	5.8	4.4	4.1	5.8	5.3	3.4	6.0	5.0	3.5
Positive AR	This paper, no PW	1.7	1.8	2.9	2.4	2.8	3.3	4.5	4.5	4.9
	This paper, PW	2.3	2.1	3.0	1.7	2.3	3.0	0.7	0.4	2.4
$\gamma \sim U[0, 0.4]$	MPP (2007)	0.3	0.5	0.8	0.1	0.1	0.1	0.0	0.0	0.0
$\varphi = 0$	IPS	5.0	4.4	3.6	4.7	3.9	3.5	3.1	3.1	3.5
	LLC	5.0	4.2	4.0	4.3	4.3	3.7	2.6	2.8	3.9
Negative AR	This paper, no PW	1.7	2.1	2.2	2.5	3.2	4.0	5.4	4.2	5.2
	This paper, PW	2.0	2.5	2.1	1.2	2.6	3.3	0.3	1.5	2.2
$\gamma_i \sim U[-0.4, 0]$	MPP (2007)	12.7	13.2	13.6	30.7	30.3	32.4	82.2	81.6	83.9
$\varphi_i = 0$	IPS	12.8	7.3	5.4	20.4	9.4	7.2	42.6	15.3	9.3
	LLC	9.0	5.8	4.2	12.0	7.0	4.0	20.8	8.5	4.3
Positive MA	This paper, no PW	1.4	1.7	2.6	2.1	3.0	3.7	4.3	4.5	4.8
	This paper, PW	2.5	2.6	3.0	2.7	3.2	4.1	1.6	3.5	4.5
$\gamma_i = 0$	MPP (2007)	0.7	0.7	0.6	0.3	0.2	0.6	0.0	0.0	0.0
$\varphi_i \sim U[0, 0.4]$	IPS	6.6	6.1	5.7	6.8	6.4	6.0	7.1	7.4	8.4
	LLC	6.1	5.9	4.7	5.4	6.0	4.6	5.2	7.2	5.6
Negative MA	This paper, no PW	1.3	2.0	2.6	2.3	3.5	3.6	4.6	4.5	4.3
	This paper, PW	1.5	2.4	3.0	0.9	3.1	4.0	0.0	1.6	3.5
$\gamma_i = 0$	MPP (2007)	17.0	16.7	17.6	40.7	44.1	40.7	93.0	91.8	93.1
$\varphi_i \sim U[-0.4, 0]$	IPS	25.0	13.8	10.2	40.4	22.6	14.3	85.0	50.0	32.9
	LLC	15.6	9.4	6.9	22.2	12.3	6.5	56.6	25.9	14.4

Note: The table reports the rejection frequency (in %) of a 5% test for a panel unit root. "This paper, no PW" refers to the common point optimal (CPO) tests proposed in this paper with no pre-whitening used when estimating long-run variances and c = 1. "This paper, PW" refers to the CPO tests in this paper with pre-whitening when estimating long-run variances. "MPP (2007)" refers to the CPO tests with c = 1 in MPP that do not allow for serial correlation. "IPS" is the *t*-bar test of Im et al. (2003) and "LLC" is the test of Levin et al. (2002).

spectral kernel and bandwidth selected in a data-based manner using the Andrews (1991) rule with no pre-whitening (PW). The second method uses pre-whitening where the appropriate model is chosen using BIC among the AR(1), MA(1), ARMA(1,1) and constant-only models. In the case of the LLC test, we follow the recommendation of Levin et al. (2002) and use a Bartlett kernel with a bandwidth equal to $3.21T^{1/3}$. Westerlund (2009) has shown that this choice gives the LLC test higher power than selecting the bandwidth in a data-dependent way. Similarly, for

Table 2. Size-adjusted power of tests - Incidental intercepts.

		$N = 10$			$N = 25$			$N = 100$		
		$T =$			$T =$			$T =$		
		100	250	500	100	250	500	100	250	500
White noise	This paper, no PW	41.6	49.3	52.9	53.0	62.6	62.8	66.4	75.2	71.7
	This paper, PW	43.2	54.8	52.5	45.1	62.2	62.6	40.5	73.0	73.4
$\gamma_i = 0$	MPP (2007)	50.5	55.1	54.1	59.4	66.4	62.5	76.7	78.7	74.6
$\varphi_i = 0$	IPS	13.2	18.1	12.9	16.8	17.5	16.9	21.6	18.6	21.3
	LLC	3.7	1.9	1.1	5.3	3.3	2.8	5.7	3.8	3.1
Positive AR	This paper, no PW	42.9	51.8	45.3	54.8	62.8	63.6	63.7	69.5	73.4
	This paper, PW	48.8	53.1	48.0	59.5	64.1	62.2	63.5	73.5	74.8
$\gamma_i \sim U[0, 0.4]$	MPP (2007)	52.0	54.5	45.4	62.4	60.7	63.9	72.2	73.1	73.7
$\varphi_i = 0$	IPS	12.0	13.0	16.1	12.4	17.2	14.8	21.3	19.5	20.7
	LLC	3.1	2.3	1.8	3.5	2.7	1.6	6.0	5.2	3.5
Negative AR	This paper, no PW	42.8	48.3	51.3	53.2	60.3	61.1	63.3	72.0	72.4
	This paper, PW	45.0	52.5	51.4	47.2	59.8	61.0	40.0	69.6	76.3
$\gamma_i = 0$	MPP (2007)	48.4	47.8	50.5	62.9	59.6	60.6	70.1	71.5	74.7
$\gamma_i \sim U[-0.4, 0]$	IPS	13.0	14.6	14.4	19.7	16.1	15.8	22.0	20.0	19.0
	LLC	5.2	2.6	1.3	6.1	2.3	2.4	7.9	4.6	3.2
Positive MA	This paper, no PW	44.2	48.0	51.3	52.8	63.2	59.8	63.6	72.8	73.0
	This paper, PW	47.6	51.8	53.5	59.0	64.5	61.2	58.4	74.7	74.8
$\gamma_i = 0$	MPP (2007)	52.2	54.0	50.9	61.3	66.4	64.2	72.3	75.8	75.2
$\varphi_i \sim U[0, 0.4]$	IPS	10.4	12.7	12.1	14.3	16.4	17.4	17.9	20.8	18.7
	LLC	2.6	1.9	1.1	4.6	2.7	1.9	5.5	4.0	3.4
Negative MA	This paper, no PW	41.4	49.4	47.5	54.8	55.5	59.9	65.9	72.7	77.9
	This paper, PW	46.7	51.8	48.7	43.0	57.6	62.2	32.6	68.1	77.8
$\gamma_i \sim U[-0.4, 0]$	MPP (2007)	48.6	48.7	46.3	57.9	57.2	58.4	68.8	70.0	72.1
$\varphi_i = 0$	IPS	15.0	16.7	14.2	21.2	15.3	18.8	23.9	26.7	22.5
	LLC	8.1	2.6	1.6	8.5	4.3	3.7	11.8	6.7	3.6

the IPS and LLC tests, the choice of lag augmentation is critical for performance. We choose this in a data-dependent way by BIC with a maximum of six lags. For both of these tests, we use the finite-sample adjustments provided in the original papers.

The size results are reported in Tables 1 and 3 for the incidental intercepts and trends cases respectively, while size-adjusted power is in Tables 2 and 4. For each of the five serial correlation specifications, each row corresponds to a different test.

In general, we see that our CPO test is conservative. This is especially true in the incidental trends case. The test is better behaved without pre-whitening in estimating the long-run variance in the incidental intercepts case, but the reverse is true in the trends case. It is evident that the

Table 3. Size of tests: incidental trends.

		$N = 10$			$N = 25$			$N = 100$		
			$T =$			$T =$			$T =$	
		100	250	500	100	250	500	100	250	500
White noise	This paper, no PW	0.0	0.1	0.5	0.0	0.6	1.5	0.0	0.5	0.9
	This paper, PW	1.0	1.0	1.4	1.8	2.3	2.7	1.4	4.1	4.5
$\gamma_i = 0$	MPP (2007)	1.7	1.3	1.4	4.6	2.5	2.9	7.2	5.1	4.4
$\varphi_i = 0$	IPS	8.2	5.5	5.2	10.6	5.6	4.4	14.9	7.5	6.4
	LLC	6.4	4.5	2.1	5.8	4.3	1.1	4.8	4.4	0.5
Positive AR	This paper, no PW	0.0	0.0	0.5	0.0	0.4	1.5	0.0	0.2	1.2
	This paper, PW	0.8	0.4	1.0	1.6	2.2	2.2	2.1	2.1	2.0
$\gamma_i \sim U[0, 0.4]$	MPP (2007)	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
$\varphi_i = 0$	IPS	5.3	3.5	4.4	4.2	3.0	4.0	3.8	2.8	3.3
	LLC	3.7	2.7	1.7	2.0	2.6	0.9	0.7	1.2	0.2
Negative AR	This paper, no PW	0.0	0.1	0.5	0.0	0.2	1.0	0.0	0.3	1.3
	This paper, PW	0.4	0.6	0.7	0.5	1.2	1.9	0.2	1.7	2.1
$\gamma_i \sim U[-0.4, 0]$	MPP (2007)	22.2	19.8	20.3	61.7	58.4	56.8	99.3	99.3	99.1
$\varphi_i = 0$	IPS	25.3	10.7	7.1	40.6	15.6	8.9	84.1	31.7	13.7
	LLC	16.9	7.6	2.8	24.5	9.4	1.6	53.8	14.5	0.5
Positive MA	This paper, no PW	0.0	0.2	0.5	0.0	0.6	1.2	0.0	0.2	1.8
	This paper, PW	1.0	1.3	1.2	3.7	3.4	2.6	6.5	6.5	5.3
$\gamma_i = 0$	MPP (2007)	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0
$\varphi_i \sim U[0, 0.4]$	IPS	8.1	6.6	6.0	10.7	7.2	6.7	12.4	11.4	8.4
	LLC	5.5	5.5	1.7	5.6	5.9	1.8	2.6	7.0	0.6
Negative MA	This paper, no PW	0.0	0.1	0.4	0.0	0.3	1.1	0.0	0.4	0.9
	This paper, PW	0.1	0.7	0.9	0.4	2.4	2.9	0.2	4.0	3.9
$\gamma_i = 0$	MPP (2007)	31.7	29.1	29.7	77.1	74.8	73.3	100.0	100.0	100.0
$\varphi_i \sim U[-0.4, 0]$	IPS	48.2	23.5	15.3	77.6	42.2	26.6	99.9	86.0	63.0
	LLC	38.1	17.7	6.1	61.1	30.1	7.0	96.6	67.8	12.2

original MPP test is not robust to the presence of serial correlation with size that can be as low as 0 or as high as 1, and that the extensions proposed here are therefore needed.

Table 2 reports results for size-adjusted power in the intercept case. Size adjustment is needed, given some of the large size distortions reported in Table 1. We see that the size-adjusted power of the CPO tests robust to serial correlation is typically lower than that of uncorrected tests, but the difference becomes smaller as N and T increase, as predicted by theory because the tests have the same asymptotic distribution. Also, we see that these tests have much higher size-adjusted power than either the LLC or IPS tests. The LLC test has poor power because it

Table 4. Size-adjusted power of tests: incidental intercepts.

		$N = 10$			$N = 25$			$N = 100$		
		$T =$			$T =$			$T =$		
		100	250	500	100	250	500	100	250	500
White noise	This paper, no PW	10.8	17.0	15.8	12.9	17.3	17.2	17.5	23.2	26.1
	This paper, PW	15.6	17.6	16.2	14.2	20.5	19.3	19.9	27.9	28.2
$\gamma_i = 0$	MPP (2007)	18.6	19.1	15.6	17.7	20.9	19.4	28.1	27.5	28.9
$\varphi_i = 0$	IPS	10.5	10.1	10.1	9.8	12.8	12.7	13.9	14.0	11.9
	LLC	8.5	6.8	4.9	9.2	7.9	5.1	11.6	8.4	5.8
Positive AR	This paper, no PW	11.0	15.6	16.4	10.9	18.1	19.5	18.1	26.0	27.5
	This paper, PW	17.1	18.7	16.5	18.8	22.3	19.8	26.2	29.9	29.5
$\gamma_i \sim U[0, 0.4]$	MPP (2007)	17.7	16.8	15.4	18.7	20.3	19.3	27.6	27.7	28.6
$\varphi = 0$	IPS	8.5	9.9	8.7	11.1	13.5	11.3	12.1	12.6	13.3
	LLC	6.2	6.6	3.9	8.5	7.0	6.0	8.6	5.6	5.0
Negative AR	This paper, no PW	10.7	15.1	15.2	14.1	16.8	18.0	15.8	23.3	23.8
	This paper, PW	13.9	16.6	16.9	16.7	20.1	19.0	20.1	28.6	25.9
$\gamma_i \sim U[-0.4, 0]$	MPP (2007)	16.5	16.6	16.3	21.8	17.5	19.2	26.5	24.3	27.0
$\varphi_i = 0$	IPS	12.0	9.6	10.4	11.2	9.9	10.9	11.9	10.9	13.8
	LLC	10.7	7.3	4.6	10.7	6.1	5.5	11.0	7.5	6.8
Positive MA	This paper, no PW	12.0	14.6	18.4	11.9	18.0	22.3	19.3	25.8	30.0
	This paper, PW	20.2	17.9	18.3	18.6	21.2	23.8	25.8	30.4	31.0
$\gamma_i = 0$	MPP (2007)	19.0	16.2	19.3	19.4	20.7	24.4	32.1	28.1	30.3
$\varphi_i \sim U[0, 0.4]$	IPS	9.1	11.2	10.3	9.9	11.2	11.0	14.0	12.6	16.0
	LLC	8.3	7.7	6.0	7.8	8.0	5.1	11.5	7.0	6.9
Negative MA	This paper, no PW	10.7	16.3	18.0	13.0	19.3	22.0	15.0	23.2	29.6
	This paper, PW	14.3	16.7	17.5	11.8	18.7	20.1	12.2	23.7	29.4
$\gamma_i = 0$	MPP (2007)	14.0	14.6	16.3	18.1	18.3	21.2	24.1	23.1	27.7
$\varphi_i \sim U[-0.4, 0]$	IPS	9.9	9.4	10.2	11.2	12.2	11.7	15.5	12.6	14.2
	LLC	11.1	6.7	5.2	11.5	8.3	5.5	12.7	7.6	6.9

corrects for bias by adjusting the numerator of the pooled OLS estimator, as pointed out by Moon and Perron (2008) and Breitung and Westerlund (2013).[3]

Table 4 presents size-adjusted power for the incidental trends case. Note that the alternative considered in this scenario is further from the unit root null than in Table 2 because of the different definition of local neighbourhoods. While the CPO tests have lower power in this case, the same conclusions remain as in the intercept case.

[3] Note that Moon and Perron (2008) incorrectly labelled their t^+ statistic as equivalent to the LLC statistic, while they should have labelled their $t^\#$ statistic (which has lower power) as equivalent to the LLC statistic.

4. CONCLUSION

In this paper we develop generalizations of the point-optimal panel unit root tests of MPP to cover the case where the error term is serially correlated. The resulting statistics have two simple modifications relative to those in MPP. First, the variance of the errors is replaced by the long-run variance. Second, the centring of the statistic is adjusted to accommodate the second-order bias induced by the correlation between the error and lagged values of the dependent variable. Simulations show that these two adjustments lead to appropriately sized tests in most cases.

ACKNOWLEDGEMENTS

We thank Vanessa Smith for raising with us questions about the performance of the original point-optimal statistics in MPP when there are serial correlated errors and about the need for possibly different correction factors in that case. B. Perron acknowledges financial support from the SSHRC and FQRSC. P. C. B. Phillips acknowledges partial support from the National Science Foundation under Grant Nos. SES-0956687 and SES-1258258.

REFERENCES

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*, 817–58.

Bai, J. and S. Ng (2004). A PANIC attack on unit roots and cointegration. *Econometrica 72*, 1127–77.

Breitung, J. and J. Westerlund (2013). Lessons from a decade of IPS and LLC. *Econometric Reviews 32*, 547–91.

Breitung, J. and M. H. Pesaran (2008). Unit roots and cointegration in panels. In L. Matyas and P. Sevestre (Eds.), *The Econometrics of Panel Data*, 279–322. Berlin: Springer.

Dufour, J. and M. King (1991). Optimal Invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors. *Journal of Econometrics 47*, 115–43.

Elliott, G., T. Rothenberg and J. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*, 813–36.

Im, K., H. Pesaran and Y. Shin (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics 115*, 53–74.

Levin, A., F. Lin and C. Chu (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics 108*, 1–24.

Moon, H. R. and B. Perron (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics 122*, 81–126.

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*, 416–59.

Phillips, P. C. B. and H. R. Moon (1999). Linear regression limit theory for nonstationary panel data. *Econometrica 67*, 1057–111.

Phillips, P. C. B. and D. Sul (2003). Dynamic panel estimation and homogeneity testing under cross-section dependence. *Econometrics Journal 6*, 217–39.

Westerlund, J. (2009). A note on the use of the LLC panel unit root test. *Empirical Economics 37*, 517–31.

APPENDIX A: PROOFS OF THE APPROXIMATIONS IN THEOREMS 2.1, 2.3 AND 2.5

We provide three appendices. Here, in Appendix A, we provide proofs of Theorems 2.1, 2.3 and 2.5 that approximate the Gaussian log-likelihood ratio statistic. In Appendix B, we provide sketches of the proofs of the limit distribution results in Theorems 2.2, 2.4 and 2.6. In Appendix C, we provide a heuristic proof of Theorem 2.7. We only provide sketches of the proofs in Appendices B and C because the details are similar to those of the corresponding theorems in MPP and can be established with only minor modifications. Throughout the appendices, M denotes a generic (finite) constant.

Proof of Theorem 2.1: Here, $\kappa = 1/2$. Let Assumption 2.1 hold and $(n/T) \to 0$. By definition, we can write

$$
\begin{aligned}
&-2L_{nT}\left(\mathbb{C}, 0, \Omega_u^{-1}\right) + 2L_{nT}\left(0, 0, \Omega_u^{-1}\right) \\
&= \sum_{i=1}^{n}\left((\Delta \underline{Y}_i + \frac{c_i}{n^\kappa T}\underline{Y}_{-1,i})'\Omega_{u,i}^{-1}(\Delta \underline{Y}_i + \frac{c_i}{n^\kappa T}\underline{Y}_{-1,i}) - (\Delta \underline{Y}_i)'\Omega_{u,i}^{-1}(\Delta \underline{Y}_i)\right) \\
&= \frac{2}{n^\kappa T}\sum_{i=1}^{n} c_i \underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\Delta \underline{Y}_i + \frac{1}{n^{2\kappa}T^2}\sum_{i=1}^{n} c_i^2 \underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\underline{Y}_{-1,i}. \qquad \text{(A.1)}
\end{aligned}
$$

Write

$$
\begin{aligned}
\frac{2}{n^\kappa T}\sum_{i=1}^{n} c_i \underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\Delta \underline{Y}_i &= \frac{1}{n^\kappa}\sum_{i=1}^{n} c_i\left(\frac{2}{T}\underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\Delta \underline{Y}_i\right) \\
&= \frac{1}{n^\kappa}\sum_{i=1}^{n} c_i\left(\frac{2}{T}\frac{\underline{Y}'_{-1,i}\Delta \underline{Y}_i}{\omega_i^2} + \frac{\sigma_i^2}{\omega_i^2} - 1\right) + \frac{1}{n^\kappa}\sum_{i=1}^{n}\eta_{iT} \qquad \text{(A.2)} \\
&= \frac{2}{n^\kappa T}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\underline{Y}'_{-1,i}\Delta \underline{Y}_i - \frac{2}{n^\kappa}\sum_{i=1}^{n} c_i\frac{\lambda_i}{\omega_i^2} + \frac{1}{n^\kappa}\sum_{i=1}^{n}\eta_{1iT},
\end{aligned}
$$

where

$$
\eta_{1iT} = 2c_i\left(\frac{1}{T}\underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\Delta \underline{Y}_i - \frac{1}{T}\frac{\underline{Y}'_{-1,i}\Delta \underline{Y}_i}{\omega_i^2} + \frac{\lambda_i}{\omega_i^2}\right),
$$

and

$$
\frac{1}{n^{2\kappa}T^2}\sum_{i=1}^{n} c_i^2 \underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\underline{Y}_{-1,i} = \frac{1}{n^{2\kappa}T^2}\sum_{i=1}^{n}\frac{c_i^2}{\omega_i^2}\underline{Y}'_{-1,i}\underline{Y}_{-1,i} + \frac{1}{n^{2\kappa}}\sum_{i=1}^{n}\eta_{2iT}, \qquad \text{(A.3)}
$$

where

$$
\eta_{2iT} = c_i^2\left(\frac{1}{T^2}\underline{Y}'_{-1,i}\Omega_{u,i}^{-1}\underline{Y}_{-1,i} - \frac{1}{\omega_i^2 T^2}\underline{Y}'_{-1,i}\underline{Y}_{-1,i}\right).
$$

In the following subsections, we show that under Assumption 2.1, as $(n/T) \to 0$,

$$
\frac{1}{n^{1/4}}\sum_{i=1}^{n}\eta_{1iT} = o_p(1), \qquad \text{(A.4)}
$$

$$
\frac{1}{n^{1/2}}\sum_{i=1}^{n}\eta_{2iT} = o_p(1). \qquad \text{(A.5)}
$$

Then, using (A.1)–(A.5) with $\kappa = 1/2$, we deduce that

$$-2L_{nT}\Big(\mathbb{C}, 0, \Omega_u^{-1}\Big) + 2L_{nT}\Big(0, 0, \Omega_u^{-1}\Big)$$

$$= -\frac{2}{n^{1/2}T}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\underline{Y}'_{-1,i}\Delta\underline{Y}_i + \frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\underline{Y}'_{-1,i}\underline{Y}_{-1,i} - \frac{2}{n^{1/2}}\sum_{i=1}^{n}c_i\frac{\lambda_i}{\omega_i^2} + o_p(1)$$

$$= -2L_{nT}\Big(\mathbb{C}, 0, \Omega\otimes I_T\Big) + 2L_{nT}\Big(0, 0, \Omega\otimes I_T\Big) - \frac{2}{\sqrt{n}}l_n'\mathbb{C}\Omega^{-1}\Lambda l_n + o_p(1),$$

as required. □

Proof of Theorem 2.3: Here, $\kappa = 1/2$. Let Assumption 2.1 hold and $(n/T^{1/2}) \to 0$. By definition, we have

$$-2\Big(\min_{\beta_0} L_{nT}\Big(\mathbb{C}, \beta_0 G_0', \Omega_u^{-1}\Big) - \min_{\beta_0} L_{nT}\Big(0, \beta_0 G_0', \Omega_u^{-1}\Big)\Big)$$

$$= \sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big) - \Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big)$$

$$- \sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G_0\Big)\Big(\Big(\Delta_{c_i}G_0\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G_0\Big)\Big)^{-1}\Big(\Delta_{c_i}G_0\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big)$$

$$-\Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta G_0\Big)\Big(\Big(\Delta G_0\Big)'\Omega_{u,i}^{-1}\Big(\Delta G_0\Big)\Big)^{-1}\Big(\Delta G_0\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big).$$

Using (A.2), (A.3), (A.4) and (A.5), we can approximate the first term as

$$\sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big) - \Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big)$$

$$= \sum_{i=1}^{n}\frac{1}{\omega_i^2}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Big(\Delta_{c_i}\underline{Y}_i\Big) - \Big(\Delta\underline{Y}_i\Big)'\Big(\Delta\underline{Y}_i\Big)\Big) - \frac{2}{n^{1/2}}\sum_{i=1}^{n}c_i\frac{\lambda_i}{\omega_i^2} + o_p(1).$$

Then, the required result for the theorem follows because

$$\sum_{i=1}^{n}\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G_0\Big)\Big(\Big(\Delta_{c_i}G_0\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G_0\Big)\Big)^{-1}\Big(\Delta_{c_i}G_0\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big)$$

$$-\sum_{i=1}^{n}\frac{1}{\omega_i^2}\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Big(\Delta_{c_i}G_0\Big)\Big(\Big(\Delta_{c_i}G_0\Big)'\Big(\Delta_{c_i}G_0\Big)\Big)^{-1}\Big(\Delta_{c_i}G_0\Big)'\Big(\Delta_{c_i}\underline{Y}_i\Big)$$

$$= o_p(1), \tag{A.6}$$

for any c_i such that $\sup_i |c_i| < M$ for some constant M. The proof of (A.6) is available in the following subsection. □

Proof of Theorem 2.5: Here, $\kappa = 1/4$. Let Assumption 2.1 hold and $(n/T^{1/4}) \to 0$. By definition, we have

$$-2\Big(\min_{\beta} L_{nT}\Big(\mathbb{C}, \beta G', \Omega_u^{-1}\Big) - \min_{\beta} L_{nT}\Big(0, \beta G', \Omega_u^{-1}\Big)\Big)$$

$$= \sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big) - \Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big)$$

$$-\sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G\Big)\Big(\Big(\Delta_{c_i}G\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G\Big)\Big)^{-1}\Big(\Delta_{c_i}G\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big)$$
$$-\Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta G\Big)\Big(\Big(\Delta G\Big)'\Omega_{u,i}^{-1}\Big(\Delta G\Big)\Big)^{-1}\Big(\Delta G\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big).$$

Using (A.2), (A.3), (A.4) and (A.5), we can approximate the first term as

$$\sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big)-\Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big)$$
$$=\sum_{i=1}^{n}\frac{1}{\omega_i^2}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Big(\Delta_{c_i}\underline{Y}_i\Big)-\Big(\Delta\underline{Y}_i\Big)'\Big(\Delta\underline{Y}_i\Big)\Big)-\frac{2}{n^{1/4}}\sum_{i=1}^{n}c_i\frac{\lambda_i}{\omega_i^2}+o_p(1).$$

Also, in the following subsection, we show that

$$\sum_{i=1}^{n}\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G\Big)\Big(\Big(\Delta_{c_i}G\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G\Big)\Big)^{-1}\Big(\Delta_{c_i}G\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big)$$
$$-\sum_{i=1}^{n}\frac{1}{\omega_i^2}\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Big(\Delta_{c_i}G\Big)\Big(\Big(\Delta_{c_i}G\Big)'\Big(\Delta_{c_i}G\Big)\Big)^{-1}\Big(\Delta_{c_i}G\Big)'\Big(\Delta_{c_i}\underline{Y}_i\Big)$$
$$=o_p(1), \tag{A.7}$$

for any c_i such that $\sup_i |c_i| < M$ for some constant M. Then, we have

$$\sum_{i=1}^{n}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G\Big)\Big(\Big(\Delta_{c_i}G\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}G\Big)\Big)^{-1}\Big(\Delta_{c_i}G\Big)'\Omega_{u,i}^{-1}\Big(\Delta_{c_i}\underline{Y}_i\Big)$$
$$-\Big(\Delta\underline{Y}_i\Big)'\Omega_{u,i}^{-1}\Big(\Delta G\Big)\Big(\Big(\Delta G\Big)'\Omega_{u,i}^{-1}\Big(\Delta G\Big)\Big)^{-1}\Big(\Delta G\Big)'\Omega_{u,i}^{-1}\Big(\Delta\underline{Y}_i\Big)\Big)$$
$$=\sum_{i=1}^{n}\frac{1}{\omega_i^2}\Big(\Big(\Delta_{c_i}\underline{Y}_i\Big)'\Big(\Delta_{c_i}G\Big)\Big(\Big(\Delta_{c_i}G\Big)'\Big(\Delta_{c_i}G\Big)\Big)^{-1}\Big(\Delta_{c_i}G\Big)'\Big(\Delta_{c_i}\underline{Y}_i\Big)$$
$$-\Big(\Delta\underline{Y}_i\Big)'\Big(\Delta G\Big)\Big(\Big(\Delta G\Big)'\Big(\Delta G\Big)\Big)^{-1}\Big(\Delta G\Big)'\Big(\Delta\underline{Y}_i\Big)\Big)+o_p(1).$$

Combining these expressions gives the required result:

$$-2\Big(\min_{\beta}L_{nT}\Big(\mathbb{C},\beta G',\Omega_u^{-1}\Big)-\min_{\beta}L_{nT}\Big(0,\beta G',\Omega_u^{-1}\Big)\Big)$$
$$=-2\Big(\min_{\beta}L_{nT}\Big(\mathbb{C},\beta G',\Omega^{-1}\otimes I_T\Big)-\min_{\beta}L_{nT}\Big(0,\beta G',\Omega^{-1}\otimes I_T\Big)\Big)$$
$$-\frac{2}{n^{1/4}}l_n'\mathbb{C}\Omega^{-1}\Lambda l_n+o_p(1). \qquad \square$$

A.1 Supplementary results

A.1.1. A useful lemma. Before we start the proof of (A.4) and (A.5), we introduce a useful technical result. When A is a matrix, we use three different norms, $\|A\|_o = \lambda_{\max}(A'A)^{1/2}$ (where $\lambda_{\max}(\cdot)$ denotes the maximum eigenvalue), $\|A\| = \mathrm{tr}(A'A)^{1/2}$ and $|A| = \sum_{i,j} |a_{ij}|$ (where a_{ij} is the (i, j)th element of A). It is well known that

$$\|A\|_o \leq \|A\| \leq |A|.$$

By definition, the covariance matrix of $\underline{U}_i$ is $\Omega_{u,i} = [\gamma_i(t-s)]_{t,s}$. Let A_i be the $(T \times T)$ matrix whose (t, s) element $a_{i,t,s}$ is ρ_i^{t-s-1}, if $t > s$, and zero, if $t \leq s$. Let

$$R_i = \omega_i \Omega_{u,i}^{-1/2} - \omega_i^{-1} \Omega_{u,i}^{1/2}. \tag{A.8}$$

LEMMA A.1. *Let Assumption 2.1 hold. Then,* $\sup_i (1/T^{1/2})\|R_i A_i\| < M$ *for some constant* M.

Proof: For the desired result, we show

$$\sup_i \frac{1}{T}\|R_i A_i\|^2 \leq M.$$

By definition,

$$\begin{aligned}
\sup_i \frac{1}{T}\|R_i A_i\|^2 &= \sup_i \frac{1}{T}\mathrm{tr}\Big(A_i' R_i' R_i A_i\Big) = \sup_i \frac{1}{T}\mathrm{tr}\Big(\omega_i^2 A_i' \Omega_{u,i}^{-1} A_i + \omega_i^{-2} A_i' \Omega_{u,i} A_i - 2A_i' A_i\Big) \\
&\leq M \sup_i \Big|\frac{1}{T}\mathrm{tr}\Big(A_i'\Big(\Omega_{u,i} - \omega_i^2\Big)A_i\Big)\Big| + M \sup_i \Big|\frac{1}{T}\mathrm{tr}\Big(A_i'\Big(\Omega_{u,i}^{-1} - \omega_i^{-2}\Big)A_i\Big)\Big| \\
&= M(I + II), \text{ say,}
\end{aligned}$$

where the inequality holds because $0 < \inf_i \omega_i^2 \leq \sup_i \omega_i^2 < \infty$ under Assumption 2.1 and by the triangle inequality.

First, we show that $I = O(1)$. Define

$$\begin{aligned}
a_{n,T,i}(k) &= \frac{1}{T}\sum_{t=1}^{T-k}\sum_{s=1}^{T} a_{i,t,s}a_{i,t+k,s} = \Big(1 - \frac{\theta_i}{n^\kappa T}\Big)^k \sum_{s=1}^{T-k}\Big(1 - \frac{\theta_i}{n^\kappa T}\Big)^{2(s-1)}\Big(\frac{t-k-s}{T}\Big) \\
a'_{n,T,i}(k) &= \sum_{s=1}^{T-k}\Big(1 - \frac{\theta_i}{n^\kappa T}\Big)^{2(s-1)}\Big(\frac{t-k-s}{T}\Big) \\
a''_{n,T,i}(k) &= \sum_{s=1}^{T}\Big(1 - \frac{\theta_i}{n^\kappa T}\Big)^{2(s-1)}\Big(\frac{t-k-s}{T}\Big) \\
a_{n,T,i}(0) &= \frac{1}{T}\Big(\sum_{t=1}^{T} a_{i,t,s}\Big)^2 = \sum_{s=1}^{T}\Big(1 - \frac{\theta_i}{n^\kappa T}\Big)^{2(s-1)}\Big(\frac{t-s}{T}\Big).
\end{aligned}$$

By adding and subtracting the terms and the triangle inequality, we can bound

$$\begin{aligned}
I &= \sup_i \Big|\frac{1}{T}\mathrm{tr}\Big(A_i'\Big(\Omega_{u,i} - \omega_i^2\Big)A_i\Big)\Big| \\
&\leq I_1 + I_2 + I_3 + I_4,
\end{aligned}$$

where

$$I_1 = \sup_i 2\Big|\sum_{k=1}^{T-1}\gamma_i(k)\Big(a_{n,T,i}(k) - a'_{n,T,i}(k)\Big)\Big|$$

$$I_2 = \sup_i 2\Big|\sum_{k=1}^{T-1}\gamma_i(k)\Big(a'_{n,T,i}(k) - a''_{n,T,i}(k)\Big)\Big|$$

$$I_3 = \sup_i 2\Big|\sum_{k=1}^{T-1}\gamma_i(k)\Big(a''_{n,T,i}(k) - a_{n,T,i}(0)\Big)\Big|$$

$$I_4 = \sup_i 2\Big|a_{n,T,i}(0)\sum_{k=T}^{\infty}\gamma_i(k)\Big|.$$

For term I_1, note that because

$$\sup_i\Big|\frac{1}{T}\sum_{s=1}^{T-k}\Big(1-\frac{\theta_i}{n^\kappa T}\Big)^{2(s-1)}\Big(\frac{t-k-s}{T}\Big)\Big| < M$$

and $\sup_i |\theta_i| < M$, we have

$$\begin{aligned}\Big|(a_{n,T,i}(k) - a'_{n,T,i}(k))\Big| &= T\Big|\Big(1-\frac{\theta_i}{n^\kappa T}\Big)^k - 1\Big|\Big|\frac{1}{T}\sum_{s=1}^{T-k}\Big(1-\frac{\theta_i}{n^\kappa T}\Big)^{2(s-1)}\Big(\frac{t-k-s}{T}\Big)\Big|\\
&\le MT\Big|\Big(1-\frac{\theta_i}{n^\kappa T}\Big)^k - 1\Big| = MT\Big|\sum_{j=1}^{k}\binom{k}{j}\Big(\frac{-\theta_i}{n^\kappa T}\Big)^j\Big|\\
&\le MT\sum_{j=1}^{k}\frac{1}{j!}\Big(\frac{|\theta_i|k}{n^\kappa T}\Big)^j = MT\frac{|\theta_i|k}{n^\kappa T}\sum_{j=1}^{k}\frac{1}{j!}\Big(\frac{|\theta_i|k}{n^\kappa T}\Big)^{j-1}\\
&\le M\frac{k}{n^\kappa}\Big(1+\sum_{j=1}^{\infty}\frac{1}{j!}\Big(\frac{M}{n^\kappa}\Big)^j\Big) = M\frac{k}{n^\kappa}\exp\Big(\frac{M}{n^\kappa}\Big),\end{aligned}$$

where the second inequality uses $\binom{k}{j} \le (k^j/j!)$, the last inequality uses $\sup_i |\theta_i| < M$ and $(k/T) \le 1$ and the equality uses the Taylor representation of the exponential function. Then,

$$I_1 \le 2\sum_{k=1}^{T-1}\gamma(k)\sup_i\Big|\Big(a_{n,T,i}(k) - a'_{n,T,i}(k)\Big)\Big| \le \frac{M}{n^\kappa}\Big(\exp\Big(\frac{M}{n^\kappa}\Big) - 1\Big)\sum_{k=1}^{T-1}\gamma(k)k = o(1).$$

For term I_2, note that

$$\begin{aligned}\Big|\Big(a'_{n,T,i}(k) - a''_{n,T,i}(k)\Big)\Big| &\le 2T\frac{1}{T}\sum_{s=T-k+1}^{T}\Big(1+\frac{|\theta_i|}{n^\kappa T}\Big)^{2(s-1)}\\
&\le 2T\int_{1-(k/T)}^{1}\exp\Big(\frac{2|\theta_i|r}{n^\kappa}\Big)\,\mathrm{d}r\\
&= \begin{cases} T\dfrac{n^\kappa}{|\theta_i|}\exp\Big(\dfrac{2|\theta_i|}{n^\kappa}\Big)\Big(1-\exp\Big(-\dfrac{2|\theta_i|}{n^\kappa}\dfrac{k}{T}\Big)\Big) & \text{for } \theta_i \ne 0\\ 2k & \text{for } \theta_i = 0\end{cases}.\\
&\le Mk,\end{aligned}$$

where the first inequality holds because $\theta_i \geq 0$ and $|((t-k-s)/T)| \leq 2$ and the last inequality holds by the mean-value theorem and $\sup_i |\theta_i| < M$. Then,

$$I_2 = \sup_i 2\Big|\sum_{k=1}^{T-1}\gamma_i(k)\left(a'_{n,T,i}(k) - a''_{n,T,i}(k)\right)\Big| \leq M\sum_{k=1}^{T-1}\gamma(k)k = O(1).$$

For term I_3, we have

$$\begin{aligned} I_3 &= \sup_i 2\Big|\sum_{k=1}^{T-1}\gamma_i(k)\left(a''_{n,T,i}(k) - a_{n,T,i}(0)\right)\Big| \\ &\leq 2\sum_{k=1}^{T-1}\gamma(k)k\sup_i\Big|\frac{1}{T}\sum_{s=1}^{T}\left(1-\frac{\theta_i}{n^\kappa T}\right)^{2(s-1)}\Big| = O(1), \end{aligned}$$

where the last line holds because $\sup_i |(1/T)\sum_{s=1}^{T}(1-(\theta_i/(n^\kappa T)))^{2(s-1)}| < M$. Finally, we have

$$\begin{aligned} I_4 &= \sup_i 2\Big|a_{n,T,i}(0)\sum_{k=T}^{\infty}\gamma_i(k)\Big| \\ &\leq \sup_i\Big|\frac{1}{T}\sum_{s=1}^{T}\left(1-\frac{\theta_i}{n^\kappa T}\right)^{2(s-1)}\left(\frac{t-s}{T}\right)\Big| T\sum_{k=T}^{\infty}\gamma(k) \\ &\leq MT\sum_{k=T}^{\infty}k^{-s} \leq M\sum_{k=T}^{\infty}k^{-s+1} = o(1), \end{aligned}$$

where the second inequality holds because

$$\sup_i\Big|\frac{1}{T}\sum_{s=1}^{T}\left(1-\frac{\theta_i}{n^\kappa T}\right)^{2(s-1)}\left(\frac{t-s}{T}\right)\Big| < M.$$

By combining terms $I_1 - I_4$, we have the required result

$$I = O(1).$$

The proof of $II = O(1)$ follows in a similar fashion and is omitted. □

Proof of (A.5): We prove the required result when $(n/T) \to 0$ and $\kappa = 1/4$. Because

$$\begin{aligned} E\Big[\frac{1}{n^{1/2}}\sum_{i=1}^{n}\eta_{2iT}\Big]^2 &= \Big(\frac{1}{n^{1/2}}\sum_{i=1}^{n}E[\eta_{2iT}]\Big)^2 + \mathrm{Var}\Big(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\eta_{2iT}\Big) \\ &\leq (n^{1/2}\sup_i|E[\eta_{2iT}]|)^2 + \sup_i \mathrm{Var}(\eta_{2iT}), \end{aligned}$$

the required result follows if we show

$$n^{1/2}\sup_i|E[\eta_{2iT}]| = o(1) \tag{A.9}$$

and

$$\sup_i \mathrm{Var}(\eta_{2iT}) = o(1). \tag{A.10}$$

For (A.9), we follow similar arguments used in proving $|E[S_1]| \to 0$ on p. 831 (in the proof of Lemma A2) of Elliott et al. (1996), and have for some constant M

$$n^{1/2} \sup_i |E[\eta_{2iT}]| \le \Big(\frac{n}{T}\Big)^{1/2} M \sup_i \Big(\frac{1}{\omega_i} \|\Omega_{u,i}\|_o \|\Omega_{u,i}^{-1}\|_o^{1/2} \frac{\|R_i A_i\|}{\sqrt{T}}\Big) \le M\Big(\frac{n}{T}\Big)^{1/2} = o(1),$$

where the second inequality holds because $0 < M_l \le \inf_i f_i(\lambda) \le \sup_i f_i(\lambda) \le M_u < \infty$ and by Lemma A.1, and the last inequality holds because $(n/T) \to 0$.

For (A.10), we also follow similar arguments to those used in proving $|\mathrm{Var}(S_1)| \to 0$ on p. 831 (in the proof of Lemma A2) of Elliott et al. (1996), and have for some constant M

$$\sup_i \mathrm{Var}(\eta_{2iT}) \le \frac{1}{T} M \sup_i \Big(\frac{1}{\omega_i^2} \|\Omega_{u,i}\|_o^2 \|\Omega_{u,i}^{-1}\|_o \frac{\|R_i A_i\|^2}{T}\Big) \le \frac{M}{T} = o(1). \qquad \square$$

Proof of (A.4): We prove the required result when $(n/T) \to 0$ and $\kappa = 1/4$. By replacing $\Delta \underline{Y}_i$ in η_{1iT} with $-(\theta_i/(n^{1/4}T))\underline{Y}_{-1,i} + \underline{U}_i$, we can decompose η_{1iT} as

$$\eta_{1iT} = \eta_{3iT} - \frac{1}{n^{1/4}} \eta_{4iT},$$

where

$$\eta_{3iT} = 2c_i \Big(\frac{1}{T} \underline{Y}_{-1,i}' \Omega_{u,i}^{-1} \underline{U}_i - \frac{1}{T} \frac{\underline{Y}_{-1,i}' \underline{U}_i}{\omega_i^2} + \frac{\lambda_i}{\omega_i^2}\Big)$$

$$\eta_{4iT} = c_i \theta_i \Big(\frac{1}{T^2} \underline{Y}_{-1,i}' \Omega_{u,i}^{-1} \underline{Y}_{-1,i} - \frac{1}{T} \frac{\underline{Y}_{-1,i}' \underline{Y}_{-1,i}}{\omega_i^2}\Big),$$

and

$$\frac{1}{n^{1/4}} \sum_{i=1}^n \eta_{1iT} = \frac{1}{n^{1/4}} \sum_{i=1}^n \eta_{3iT} - \frac{1}{n^{1/2}} \sum_{i=1}^n \eta_{4iT}.$$

First, similar arguments to those in the proof of (A.5) lead to

$$\frac{1}{n^{1/2}} \sum_{i=1}^n \eta_{4iT} = o_p(1).$$

Then, the required result follows if

$$\frac{1}{n^{1/4}} \sum_{i=1}^n \eta_{3iT} = o_p(1),$$

which follows if

$$E[\frac{1}{n^{1/4}} \sum_{i=1}^n \eta_{3iT}]^2 = o(1).$$

Note that

$$E\Big[\frac{1}{n^{1/4}} \sum_{i=1}^n \eta_{3iT}\Big]^2 = \Big(\frac{1}{n^{1/4}} \sum_{i=1}^n E[\eta_{3iT}]\Big)^2 + \mathrm{Var}\Big(\frac{1}{n^{1/4}} \sum_{i=1}^n \eta_{3iT}\Big)$$
$$\le (n^{3/4} \sup_i |E[\eta_{3iT}]|)^2 + n^{1/2} \sup_i \mathrm{Var}(\eta_{3iT}),$$

Using similar arguments to those used for the proof of $\sup_i \mathrm{Var}(\eta_{2iT}) = O(1/T)$, we can show that

$$n^{1/2} \sup_i \mathrm{Var}(\eta_{3iT}) = n^{1/2} O\Big(\frac{1}{T}\Big) = o(1).$$

For $n^{3/4} \sup_i |E[\eta_{3iT}]| = o(1)$, we show

$$\sup_i |E[\eta_{3iT}]| = O\Big(\frac{1}{T}\Big). \tag{A.11}$$

Because $(n/T) \to 0$, the desired result follows. Because $\underline{Y}_{-1,i} = A_i \underline{U}_i$, we have

$$\eta_{3iT} = 2c_i \Big(\frac{1}{T} A_i' \underline{U}_i' \Omega_{u,i}^{-1} \underline{U}_i - \frac{1}{T} \frac{A_i' \underline{U}_i' \underline{U}_i}{\omega_i^2} + \frac{\lambda_i}{\omega_i^2}\Big).$$

Because $\mathrm{tr}(A_i) = 0$, we have

$$\begin{aligned}
E[\eta_{3iT}] &= 2c_i \frac{1}{T}\Big(\mathrm{tr}\Big(A_i\Big) - \frac{1}{\omega_i^2}\mathrm{tr}\Big(\Omega_{u,i} A_i\Big)\Big) + 2c_i \frac{\lambda_i}{\omega_i^2} \\
&= \frac{-2c_i}{\omega_i^2}\Big(\sum_{k=1}^{T} \gamma_i(k)\Big(1 - \frac{k}{T}\Big)\rho_i^{k-1} - \sum_{k=1}^{\infty} \gamma_i(k)\Big) \\
&= \frac{-2c_i}{\omega_i^2}\Big(\sum_{k=1}^{T} \gamma_i(k)\Big(\rho_i^{k-1} - 1\Big) - \frac{1}{T}\sum_{k=1}^{T} k\gamma_i(k)\rho_i^{k-1} - \sum_{k=T+1}^{\infty} \gamma_i(k)\Big) \\
&= I + II + III, \text{ say.}
\end{aligned}$$

For term I, we can bound

$$\begin{aligned}
0 \le \Big|1 - \rho_i^{k-1}\Big| &= \Big|1 - \Big(1 - \frac{\theta_i}{n^\kappa T}\Big)^{k-1}\Big| \le \sum_{j=1}^{k-1} \binom{k-1}{j} \Big(\frac{|\theta_i|}{n^\kappa T}\Big)^j \\
&\le \sum_{j=1}^{k-1} \frac{1}{j!}\Big(\frac{|\theta_i|(k-1)}{n^\kappa T}\Big)^j = \frac{|\theta_i|(k-1)}{n^\kappa T} \sum_{j=1}^{k-1} \frac{1}{j!}\Big(\frac{|\theta_i|(k-1)}{n^\kappa T}\Big)^{j-1} \\
&\le \frac{M(k-1)}{n^\kappa T} \sum_{j=1}^{k-1} \frac{1}{j!}\Big(\frac{M}{n^\kappa}\Big)^{j-1} \le \frac{M(k-1)}{n^\kappa T}\Big(1 + \sum_{j=1}^{\infty} \frac{1}{j!}\Big(\frac{M}{n^\kappa}\Big)^j\Big) \\
&\le \frac{M(k-1)}{n^\kappa T}\Big(\exp\Big(\frac{M}{n^\kappa}\Big)\Big).
\end{aligned}$$

Then, for some constant $M > 0$, we can bound

$$|I| = \sup_i \frac{2c_i}{\omega_i^2} \sum_{k=1}^{T} |\gamma_i(k)|\Big|\Big(1 - \rho_i^{k-1}\Big)\Big| \le \frac{M}{n^\kappa T}\Big(\exp\Big(\frac{M}{n^\kappa}\Big)\Big) \sum_{k=1}^{T} \gamma(k)k = o\Big(\frac{1}{T}\Big),$$

as required. Next,

$$|II| \le \frac{M}{T} \sum_{k=1}^{T} k\gamma(k) = O\Big(\frac{1}{T}\Big),$$

and, by Assumption 2.1(c),

$$|III| \leq M \sum_{k=T+1}^{\infty} \gamma(k) \leq M \sum_{k=T+1}^{\infty} k^{-s}$$

$$\leq M(T+1)^{-s+1} = o\Big(\frac{1}{T}\Big),$$

because $s > 2$, as required. □

A.1.2. More preliminary results. In this section, $\kappa = 1/4$ and let Assumption 2.1 hold. Define Φ_i to be the $(T \times T)$ matrix whose $(r, s)^{\text{th}}$ element is $\phi_i(r-s)$, where $\phi_i(k)$ is defined in Assumption 2.1.

Define $\tilde{G} = [\tilde{G}_0, \tilde{G}_1] = [G_0, G_1](\text{diag}(\sqrt{T}, 1))$. Direct calculations show that

$$\Delta_{c_i}\tilde{G}_0 = \Big(T^{1/2}, \frac{c_i}{n^{1/4}T^{1/2}}, \ldots, \frac{c_i}{n^{1/4}T^{1/2}}\Big)',$$

$$\Delta_{c_i}\tilde{G}_1 = \Big(0, 1 + \frac{c_i}{n^{1/4}}\frac{1}{T}, \ldots, 1 + \frac{c_i}{n^{1/4}}\frac{t-1}{T}, \ldots, 1 + \frac{c_i}{n^{1/4}}\frac{T-1}{T}\Big)',$$

$$\frac{1}{T}\Big(\Delta_{c_i}\tilde{G}\Big)'\Big(\Delta_{c_i}\tilde{G}\Big)$$

$$= \begin{pmatrix} 1 + \frac{c_i^2}{n^{1/2}T}\frac{T-1}{T} & \frac{1}{T^{1/2}}\Big(\frac{c_i}{n^{1/4}} + \frac{c_i^2}{n^{1/2}}\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}\Big) \\ \frac{1}{T^{1/2}}\Big(\frac{c_i}{n^{1/4}} + \frac{c_i^2}{n^{1/2}}\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}\Big) & \frac{1}{T}\sum_{t=2}^{T}\Big(1 + \frac{c_i}{n^{1/4}}\frac{t-1}{T}\Big)^2 \end{pmatrix},$$

and

$$\frac{1}{T^{1/2}}(\Delta_{c_i}\tilde{G})'(\Delta_{c_i}\underline{Y}_i) = \Big(y_{i1} + \frac{c_i}{n^{1/4}T^{1/2}}\frac{1}{T^{1/2}}(y_{iT} - y_{i1}) + \frac{c_i^2}{n^{1/2}T^{1/2}}\frac{1}{T^{3/2}}\sum_{t=2}^{T} y_{it-1}$$

$$\times \frac{1}{T^{1/2}}(y_{iT} - y_{i1}) + \frac{c_i}{n^{1/4}T^{1/2}}\Big(y_{iT} - \frac{1}{T}(y_{iT} + y_{i0})\Big)$$

$$+ \frac{c_i^2}{n^{1/2}}\frac{1}{T^{3/2}}\sum_{t=2}^{T}\frac{t-1}{T}y_{it-1}\Big). \qquad \text{(A.12)}$$

Define

$$b_{n,T,i}^{jl}(k) = \frac{1}{T}\sum_{t=2}^{T-k}\Big[(\Delta_{c_i}\tilde{G}_j)_t(\Delta_{c_i}\tilde{G}_l)_{t+k} + (\Delta_{c_i}\tilde{G}_l)_t(\Delta_{c_i}\tilde{G}_j)_{t+k}\Big],$$

where $(x)_t$ is the tth element of the vector x and $j, l = 0, 1$.

LEMMA A.2. *(a)* $\sup_i |b_{n,T,i}^{01}(k)| \leq (M/T^{1/2})$ *for all* k*; (b)* $\sup_i |b_{n,T,i}^{11}(k) - b_{n,T,i}^{11}(0)| \leq M(k/T)$ *for some finite constant* M.

Proof: Part (a). By definition, for $k = 0$,

$$\sup_i |b_{n,T,i}^{01}(0)| = 2\sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_0)'(\Delta_{c_i}\tilde{G}_1)|$$

$$= \frac{2}{T^{1/2}}\sup_i |\frac{c_i}{n^{1/4}} + \frac{c_i^2}{n^{1/2}}\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}| \leq \frac{M}{n^{1/4}T^{1/2}}.$$

For $k \geq 1$, we have

$$\sup_i |b^{01}_{n,T,i}(k)| = \sup_i |\frac{1}{T}\sum_{t=2}^{T-k}\Big[\left(\Delta_{c_i}\tilde{G}_0\right)_t\left(\Delta_{c_i}\tilde{G}_1\right)_{t+k} + \left(\Delta_{c_i}\tilde{G}_1\right)_t\left(\Delta_{c_i}\tilde{G}_0\right)_{t+k}\Big]| \leq \frac{M}{T^{1/2}},$$

as required.

Part (b). By definition,

$$\begin{aligned}
&\sup_i |b^{11}_{n,T,i}(k) - b^{11}_{n,T,i}(0)| \\
&\leq 2\frac{1}{T}\sum_{t=2}^{T-k}\left|\left(1+\frac{c_i}{n^{1/4}T}\frac{t-1}{T}\right)\left(1+\frac{c_i}{n^{1/4}T}\frac{t+k-1}{T}\right) - \left(1+\frac{c_i}{n^{1/4}T}\frac{t-1}{T}\right)^2\right| \\
&\quad + 2\frac{1}{T}\sum_{t=T-k+1}^{T}\left(1+\frac{c_i}{n^{1/4}T}\frac{t-1}{T}\right)^2 \\
&\leq \frac{M}{n^{1/4}}\frac{k}{T} + M\frac{k}{T},
\end{aligned}$$

as required. □

LEMMA A.3. *Suppose that x_i and z_i are $T-$ vectors such that $\sup_{i,t}|z_{it}|$ is bounded, where z_{it} is the $(t)^{th}$ element of z_i. Then, (a) $\sup_i |(1/T)x_i'(\Omega^{-1}_{u,i} - \Phi_i)z_i| = O((\sup_i \|x_i\|)/T)$; (b) $\sup_i (1/T)\|R_i(\Delta_{c_i}\tilde{G}_1)\|^2 = O(1/T^{1/2})$, where R_i is defined in (A.8).*

Proof: Part (a). The proof is similar to that of Lemma A1 of Elliott et al. (1996) and is omitted.

Part (b). We replace A_i in the proof of Lemma A.1 with $(\Delta_{c_i}\tilde{G}_1)$. Then, the required result follows if we show

$$\text{(b1)} \qquad \sup_i |\frac{1}{T}\left(\Delta_{c_i}\tilde{G}_1\right)'\left(\Omega_{u,i} - \omega_i^2\right)\left(\Delta_{c_i}\tilde{G}_1\right)| = O\Big(\frac{1}{T}\Big)$$

$$\text{(b2)} \qquad \sup_i |\frac{1}{T}\left(\Delta_{c_i}\tilde{G}_1\right)'\left(\Omega^{-1}_{u,i} - \omega_i^{-2}\right)\left(\Delta_{c_i}\tilde{G}_1\right)| = O\Big(\frac{1}{T^{1/2}}\Big).$$

For Part (b1), by definition, we have

$$\begin{aligned}
&\sup_i\left|\frac{1}{T}\Big(\Delta_{c_i}\tilde{G}_1\Big)'\Big(\Omega_{u,i} - \omega_i^2\Big)\Big(\Delta_{c_i}\tilde{G}_1\Big)'\right| \\
&= \sup_i\left|\sum_{k=1}^{T-1}\gamma_i(k)\Big(b^{11}_{n,T,i}(k) - b^{11}_{n,T,i}(0)\Big) - \Big(\sum_{k=T}^{\infty}\gamma_i(k)\Big)b^{11}_{n,T,i}(0)\right| \\
&\leq \sup_i\left|\sum_{k=1}^{T-1}\gamma_i(k)\Big(b^{11}_{n,T,i}(k) - b^{11}_{n,T,i}(0)\Big)\right| + \sup_i\left|\Big(\sum_{k=T}^{\infty}\gamma_i(k)\Big)b^{11}_{n,T,i}(0)\right|.
\end{aligned}$$

By Lemma A.2(b), the first term is bounded by

$$\sum_{k=1}^{T-1}\gamma(k)\sup_i|b^{11}_{n,T,i}(k) - b^{11}_{n,T,i}(0)| \leq M\frac{1}{T}\sum_{k=1}^{T-1}\gamma(k)k = O\Big(\frac{1}{T}\Big),$$

as required. Under Assumption 2.1(c), the second term is bounded by

$$\Big(\sum_{k=T}^{\infty} k^{-s}\Big) \sup_i |b_{n,T,i}^{11}(0)| \leq o\Big(\frac{1}{T}\Big),$$

as required.

For Part (b2), we have

$$\sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_1)'(\Omega_{u,i}^{-1} - \omega_i^{-2})(\Delta_{c_i}\tilde{G}_1)|$$

$$\leq \sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_1)'(\Omega_{u,i}^{-1} - \Phi_i)(\Delta_{c_i}\tilde{G}_1)| + \sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_1)'(\Phi_i - \omega_i^{-2})(\Delta_{c_i}\tilde{G}_1)|.$$

By Part (a), we have

$$\sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_1)'(\Omega_{u,i}^{-1} - \Phi_i)(\Delta_{c_i}\tilde{G}_1)| \leq O\Bigg(\sup_i \frac{\|\Delta_{c_i}\tilde{G}_1\|}{T}\Bigg) = O\Bigg(\frac{1}{T^{1/2}}\Bigg).$$

Using a similar argument to that used in the proof of Part (b1), we can bound the second term by

$$\sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_1)'(\Phi_i - \omega_i^{-2})(\Delta_{c_i}\tilde{G}_1)| \leq O(\frac{1}{T}).$$

Combining these, we have the required result for Part (b2). □

For $\mathbb{C} = \mathrm{diag}(c_1, \ldots, c_n)$, we define

$$A_{iT}(\mathbb{C}) = \frac{1}{T^{1/2}}(\Delta_{c_i}\tilde{G})'\Omega_{u,i}^{-1}(\Delta_{c_i}\underline{Y}_i),\ \ A_{iT}^*(\mathbb{C}) = \frac{1}{\omega_i^2}\frac{1}{T^{1/2}}(\Delta_{c_i}\tilde{G})'(\Delta_{c_i}\underline{Y}_i),$$

$$B_{iT}(\mathbb{C}) = \frac{1}{T}(\Delta_{c_i}\tilde{G})'\Omega_{u,i}^{-1}(\Delta_{c_i}\tilde{G}), \quad B_{iT}^*(\mathbb{C}) = \frac{1}{\omega_i^2}\frac{1}{T}(\Delta_{c_i}G)'(\Delta_{c_i}G),$$

$$\tilde{B}_{iT}(\mathbb{C}) = \mathrm{diag}(B_{11,iT}(\mathbb{C}), B_{22,iT}(\mathbb{C})), \quad \tilde{B}_{iT}^*(\mathbb{C}) = \mathrm{diag}(B_{11,iT}^*(\mathbb{C}), B_{22,iT}^*(\mathbb{C})).$$

We define $B_{kl,iT}(\mathbb{C})$ to be the (k, l)th element of $B_{iT}(\mathbb{C})$ and $A_{k,iT}(\mathbb{C})$ to be the kth element of $A_{k,iT}(\mathbb{C})$, where $k, l = 1, 2$. Similarly, we define $A_{k,iT}^*(\mathbb{C})$ and $B_{kl,iT}^*(\mathbb{C})$.

LEMMA A.4. *Under Assumptions 2.1, the following hold: (a)* $\sup_i |B_{12,iT}(\mathbb{C})|, \sup_i |B_{12,iT}^*(\mathbb{C})| = O(1/T^{1/2})$*; (b)* $\sup_i \|B_{iT}(\mathbb{C}) - \tilde{B}_{iT}(\mathbb{C})\|,\ \sup_i \|B_{iT}^*(\mathbb{C}) - \tilde{B}_{iT}^*(\mathbb{C})\| = O(1/T^{1/2})$*; (c)* $\sup_i \|B_{iT}(\mathbb{C})^{-1}\|$, $\sup_i \|\tilde{B}_{iT}(\mathbb{C})^{-1}\| \leq M$*; (d)* $\sup_i E\|A_{iT}(\mathbb{C})\|^2,\ \sup_i E\|A_{iT}^*(\mathbb{C})\|^2 \leq M$*; (e)* $\sup_i |B_{11,iT}(\mathbb{C}) - B_{11,iT}(0)| = O(1/(n^{1/4}T^{1/2}))$*; (f)* $\sup_i |B_{22,iT}(\mathbb{C}) - B_{22,iT}^*(\mathbb{C})| = O(1/T^{1/2})$*; (g)* $\sup_i E|A_{2,iT}(\mathbb{C}) - A_{2,iT}^*(\mathbb{C})|^2 = O(1/T^{1/2})$.

Proof: Part (a). A direct calculation shows that $\sup_i |B_{12,iT}^*(\mathbb{C})| = O(1/T^{1/2})$. We bound $\sup_i |B_{12,iT}(\mathbb{C})|$ by

$$\sup_i |B_{12,iT}(\mathbb{C})| = \sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_0)'\Omega_{u,i}^{-1}(\Delta_{c_i}\tilde{G}_1)|$$

$$\leq \sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_0)'(\Omega_{u,i}^{-1} - \Phi_i)(\Delta_{c_i}\tilde{G}_1)|$$

$$+ \sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_0)'\Big(\Phi_i - \frac{1}{\omega_i^2}\Big)(\Delta_{c_i}\tilde{G}_1)|$$

$$+ \sup_i |\frac{1}{T}\frac{1}{\omega_i^2}(\Delta_{c_i}\tilde{G}_0)'(\Delta_{c_i}\tilde{G}_1)|.$$

By Lemma A.3(a), we have

$$\sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_0)'(\Omega_{u,i}^{-1} - \Phi_i)(\Delta_{c_i}\tilde{G}_1)| = O\left(\sup_i \frac{\|\Delta_{c_i}\tilde{G}_0\|}{T}\right) = O\left(\frac{1}{T^{1/2}}\right).$$

By Lemma A.2 and Assumption 2.1, we have

$$\sup_i |\frac{1}{T}(\Delta_{c_i}\tilde{G}_0)'\left(\Phi_i - \frac{1}{\omega_i^2}\right)(\Delta_{c_i}\tilde{G}_1)| = O\left(\frac{1}{T^{1/2}}\right).$$

Finally, the last term is

$$\sup_i |\frac{1}{T}\frac{1}{\omega_i^2}(\Delta_{c_i}\tilde{G}_0)'(\Delta_{c_i}\tilde{G}_1)| = \sup_i |B_{12,iT}^*(\mathbb{C})| = O\left(\frac{1}{T^{1/2}}\right),$$

as required.

Part (b) is an immediate corollary of Part (a).

Part (c). First, note that under Assumption 2.1 we have

$$\begin{aligned}0 < M_l \le \inf_i B_{kk,iT}(\mathbb{C}) &= \|\frac{1}{T}(\Delta_{c_i}\tilde{G}_{k-1})\|^2 \frac{1}{\sup_i \lambda_{\max}(\Omega_{u,i})} \\ \le B_{kk,iT}(\mathbb{C}) \le \sup_i B_{kk,iT}(\mathbb{C}) &= \|\frac{1}{T}(\Delta_{c_i}\tilde{G}_{k-1})\|^2 \frac{1}{\inf_i \lambda_{\min}(\Omega_{u,i})} \le M_u < \infty,\end{aligned}$$

where $k = 1, 2$. It follows immediately that

$$\sup_i \|\tilde{B}_{iT}(\mathbb{C})^{-1}\| \le \frac{1}{\inf_i B_{11,iT}(\mathbb{C})} + \frac{1}{\inf_i B_{22,iT}(\mathbb{C})} \le M,$$

as required. Also, the desired result follows because

$$\begin{aligned}\sup_i \|B_{iT}(\mathbb{C})^{-1}\| &= \sup_i \|\frac{1}{\det(B_{iT}(\mathbb{C}))}\begin{pmatrix} B_{22,iT}(\mathbb{C}) & -B_{12,iT}(\mathbb{C}) \\ -B_{12,iT}(\mathbb{C}) & B_{11,iT}(\mathbb{C})\end{pmatrix}\| \\ &\le \frac{\sup_i \|B_{iT}(\mathbb{C})\|}{\inf_i B_{11,iT}(\mathbb{C}) \inf_i B_{22,iT}(\mathbb{C}) - \sup_i B_{12,iT}(\mathbb{C})^2} \\ &= \frac{\sup_i \|\tilde{B}_{iT}(\mathbb{C})\| + o(1)}{\inf_i B_{11,iT}(\mathbb{C}) \inf_i B_{22,iT}(\mathbb{C}) + o(1)} \le M,\end{aligned}$$

where the second equality holds by Part (a).

Part (d). The desired result $\sup_i E\|A_{iT}^*(\mathbb{C})\|^2 \le M$ follows from (A.12) and by direct calculation. For the second desired result, note that $E\|A_{iT}(\mathbb{C})\|^2 = E\|A_{1,iT}(\mathbb{C})\|^2 + E\|A_{2,iT}(\mathbb{C})\|^2$. First, $\sup_i E\|A_{2,iT}(\mathbb{C})\|^2 \le M$ since $E\|A_{2,iT}(\mathbb{C})\|^2 \le 2E\|A_{2,iT}^*(\mathbb{C})\|^2 + 2E\|A_{2,iT}(\mathbb{C}) - A_{2,iT}^*(\mathbb{C})\|^2 \le M$ by $\sup_i E\|A_{iT}^*(\mathbb{C})\|^2 \le M$ and by Part (g), which we prove later. Next, by definition,

$$\begin{aligned}A_{1,iT}(\mathbb{C}) &= \frac{1}{T^{1/2}}(\Delta_{c_i}\tilde{G}_0)'\Omega_{u,i}^{-1}(\Delta_{c_i}\underline{Y}_i) \\ &= \frac{(c_i - \theta_i)}{n^{1/4}}\frac{1}{T^{3/2}}(\Delta_{c_i}\tilde{G}_0)'\Omega_{u,i}^{-1}\underline{Y}_{-1,i} + \frac{1}{T^{1/2}}(\Delta_{c_i}\tilde{G}_0)'\Omega_{u,i}^{-1}\underline{U}_i \\ &= I_i + II_i, \text{ say.}\end{aligned}$$

Because $\underline{Y}_{-1,i} = A_i\underline{U}_i$, where A_i is defined above Lemma A.1, we have

$$\sup_i E[I_i^2] = \sup_i E\left[\frac{(c_i - \theta_i)^2}{n^{1/2}\omega_i^2}\frac{1}{T^3}(\Delta_{c_i}\tilde{G}_0)'\Omega_{u,i}^{-1}A_i\Omega_{u,i}A_i\Omega_{u,i}^{-1}(\Delta_{c_i}\tilde{G}_0)\right]$$

$$\leq M \frac{1}{n^{1/2}} \left(\sup_i \| \frac{\Delta_{c_i} \tilde{G}_0}{T^{1/2}} \|^2 \right) \left(\sup_i \| \frac{A_i}{T} \|^2 \right) (\sup_i \| \Omega_{u,i}^{-1} \|_o^2) (\sup_i \| \Omega_{u,i} \|_o)$$

$$= O\left(\frac{1}{n^{1/2}} \right) = o(1),$$

and

$$\sup_i E[II_i^2] = \sup_i \frac{1}{T} (\Delta_{c_i} \tilde{G}_0)' \Omega_{u,i}^{-1} (\Delta_{c_i} \tilde{G}_0)$$

$$\leq \left(\sup_i \| \frac{\Delta_{c_i} \tilde{G}_0}{T^{1/2}} \|^2 \right) (\sup_i \| \Omega_{u,i}^{-1} \|_o) = O(1).$$

Therefore, we have

$$\sup_i E \| A_{1,iT}^*(\mathbb{C}) \|^2 \leq M,$$

as required.

Part (e). Note that

$$B_{11,iT}(\mathbb{C}) - B_{11,iT}(0) = \frac{1}{T} (\Delta_{c_i} \tilde{G}_0 - \Delta \tilde{G}_0)' \Omega_{u,i}^{-1} (\Delta_{c_i} \tilde{G}_0 - \Delta \tilde{G}_0)$$

$$- \frac{2}{T} (\Delta_{c_i} \tilde{G}_0 - \Delta \tilde{G}_0)' \Omega_{u,i}^{-1} (\Delta \tilde{G}_0).$$

The required result follows because

$$\sup_i | B_{11,iT}(\mathbb{C}) - B_{11,iT}(0) | \leq \frac{1}{T} (\sup_i \| \Delta_{c_i} \tilde{G}_0 - \Delta \tilde{G}_0 \|^2) (\sup_i \| \Omega_{u,i}^{-1} \|_o)$$

$$+ \frac{2}{T} (\sup_i \| \Delta_{c_i} \tilde{G}_0 - \Delta \tilde{G}_0 \|) (\sup_i \| \Omega_{u,i}^{-1} \|_o) \| \Delta \tilde{G}_0 \|$$

$$= \frac{1}{T} O\left(\frac{1}{n^{1/2}} \right) O(1) + \frac{1}{T} O\left(\frac{1}{n^{1/4}} \right) O(1) O(T^{1/2}) = O\left(\frac{1}{n^{1/4} T^{1/2}} \right),$$

as required.

Part (f) follows by Lemma A.3(b2).

Part (g). By definition, we have

$$\begin{aligned} & A_{2,iT}(\mathbb{C}) - A_{2,iT}^*(\mathbb{C}) \\ & = \frac{(c_i - \theta_i)}{n^{1/4}} \frac{1}{T^{3/2}} (\Delta_{c_i} \tilde{G}_1)' \Big(\Omega_{u,i}^{-1} - \frac{1}{\omega_i^2} \Big) \underline{Y}_{-1,i} + \frac{1}{T^{1/2}} (\Delta_{c_i} \tilde{G}_1)' \Big(\Omega_{u,i}^{-1} - \frac{1}{\omega_i^2} \Big) \underline{U}_i \\ & = \frac{(c_i - \theta_i)}{n^{1/4} \omega_i} \frac{1}{T^{3/2}} \Big(\Big(\Delta_{c_i} \tilde{G}_1 \Big)' R_i \Big) \Omega_{u,i}^{-1/2} A_i \underline{U}_i + \frac{1}{\omega_i T^{1/2}} \Big(\Big(\Delta_{c_i} \tilde{G}_1 \Big)' R_i \Big) \Omega_{u,i}^{-1/2} \underline{U}_i \\ & = I_i + II_i, \text{ say.} \end{aligned}$$

Here, the second equality holds because $\underline{Y}_{-1,i} = A_i \underline{U}_i$, where A_i is defined above Lemma A.1, and R_i is defined in (A.8) :

$$\sup_i E[I_i^2] = \sup_i E\left[\frac{(c_i - \theta_i)^2}{n^{1/2} \omega_i^2} \frac{1}{T^3} ((\Delta_{c_i} \tilde{G}_1)' R_i) \Omega_{u,i}^{-1/2} A_i \Omega_{u,i} A_i \Omega_{u,i}^{-1/2} (R_i' (\Delta_{c_i} \tilde{G}_1)) \right]$$

$$\leq M \frac{1}{n^{1/2}} (\sup_i \frac{1}{T} \| (\Delta_{c_i} \tilde{G}_1)' R_i \|^2) (\sup_i \| \frac{A_i}{T} \|^2) (\sup_i \| \Omega_{u,i}^{-1} \|_o) (\sup_i \| \Omega_{u,i} \|_o)$$

$$= O\left(\frac{1}{n^{1/2}T^{1/2}}\right)$$

and

$$\sup_i E[II_i^2] = \sup_i \frac{1}{\omega_i^2} \sup_i \frac{1}{T} \|(\Delta_{c_i} \tilde{G}_1)' R_i\|^2 = O\left(\frac{1}{T^{1/2}}\right).$$

Combining the bounds of $\sup_i E[I_i^2]$ and $\sup_i E[II_i^2]$, we have the desired result for Part (g). □

Proof of (A.6): The required result follows if we show

$$\sum_{i=1}^n [A_{1,iT}(\mathbb{C})^2 B_{11,iT}(\mathbb{C})^{-1} - A_{1,iT}(0)^2 B_{11,iT}(0)^{-1}] = o_p(1),$$

$$\sum_{i=1}^n [A_{1,iT}^*(\mathbb{C})^2 B_{11,iT}^*(\mathbb{C})^{-1} - A_{1,iT}^*(0)^2 B_{11,iT}^*(0)^{-1}] = o_p(1).$$

Note that

$$\begin{aligned} &|\sum_{i=1}^n [A_{1,iT}(\mathbb{C})^2 B_{11,iT}(\mathbb{C})^{-1} - A_{1,iT}(0)^2 B_{11,iT}(0)^{-1}]| \\ &\leq \sum_{i=1}^n |A_{1,iT}(\mathbb{C})^2 (B_{11,iT}(\mathbb{C})^{-1} - B_{11,iT}(0)^{-1})| \\ &\quad + \sum_{i=1}^n |A_{1,iT}(\mathbb{C})^2 - A_{1,iT}(0)^2| B_{11,iT}(0)^{-1}. \end{aligned}$$

The first term is bounded by

$$\begin{aligned} &n(\sup_i B_{11,iT}(\mathbb{C})^{-1})(\sup_i B_{11,iT}(0)^{-1}) \sup_i |B_{11,iT}(\mathbb{C}) - B_{11,iT}(0)| \left(\frac{1}{n}\sum_{i=1}^n A_{1,iT}(\mathbb{C})^2\right) \\ &= nO(1)O(1)O\left(\frac{1}{n^{1/4}T^{1/2}}\right)O_p(1) = O_p\left(\frac{n^{3/4}}{T^{1/2}}\right) = o_p(1), \end{aligned}$$

where the first equality holds by Lemma A.4(c)–(e) and the last equality holds because $(n/T^{1/2}) = o(1)$. The second term is bounded by

$$\begin{aligned} &n\left(\frac{1}{n}\sum_{i=1}^n (A_{1,iT}(\mathbb{C}) - A_{1,iT}(0))^2\right)^{1/2} \left(\frac{1}{n}\sum_{i=1}^n (A_{1,iT}(\mathbb{C}) + A_{1,iT}(0))^2\right)^{1/2} \sup_i B_{11,iT}(0)^{-1} \\ &= nO_p\left(\frac{1}{n^{1/8}T^{1/2}}\right)O_p(1)O(1) = O_p\left(\frac{n^{7/8}}{T^{1/2}}\right) = o_p(1), \end{aligned}$$

where the first equality holds by Lemma A.4(c), (d) and $\sup_i E[A_{1,iT}(\mathbb{C}) - A_{1,iT}(0)]^2 = O(1/(n^{1/2}T))$, and the last equality holds because $(n/T^{1/2}) = o(1)$. Combining these two, we have the required result

$$\sum_{i=1}^n (A_{1,iT}(\mathbb{C})^2 B_{11,iT}(\mathbb{C})^{-1} - A_{1,iT}(0)^2 B_{11,iT}(0)^{-1}) = o_p(1).$$

The second term follows in similar fashion and we omit it. □

Proof of (A.7): The required result follows, if we show

$$\sum_{i=1}^{n}(A_{iT}(\mathbb{C})'B_{iT}(\mathbb{C})^{-1}A_{iT}(\mathbb{C}) - A_{iT}(0)B_{iT}(0)^{-1}A_{iT}(0))$$
$$-\sum_{i=1}^{n}(A^*_{iT}(\mathbb{C})'B^*_{iT}(\mathbb{C})^{-1}A^*_{iT}(\mathbb{C}) - A^*_{iT}(0)'B^*_{iT}(0)^{-1}A^*_{iT}(0))$$
$$= o_p(1),$$

which is established by the following three steps.

Step 1. We show

$$\sum_{i=1}^{n}(A_{iT}(\mathbb{C})'B_{iT}(\mathbb{C})^{-1}A_{iT}(\mathbb{C}) - A_{iT}(0)B_{iT}(0)^{-1}A_{iT}(0))$$
$$-\sum_{i=1}^{n}(A^*_{iT}(\mathbb{C})'B^*_{iT}(\mathbb{C})^{-1}A^*_{iT}(\mathbb{C}) - A^*_{iT}(0)'B^*_{iT}(0)^{-1}A^*_{iT}(0))$$
$$= \sum_{i=1}^{n}(A_{iT}(\mathbb{C})'\tilde{B}_{iT}(\mathbb{C})^{-1}A_{iT}(\mathbb{C}) - A_{iT}(0)\tilde{B}_{iT}(0)^{-1}A_{iT}(0))$$
$$-\sum_{i=1}^{n}(A^*_{iT}(\mathbb{C})'\tilde{B}^*_{iT}(\mathbb{C})^{-1}A^*_{iT}(\mathbb{C}) - A^*_{iT}(0)'\tilde{B}^*_{iT}(0)^{-1}A^*_{iT}(0)) + o_p(1).$$

Step 2. By (A.6), we have

$$\sum_{i=1}^{n}(A_{iT}(\mathbb{C})'\tilde{B}_{iT}(\mathbb{C})^{-1}A_{iT}(\mathbb{C}) - A_{iT}(0)\tilde{B}_{iT}(0)^{-1}A_{iT}(0))$$
$$-\sum_{i=1}^{n}(A^*_{iT}(\mathbb{C})'\tilde{B}^*_{iT}(\mathbb{C})^{-1}A^*_{iT}(\mathbb{C}) - A^*_{iT}(0)'\tilde{B}^*_{iT}(0)^{-1}A^*_{iT}(0))$$
$$= \sum_{i=1}^{n}(A_{2,iT}(\mathbb{C})^2B_{22,iT}(\mathbb{C})^{-1} - A^*_{2,iT}(\mathbb{C})^2B^*_{22,iT}(\mathbb{C})^{-1})$$
$$-\sum_{i=1}^{n}(A_{2,iT}(0)^2B_{22,iT}(0)^{-1} - A^*_{2,iT}(0)^2B^*_{22,iT}(0)^{-1}) + o_p(1).$$

Step 3. We show

$$\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C})^2B_{22,iT}(\mathbb{C})^{-1} - A^*_{2,iT}(\mathbb{C})^2B^*_{22,iT}(\mathbb{C})^{-1}) = o_p(1)$$
$$\sum_{i=1}^{n}(A_{2,iT}(0)^2B_{22,iT}(0)^{-1} - A^*_{2,iT}(0)^2B^*_{22,iT}(0)^{-1}) = o_p(1).$$

□

Proof of Step 1: Note that because $B^*_{iT}(\mathbb{C})$ is a diagonal matrix,

$$\sum_{i=1}^{n}A^*_{iT}(\mathbb{C})'(B^*_{iT}(\mathbb{C})^{-1} - \tilde{B}^*_{iT}(\mathbb{C})^{-1})A^*_{iT}(\mathbb{C}) = 0.$$

Then, the required result for Step 1 follows if we show

$$\sum_{i=1}^{n} A_{iT}(\mathbb{C})'(B_{iT}(\mathbb{C})^{-1} - \tilde{B}_{iT}(\mathbb{C})^{-1})A_{iT}(\mathbb{C}) = o_p(1).$$

The required result follows because

$$\begin{aligned}
&|\sum_{i=1}^{n} A_{iT}(\mathbb{C})'(B_{iT}(\mathbb{C})^{-1} - \tilde{B}_{iT}(\mathbb{C})^{-1})A_{iT}(\mathbb{C})| \\
&= |\sum_{i=1}^{n} A_{iT}(\mathbb{C})'(B_{iT}(\mathbb{C})^{-1}(\tilde{B}_{iT}(\mathbb{C}) - B_{iT}(\mathbb{C}))\tilde{B}_{iT}(\mathbb{C})^{-1})A_{iT}(\mathbb{C})| \\
&\le \sum_{i=1}^{n} \|A_{iT}(\mathbb{C})\|^2 \|B_{iT}(\mathbb{C})^{-1}\| \|\tilde{B}_{iT}(\mathbb{C})^{-1}\| \|\tilde{B}_{iT}(\mathbb{C}) - B_{iT}(\mathbb{C})\| \\
&\le n(\sup_i \|B_{iT}(\mathbb{C})^{-1}\|)(\sup_i \|\tilde{B}_{iT}(\mathbb{C})^{-1}\|)(\sup_i \|\tilde{B}_{iT}(\mathbb{C}) - B_{iT}(\mathbb{C})\|)(\frac{1}{n}\sum_{i=1}^{n} \|A_{iT}(\mathbb{C})\|^2) \\
&= nO(1)O(1)O\left(\frac{1}{T^{1/2}}\right)O_p(1) = O_p\left(\frac{n}{T^{1/2}}\right) = o_p(1),
\end{aligned}$$

where the last line holds by Lemma A.4(b)–(d) and the condition $(n/T^{1/4}) \to 0$. □

Proof of Step 3: We show

$$\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C})^2 B_{22,iT}(\mathbb{C})^{-1} - A^*_{2,iT}(\mathbb{C})^2 B^*_{22,iT}(\mathbb{C})^{-1}) = o_p(1).$$

The other required result $\sum_{i=1}^{n}(A_{2,iT}(0)^2 B_{22,iT}(0)^{-1} - A^*_{2,iT}(0)^2 B^*_{22,iT}(0)^{-1}) = o_p(1)$ follows in similar fashion and we omit the derivation. Note that

$$\begin{aligned}
&|\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C})^2 B_{22,iT}(\mathbb{C})^{-1} - A^*_{2,iT}(\mathbb{C})^2 B^*_{22,iT}(\mathbb{C})^{-1})| \\
&\le |\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C})^2 - A^*_{2,iT}(\mathbb{C})^2)B_{22,iT}(\mathbb{C})^{-1}| \\
&+ |\sum_{i=1}^{n} A^*_{2,iT}(\mathbb{C})^2(B_{22,iT}(\mathbb{C})^{-1} - B^*_{22,iT}(\mathbb{C})^{-1})|.
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
&|\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C})^2 - A^*_{2,iT}(\mathbb{C})^2)B_{22,iT}(\mathbb{C})^{-1}| \le n\Big(\frac{1}{n}\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C}) - A^*_{2,iT}(\mathbb{C}))^2\Big)^{1/2} \\
&\times\Big(\frac{1}{n}\sum_{i=1}^{n}(A_{2,iT}(\mathbb{C}) + A^*_{2,iT}(\mathbb{C}))^2\Big)^{1/2} \sup_i B_{22,iT}(\mathbb{C})^{-1} \\
&= nO_p\left(\frac{1}{T^{1/4}}\right)O_p(1)O(1) = O_p\left(\frac{n}{T^{1/4}}\right) = o_p(1),
\end{aligned}$$

where the first equality holds by Lemma A.4(c), (d) and (g) and the last equality holds by the condition $(n/T^{1/4}) \to 0$. For the second term, note that

$$
\begin{aligned}
&|\sum_{i=1}^{n}(A^*_{2,iT}(\mathbb{C})^2(B_{22,iT}(\mathbb{C})^{-1} - B^*_{22,iT}(\mathbb{C})^{-1}))| \\
&= |\sum_{i=1}^{n}(A^*_{2,iT}(\mathbb{C})^2 B_{22,iT}(\mathbb{C})^{-1}(B_{22,iT}(\mathbb{C}) - B^*_{22,iT}(\mathbb{C}))B^*_{22,iT}(\mathbb{C})^{-1})| \\
&\leq n(\frac{1}{n}\sum_{i=1}^{n} A^*_{2,iT}(\mathbb{C})^2)(\sup_i B_{22,iT}(\mathbb{C})^{-1})(\sup_i B^*_{22,iT}(\mathbb{C})^{-1}) \sup_i |B_{22,iT}(\mathbb{C}) - B^*_{22,iT}(\mathbb{C})| \\
&= n O_p(1)O(1)O(1)O(\frac{1}{T^{1/2}}) = O_p(\frac{n}{T^{1/2}}) = o_p(1),
\end{aligned}
$$

where the second equality holds by Lemma A.4(c), (d) and (f), and the last equality holds by the condition $(n/T^{1/4}) \to 0$. Then, we have all of the desired results for Part (c). □

APPENDIX B: PROOFS OF THE LIMIT DISTRIBUTION RESULTS, THEOREMS 2.2, 2.4 AND 2.6

In this section, we provide proofs of Theorems 2.2, 2.4 and 2.6. These proofs are very similar to the proofs of the corresponding results in MPP and therefore we provide just an outline of the proofs here.

Proof of Theorem 2.2: Because $\Delta y_{it} = -(\theta_i/(n^{1/2}T))y_{it-1} + u_{it}$, we can write

$$
\begin{aligned}
V_{nT}(\mathbb{C}) &= \sum_{i=1}^{n}\frac{1}{\omega_i^2}(y_{i1}^2 + \sum_{t=2}^{T}(\Delta_{c_i} y_{it})^2) - \sum_{i=1}^{n}\frac{1}{\omega_i^2}(y_{i1}^2 + \sum_{t=2}^{T}(\Delta y_{it})^2) - \frac{2}{n^{1/2}}\sum_{i=1}^{n} c_i\frac{\lambda_i}{\omega_i^2} - \frac{1}{2}\mu_{c,2} \\
&= \frac{2}{n^{1/2}T}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\sum_{t=2}^{T}\Delta y_{it} y_{it-1} + \frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i^2}{\omega_i^2}\sum_{t=2}^{T} y_{it-1}^2 - \frac{2}{n^{1/2}}\sum_{i=1}^{n} c_i\frac{\lambda_i}{\omega_i^2} - \frac{1}{2}\mu_{c,2} \\
&= -\frac{2}{nT^2}\sum_{i=1}^{n}\frac{c_i\theta_i}{\omega_i^2}\sum_{t=2}^{T} y_{it-1}^2 + \frac{2}{n^{1/2}T}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\sum_{t=2}^{T} u_{it} y_{it-1} - \frac{2}{n^{1/2}}\sum_{i=1}^{n} c_i\frac{\lambda_i}{\omega_i^2} \\
&\quad + \frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i^2}{\omega_i^2}\sum_{t=2}^{T} y_{it-1}^2 - \frac{1}{2}\mu_{c,2}.
\end{aligned}
$$

Direct calculation shows that under the assumptions of the theorem, we have the following joint limits

$$
-\frac{2}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{c_i\theta_i}{\omega_i^2} y_{it-1}^2 \xrightarrow{p} -E[c_i\theta_i],
$$

$$
\frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i^2}{\omega_i^2}\sum_{t=1}^{T} y_{it-1}^2 \xrightarrow{p} \frac{1}{2}\mu_{c,2},
$$

and the central limit theorem (e.g. Moon and Phillips, 1999)

$$
\frac{2}{n^{1/2}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\left(\frac{1}{T}\sum_{t=1}^{T} u_{it} y_{it-1} - \lambda_i\right) \to N(0, 2\mu_{c,2}),
$$

thereby giving the required result. □

Proof of Theorem 2.4: For the required result of the theorem, it is enough to show that

$$V_{fe1,nT}(\mathbb{C}) = V_{nT}(\mathbb{C}) + o_p(1).$$

Let $\hat{b}_{0i}(c_i) = (\Delta_{c_i} G_0' \Delta_{c_i} G_0)^{-1}(\Delta_{c_i} G_0' \Delta_{c_i} \underline{Z}_i)$. Then, $\underline{Z}_i - G_0\hat{b}_{0i}(c_i) = \underline{Y}_i - G_0(\hat{b}_{0i}(c_i) - b_{0i})$, and we can rewrite $V_{fe1,nT}(\mathbb{C})$ as

$$\begin{aligned} V_{fe1,nT}(\mathbb{C}) = \sum_{i=1}^{n} \frac{1}{\omega_i^2}\Big(&(\Delta_{c_i}\underline{Y}_i - \Delta_{c_i}G_0(\hat{b}_{0i}(c_i) - b_{0i}))'(\Delta_{c_i}\underline{Y}_i - \Delta_{c_i}G_0(\hat{b}_{0i}(c_i) - b_{0i})) \\ &-(\Delta\underline{Y}_i - \Delta G_0(\hat{b}_{0i}(c_i) - b_{0i}))'(\Delta\underline{Y}_i - \Delta G_0(\hat{b}_{0i}(c_i) - b_{0i}))\Big) \\ &-\frac{2}{n^{1/2}}\sum_{i=1}^{n} c_i \frac{\lambda_i}{\omega_i^2} - \frac{1}{2}\mu_{c,2} = V_{nT}(\mathbb{C}) + V_{fe11,nT}(\mathbb{C}), \end{aligned}$$

where

$$\begin{aligned} V_{fe11,nT}(\mathbb{C}) = \sum_{i=1}^{n}\frac{1}{\omega_i^2}\Big(&(\Delta\underline{Y}_i'\Delta G_0)(\Delta G_0'\Delta G_0)^{-1}(\Delta G_0'\Delta\underline{Y}_i) \\ &-(\Delta_{c_i}\underline{Y}_i'\Delta_{c_i}G_0)(\Delta_{c_i}G_0'\Delta_{c_i}G_0)^{-1}(\Delta_{c_i}G_0'\Delta_{c_i}\underline{Y}_i)\Big). \end{aligned}$$

We can follow the proof in MPP (pp. 449–450) and deduce that

$$V_{fe11,nT}(\mathbb{C}) = o_p(1),$$

as $n, T \to \infty$ with $(n/T) \to 0$, which proves the desired result. □

Proof of Theorem 2.6: The required result is a consequence of Lemmata A.5 and A.6. □

LEMMA A.5. *Let Assumption 2.1 hold. Then, as $n, T \to \infty$ with $(n/T) \to 0$, we have*

$$\begin{aligned} V_{fe2,nT}(\mathbb{C}) = &\frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\left(\frac{2}{T}\sum_{t=2}^{T}\Delta y_{it}y_{it-1} - \left(\frac{y_{iT}}{\sqrt{T}}\right)^2 + \left(\frac{y_{i1}}{\sqrt{T}}\right)^2 + \sigma_i^2\right) + \frac{1}{n^{1/2}}\sum_{i=1}^{n} \\ &\times\frac{c_i^2}{\omega_i^2}\left(\frac{1}{T^2}\sum_{t=2}^{T}y_{it-1}^2 - 2\left(\frac{y_{iT}}{\sqrt{T}}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=2}^{T}\frac{t}{T}y_{it-1}\right) + \frac{1}{3}\left(\frac{y_{iT}}{\sqrt{T}}\right)^2 + \omega_i^2\omega_{p2T}\right) + \frac{1}{n}\sum_{i=1}^{n} \\ &\times\frac{c_i^4}{\omega_i^2}\left(-\left(\frac{1}{T\sqrt{T}}\sum_{t=2}^{T}\frac{t}{T}y_{it-1}\right)^2 + \frac{2}{3}\left(\frac{y_{iT}}{\sqrt{T}}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=2}^{T}\frac{t}{T}y_{it-1}\right) - \frac{1}{9}\left(\frac{y_{iT}}{\sqrt{T}}\right)^2 + \omega_i^2\omega_{p4T}\right) \\ &+o_p(1). \end{aligned}$$

Proof: The proof is similar to the proof of Lemma 11 of MPP and is omitted. □

LEMMA A.6. *Let Assumption 2.1 hold. Then, as $n, T \to \infty$ with $(n/T) \to 0$, the following hold:*

$$(a)\quad \frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\left(\frac{2}{T}\sum_{t=1}^{T}\Delta y_{it}y_{it-1} - \left(\frac{y_{iT}}{\sqrt{T}}\right)^2 + \left(\frac{y_{i1}}{\sqrt{T}}\right)^2 + \sigma_i^2\right) = o_p(1);$$

$$(b)\qquad \frac{1}{n^{1/2}}\sum_{i=1}^{n}\frac{c_i^2}{\omega_i^2}\Bigg(\bigg(\frac{1}{T^2}\sum_{t=2}^{T}y_{it-1}^2-\omega_i^2\frac{1}{T}\sum_{t=2}^{T}\frac{t-1}{T}\bigg)+\frac{1}{3}\bigg(\Big(\frac{y_{iT}}{\sqrt{T}}\Big)^2-\omega_i^2\bigg)$$

$$-\bigg(2\Big(\frac{y_{iT}}{\sqrt{T}}\Big)\Big(\frac{1}{T\sqrt{T}}\sum_{t=2}^{T}\frac{t}{T}y_{it-1}\Big)-\omega_i^2\frac{2}{T}\sum_{t=2}^{T}\Big(\frac{t}{T}\Big)\Big(\frac{t-1}{T}\Big)\bigg)\Bigg)$$

$$\to N\Big(-\frac{1}{90}E[c_i^2\theta_i^2],\ \frac{1}{45}E[c_i^4]\Big);$$

$$(c)\qquad \frac{1}{n}\sum_{i=1}^{n}\frac{c_i^4}{\omega_i^2}\bigg(-\Big(\frac{1}{T\sqrt{T}}\sum_{t=2}^{T}\frac{t}{T}y_{it-1}\Big)^2+\frac{2}{3}\Big(\frac{y_{iT}}{\sqrt{T}}\Big)\Big(\frac{1}{T\sqrt{T}}\sum_{t=2}^{T}\frac{t}{T}y_{it-1}\Big)$$

$$-\frac{1}{9}\Big(\frac{y_{iT}}{\sqrt{T}}\Big)^2+\omega_i^2\omega_{p4T}\bigg)=o_p(1).$$

Proof: The proofs of Parts (b) and (c) are similar to those of Lemma 12(b) and (c) of MPP, and are omitted.

Part (a). First, note from

$$y_{it}^2-y_{it-1}^2=(\rho_i^2-1)y_{it-1}^2+2\rho_i y_{it-1}u_{it}+u_{it}^2 \text{ for } t\geq 2,$$

that

$$\Big(\frac{y_{iT}}{\sqrt{T}}\Big)^2-\Big(\frac{y_{i1}}{\sqrt{T}}\Big)^2=(\rho_i^2-1)\frac{1}{T}\sum_{t=2}^{T}y_{it-1}^2+2\rho_i\frac{1}{T}\sum_{t=2}^{T}y_{it-1}u_{it}+\frac{1}{T}\sum_{t=2}^{T}u_{it}^2.$$

Because $\Delta y_{it}=(\rho_i-1)y_{it-1}+u_{it}$, we have

$$2\frac{1}{T}\sum_{t=2}^{T}\Delta y_{it}y_{it-1}=2(\rho_i-1)\frac{1}{T}\sum_{t=2}^{T}y_{it-1}^2+2\frac{1}{T}\sum_{t=2}^{T}y_{it-1}u_{it}.$$

Then,

$$\frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\bigg(\frac{2}{T}\sum_{t=2}^{T}\Delta y_{it}y_{it-1}-\Big(\frac{y_{iT}}{\sqrt{T}}\Big)^2+\Big(\frac{y_{i1}}{\sqrt{T}}\Big)^2+\sigma_i^2\bigg)$$

$$=\frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\bigg(-(\rho_i-1)^2\frac{1}{T}\sum_{t=2}^{T}y_{it-1}^2+2(1-\rho_i)\frac{1}{T}\sum_{t=1}^{T}y_{it-1}u_{it}-\Big(\frac{1}{T}\sum_{t=1}^{T}u_{it}^2-\sigma_i^2\Big)\bigg).$$

Under the assumptions of the lemma,

$$\frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}(\rho_i-1)^2\frac{1}{T}\sum_{t=2}^{T}y_{it-1}^2=\frac{n^{1/4}}{T}\bigg(\frac{1}{n}\sum_{i=1}^{n}c_i\theta_i^2\Big(\frac{1}{T^2\omega_i^2}\sum_{t=2}^{T}y_{it-1}^2\Big)\bigg)=O_p\Big(\frac{n^{1/4}}{T}\Big),$$

$$\frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}(1-\rho_i)\frac{1}{T}\sum_{t=2}^{T}y_{it-1}u_{it}=\frac{1}{T}\frac{2}{n^{1/2}}\sum_{i=1}^{n}\theta_i\Big(\frac{1}{T\omega_i^2}\sum_{t=2}^{T}y_{it-1}u_{it}\Big)=O_p\Big(\frac{n^{1/2}}{T}\Big),$$

and

$$\frac{1}{n^{1/4}}\sum_{i=1}^{n}\frac{c_i}{\omega_i^2}\left(\frac{1}{T}\sum_{t=2}^{T}u_{it}^2-\sigma_i^2\right)=O_p\left(\frac{n^{1/4}}{T^{1/2}}\right),$$

leading to the required result for Part (a). □

APPENDIX C: PROOF OF THEOREM 2.7

Proof of Theorem 2.7: We provide a sketch of the proof. Note that under Assumption 2.2, the following hold:

$$\sup_i|\hat{\omega}_i^2-\omega_i^2|,\quad \sup_i|\hat{\lambda}_i^2-\lambda_i^2|,\quad \sup_i|\hat{\sigma}_i^2-\sigma_i^2|=o_p(1);$$

$$\sum_{i=1}^{n}\left(\hat{\omega}_i^2-\omega_i^2\right)^2,\quad \sum_{i=1}^{n}\left(\hat{\lambda}_i^2-\lambda_i^2\right)^2,\quad \sum_{i=1}^{n}\left(\hat{\sigma}_i^2-\sigma_i^2\right)^2=o_p(1).$$

Define $\hat{\imath}=\arg\min_{i\in\{1,\ldots,n\}}\hat{\omega}_i^2$ and $i^*=\arg\min_{i\in\{1,\ldots,n\}}\omega_i^2$ Then,

$$\inf_i\hat{\omega}_i^2-\inf_i\omega_i^2\geq\hat{\omega}_{\hat{\imath}}^2-\omega_{\hat{\imath}}^2\geq-\sup_i|\hat{\omega}_i^2-\omega_i^2|=o_p(1)$$

and

$$\inf_i\hat{\omega}_i^2-\inf_i\omega_i^2\leq\hat{\omega}_{i^*}^2-\omega_{i^*}^2\leq\sup_i|\hat{\omega}_i^2-\omega_i^2|=o_p(1).$$

Because $\inf_i\omega_i^2>0$ under Assumption 2.1, we have

$$\inf_i\hat{\omega}_i^2=\inf_i\omega_i^2+o_p(1)>0$$

with probability approaching one. These imply that $\hat{\omega}_i^2$ satisfies the properties in Lemmata 8, 10 and 14 of MPP, while $\hat{\lambda}_i^2$ and $\hat{\sigma}_i^2$ satisfy the properties in Lemmata 8(a), 8(b), 10(a) and 14(a)–(d) of MPP. The desired results follow by similar arguments to those used in Theorems 8, 10 and 15 of MPP. □

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Replication Files