Econometric Theory, **0**, 2024, 1–26.
doi:10.1017/S0266466624000112

ASYMPTOTICALLY UNIFORMLY MOST POWERFUL TESTS FOR UNIT ROOTS IN GAUSSIAN PANELS WITH CROSS-SECTIONAL DEPENDENCE GENERATED BY COMMON FACTORS

OLIVER WICHERT
CHARLES RIVER ASSOCIATES

I. GAIA BECHERI
INDEPENDENT RESEARCHER

FEIKE C. DROST
TILBURG UNIVERSITY

RAMON VAN DEN AKKER
TILBURG UNIVERSITY

This paper considers testing for unit roots in Gaussian panels with cross-sectional dependence generated by common factors. Within our setup, we can analyze restricted versions of the two prevalent approaches in the literature, that of Moon and Perron (2004, *Journal of Econometrics* 122, 81–126), who specify a factor model for the innovations, and the PANIC setup proposed in Bai and Ng (2004, *Econometrica* 72, 1127–1177), who test common factors and idiosyncratic deviations separately for unit roots. We show that both frameworks lead to locally asymptotically normal experiments with the *same* central sequence and Fisher information. Using Le Cam's theory of statistical experiments, we obtain the local asymptotic power envelope for unit-root tests. We show that the popular Moon and Perron (2004, *Journal of Econometrics* 122, 81–126) and Bai and Ng (2010, *Econometric Theory* 26, 1088–1114) tests only attain the power envelope in case there is no heterogeneity in the long-run variance of the idiosyncratic components. We develop a new test which is asymptotically uniformly most powerful irrespective of possible heterogeneity in the long-run variance of the idiosyncratic components. Monte Carlo simulations corroborate our asymptotic results and document significant gains in finite-sample power if the variances of the idiosyncratic shocks differ substantially among the cross-sectional units.

Address correspondence to Feike C. Drost, Econometrics Group, CentER, Tilburg University, Tilburg, The Netherlands; e-mail: F.C.Drost@tilburguniversity.edu

1. INTRODUCTION

Testing for unit roots is an important aspect of time series and panel data analysis. See, for example, the monographs Patterson (2011, 2012) and Choi (2015) for overviews. A well-known problem with univariate unit-root tests is their low power. In the last two decades, increased data availability led to the development of panel unit-root tests that increase the statistical power by exploiting the cross-sectional data dimension. The "first generation" of panel unit-root tests imposes the panel observations Z_{it} to be independent over panel units i. Surveys of this literature are provided by Banerjee (1999), Baltagi and Kao (2000), Choi (2006), Breitung and Pesaran (2008), and Westerlund and Breitung (2013). Moon, Perron, and Phillips (2007), Becheri, Drost, and Van den Akker (2015a), Moon, Perron, and Phillips (2014), and Juodis and Westerlund (2019) studied optimal testing for unit roots for first-generation frameworks, when the driving innovations are Gaussian. O'Connell (1998) and Gutierrez (2006) showed that presence of cross-sectional dependence typically leads to invalidity of "first-generation tests." For this reason, a "second generation" of models and tests has been introduced. To the best of our knowledge, optimality of unit-root tests has not yet been studied for second-generation models.

To (partly) fill this gap, we focus on two widely used setups for second-generation panel unit-root tests: the "PANIC" framework of Bai and Ng (2004) and the framework of Moon and Perron (2004) ("MP"). We introduce the following data-generating process (DGP): the observations Z_{it}, $i = 1, \dots, n$ and $t = 1, \dots, T$, are assumed to be generated by the components specification:

$$Z_{it} = m_i + Y_{it}, \tag{1}$$

$$Y_{it} = \sum_{k=1}^{K} \lambda_{ki} F_{kt} + E_{it}, \tag{2}$$

$$E_{it} = \rho E_{i,t-1} + \eta_{it}, \tag{3}$$

$$F_{kt} = \rho_k F_{k,t-1} + f_{kt}, \tag{4}$$

where λ_{ki} is the loading of *unobserved* factor F_{kt} on panel unit i, the m_i are fixed effects, and the innovations $\{\eta_{it}\}$ and $\{f_{kt}\}$ are assumed to be mutually independent, Gaussian, stationary time series. Section 2.2 discusses the precise assumptions.

For $\rho_k = 1, k = 1, \dots, K,$ we obtain a restricted DGP that falls in the PANIC framework. And, with $\rho_k = \rho, k = 1, \dots, K,$ we can rewrite (2)–(4) as

$$Y_{it} = \rho Y_{i,t-1} + \varepsilon_{it} \text{ and } \varepsilon_{it} = \sum_{k=1}^{K} \lambda_{ki} f_{kt} + \eta_{it}, \tag{5}$$

which corresponds to Displays (1)–(3) in MP. Note that MP uses an autoregressive structure with the factors appearing in the innovations ε_{it} in (5), whereas the factors are part of the "mean specification," i.e., (2), in the PANIC setup. Consequently, the PANIC framework allows for nonstationarity of Z_{it} generated by the factors F_{kt} and for nonstationarity generated by the idiosyncratic components E_{it}, while the factors and the idiosyncratic components have the same order of integration in the MP framework.

Following Bai and Ng (2010), Pesaran, Smith, and Yamagata (2013), and Westerlund (2015), when considering the PANIC framework, we focus on testing for unit roots in the idiosyncratic components, i.e., $\mathrm{H}_0 : \rho = 1$ versus $\mathrm{H}_a : \rho < 1$. Note that, under the null hypothesis, the model equations of both models coincide. The main restrictions on the DGP considered here are the absence of idiosyncratic deterministic trends and the assumption of Gaussian innovations $\{\eta_{it}\}$. In Section 7, we discuss to what extent we expect that these assumptions can be relaxed. As another extension, it is possible to allow for heterogeneity in the alternatives (i.e., to replace ρ by ρ_i under the alternative hypothesis; see Remark 2.3).

We show that in cases where the nuisance parameters are known, the MP experiment is *locally asymptotically normal* (LAN) when $n, T \to \infty$ (jointly). This means that the limit experiment, in the Le Cam sense, is a simple Gaussian shift experiment (see, for example, Van der Vaart, 2000). We further establish that the PANIC experiment for the idiosyncratic parts, in cases where the nuisance parameters are known, is also LAN with *the same central sequence and Fisher information* as for the MP experiment.

The LAN results imply that *for any test satisfying a mild regularity condition, it suffices to determine its asymptotic size and local power in one of the frameworks, since the same results automatically hold for the other one.* To the best of our knowledge, even for the well-studied tests proposed in Moon and Perron (2004) and Bai and Ng (2010), the literature only conducted specific local asymptotic power analyses without noting the power implications for other local alternatives.

The LAN results, which are based on known nuisance parameters, directly yield an upper bound, which is the same for PANIC and MP, to the local asymptotic power of unit-root tests. We demonstrate that we can attain this upper bound also for the case the $O(n)$ nuisance parameters are unknown. In other words, we establish *adaptivity*: the obtained upper bound yields the *local asymptotic power envelope*.

On comparing the local asymptotic power functions of the tests proposed in Bai and Ng (2010) and Moon and Perron (2004) to the power envelope, it is seen that these tests are optimal only in cases where there is no heterogeneity in the long-run variances of the idiosyncratic components. We propose a new test that is asymptotically uniformly most powerful (UMP). This test is also valid (under suitable moment conditions) in non-Gaussian settings. We report numerical asymptotic powers for commonly encountered amounts of heterogeneity and use Monte Carlo experiments to show that the new test also compares favorably to existing tests in finite samples. These results extend the work on optimality on first-generation frameworks (Moon et al., 2007; Becheri et al., 2015a; Moon et al., 2014; Juodis and Westerlund, 2019), who considered optimal testing for unit roots in Gaussian panels without cross-sectional dependence, to the second-generation models.

The paper is organized as follows. Section 2 presents and discusses the precise assumptions we impose. Section 3 derives the common approximation to the local likelihood ratios in the two experiments and derives its limiting distribution. Section 4 introduces our new UMP test based on the limit experiment. Section 5 computes the local asymptotic power functions of the tests proposed in Moon and Perron (2004) and Bai and Ng (2010), and Section 6 compares their asymptotic and finite-sample power to those of the new UMP test. Section 7 concludes. All proofs are organized in several appendixes that also contain additional results on finite-sample performances. These appendixes are available as the Supplementary Material.

2. NOTATION AND ASSUMPTIONS

2.1. Matrix Notation

Before we introduce our assumptions, we introduce some notation in order to write the model in matrix form. We write I_n and I_T for identity matrices of dimension n and T, respectively, while ι denotes a T-vector of ones. Introduce the n-vectors $\lambda_k = (\lambda_{k1},\ldots,\lambda_{kn})'$, $k = 1,\ldots,K$, and the $n\times K$ matrix $\Lambda = (\lambda_1,\ldots,\lambda_K)$. Collect the observations as $Z = (Z_{11}, Z_{12},\ldots,Z_{1T},\ldots,Z_{n1},\ldots,Z_{nT})'$. We also write $Z_{-1} = (Z_{10}, Z_{11},\ldots, Z_{1,T-1},\ldots,Z_{n0},\ldots,Z_{n,T-1})'$, $\Delta Z = Z - Z_{-1}$, and define ε, η, E, E_{-1}, ΔE, Y, Y_{-1}, and ΔY analogously. Write $m = (m_1,\ldots,m_n)'$, $\eta_i = (\eta_{i1},\ldots,\eta_{iT})'$, $i = 1,\ldots,n$, $f_k = (f_{k1},\ldots,f_{kT})'$, $k = 1,\ldots,K$, and denote their corresponding covariance matrices by $\Sigma_{f,k} = \operatorname{var} f_k \in \mathbb{R}^{T\times T}$ and

$$\Sigma_\eta = \operatorname{diag}(\Sigma_{\eta,1},\ldots,\Sigma_{\eta,n}), \text{ with } \Sigma_{\eta,i} = \operatorname{var} \eta_i \in \mathbb{R}^{T\times T}.$$

The long-run variances of $\{f_{kt}\}$ and $\{\eta_{it}\}$ (see Remark 2.2 below) are denoted by $\omega^2_{f,k}$ and $\omega^2_{\eta,i}$, respectively. In addition, we define the approximate long-run variances $\omega^2_{f,k,T} = \iota'\Sigma_{f,k}\iota/T$ and $\omega^2_{\eta,i,T} = \iota'\Sigma_{\eta,i}\iota/T$. For a given T, these ignore the contribution of any autocovariances further than T apart. We will use the approximate long-run variances to simplify notation and the structure of our proofs. We add the subscript T to the approximate versions to emphasize the difference and define

$$\Omega_\eta = \operatorname{diag}(\omega^2_{\eta,1,T},\ldots,\omega^2_{\eta,n,T}) \text{ and } \Omega_F = \operatorname{diag}(\omega^2_{f,1,T},\ldots,\omega^2_{f,K,T}).$$

In addition to this "vectorized" notation, it will also be useful to consider the observations as $T\times n$ matrices. Thus, let $\tilde\eta = (\eta_1,\ldots,\eta_n)$, and define $\tilde\varepsilon$, $\tilde Y$, $\tilde Z$, $\tilde E$, $\tilde f = (f_1,\ldots,f_K)$, and $\tilde F$ analogously. With this notation, (5) can be rewritten as

$$\tilde\varepsilon = \tilde f\Lambda' + \tilde\eta, \tag{6}$$

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

while for the vectorized versions, we have

$$\varepsilon = \sum_{k=1}^{K} \lambda_k \otimes f_k + \eta.$$

Finally, we introduce the $T \times T$ matrix A by $A_{st} := 1$ if $s > t$ and 0 otherwise, and we put $\mathcal{A} := I_n \otimes A \in \mathbb{R}^{nT \times nT}$, i.e.,

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \dots & 1 & 0 \end{pmatrix} \text{ and } \mathcal{A} = \begin{pmatrix} A & 0_{T\times T} & \dots & 0_{T\times T} \\ 0_{T\times T} & A & \dots & 0_{T\times T} \\ \vdots & \ddots & \ddots & 0_{T\times T} \\ 0_{T\times T} & \dots & 0_{T\times T} & A \end{pmatrix}.$$

The matrix A can be considered a cumulative sum operator, and premultiplying the vectorized panel with $\mathcal{A}$ takes the cumulative sum in the time direction for each panel unit. It is also related to "approximate one-sided long-run variances," which we can define by $\delta_{\eta,i,T} = \mathrm{tr}[A\Sigma_{\eta,i}/T]$ and $\delta_{f,k,T} = \mathrm{tr}[A\Sigma_{f,k}/T]$. Note $A + A' = \iota\iota' - I_T$, so that, analogous to the long-run variances, we have $2\delta_{\eta,i,T} = \omega^2_{\eta,i,T} - \gamma_{\eta,i}(0)$.

2.2. Assumptions

Now we can formally state the full specifications of our DGPs in Equations (1)–(4). The distributional assumptions on the time series of the factors $\{f_{kt}\}$ and idiosyncratic shocks $\{\eta_{it}\}$ are given in Assumption 2.1, and we formulate the assumptions on the (deterministic) factor loadings λ_{ki} in Assumption 2.2. Assumption 2.3 states the assumption on the initial values E_{i0} and F_{k0}. Assumption 2.4 specifies the joint asymptotics we consider in this paper. Finally, Assumption 2.5 differentiates between the two setups discussed in Section 2.

Assumption 2.1.

(a) Each factor innovation, indexed $k = 1, \dots, K$, is a zero-mean ergodic stationary time series $\{f_{kt}\}$ independent of the other factors and all idiosyncratic parts η_{it}. Its autocovariance function $\gamma_{f,k}$ satisfies

$$\sum_{m=-\infty}^{\infty} (|m|+1)|\gamma_{f,k}(m)| < \infty$$

and is such that the variance of each factor innovation $\{f_{kt}\}$ is strictly positive.

(b) For each panel unit $i \in \mathbb{N}$, the idiosyncratic part $\{\eta_{it}\}$ is a Gaussian zero-mean stationary time series independent of the other idiosyncratic parts and all factors. The autocovariance function $\gamma_{\eta,i}$ satisfies

$$\sup_{i\in\mathbb{N}} \sum_{m=-\infty}^{\infty} (|m|+1)|\gamma_{\eta,i}(m)| < \infty \tag{7}$$

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

and is such that the eigenvalues of the $T \times T$ covariance matrices are uniformly bounded away from zero, i.e., $\inf_{i,T} \lambda_{\min}\left(\Sigma_{\eta,i}\right) > 0$.

Remark 2.1. The imposed restrictions on serial correlation are sometimes phrased in terms of spectral densities. Note that our assumption on the boundedness of the eigenvalues is implied by the spectral density being uniformly bounded away from zero (see, for example, Proposition 4.5.3 in Brockwell and Davis, 1991). Similarly, they are sometimes phrased in terms of linear processes on which analogous assumptions are imposed (see, for example, Assumption C in Bai and Ng, 2004 and Assumption 2 in Moon and Perron, 2004). Finally, note that a collection of causal ARMA processes satisfies Assumption 2.1 if the roots are uniformly bounded away from the unit circle.

Remark 2.2. Note that, under Assumption 2.1, the long-run variances of the $\{\eta_{it}\}$, $\omega_{\eta,i}^2$, are also uniformly bounded and uniformly bounded away from zero. The former directly follows from (7), whereas the latter follows from $\omega_{\eta,i}^2 = \lim_{T\to\infty} \frac{1}{T}\iota'\Sigma_{\eta,i}\iota \geq \lim_{T\to\infty} \frac{1}{T}\lambda_{\min}\left(\Sigma_{\eta,i}\right)\iota'\iota \geq \inf_{i,T}\lambda_{\min}\left(\Sigma_{\eta,i}\right) > 0$. Moreover, the one-sided long-run variances

$$\delta_{\eta,i} = \sum_{m=1}^{\infty} \gamma_{\eta,i}(m) = \frac{1}{2}\left(\omega_{\eta,i}^2 - \gamma_{\eta,i}(0)\right), \quad i \in \mathbb{N},$$

are also well defined.

As already announced, we also need to impose some stability on the factor loadings λ_{ki}, which we assume to be fixed. Assumption 2.2 is standard in the literature (cf. Assumption A in Bai and Ng, 2004 or Assumption 6 in Moon and Perron, 2004). It is commonly referred to as the factors being “strong.”

Assumption 2.2. There exists a positive definite $K \times K$ matrix Ψ_Λ such that $\lim_{n\to\infty} \frac{1}{n}\Lambda'\Lambda = \Psi_\Lambda$. Moreover, $\max_{k=1,\dots,K} \sup_{i\in\mathbb{N}} |\lambda_{ki}| < \infty$.

For univariate time series, it is known (see, for example, Müller and Elliott, 2003) that the initial value can have a non-negligible impact on the asymptotic behavior of unit-root tests. Our assumption on the initial values is as follows.

Assumption 2.3. We assume zero starting values: $E_{i0} = 0$ and $F_{k0} = 0$.

We refer to Section 6.2 in Moon et al. (2007) for a discussion on why relaxing initial conditions can be problematic in a panel context and do not pursue this issue further, except by noting that our tests are invariant with respect to the m_i.

Assumption 2.4 below specifies the asymptotic framework we consider throughout this paper. We follow Moon and Perron (2004), Bai and Ng (2010), and Westerlund (2015) in considering large “macro panels,” where both n and T go to infinity, but T will be the larger dimension. We derive all our results using joint asymptotics, which yields more robust results than taking sequential limits where first $T \to \infty$ and subsequently $n \to \infty$.

Assumption 2.4. We consider joint asymptotics, in the Phillips and Moon (1999) sense, with $n/T \to 0$.

Assumption 2.5 below specifies that we either operate in the PANIC (case (a)) or in the MP (case (b)) framework. In the PANIC framework, we allow the long-run variance of the factor innovations to be zero, so that we consider both integrated and stationary factors. This is ruled out in the MP case, in which the factors have the same order of integration as the idiosyncratic parts.

Assumption 2.5. One of the below holds:

(a) For each factor $F_k, k = 1, \dots, K,$ we have $\rho_k = 1$.
(b) For each factor $k = 1, \dots, K$, we have $\rho_k = \rho$. Moreover, $\{f_{kt}\}$ is *Gaussian* and its long-run variance exists and is strictly positive.

We phrase our hypotheses about ρ in Equations (1)–(4) using the following local parameterization.

Assumption 2.6. We use, in (1)–(4), the following local parameterization for ρ in (3):

$$\rho = \rho^{(n,T)} = 1 + \frac{h}{\sqrt{n}T}.$$

As shown below, these rates indeed lead to contiguous alternatives, which allow us to obtain the (local) power of our tests. The unit-root hypothesis can be reformulated in terms of the "local parameter" h:

$$\mathrm{H}_0 : h = 0 \text{ versus } \mathrm{H}_a : h < 0.$$

Remark 2.3. The main setup does not allow for "heterogeneous alternatives" with ρ different across panel units. However, for the case without factors, Becheri et al. (2015a) prove that unobserved heterogeneity in the autoregressive process has no impact on the power envelope or optimal tests. This result can be easily generalized to the current model in case we use, in (1)–(4), the following local parameterization, which allows for heterogeneity, instead of ρ in (3):

$$\rho_i^{(n,T)} = 1 + \frac{h}{\sqrt{n}T} U_i,$$

where $U_1, \dots, U_n$ are unobserved i.i.d. variables with mean 1. The asymptotic power envelope only depends on h and not on the distribution of U_i. In Section 7, we provide a sketch of the proof.

3. LIMIT EXPERIMENT AND POWER ENVELOPE

In this section, we show that likelihood ratios related to the unit-root hypothesis, for the MP and for the PANIC framework, exhibit the same local asymptotic expansion. For both setups, we consider the likelihood ratio for observing Z_{it} in case ρ is the only unknown parameter. Hence, the number of factors K, the factor

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

loadings λ_{ki}, the autocovariance functions, and the fixed effects m_i are considered as known in this section. We will first show, for each model separately, that its likelihood ratio satisfies an expansion, under the null hypothesis, of the form

$$\log \frac{\mathrm{dP}_{h,n,T}}{\mathrm{dP}_{0,n,T}} = h\Delta_{n,T} - h^2 J/2 + o_p(1)$$

with Fisher information $J = 1/2$ and where $\Delta_{n,T}$ will be defined in Lemma 3.1.

In Section 3.3, we consider the limiting distribution of their common central sequence $\Delta_{n,T}$ and will conclude that both experiments enjoy the LAN property. This result allows us to treat the two setups jointly. It also enables us to obtain several important/main results in cases where the nuisance parameters are unknown. First, it yields an upper bound to the local asymptotic powers of tests because the testing problem with unknown nuisance parameters is more complex than the one without. Second, the Gaussian MP and PANIC experiments appear to be *adaptive* with respect to the nuisance parameters since, in Section 4, we propose a new, feasible test (not depending on the unknown nuisance structures) attaining the upper bound derived in this section. Hence, our new test is locally asymptotically UMP also in cases where the nuisance parameters are unknown. Third, the LAN results allow us to show that any test, satisfying a mild regularity condition, has the same, typically nonoptimal, local asymptotic power function under both DGPs.

Remark 3.1. For unit-root problems in (univariate) time series, limit experiment theory has been exploited by, among others, Jansson (2008) and Zhou, Van den Akker, and Werker (2019). That limit experiment is of the Locally Asymptotically Brownian Functional type for which asymptotically UMP tests do not exist. Also, in our case, the central sequence could be written as an (approximate) stochastic integral. However, we obtain an additional sum across panel units. Combined with a CLT-type argument, but now in the more complicated joint (n,T)-convergence case, this sum is the intuition for the Gaussian limits we obtain in this panel setting.

3.1. Expanding the Likelihood in the PANIC Setup

For the PANIC case, we will assume, in this subsection, that the factors F_{kt} are observed. Just as for the nuisance parameters, we show in Section 4 that the resulting likelihood ratio can still be approximated by an observable version (up to a negligible term). This result implies that observing the factors will not lead to an increase in local asymptotic power *for the PANIC framework*. This appears to be a surprising result. Indeed, Moon et al. (2014) derived the power envelope for a first-generation DGP that basically corresponds to PANIC with observed factors. Our analysis implies that, for the PANIC framework, the same power envelope applies. We stress that for the MP setting, the situation is different: Becheri, Drost, and Van den Akker (2015b) report higher powers in cases where factors are observed and Juodis and Westerlund (2019) show power gains when covariates correlated to the innovations are observed.

Denote the joint law of F and Z under Assumptions 2.1–2.3, 2.5(a), and 2.6 by $\mathrm{P}^{\mathrm{PANIC}}_{h,n,T}$. Using $\eta \sim N(0,\Sigma_\eta)$ and $\eta = \Delta E - hE_{-1}/(\sqrt{n}T)$, we obtain the log-likelihood ratio

$$\log \frac{d\mathrm{P}^{\mathrm{PANIC}}_{h,n,T}}{d\mathrm{P}^{\mathrm{PANIC}}_{0,n,T}} = \frac{h}{\sqrt{n}T}\Delta E' \mathcal{A}' \Sigma_\eta^{-1} \Delta E - \frac{h^2}{2nT^2}\Delta E' \mathcal{A}' \Sigma_\eta^{-1} \mathcal{A} \Delta E$$

$$=: h\Delta^{\mathrm{PANIC}}_{n,T} - \frac{1}{2}h^2 J^{\mathrm{PANIC}}_{n,T}.$$

Note, from (6), $\Delta\tilde{E} = \Delta\tilde{Y} - \Delta\tilde{F}\Lambda'$, implying ΔE is indeed observable in this PANIC framework (with observed factors as considered here). Moreover, under $\mathrm{P}^{\mathrm{PANIC}}_{0,n,T}$, $\Delta E = \eta$. We now show that we can replace variances by long-run variances, to obtain a more tractable version of the central sequence and empirical Fisher information.

LEMMA 3.1. *Suppose that Assumptions 2.1–2.4, 2.5(a), and 2.6 hold. Then we have, under* $\mathrm{P}^{PANIC}_{0,n,T}$, $(\Delta^{PANIC}_{n,T}, J^{PANIC}_{n,T}) = (\Delta_{n,T}, \frac{1}{2}) + o_p(1)$, *where*

$$\Delta_{n,T} = \frac{1}{\sqrt{n}T}\Delta E' \mathcal{A}' \Psi_\eta^{-1} \Delta E - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\delta_{\eta,i,T}}{\omega^2_{\eta,i,T}}, \text{ with } \Psi_\eta^{-1} = \Omega_\eta^{-1} \otimes I_T.$$

Remark 3.2. The simplified central sequence $\Delta_{n,T}$ is the result of substituting Σ_η^{-1} by Ψ_η^{-1}. To obtain the correct centering, a correction term involving the one-sided long-run variance is needed for each panel unit. This is analogous to the univariate case (see Elliott, Rothenberg, and Stock, 1996) and arises due to the fact that, contrary to $\Sigma_\eta^{-1/2}\Delta E$, $\Psi_\eta^{-1/2}\Delta E$ exhibits serial correlation.

3.2. Expanding the Likelihood in the Moon and Perron (2004) Setup

Let us denote the law of Z under Assumptions 2.1–2.3, 2.5(b), and 2.6 by $\mathrm{P}^{\mathrm{MP}}_{h,n,T}$. Then the log-likelihood ratio of $\mathrm{P}^{\mathrm{MP}}_{h,n,T}$ with respect to $\mathrm{P}^{\mathrm{MP}}_{0,n,T}$ is given by, using $\varepsilon \sim N(0,\Sigma_\varepsilon)$ and $\varepsilon = \Delta Y - hY_{-1}/(\sqrt{n}T)$,

$$\log \frac{d\mathrm{P}^{\mathrm{MP}}_{h,n,T}}{d\mathrm{P}^{\mathrm{MP}}_{0,n,T}} = \frac{h}{\sqrt{n}T}\Delta Y' \mathcal{A}' \Sigma_\varepsilon^{-1} \Delta Y - \frac{h^2}{2nT^2}\Delta Y' \mathcal{A}' \Sigma_\varepsilon^{-1} \mathcal{A} \Delta Y$$

$$=: h\Delta^{\mathrm{MP}}_{n,T} - \frac{1}{2}h^2 J^{\mathrm{MP}}_{n,T}.$$

In this more complicated model, we simplify the central sequence and also the Fisher information in two steps. The first is analogous to the approximation in the PANIC setup, i.e., we replace variances by long-run variances. Note that thanks to our independence assumptions, the $nT \times nT$ covariance matrix of the ε can be written as

$$\Sigma_\varepsilon = \text{var}\, \varepsilon = \sum_{k=1}^{K} \left(\lambda_k \lambda_k' \otimes \Sigma_{f,k}\right) + \Sigma_\eta. \tag{8}$$

Replacing $\Sigma_{f,k}$ by $\omega^2_{f,k,T} I_T$ and $\Sigma_{\eta,i}$ by $\omega^2_{\eta,i,T} I_T$ in (8), we obtain the simplified versions of central sequence

$$\tilde{\Delta}^{\text{MP}}_{n,T} := \frac{1}{\sqrt{nT}} \Delta Y' \mathcal{A}' \Psi_\varepsilon^{-1} \Delta Y - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\delta_{\eta,i,T}}{\omega^2_{\eta,i,T}},$$

where the $nT \times nT$ matrix Ψ_ε is defined by

$$\Psi_\varepsilon := \psi_\varepsilon \otimes I_T := \left(\Lambda \Omega_F \Lambda' + \Omega_\eta\right) \otimes I_T, \tag{9}$$

with $\Omega_\eta = \text{diag}(\omega^2_{\eta,1,T}, \ldots, \omega^2_{\eta,n,T})$ and $\Omega_F = \text{diag}(\omega^2_{f,1,T}, \ldots, \omega^2_{f,K,T})$. The following lemma demonstrates that applying these replacements to the central sequence and Fisher information do not affect their asymptotic behavior.

LEMMA 3.2. *Suppose that Assumptions 2.1–2.4, 2.5(b), and 2.6 hold. Then we have, under* $\mathrm{P}^{MP}_{0,n,T}$, $(\Delta^{MP}_{n,T}, J^{MP}_{n,T}) = (\tilde{\Delta}^{MP}_{n,T}, \frac{1}{2}) + o_p(1)$.

Remark 3.3. In the MP case, the covariance matrix that is approximated by long-run variances is not block diagonal. Therefore, contrary to Lemma 3.1, the proof of Lemma 3.2 exploits the assumption that $n/T \to 0$.

Exploiting the Sherman–Morrison–Woodbury formula, we obtain

$$\Psi_\varepsilon^{-1} = \psi_\varepsilon^{-1} \otimes I_T = \left(\Omega_\eta^{-1} - \Omega_\eta^{-1} \Lambda \left(\Omega_F^{-1} + \Lambda' \Omega_\eta^{-1} \Lambda\right)^{-1} \Lambda' \Omega_\eta^{-1}\right) \otimes I_T. \tag{10}$$

Note that removing Ω_F^{-1} from (10) yields a projection matrix corresponding to "projecting out the factors." Thus, basing a central sequence on such a projection matrix would simplify approximating it based on observables by removing the need to estimate Ω_F^{-1} and, more importantly, by ensuring that the factors are projected out. The next lemma shows that using such a projection version ψ_ε^{*-1} of ψ_ε^{-1} in the central sequence does not change its asymptotic behavior.

LEMMA 3.3. *Suppose that Assumptions 2.1–2.4, 2.5(b), and 2.6 hold. Then we have, under* $\mathrm{P}^{MP}_{0,n,T}$, $\tilde{\Delta}^{MP}_{n,T} = \Delta^*_{n,T} + o_p(1)$, *where*

$$\Delta^*_{n,T} = \frac{1}{\sqrt{nT}} \Delta Y' \mathcal{A}' (\psi_\varepsilon^{*-1} \otimes I_T) \Delta Y - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\delta_{\eta,i,T}}{\omega^2_{\eta,i,T}}, \text{ with}$$

$$\psi_\varepsilon^{*-1} = \Omega_\eta^{-1} - \Omega_\eta^{-1} \Lambda \left(\Lambda' \Omega_\eta^{-1} \Lambda\right)^{-1} \Lambda' \Omega_\eta^{-1}. \tag{11}$$

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

3.3. Asymptotic Normality

Having simplified each framework's central sequence and Fisher information separately, we are now ready to show that they are asymptotically equivalent and the central sequences converge to a normal distribution. We begin this section by showing that the central sequence in the MP framework is asymptotically equivalent to the one in the PANIC framework.

LEMMA 3.4. *Suppose that Assumptions 2.1–2.6 hold. Then we have, under* $\mathrm{P}^{PANIC}_{0,n,T}$ *and* $\mathrm{P}^{MP}_{0,n,T}$, $\Delta^*_{n,T} = \Delta_{n,T} + o_P(1)$.

Finally, we consider the weak limit of the central sequence $\Delta_{n,T}$ (and therefore also of $\Delta^*_{n,T}$), showing that both experiments are LAN.

PROPOSITION 3.1. *Suppose that Assumptions 2.1–2.6 hold. Then we have, under* $\mathrm{P}^{PANIC}_{0,n,T}$ *and* $\mathrm{P}^{MP}_{0,n,T}$, $\Delta_{n,T} \xrightarrow{d} N(0,J)$ *with* $J = \frac{1}{2}$.

Remark 3.4. Under the null hypothesis, the model equations of both models coincide. Hence, the additional distributional Assumption 2.5(b) implies that under the null, the MP framework is a special case of the PANIC framework. Therefore, it is sufficient to show the desired convergence for $\mathrm{P}^{\mathrm{PANIC}}_{0,n,T}$. This principle applies to all calculations under the null hypothesis. As the central sequences are equal as well and thanks to the LAN result below, it even extends to many calculations under alternatives, through Le Cam's third lemma.

Proposition 3.1 is an important result as it establishes that the unit-root testing problem in both models is LAN, i.e., it is asymptotically equivalent to testing $h = 0$ against $h < 0$ based on one observation $X \sim N(Jh, J)$. This equivalence prescribes how to perform asymptotically optimal inference and yields the asymptotic local power envelope and the power functions of various test statistics. The asymptotic representation theorem (see, for example, Van der Vaart, 2000, Chap. 9) implies that in our framework no unit-root test can have higher power than the optimal test in the limit experiment. This best test is clearly rejecting for small values of X, leading to a power (for a level-α test) of $\Phi(\Phi^{-1}(\alpha) - J^{1/2}h)$. Thus, with $J = 1/2$, this constitutes the power envelope for our unit-root testing problems:

COROLLARY 3.1. *Suppose that Assumptions 2.1–2.4, 2.5(a), and 2.6 hold. Let* $\phi_{n,T} = \phi_{n,T}(Z_{11}, \ldots, Z_{nT})$ *be a sequence of tests and denote their powers, under* $\mathrm{P}^{PANIC}_{h,n,T}$, *by* $\pi_{n,T}(h)$. *If the sequence* $\phi_{n,T}$ *is asymptotically of level* $\alpha \in (0,1)$, *i.e.,* $\limsup_{n,T\to\infty} \pi_{n,T}(0) \leq \alpha$, *we have, for all* $h \leq 0$,

$$\limsup_{n,T\to\infty} \pi_{n,T}(h) \leq \Phi\left(\Phi^{-1}(\alpha) - \frac{h}{\sqrt{2}}\right).$$

Replacing Assumption 2.5(a) by Assumption 2.5(b), the same bound applies to powers under $\mathrm{P}^{MP}_{h,n,T}$.

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

The above power envelope would be reached by any of our previously introduced central sequences. This always holds in LAN experiments and follows from Le Cam's third lemma (see, for example, Van der Vaart, 2000, Chap. 6). In the next section, we show that we can approximate these central sequences based on observables, yielding a feasible test that attains the asymptotic power envelope. Such a result is nontrivial and requires, roughly speaking, that the information matrix is block-diagonal. Even in simple parametric problems this is an exception, let alone in our current case with $O(n)$ nuisance parameters.

Remark 3.5. Note that the level of the local asymptotic power envelope only depends on the (local) deviation to the unit root. The power loss attributed by Moon and Perron (2004) and Westerlund (2015) to the heteroskedasticity in η_{it} is thus a feature of the test statistics under consideration, rather than of the MP and PANIC models.

4. AN ASYMPTOTICALLY UMP TEST

In the previous section, we derived a testing procedure that reaches the power envelope for the unit-root testing problem. This test, however, is not feasible when the nuisance parameters are unknown. In this section, we demonstrate how to estimate the nuisance parameters to obtain a feasible version that also attains the power envelope. We provide a feasible version of $\Delta^*_{n,T}$, which is motivated by the likelihood ratio in the MP experiment. As (11) projects out the factors, basing our feasible version on $\Delta^*_{n,T}$ instead of $\Delta_{n,T}$ spares us the approximation of the idiosyncratic parts.

Recalling our LAN results in Section 3 and that the central sequences are asymptotically equivalent across the two setups (see Lemma 3.4), it is clear that a feasible version of $\Delta^*_{n,T}$ would be optimal. Therefore, we show that replacing all nuisance parameters with estimates does not change the limiting behavior of $\Delta^*_{n,T}$. Specifically, we need estimates $\hat{\Lambda}$ of the factor loadings, as well as estimates $\hat{\delta}_{\eta,i}$ and $\hat{\omega}^2_{\eta,i}$ of the (one-sided) long-run variances of each idiosyncratic part. The feasible test statistic is then

$$\hat{\Delta}_{n,T} = \frac{1}{\sqrt{n}T}\sum_{t=2}^{T}\sum_{s=2}^{t-1}\Delta Z'_{\cdot,s}\hat{\psi}_\varepsilon^{-1}\Delta Z_{\cdot,t} - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\hat{\delta}_{\eta,i}}{\hat{\omega}^2_{\eta,i}}, \text{ where} \tag{12}$$

$$\hat{\psi}_\varepsilon^{-1} := \hat{\Omega}_\eta^{-1} - \hat{\Omega}_\eta^{-1}\hat{\Lambda}(\hat{\Lambda}'\hat{\Omega}_\eta^{-1}\hat{\Lambda})^{-1}\hat{\Lambda}'\hat{\Omega}_\eta^{-1}. \tag{13}$$

Assumption 4.1. Let $\hat{\delta}_{\eta,i}$, $\hat{\omega}^2_{\eta,i}$, and $\hat{\Lambda}$ be estimators of $\delta_{\eta,i}$, $\omega^2_{\eta,i}$, and Λ satisfying, under $\mathrm{P}^{\mathrm{MP}}_{0,n,T}$ and $\mathrm{P}^{\mathrm{PANIC}}_{0,n,T}$,

1. $\max_{i=1,\dots,n}\mathrm{E}|\hat{\delta}_{\eta,i} - \delta_{\eta,i}|^2 = o(1/n)$,
2. $\max_{i=1,\dots,n}\mathrm{E}|\hat{\omega}^2_{\eta,i} - \omega^2_{\eta,i}|^2 = o(1/n)$, and
3. for a $K \times K$ matrix H_K satisfying $\|H_K\|_F = O_p(1)$ and $\left\|H_K^{-1}\right\|_F = O_p(1)$, we have $\left\|\Lambda H_K - \hat{\Lambda}\right\|_F = o_p(1)$.

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

Under suitable restrictions on the bandwidth and the kernel, conditions Items 1 and 2 hold for kernel spectral density estimates (see Remark 2.9 in Moon et al., 2014). Items 3, on the other hand, is stronger than the results in Moon and Perron (2004), so we show in Lemma 4.1 that it indeed holds under our assumptions.

LEMMA 4.1. *Let* $\bar{\Lambda}$ *be* $\sqrt{n}$ *times the* $n \times K$ *matrix containing the* K *orthonormal eigenvectors corresponding to the* K *largest eigenvalues of* $\frac{\Delta\tilde{Z}'\Delta\tilde{Z}}{nT}$. *Take* $\hat{\Lambda} = \frac{\Delta\tilde{Z}'\Delta\tilde{Z}}{nT}\bar{\Lambda}$. *There exists a* $K \times K$ *matrix* H_K *such that, under* $\mathrm{P}^{MP}_{0,n,T}$ *and* $\mathrm{P}^{PANIC}_{0,n,T}$, $\left\|\Lambda H_K - \hat{\Lambda}\right\|_F = o_p(1)$ *and both* $\|H_K\|_F$ *and* $\left\|H_K^{-1}\right\|_F$ *are* $O_p(1)$.

Remark 4.1. These factor estimates are the same as those used in Moon and Perron (2004) and correspond to factor estimates based on classical principal component analysis. We adapt the proof of Moon and Perron (2004), who have demonstrated $\left\|\Lambda H_K - \hat{\Lambda}\right\|_F = O_p(1)$, but we treat one term differently (see Remark A.2 in the Supplementary Material).

Remark 4.2. The factors are only identified up to a "rotation" H_K. Note that $\Delta^*_{n,T}$ is (indeed) invariant under such rotations, as ψ_ε^{*-1} also equals

$$\Omega_\eta^{-1} - \Omega_\eta^{-1}\Lambda H_K\left(H_K'\Lambda'\Omega_\eta^{-1}\Lambda H_K\right)^{-1}H_K'\Lambda'\Omega_\eta^{-1}.$$

LEMMA 4.2. *Under Assumptions 2.1–2.4, 2.6, and 4.1, we have, under* $\mathrm{P}^{MP}_{0,n,T}$ *and* $\mathrm{P}^{PANIC}_{0,n,T}$, $\hat{\Delta}_{n,T} = \Delta^*_{n,T} + o_p(1)$.

Although Lemma 4.2 only concerns adaptivity under the null hypothesis H_0, we can use Le Cam's First Lemma to obtain that, thanks to contiguity, also under $\mathrm{P}^{\mathrm{MP}}_{h,n,T}$ or $\mathrm{P}^{\mathrm{PANIC}}_{h,n,T}$, $\hat{\Delta}_{n,T}$ has the same limiting distribution as $\Delta^*_{n,T}$, so that tests based on $\hat{\Delta}_{n,T}$ will be UMP. Formally, the size and power properties of our optimal test follow from the following theorem.

THEOREM 4.1. *Let* $t_{UMP} = \sqrt{2}\hat{\Delta}_{n,T}$. *Under Assumptions 2.1–2.6 and 4.1, we have, under* $\mathrm{P}^{MP}_{h,n,T}$ *and* $\mathrm{P}^{PANIC}_{h,n,T}$,

$$t_{UMP} \xrightarrow{d} N\left(\frac{h}{\sqrt{2}}, 1\right).$$

Rejecting H_0 *for* $t_{UMP} \le \Phi^{-1}(\alpha)$, $\alpha \in (0,1)$, *leads to an asymptotic power of* $\Phi\left(\Phi^{-1}(\alpha) - \frac{h}{\sqrt{2}}\right)$, *implying that* t_{UMP} *is asymptotically UMP.*

Remark 4.3. The asymptotic size of our test can also be obtained under much weaker assumptions not exploiting Gaussianity (see Remarks A.1 and A.3 in the Supplementary Material). In such a situation, our test is still valid although perhaps

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

nonoptimal. For optimal inference with non-Gaussian innovations, a new analysis of the likelihood ratio would be needed, but this is not feasible here.

Remark 4.4. Note that the limiting distribution of t_{UMP}, both under the null hypothesis and under local alternatives, does not depend on the autocorrelations or the heterogeneity of the long-run variances. This shows that the decrease in asymptotic power attributed to these features, for example, in Remark 2 of Westerlund (2015), was due to the specific tests under consideration rather than being a feature of the unit-root testing problem.

Remark 4.5. Note that $\hat{\Delta}_{n,T}$ only involves differenced data, so that our test is invariant with respect to the incidental intercepts m_i.

Here is one way to obtain the UMP test in practice:

1. Compute an estimator $\hat{K}$ of the number of common factors on the basis of the observations $\Delta Z_{\cdot t}$, $t = 2, \ldots, T$, using information criteria from Bai and Ng (2002). As $(n, T \to \infty)$, these criteria select the correct number of factors with probability 1. Therefore, we can treat the number of factors as known in our asymptotic analyses.
2. Use the observations $\Delta Z_{\cdot t}$, $t = 2, \ldots, T$, and $\hat{K}$ to determine the factor loadings $\hat{\Lambda}$ and the factor residuals $\hat{\eta}_{\cdot t}$, $t = 2, \ldots, T$, using principal components.
3. Determine estimates $\hat{\omega}^2_{\eta,i}$ of $\omega^2_{\eta,i}$ and $\hat{\delta}_{\eta,i}$ of $\delta_{\eta,i}$ from $\hat{\eta}_{\cdot t}$, $t = 2, \ldots, T$, using kernel spectral density estimates. Let $\hat{\Omega} = \text{diag}(\hat{\omega}^2_{\eta,1}, \ldots, \hat{\omega}^2_{\eta,n})$.
4. Calculate the estimated central sequence $\hat{\Delta}_{n,T}$ as in (12) and reject when $t_{\text{UMP}} = \sqrt{2}\hat{\Delta}_{n,T} \leq \Phi^{-1}(\alpha)$. Alternatively, based on small sample considerations, also estimate the empirical Fisher information

$$\hat{J}_{n,T} := \frac{1}{nT^2} \sum_{t=2}^{T} \sum_{s=2}^{t-1} \Delta Z'_{\cdot,s} \hat{\psi}_{\varepsilon}^{-1} \sum_{u=2}^{t-1} \Delta Z_{\cdot,u},$$

and reject the null hypothesis when $t^{\text{emp}}_{\text{UMP}} := \hat{\Delta}_{n,T} / \sqrt{\hat{J}_{n,T}} \leq \Phi^{-1}(\alpha)$.

Remark 4.6. Although the UMP test t_{UMP} does not require a complicated estimate of the known $J = 1/2$, it can be undersized in small samples, whereas the empirical version $t^{\text{emp}}_{\text{UMP}}$ behaves very well in most DGPs, both in terms of size and power. Thus, we recommend to use the $t^{\text{emp}}_{\text{UMP}}$ in small samples. See Section 6 for details.

5. COMPARING POWERS ACROSS TESTS AND FRAMEWORKS

This section derives the asymptotic powers of commonly used tests in both the Moon and Perron (2004) and the Bai and Ng (2004) frameworks. We start by formalizing our observation that local powers are equal across the two frameworks.

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

COROLLARY 5.1. *Let $t_{n,T}$ be a test statistic that, under $\mathrm{P}^{PANIC}_{0,n,T}$, converges in distribution jointly with $\Delta_{n,T}$. Then, for all $x \in \mathbb{R}$, and all h,*

$$\lim_{(n,T\to\infty)} \mathrm{P}^{MP}_{h,n,T}[t_{n,T} \leq x] = \lim_{(n,T\to\infty)} \mathrm{P}^{PANIC}_{h,n,T}[t_{n,T} \leq x].$$

If, more specifically, $t_{n,T} \overset{\mathrm{P}^{PANIC}_{0,n,T}}{\to} N(0,1)$ and if $t_{n,T}$ and $\Delta_{n,T}$ are jointly asymptotically normal under $\mathrm{P}^{PANIC}_{0,n,T}$ with asymptotic covariance $\sigma_{\Delta,t}$, its limiting distribution under local alternatives is given by

$$t_{n,T} \overset{\mathrm{P}^{PANIC}_{h,n,T}}{\to} N(h\sigma_{\Delta,t},1), \text{ and } t_{n,T} \overset{\mathrm{P}^{MP}_{h,n,T}}{\to} N(h\sigma_{\Delta,t},1).$$

Once again, our result on the asymptotic equivalence of the two experiments allows us to obtain results for both frameworks at the same time. By demonstrating the joint normality under the null as in Corollary 5.1, we obtain simple proofs of the powers of commonly used tests in these frameworks, without ever relying on triangular array calculations. To show the elegance of this approach, we include here the full argument for the first part of this corollary. The second part follows immediately from a more specific version of Le Cam's third lemma, which directly prescribes the desired normal distribution under alternatives. We can use this simple way to obtain powers under local alternatives thanks to our LAN results of Section 3.

Denote the weak limit of $(t_{n,T},\Delta_{n,T})$ under $\mathrm{P}^{\mathrm{MP}}_{0,n,T}$ by (t,Δ). Thanks to our results in Section 3, both $(t_{n,T}, \frac{d\mathrm{P}^{\mathrm{PANIC}}_{h,n,T}}{d\mathrm{P}^{\mathrm{PANIC}}_{0,n,T}})$ and $(t_{n,T}, \frac{d\mathrm{P}^{\mathrm{MP}}_{h,n,T}}{d\mathrm{P}^{\mathrm{MP}}_{0,n,T}})$ converge in distribution to $(t, \exp(h\Delta - h^2/4))$. By a general form of Le Cam's third lemma, the distribution of $t_{n.T}$ under local alternatives only depends on this joint limiting law and is thus equal across the two frameworks (see Theorem 6.6 in Van der Vaart, 2000).

Remark 5.1. The equality of powers across the two frameworks applies to the practically relevant case of the factors being unobserved. In the PANIC setting, observing the factors does not yield any additional power. This is sharp contrast to other DGPs, used in the literature on panel unit roots, where observing factors or correlated covariates does yield additional power (see, for example, Pesaran et al., 2013; Becheri et al., 2015b; Juodis and Westerlund, 2019).

Before we apply Corollary 5.1 to derive asymptotic powers, we first describe the relevant test statistics in some detail. We focus on the tests proposed in Bai and Ng (2010) ("BN tests") and Moon and Perron (2004) ("MP tests"). Following these papers, we denote

$$\omega^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \omega^2_{\eta,i}, \quad \phi^4 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \left(\omega^2_{\eta,i}\right)^2, \quad \delta = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \delta_{\eta,i},$$

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

all assumed to be positive, and their estimated counterparts

$$\hat{\omega}^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{\omega}^2_{\eta,i}, \quad \hat{\phi}^4 = \frac{1}{n}\sum_{i=1}^{n} \left(\hat{\omega}^2_{\eta,i}\right)^2, \text{ and } \hat{\delta} = \frac{1}{n}\sum_{i=1}^{n} \hat{\delta}_{\eta,i}.$$

Finally, we define $\omega^4 = (\omega^2)^2$ and $\hat{\omega}^4 = (\hat{\omega}^2)^2$.

Both the MP and BN tests rely on a two-stage procedure. In the first stage, the unobserved idiosyncratic innovations E are estimated. Subsequently, a pooled regression procedure is used to estimate the (pooled) autoregression parameter. This pooled estimator is then used to construct a t-test. The main difference between the MP and the BN procedures lies in the way the idiosyncratic innovations are estimated.

Bai and Ng (2010) propose to estimate the idiosyncratic errors E by the PANIC approach introduced in Bai and Ng (2004), which in turn relies on principal component analysis applied to the differences ΔY_{it}. Denoting this estimator of E_i by $\hat{E}_i$, the BN tests are

$$P_a = \frac{\sqrt{n}T(\hat{\rho}^+ - 1)}{\sqrt{2\hat{\phi}^4/\hat{\omega}^4}} \text{ and}$$

$$P_b = \sqrt{n}T(\hat{\rho}^+ - 1)\sqrt{\frac{1}{nT^2}\sum_{i=1}^{n} \hat{E}'_{-1,i}\hat{E}_{-1,i}\frac{\hat{\omega}^2}{\hat{\phi}^4}}, \text{ where}$$

$$\hat{\rho}^+ = \frac{\sum_{i=1}^{n} \hat{E}'_{-1,i}\hat{E}_i - nT\hat{\delta}}{\sum_{i=1}^{n} \hat{E}'_{-1,i}\hat{E}_{-1,i}}$$

is a bias-corrected pooled estimator for the autoregressive coefficients.

Remark 5.2. Recall that $t^{\text{emp}}_{\text{UMP}}$ is a modification of t_{UMP} that replaces the asymptotic Fisher Information $J = 1/2$, with its finite-sample equivalent in the MP setup, $\tilde{J}^{\text{MP}}_{n,T}$. The resulting statistics can be considered a version of P_b: In the case of homogeneous long-run variances, inserting the true long-run variances into $t^{\text{emp}}_{\text{UMP}}$ yields P_b. Conversely, $t^{\text{emp}}_{\text{UMP}}$ is a version of P_b that takes into account the heterogeneity in the long-run variances.

The MP tests are based on a different estimator of ρ. The idiosyncratic components E_i are estimated by projecting the data on the space orthogonal to the common factors. Let $\hat{\Lambda}$ be a consistent estimators for Λ as defined in Moon and Perron (2004, pp. 89–90), and $Y_{\cdot,t} = (Y_{1t}, \ldots, Y_{nt})'$. Then the MP test statistics are given by

$$t_a = \frac{\sqrt{n}T(\rho^+_{\text{pool}} - 1)}{\sqrt{2\hat{\phi}^4/\hat{\omega}^4}}, \text{ and}$$

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

$$t_b = \sqrt{n}T(\rho^+_{\text{pool}} - 1)\sqrt{\frac{1}{nT^2}\sum_{t=1}^{T} Y'_{\cdot,t-1} Q_{\hat{\gamma}} Y_{\cdot,t-1} \frac{\hat{\omega}^2}{\hat{\phi}^4}}, \text{ where}$$

$$\rho^+_{\text{pool}} = \frac{\sum_{t=1}^{T} Y'_{\cdot,t} Q_{\hat{\gamma}} Y_{\cdot,t-1} - nT\hat{\delta}}{\sum_{t=1}^{T} Y'_{\cdot,t-1} Q_{\hat{\gamma}} Y_{\cdot,t-1}}, \text{ and } Q_{\hat{\gamma}} = I - \hat{\Lambda}(\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}'.$$

We are now ready to compute the asymptotic behavior of the MP and BN tests under local alternatives by an application of Corollary 5.1. The power of the MP tests in the MP framework has been derived in Moon and Perron (2004) and that of the BN tests in the PANIC framework has been derived in Westerlund (2015). Given our LAN result, we can provide simple independent proofs of these results. These rely on the second part of Corollary 5.1; we demonstrate the required joint asymptotic normality in the Supplementary Material. More importantly, our approach also leads to new results, namely the asymptotic powers of the MP test in the PANIC framework and the asymptotic powers of the BN tests in the MP framework. In fact, those results can be considered an immediate consequence of the first part of Corollary 5.1 and the existing power results in the literature.

PROPOSITION 5.1. *Suppose that Assumptions 2.1–2.6 and 4.1 hold. Then, under* $\mathrm{P}^{PANIC}_{h,n,T}$ *or* $\mathrm{P}^{MP}_{h,n,T}$*, as* $(n,T\to\infty)$*, the test statistics* P_a, P_b, t_a*, and* t_b *all converge in distribution to a normal distribution with mean* $h\sqrt{\frac{\omega^4}{2\phi^4}}$ *and variance one. Rejecting for small values of any of these statistics leads to an asymptotic power for a level-*α *test of* $\Phi(\Phi^{-1}(\alpha) - h\sqrt{\frac{\omega^4}{2\phi^4}})$ *in both frameworks.*

Remark 5.3. It turns out that the powers are equal, no matter which test statistic and which framework is considered. We have discussed in some detail that, for a given test, the equality of powers across frameworks is a general phenomenon. The fact that in each framework, the power of the MP tests is equal to that of the BN tests, on the other hand, is a "coincidence." Originally, the MP tests have been developed for the MP experiment, whereas the BN tests are designed for the PANIC experiment. It has been noted in Bai and Ng (2010) that the MP tests are valid in terms of size in the PANIC setup for testing the idiosyncratic component of the innovation for a unit root but their (local and asymptotic) power in the PANIC framework has not been considered. More discussion on the use of the MP tests in the PANIC setup can be found in Bai and Ng (2010) and Gengenbach, Palm, and Urbain (2010). Similarly, to the best of our knowledge, there are no studies on the local asymptotic power of the BN tests in the MP framework.

The Cauchy–Schwarz inequality implies $\frac{\omega^4}{\phi^4} \le 1$; thus, Proposition 5.1 shows that, in general, the local asymptotic power of the MP and BN tests lies below the power envelope. In fact, they are all asymptotically UMP only when $\frac{\omega^4}{\phi^4} = 1$. This condition is satisfied when the long-run variances of the idiosyncratic shocks η_{it} are

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

homogeneous across i. The proposed test t_{UMP} is asymptotically UMP irrespective of possible heterogeneity. In Section 6, we assess whether the asymptotic power gains, compared to the MP and BN tests, are also reflected in finite samples for realistic parametric settings.

Remark 5.4. It should be noted that the results in this section depend on the localizing rate that is used. As stated in Assumption 2.6, this paper uses the contiguity rate $T^{-1}n^{-1/2}$ as localizing rate and obtain local asymptotic powers at the distance $hT^{-1}n^{-1/2}$. For h large the local power is close to one. For other localizing rates, alternative approaches are needed. For example, Phillips and Magdalinos (2007) have proposed, for univariate time series with a (near) unit root, to use different localizing rates—representing moderate deviations from a unit root—in order to get insight into the discontinuities between stationary, unit-root, and explosive autoregressions. Yamamoto and Horie (2023) have used such moderately local to unity rates in order to analyze the power of right-tailed versions of the PANIC tests when the common and/or the idiosyncratic components are moderately explosive.

6. SIMULATION RESULTS

This section reports the results of a Monte Carlo study with three main goals: first, to assess the finite-sample performance of our proposed test t_{UMP}; second, to see how the asymptotic equivalence between the Moon and Perron (2004) and PANIC setups is reflected in finite samples; and, finally, to check the robustness of our results to deviations from our assumptions.

6.1. The DGPs

We generate the data from Equations (1)–(4) with $m_i = 0$. Recall that our tests are invariant with respect to m_i. Using sample sizes $n = 25, 50, 100$ and $T = n, 2n, 4n$, we simulate both the MP and the PANIC setups. Recall that, for a local alternative h, we take $\rho = 1 + \frac{h}{\sqrt{n}T}$ in both setups. In the MP case, we also set $\rho_k = \rho$, whereas in the PANIC case, we set $\rho_k = 1$ under the null and all alternatives. The factor loadings Λ are drawn from a normal distribution with mean $K^{-1/2}$ and covariance matrix $K^{-1}I_K$. As done in Moon and Perron (2004), we scale by $\sqrt{K}$ to ensure the contribution of the factors is comparable across specifications. Most of the simulations are run with $K = 1$, but we also explore what happens with more factors. Throughout this section, we assume the number of factors to be known. This number can be estimated consistently, so this makes no difference for the asymptotic analysis. See, for example, Section 2.3 in Moon and Perron (2004) and Section 5 in Bai and Ng (2010) for a discussion of this issue. For the innovation processes f_{kt} and η_{it}, we examine Gaussian i.i.d., MA(1), and AR(1) processes. We fix the MA or AR parameter at 0.4 and set the variance such that the long-run variances of the f_{kt} equal one, and the long-run variance of the η_{it} is ω_i^2. The ω_i^2 are drawn i.i.d. from a lognormal distribution whose parameters are chosen to

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

match different values of ω^4/ϕ^4 and a mean of one. Recall from Section 5 that the asymptotic relative efficiency of the existing tests compared to our UMP test depends on the heterogeneity of the long-run variances and more specifically on the ratio ω^4/ϕ^4. Therefore, the sample size at which it becomes worthwhile to estimate the heterogeneous long-run variances (i.e., use the asymptotically UMP tests suggested here) mainly depends on this ratio. We present simulation results for $\sqrt{\omega^4/\phi^4}$ between 0.6 and 1, where lower values indicate more heterogeneity. A cursory look at a few typical applications reveals that these ratios are mostly between 0.6 and 0.8 and match the skewed nature of the lognormal distribution.

6.1.1. *The Test Statistics.* In addition to the tests proposed in Section 4, t_{UMP} and $t_{\text{UMP}}^{\text{emp}}$, we consider the MP tests of Moon and Perron (2004) and the BN tests of Bai and Ng (2010). However, the powers and sizes of the (MP) t_b and (BN) P_b tests were very similar also in finite samples, so we only report results for P_b. We omit the comparison with P_a and t_a since they tend to show large biases in terms of size (see, for example, the Monte Carlo studies in Gengenbach et al. (2010) and Bai and Ng (2010)).

The sizes of all considered tests are highly sensitive to estimation of the (one-sided) long-run variances. We have considered a variety of methods, for example, using a Bartlett or quadratic spectral kernel and selection of the bandwidth according to the Newey and West (1994) or the Andrews (1991) rule with/without various forms of prewhitening. Whereas the differences from using different kernels are small, the selection of both the bandwidth and the prewhitening are essential. Our preferred method employs a Bartlett kernel with prewhitening. As in Moon et al. (2014), the prewhitening model is selected based on the BIC between four simple ARMA models. There is a size-power tradeoff between using the Andrews (1991) and the Newey and West (1994) bandwidth selection: the Andrews (1991) bandwidth leads to higher powers for the smallest sample sizes, but an oversized test when the innovations have a strong MA component. The decision about which bandwidth to use thus depends on the preferences of the researcher. In this section, all results are based on the Andrews (1991) bandwidth. However, the sizes and powers based on the Newey and West (1994) bandwidth can be found in the Supplementary Material.

6.2. Sizes

Table 1 reports the sizes, for the setting $\sqrt{\omega^4/\phi^4} = 0.8$, of our tests for the baseline DGP based on the Andrews bandwidth. Table B.1 in the Supplementary Material contains the sizes for $\sqrt{\omega^4/\phi^4} \in \{0.6, 1\}$. The Supplementary Material also considers many other specifications. Recall that the sizes depend considerably on how the long-run variances are estimated. Using the method described above, the sizes of $t_{\text{UMP}}^{\text{emp}}$ are reasonable across most DGPs and generally comparable to those of P_b. t_{UMP}, on the other hand, is undersized in many specifications, so that we focus on its empirical version $t_{\text{UMP}}^{\text{emp}}$ in the remainder. Only in the MA(1)

TABLE 1. Sizes (in percent) of nominal 5% level tests with no heterogeneity in the alternatives and $\sqrt{\omega^4/\phi^4} = 0.8$. Based on 1,000,000 replications. Andrews Bandwidth.

		i.i.d.			AR(1)			MA(1)		
n	T	t_{UMP}	$t_{\text{UMP}}^{\text{emp}}$	P_b	t_{UMP}	$t_{\text{UMP}}^{\text{emp}}$	P_b	t_{UMP}	$t_{\text{UMP}}^{\text{emp}}$	P_b
25	25	0.9	3.1	3.5	1.8	4.3	4.7	2.4	6.7	6.4
25	50	1.8	5.1	4.6	1.7	4.4	4.0	3.2	8.3	7.2
25	100	2.3	5.8	5.2	2.2	5.3	4.6	3.9	9.3	7.8
50	50	2.4	4.6	4.2	2.4	4.2	4.2	5.1	9.3	8.3
50	100	3.0	5.4	4.8	2.6	4.6	4.3	5.9	10.1	8.5
50	200	3.3	5.7	5.2	3.1	5.2	4.7	5.0	8.4	7.1
100	100	3.5	5.1	4.6	3.1	4.4	4.4	8.7	12.3	10.4
100	200	3.8	5.5	5.0	3.3	4.7	4.5	6.6	9.2	7.9
100	400	3.9	5.5	5.1	3.9	5.5	5.0	4.7	6.6	5.9

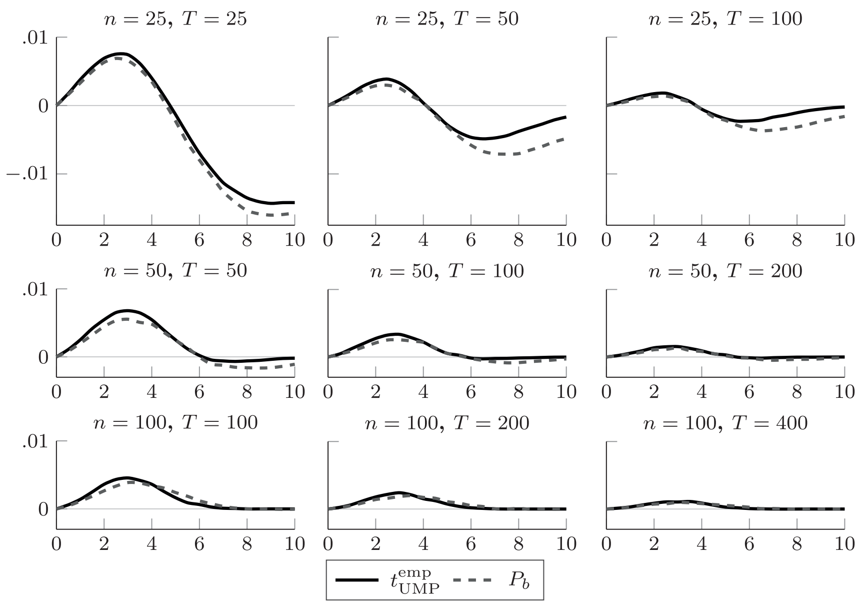

FIGURE 1. Difference between powers in the MP vs. the PANIC framework as a function of $-h$ with i.i.d. factor innovations and i.i.d. idiosyncratic parts and $\sqrt{\omega^4/\phi^4} = 0.8$. Based on 1,000,000 replications.

example, both $t_{\text{UMP}}^{\text{emp}}$ and P_b are oversized ($t_{\text{UMP}}^{\text{emp}}$ is more oversized for the smallest sample sizes and marginally less oversized in the larger ones). Thus, when a strong MA component is suspected, we recommend to use tests based on the Newey and West (1994) bandwidth. Generally, the Newey and West (1994) bandwidth

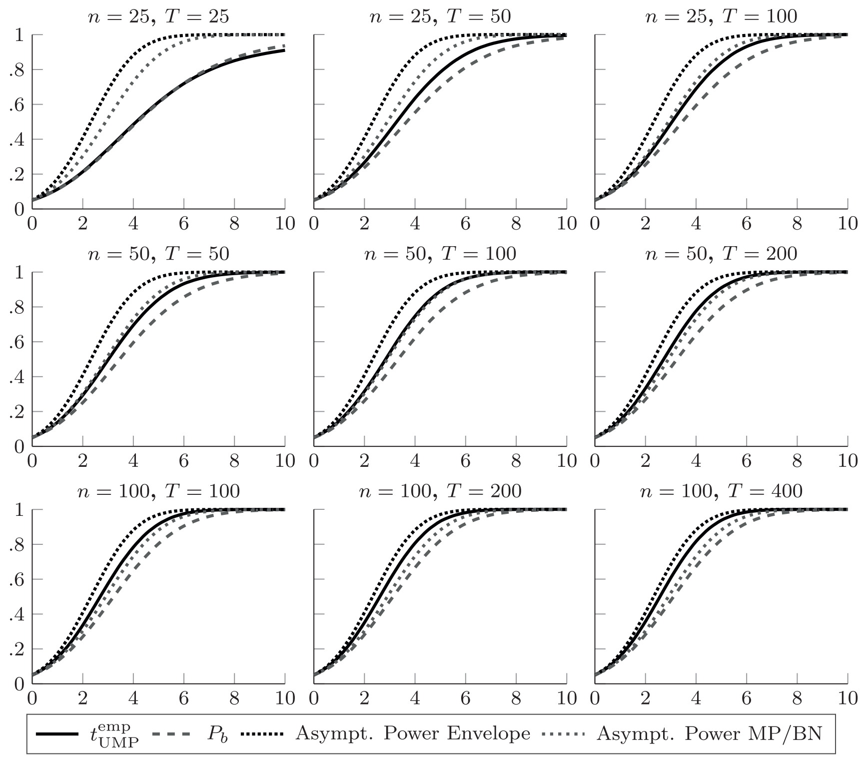

FIGURE 2. Size-corrected power of unit-root tests as a function of $-h$ for varying sample sizes in the PANIC framework with i.i.d. factor innovations and i.i.d. idiosyncratic parts and $\sqrt{\omega^4/\phi^4} = 0.8$. Based on 100,000 replications.

provides better sizes, especially in the MA case. However, small sample powers are slightly lower. Both sizes and powers based on the Newey and West (1994) bandwidth can be found in the Supplementary Material.

6.3. Powers

We start this subsection by investigating the finite-sample differences between the MP and the PANIC setups. Recall that we have shown that the asymptotic, local power functions are the same and that (under some regularity conditions) all tests have the same asymptotic power in the MP framework as they do in the PANIC framework. Figure 1 compares the powers of $t^{\text{emp}}_{\text{UMP}}$ and P_b across the two frameworks. Indeed, also in small samples, the powers are very similar. Moreover, both a larger n and a larger T contribute to reduce the difference. When the factor is stationary under the alternative hypothesis, the difference is considerably smaller still. Noting the small scale on the y-axis in these plots, in the remainder, we will

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

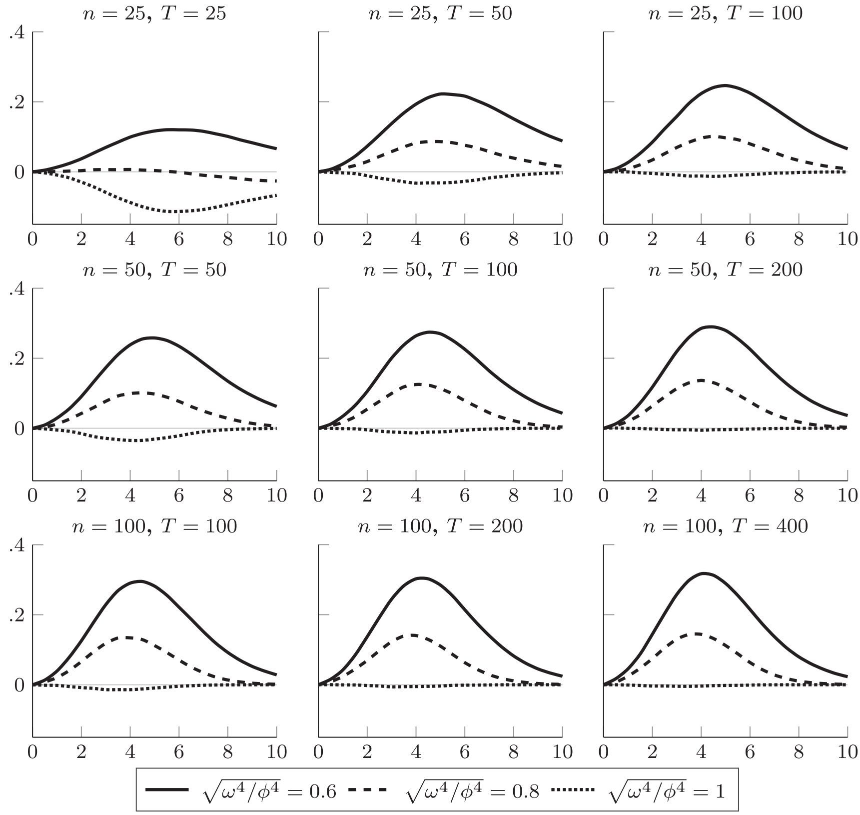

FIGURE 3. (Size-corrected) power gains from using $t_{\text{UMP}}^{\text{emp}}$ over P_b for varying values of $\sqrt{\omega^4/\phi^4}$ and sample sizes in the PANIC framework with i.i.d. factor innovations and i.i.d. idiosyncratic parts. Based on 100,000 replications.

only present results for the PANIC framework, as the lines would otherwise be mostly indistinguishable.

We now turn to comparing the performance of the UMP tests to existing ones. As discussed in Section 4, we need to estimate the individual long-run variance of each idiosyncratic part in order to attain the power envelope. Of course, this becomes easier with a larger time series dimension and is more beneficial when the long-run variances differ substantially between series.

Figure 2 presents the baseline power results for a medium amount of heterogeneity ($\sqrt{\omega^4/\phi^4} = 0.8$). It is evident that even for relatively small samples using the optimal test pays off: except for $n = T = 25$, the power of $t_{\text{UMP}}^{\text{emp}}$ is uniformly higher than that of P_b.

Next, Figure 3 presents the power difference between the optimal test and P_b for varying degrees of heterogeneity. As expected, the higher the amount of

https://doi.org/10.1017/S0266466624000112 Published online by Cambridge University Press

heterogeneity, the more beneficial it is to use the optimal test, also in finite samples. In the case of perfect homogeneity, the losses from estimating individual long-run variances are minor, except for the $n = T = 25$ case.

In the Supplementary Material, we investigate the effects of serial correlation and multiple factors. Qualitatively, the power results are not affected by these variations in the DGP. We also consider the robustness of our results to deviations of our assumptions: we consider the power against heterogeneous alternatives and investigate the effects of non-Gaussian innovations.

7. CONCLUSION AND DISCUSSION

This paper shows that restricted versions of the MP and PANIC frameworks are equivalent, for unit-root testing, from a local and asymptotic point of view. Using the underlying LAN result, the local asymptotic power envelope for the MP and PANIC frameworks readily follows. We show that the tests proposed in Moon and Perron (2004) and Bai and Ng (2010) only attain this bound in cases where the long-run variances of the idiosyncratic component are sufficiently homogeneous. We develop an asymptotically UMP test; a Monte Carlo study demonstrates that this test also improves on existing tests for finite samples.

To obtain the local and asymptotic equivalence of the MP and PANIC frameworks, we need to impose some restrictions. First, we assume that the driving innovations are Gaussian. Second, we do not allow for (incidental) trends. And third, we impose the deviations to the unit root, under the alternative hypothesis, to be the same for all panel units.

The Gaussianity facilitates a relatively easy proof of the LAN result and it seems to be rather difficult to generalize this assumption. Indeed, only recently a semiparametric analysis has been conducted for a first-generation framework (see Van den Akker, Werker, and Zhou, 2023). For the proposed asymptotically UMP test, we stress that Gaussianity is not required for its validity.

To allow for incidental trends, the proper strategy seems to be to first determine the maximal invariant (i.e., determine which transformation of the observations is invariant with respect to transformations of the form $Y_{it} \mapsto a_i + b_i t + Y_{it}$). The next step is to determine the likelihood of this maximal invariant and to analyze if a LAN type of expansion holds true (at conjectured localizing rate $n^{-1/4}T^{-1}$ in line with Moon et al. (2007)). As this expansion will be different from the one in this paper, this steps needs its own proofs. If a LAN expansion can indeed be obtained, similar steps as in this paper are expected to provide asymptotically optimal tests for the unit-root hypothesis.

In view of Becheri et al. (2015a), we do not expect that imposing constant deviations to the unit root, under the alternative hypothesis, affects our main results. The Monte Carlo results seem to confirm this conjecture for finite samples. Here, we give an outline to obtain the same limiting experiment in cases where we use, in (1)–(4), instead of ρ the following local parameterization with heterogeneous alternatives:

$$\rho_i = \rho_i^{(n,T)} = 1 + \frac{h}{\sqrt{n}T} U_i,$$

where $U_1, \ldots, U_n$ are i.i.d. with mean 1. The log-likelihood ratio in this extended experiment, where the factors F and the perturbations U are also observed, is given by

$$\frac{h}{\sqrt{n}} \sum_{i=1}^{n} U_i \frac{\Delta E_i' A' \Sigma_{\eta,i}^{-1} \Delta E_i}{T} - \frac{h^2}{n} \sum_{i=1}^{n} U_i^2 \frac{\Delta E_i' A' \Sigma_{\eta,i}^{-1} \Delta E_i}{T^2}$$

$$\equiv \frac{h}{\sqrt{n}} \sum_{i=1}^{n} U_i X_{ni} - \frac{h^2}{n} \sum_{i=1}^{n} U_i^2 J_{ni}. \qquad (*)$$

So, the log-likelihood ratio expansion for the setting with observed heterogeneity has exactly the same structure as in Theorem 3.1 in Becheri et al. (2015a). With $U_i = 1$, we just get the expression on page 9. Using the derivations in the proofs of Lemma 3.1 and Lemma A.5 in the Supplementary Material, one readily verifies $\frac{1}{n}\sum_{i=1}^{n} X_{ni}^2 \xrightarrow{P} \frac{1}{2}$ and $\frac{1}{n}\sum_{i=1}^{n} J_{ni} \xrightarrow{P} \frac{1}{2}$. Assume additionally, $\max_{i=1,\ldots,n} \frac{X_{ni}}{\sqrt{n}} = o_P(1)$ and $\max_{i=1,\ldots,n} \frac{J_{ni}}{n} = o_P(1)$ (CLT and WLLN results are still applicable for the various panel units, so the panel units are not too different). Finally, assume the existence of the moment generating function of the U_i's. Then, according to Theorem 3.1 of Becheri et al. (2015a), the log-likelihood ratio in the extended experiment, where the factors F are still observed but the heterogeneity due to the perturbations U are unknown, is given by, up to a $o_p(1)$-term, to $(*)$ with U_i replaced by 1. Hence, we obtain the same limit experiment, and hence power envelope, as in the homogeneous setting.

This paper, in line with the setup in Moon and Perron (2004), Moon et al. (2007, 2014), and Bai and Ng (2010), focuses on "large n and T." In practice, however, one commonly encounters relatively short time series and one could expect that asymptotic results based on T fixed and $n \to \infty$ provide better approximations to finite-sample distributions. An extension of the Nickell bias for cross-sectionally dependent panels with a unit root with "T small, n large" is discussed in Phillips and Sul (2003); Phillips and Sul (2007) developed for this setting, a bias-corrected least-squares estimator using mean unbiased functions. It is still an interesting open question, outside the scope of this paper and left for further research, whether it is possible to use the proof lines of this paper in deriving optimality of (existing) unit-root tests for this setting with T fixed and $n \to \infty$.

SUPPLEMENTARY MATERIAL

Wichert, O., I.G. Becheri, F.C. Drost, and R. Van den Akker (2024): Supplement to "Asymptotically uniformly most powerful tests for unit roots in Gaussian panels with cross-sectional dependence generated by common factors," Econometric Theory Supplementary Material. To view, please visit: https://doi.org/10.1017/S0266466624000112

REFERENCES

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59, 817–858.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70, 191–221.

Bai, J., & Ng, S. (2004). A PANIC attack on unit roots and cointegration. *Econometrica*, 72, 1127–1177.

Bai, J., & Ng, S. (2010). Panel unit root tests with cross-section dependence: A further investigation. *Econometric Theory*, 26, 1088–1114.

Baltagi, B. H., & Kao, C. (2000). Nonstationary panels, cointegration in panels and dynamic panels: A survey. In B. H. Baltagi, T. B. Fomby, & R. C. Hill (Eds.), *Nonstationary panels, panel cointegration, and dynamic panels* (pp. 7–51). Emerald.

Banerjee, A. (1999). Panel data unit roots and cointegration: An overview. *Oxford Bulletin of Economics and Statistics*, 61, 607–629.

Becheri, I. G., Drost, F. C., & Van den Akker, R. (2015a). Asymptotically UMP panel unit root tests—The effect of heterogeneity in the alternatives. *Econometric Theory*, 31, 539–559.

Becheri, I. G., Drost, F. C., & Van den Akker, R. (2015b). Unit root tests for cross-sectionally dependent panels: The influence of observed factors. *Journal of Statistical Planning and Inference*, 160, 11–22.

Breitung, J., & Pesaran, M. H. (2008). Unit roots and cointegration in panels. In L. Mátyás, & P. Sevestre (Eds.), *The econometrics of panel data* (pp. 279–322). Springer.

Brockwell, P. J., & Davis, R. A. (1991). *Time series: Theory and methods*. Springer Series in Statistics. Springer.

Choi, I. (2006). Nonstationary panels. In H. Hassani, T. C. Mills, & K. Patterson (Eds.), *Palgrave handbook of econometrics* (pp. 511–539). Palgrave Macmillan.

Choi, I. (2015). *Almost all about unit roots: Foundations, developments, and applications*. Cambridge University Press.

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

Gengenbach, C., Palm, F. C., & Urbain, J. P. (2010). Panel unit root tests in the presence of cross-sectional dependencies: Comparison and implications for modelling. *Econometric Reviews*, 29, 111–145.

Gutierrez, L. (2006). Panel unit-root tests for cross-sectionally correlated panels: A Monte Carlo comparison. *Oxford Bulletin of Economics and Statistics*, 68, 519–540.

Jansson, M. (2008). Semiparametric power envelopes for tests of the unit root hypothesis. *Econometrica*, 76, 1103–1142.

Juodis, A., & Westerlund, J. (2019). Optimal panel unit root testing with covariates. *Econometrics Journal*, 22, 57–72.

Moon, H. R., Perron, B., & Phillips, P. C. B. (2007). Incidental trends and the power of panel unit root tests. *Journal of Econometrics*, 141, 416–459.

Moon, H. R., Perron, B., & Phillips, P. C. B. (2014). Point-optimal panel unit root tests with serially correlated errors. *Econometrics Journal*, 17, 338–372.

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

Müller, U. K., & Elliott, G. (2003). Tests for unit roots and the initial condition. *Econometrica*, 71, 1269–1286.

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61, 631–653.

O'Connell, P. (1998). The overvaluation of purchasing power parity. *Journal of International Economics*, 44, 1–19.

Patterson, K. (2011). *Unit root tests in time series, Volume 1: Key concepts and problems*. Palgrave Macmillan.

Patterson, K. (2012). *Unit root tests in time series, Volume 2: Extensions and developments*. Palgrave Macmillan.

Pesaran, M. H., Smith, L. V., & Yamagata, T. (2013). Panel unit root tests in the presence of a multifactor error structure. *Journal of Econometrics*, 175, 94–115.

Phillips, P. C. B., & Magdalinos, T. (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics*, 136, 115–130.

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

Phillips, P. C. B., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, 6, 217–259.

Phillips, P. C. B., & Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *Journal of Econometrics*, 137, 162–188.

Van den Akker, R., Werker, B. J. M., & Zhou, B. (2023). *Hybrid rank-based panel unit root tests*. Forthcoming in: Festschrift in honour of Marc Hallin. Available at SSRN: https://ssrn.com/abstract=4613034

Van der Vaart, A. W. (2000). *Asymptotic statistics*. Cambridge University Press.

Westerlund, J. (2015). The power of PANIC. *Journal of Econometrics*, 185, 495–509.

Westerlund, J., & Breitung, J. (2013). Lessons from a decade of IPS and LLC. *Econometric Reviews*, 32, 547–591.

Yamamoto, Y., & Horie, T. (2023). A cross-sectional method for right-tailed panic tests under a moderately local to unity framework. *Econometric Theory*, 39, 389–411.

Zhou, B., Van den Akker, R., & Werker, B. J. M. (2019). Semiparametrically optimal hybrid rank tests for unit roots. *Annals of Statistics*, 47, 2601–2638.