Econometric Theory, **24**, 2008, 88–108. Printed in the United States of America.
DOI: 10.1017/S0266466608080067

TESTING FOR UNIT ROOTS IN PANELS WITH A FACTOR STRUCTURE

JÖRG BREITUNG
University of Bonn

SAMARJIT DAS
Indian Statistical Institute

This paper considers various tests of the unit root hypothesis in panels where the cross-section dependence is due to common dynamic factors. Three situations are studied. First, the common factors and idiosyncratic components may both be nonstationary. In this case test statistics based on generalized least squares (GLS) possess a standard normal limiting distribution, whereas test statistics based on ordinary least squares (OLS) are invalid. Second, if the common component is $I(1)$ and the idiosyncratic component is stationary (the case of cross-unit cointegration), then both the OLS and the GLS statistics fail. Finally, if the idiosyncratic components are $I(1)$ but the common factors are stationary, then the OLS-based test statistics are severely biased, whereas the GLS-based test statistics are asymptotically valid in this situation. A Monte Carlo study is conducted to verify the asymptotic results.

1. INTRODUCTION

Panel unit root tests are proposed to improve the power of standard univariate unit root tests. However, an important problem with the application of panel data methods to regional data is that the data usually exhibit substantial cross-sectional dependence. In recent years, different approaches have been suggested to cope with such situations. Chang (2002, 2004) and Harvey and Bates (2003) assume that the errors are cross-correlated with an arbitrary covariance matrix and present the limiting distribution as the number of time periods T tends to infinity and the number of cross-section units N is fixed. Breitung and Das (2005) assume weak error dependence that can be characterized by the fact that all eigenvalues of the error covariance matrix are bounded as the number of cross-section units tend to infinity. Using a sequential limit theory where $T \to \infty$ is followed by $N \to \infty$, they show that the ordinary least squares (OLS)

The research for this paper was carried out within research project "Unit roots and cointegration in panel data" financed by the German Research Association (DFG). We thank Paulo Rodrigues and two anonymous referees for helpful comments and suggestions. Address correspondence to Jörg Breitung, University of Bonn, Institute of Econometrics, Adenauerallee 24-42, 53113 Bonn, Germany; e-mail: breitung@uni-bonn.de.

 © 2008 Cambridge University Press 0266-4666/08 $15.00

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

t-statistic with "panel corrected standard errors" (PCSE) (Jönsson, 2005) and the generalized least squares (GLS) *t*-statistic possess a standard normal limit distribution.

Choi (2002), Phillips and Sul (2003), Bai and Ng (2004), Moon and Perron (2004), and Pesaran (2005) consider a strong form of dependence that is due to common factors. In this case the largest eigenvalue of the covariance matrix tends to infinity as the cross-section dimension (N) increases.[1]

Recently, Banerjee, Marcellino, and Osbat (2005) have studied the performance of so-called first generation panel data unit root tests that ignore the cross-section dependence among the panel units. Their findings suggest that in this situation standard panel unit root tests suffer from severe size distortions. In this paper, we confirm their findings by employing a joint limit theory to analyze the asymptotic properties of first generation panel unit root tests. Moreover, we also consider second generation panel unit root tests suggested by Harvey and Bates (2003), Chang (2004), Moon and Perron (2004), and Breitung and Das (2005). In contrast to the approach suggested by Bai and Ng (2004), these test procedures are based on the "reduced form representation" of the factor model. It is well known (e.g., Harvey, 1989) that unobserved component models (the structural form) possess an ARIMA(p,d,q) representation (the reduced form). The Panel Analysis of Nonstationarity in Idiosyncratic and Common Components (PANIC) approach suggested by Bai and Ng (2004) is based on two steps. First, the structural factor model is estimated by principal components. In a second step, the estimated factors and idiosyncratic components are tested by using cointegration and panel unit root tests. If the tests indicate that the factors and idiosyncratic components are stationary, it is concluded that at least some time series in the panel are stationary. If one or both tests accept the null hypothesis, the joint hypothesis that all time series are $I(1)$ cannot be rejected. Using Monte Carlo simulations, Gengenbach, Palm, and Urbain (2005) investigate the small-sample properties of the test procedure proposed by Bai and Ng (2004) and the second generation panel unit root tests of Moon and Perron (2004) and Pesaran (2005). They observe that the latter tests tend to (wrongly) indicate stationary series if the common factor is $I(1)$ and the idiosyncratic component is $I(0)$. In this paper we explain this finding by resorting to asymptotic theory.

In our theoretical analysis we study four panel unit root tests in situations where the cross-section dependence is due to common factors. The OLS *t*-statistic is a first generation panel unit root test that simply ignores possible cross-section dependence. The robust OLS *t*-statistic suggested by Breitung and Das (2005) and Jönsson (2005) was found to perform well in the case of weak cross-section dependence. In this paper we show that in the presence of common factors the test has a nonstandard limiting distribution. We also study test statistics based on suitable transformations of the variables. Two different approaches are considered. First, the inverse of the estimated residual covariance matrix is used to compute the GLS *t*-statistic. The second approach

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

employs a transformation of the data that eliminates the common factors. Such test procedures are suggested by Phillips and Sul (2003) and Moon and Perron (2004).[2]

These test statistics are used to test the null hypothesis that all time series in the panel are $I(1)$. This null hypothesis may be the result of three different situations. In case A we assume that both the common and idiosyncratic components are nonstationary. In case B it is assumed that the common factors are $I(1)$ and the idiosyncratic components are $I(0)$, i.e., the time series are cointegrated across the panel units, whereas in case C the common factors are $I(0)$ and the idiosyncratic components are $I(1)$. In this paper we study the asymptotic properties of alternative panel unit root tests in these three situations. We show that no test procedure is able to cope with case B, where the series are cross-cointegrated. In the other cases at least some of the tests are valid.

The rest of the paper is organized as follows. In Section 2, the details of the panel data setup are presented, and in Section 3 the test statistics are defined. The asymptotic properties of the test statistics are analyzed in Sections 4–6. The results of some Monte Carlo experiments are presented in Section 7. Section 8 concludes.

Finally, some notational conventions used in this paper. We denote by $W_1(a), \ldots, W_N(a)$ a collection of N independent standard Brownian motions defined on the interval $[0,1]$. Following Phillips and Moon (1999) we indicate the joint limit by $(N,T \to \infty)$. Weak convergence is denoted by $\Rightarrow$, and $\xrightarrow{p}$ means convergence in probability.

2. THE TESTING PROBLEM

Consider a collection of time series $\{y_{i0}, \ldots, y_{iT}\}_{i=1,\ldots,N}$ that is generated by the composite process

$$y_{it} = \gamma_i' f_t + u_{it}, \tag{1}$$

$$f_t = A f_{t-1} + v_t, \tag{2}$$

$$u_{it} = \theta u_{i,t-1} + \varepsilon_{it}, \tag{3}$$

where f_t is an $r \times 1$ vector of unobservable common factors, γ_i is an $r \times 1$ vector of nonrandom factor loadings, $\Gamma = [\gamma_1, \ldots, \gamma_N]'$ is the $N \times r$ matrix of factor loadings, and u_{it} is an idiosyncratic error component. The eigenvalues of the $r \times r$ matrix A are on or inside the unit circle. Specifically, we will focus on the special case $A = \rho I_r$ with $|\rho| \leq 1$.[3]

To simplify the exposition we focus on the case without deterministic terms and short-run dynamics. The inclusion of constants, time trends, and short-run dynamics is straightforward (cf. Breitung and Das, 2005; Breitung and Pesaran, 2007) and will be briefly discussed in Section 3.

For the errors v_t and ε_{it} we make the following assumption.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

Assumption 1. The vector $\varepsilon_t = [\varepsilon_{1t}, \ldots, \varepsilon_{Nt}]'$ is independent and identically distributed (i.i.d.) with $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$, and $\Psi = \lim_{N\to\infty} N^{-1}\Gamma'\Gamma$ a positive definite matrix. The $r \times 1$ error vector v_t is i.i.d. with $E(v_t) = 0$, $E(v_t v_t') = I_r$, and $E(\varepsilon_t v_t') = 0$. Furthermore, $E(\varepsilon_{it}^4) < \infty$ for all i and t, and $E(v_{jt}^2 v_{kt}^2) < \infty$ for all j, k, t.

Remark 1. Phillips and Sul (2003), Moon and Perron (2004), and Pesaran (2005) consider the special case with $A = \theta I_r$ such that the data generating process (1)–(3) reduces to

$$y_{it} = \theta y_{i,t-1} + \gamma_i' v_t + \varepsilon_{it}.$$

In this special case the common factors and the idiosyncratic components are generated by similar autoregressive processes. In our analysis we follow Bai and Ng (2004), who assume that the common factors and the idiosyncratic components may be generated by different processes.

In contrast to the PANIC approach suggested by Bai and Ng (2004), we consider unit root tests based on the "reduced form" autoregression

$$\Delta y_{it} = \phi y_{i,t-1} + e_{it}. \tag{4}$$

In general, $\Delta y_{it} = \gamma_i' \Delta f_t + \Delta u_{it}$ is a linear combination of $r + 1$ ARMA(1,1) processes, and therefore Δy_{it} has an autoregressive moving average (ARMA) representation given by $\alpha_i(L)\Delta y_{it} = \beta_i(L)\xi_{it}$, where $\alpha_i(L)$ and $\beta_i(L)$ are lag polynomials of order $r + 1$ and $\xi_{it} = \Delta y_{it} - E(\Delta y_{it} | \Delta y_{i,t-1}, \Delta y_{i,t-2}, \ldots)$. Under the null hypothesis $y_{it} \sim I(1)$ all roots of the polynomials $\alpha_i(L)$ and $\beta_i(L)$ are outside the unit circle for all $i = 1, \ldots, N$.

An important advantage of testing the reduced form (4) is that no factor representation needs to be estimated. Furthermore, the multiple testing problem resulting from separate tests of the factors and idiosyncratic components is avoided.

3. TEST STATISTICS

Following Levin, Lin, and Chu (2002) the test of the unit root hypothesis $\phi = 0$ against the homogeneous alternative $\phi < 0$ is based on the pooled OLS regression

$$\Delta y_t = \phi y_{t-1} + e_t, \qquad t = 1, 2, \ldots, T, \tag{5}$$

where $\Delta y_t = [\Delta y_{1t}, \ldots, \Delta y_{Nt}]'$, $y_{t-1} = [y_{1,t-1}, \ldots, y_{N,t-1}]'$, and $e_t = [e_{1t}, \ldots, e_{Nt}]'$. The pooled OLS t-statistic is

$$t_{ols} = \frac{\sum_{t=1}^{T} y_{t-1}' \Delta y_t}{\hat{\sigma} \sqrt{\sum_{t=1}^{T} y_{t-1}' y_{t-1}}}, \tag{6}$$

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

where $\hat{\sigma}^2 = (NT)^{-1} \sum_{t=1}^{T} \hat{e}_t' \hat{e}_t$ and $\hat{e}_t = \Delta y_t - \hat{\phi} y_{t-1}$ denotes the residual vector of regression (5).

Following Breitung and Das (2005) and Jönsson (2005) we will also consider the robust t-statistic that employs PCSE. The test statistic results as

$$t_{rob} = \frac{\sum_{t=1}^{T} y_{t-1}' \Delta y_t}{\sqrt{\sum_{t=1}^{T} y_{t-1}' \hat{\Omega} y_{t-1}}}, \tag{7}$$

where

$$\hat{\Omega} = \frac{1}{T} \sum_{t=1}^{T} \hat{e}_t \hat{e}_t'. \tag{8}$$

As the GLS estimator of ϕ in (5) is more efficient than the OLS estimator, a more powerful test can be constructed based on the GLS estimator of ϕ. The GLS t-statistic is given by

$$t_{gls} = \frac{\sum_{t=1}^{T} y_{t-1}' \hat{\Omega}^{-1} \Delta y_t}{\sqrt{\sum_{t=1}^{T} y_{t-1}' \hat{\Omega}^{-1} y_{t-1}}} \tag{9}$$

(cf. Harvey and Bates, 2003; Breitung and Das, 2005). It is important to note that this test statistic can only be computed if $T > N$ as otherwise the estimated covariance matrix $\hat{\Omega}$ is singular. This is an important drawback compared to the OLS-based test statistics.

An attractive feature of the GLS t-statistic is that it does not require an assumption about the structure of the covariance matrix (e.g., the number of common factors). On the other hand, this generality gives rise to poor small-sample properties (see Section 7). To improve the small-sample properties of the test it is therefore desirable to impose some structure on the covariance matrix. Assume that the innovations admit a (strict) factor structure such that $\Omega = \Gamma\Gamma' + \Sigma$. Using the well-known result on the inverse of a sum of two matrices, we obtain

$$\Omega^{-1} = \Sigma^{-1} - \Sigma^{-1}\Gamma(I_r + \Gamma'\Sigma^{-1}\Gamma)^{-1}\Gamma'\Sigma^{-1}. \tag{10}$$

To estimate Ω^{-1} consistent estimators of Γ and Σ are required. Following Bai and Ng (2004) and Moon and Perron (2004) a principal component approach can be adopted. Let $\hat{V}_r = [\hat{\nu}_1, \ldots, \hat{\nu}_r]$ denote the matrix of r eigenvectors (ν_j, $j = 1, \ldots, r$) of $\hat{\Omega}$ associated with the r largest eigenvalues. The matrix Γ is estimated as $\hat{\Gamma} = \hat{V}_r \hat{\Lambda}_r^{1/2}$, where $\hat{\Lambda}_r$ is a diagonal matrix with the largest r

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

eigenvalues on the leading diagonal and $\hat{\Sigma}$ is a diagonal matrix with the diagonal elements of the matrix $(I_N - \hat{V}_r \hat{V}_r')\hat{\Omega}(I_N - \hat{V}_r \hat{V}_r')$ on the leading diagonal.

Another approach to deal with cross-section dependence due to common factors is the elimination of the common factors by employing a transformation matrix Q with the property $Q\Gamma = 0$. From the transformed regression

$$\Delta\zeta_t = \phi^* \zeta_{t-1} + \eta_t \tag{11}$$

with $\zeta_t = Qy_t$ and $\eta_t = Q\varepsilon_t$ we obtain the t-statistic for the hypothesis $\phi^* = 0$ as

$$t_Q = \frac{\sum_{t=1}^{T} \zeta_{t-1}' \Delta\zeta_t}{\sqrt{\sum_{t=1}^{T} \zeta_{t-1}' \zeta_{t-1}}}. \tag{12}$$

Phillips and Sul (2003) suggest using an estimated version of the matrix

$$Q_{PS} = (\Gamma_\perp' \Sigma \Gamma_\perp)^{-1/2} \Gamma_\perp',$$

where $\Gamma_\perp$ is the orthogonal complement of Γ, which can be estimated by using a least-squares approach. Moon and Perron (2004) employ the estimated analogue of the matrix

$$Q_{MP} = I_N - V_r V_r', \tag{13}$$

where $V_r = [\nu_1, \ldots, \nu_r]$ is the matrix of r eigenvectors $(\nu_j,\ j = 1, \ldots, r)$ associated with the largest r eigenvalues of Ω. If the transformation matrix Q_{MP} is used to construct ζ_t, the resulting statistic $t_{Q_{MP}}$ depends on the variances $\sigma_1^2, \ldots, \sigma_N^2$. As shown by Moon and Perron (2004) the statistic $t_{Q_{MP}}$ has to be multiplied by the factor $\sqrt{\sum_{i=1}^N \sigma_i^2}/\sqrt{\sum_{i=1}^N \sigma_i^4}$ to obtain an asymptotically similar test statistic. If $\sigma_1^2 = \cdots = \sigma_N^2 = \sigma^2$ then this factor simplifies to $1/\sigma$, and the ordinary form of the t-statistic results.

So far we have assumed that $E(y_{it}) = 0$ for all t. In empirical practice, however, the inclusion of individual specific constants and time trends is important. Levin et al. (2002) have suggested eliminating individual specific constants by running a pooled OLS regression of the transformed series $\Delta y_{it}^* = \Delta y_{it} - \Delta\bar{y}_i$ on $y_{i,t-1}^* = y_{i,t-1} - \bar{y}_{i,-1}$, where $\Delta\bar{y}_i = T^{-1} \sum_{t=1}^T \Delta y_{it}$ and $\bar{y}_{i,-1} = T^{-1} \sum_{t=1}^T y_{i,t-1}$. To correct for the negative mean of the t-statistic, the numerator and denominator of the test statistic are adjusted by applying correction terms. However, this approach is valid only if the test of Levin et al. (2002) is applied to the GLS transformed series $\hat{\Omega}^{-1/2} y_t$ in case A (see Section 4). In all other cases, the correction terms that have to be applied to the t-statistics depend on nuisance parameters, i.e., the elements of the covariance matrix Ω.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

To sidestep this difficulty, Breitung and Das (2005) suggest subtracting the first observation to remove the individual specific constants.[4] In this case, it is straightforward to show that our asymptotic results carry over to the test statistics computed by using $y_{it}^* = y_{it} - y_{i0}$ instead of the original series y_{it}.

4. CASE A: BOTH COMPONENTS ARE *I*(1)

First we analyze the asymptotic properties of the alternative tests by letting $A = I_r$ and $\theta = 1$; i.e., the common and idiosyncratic components are nonstationary. This case is similar to the situation considered in Phillips and Sul (2003), Moon and Perron (2004), and Pesaran (2005). From (1)–(3) and Assumption 1 it follows that under the null hypothesis

$$H_0^{(a)}: A = I_r \quad \text{and} \quad \theta = 1$$

the reduced form representation (4) implies $\phi = 0$ and $e_{it} = \gamma_i' v_t + \varepsilon_{it}$ is white noise. Because $E(e_{it} e_{jt}) = \gamma_i' \gamma_j \neq 0$, in general, the pooled OLS statistic t_{ols} is invalid. Using a sequential limit theory, Breitung and Das (2005) show that if all eigenvalues of the covariance matrix Ω are bounded, this test statistic is distributed as $\mathcal{N}(0, v_\Omega)$, where $v_\Omega = \lim_{N\to\infty} \mathrm{tr}(\Omega^2/N)/[\mathrm{tr}(\Omega/N)]^2$. However, if Ω has a factor structure, then r eigenvalues are $O(N)$, and therefore the limiting distribution is no longer standard normal.

In the following theorem the limiting null distributions of the test statistics are given.

THEOREM 1. *Assume that y_{it} is generated as in (1)–(3) with $A = I_r$ and $\theta = 1$. Let $\lambda_i^* = \lim_{N\to\infty} N^{-1}\lambda_i$ denote the limit of the ith (ordered) eigenvalue of the covariance matrix $\Omega = E(e_t e_t')$.*

(i) As $(N, T \to \infty)$ and $N/T \to \infty$ we have

$$N^{-1/2} t_{ols} \Rightarrow \frac{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a)}{\bar{\omega} \sqrt{\frac{1}{2}\bar{\sigma}^2 + \sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da}},$$

where $\bar{\sigma}^2 = \lim_{N\to\infty} \mathrm{tr}(\Sigma/N)$ and $\bar{\omega}^2 = \lim_{N\to\infty} \mathrm{tr}(\Omega/N)$,

$$t_{rob} \Rightarrow \frac{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a)}{\sqrt{\sum_{i=1}^{r} \lambda_i^{*2} \int_0^1 W_i(a)^2\, da}},$$

and t_Q has a standard normal limiting distribution.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

(ii) If $(N,T \to \infty)$ *and* $N^2/T \to 0$ *the test statistic* t_{gls} *has a standard normal limiting distribution.*

Remark 2. The limiting distribution of $N^{-1/2}t_{ols}$ implies that for the single-factor model with $r = 1$ the actual size of the OLS statistic tends to the probability

$$\lim_{(T,N\to\infty)} P(t_{ols} < c_\alpha) = \lim_{(T,N\to\infty)} P\left(t_{ols}/\sqrt{N} < c_\alpha/\sqrt{N}\right)$$
$$= P\left(\int_0^1 W_1(a)\, dW_1(a) < 0\right) \simeq 0.6827,$$

where c_α is the critical value of a standard normal distribution with significance level α (cf. Evans and Savin, 1981). Thus, the OLS t-statistic is severely biased.

Remark 3. For the single-factor model with $r = 1$ the limiting distribution of the robust OLS statistic is identical to the limiting distribution of the Dickey–Fuller test statistic. Indeed, in this case, the test is asymptotically equivalent to a Dickey–Fuller test applied to the first principal component $\hat{\nu}_1' y_t$, where $\hat{\nu}_1$ is the eigenvector associated with the largest eigenvalue of $\hat{\Omega}$. Thus, the limiting behavior of this test statistic is dominated by the common factor, and the idiosyncratic components do not affect the limiting distribution.

Remark 4. For the (feasible) GLS statistic t_{gls} given in (9) the relative rates of N and T are important. This is due to the fact that replacing Ω by $\hat{\Omega}$ implies an additional term $O_p(NT^{-1/2})$ and therefore the assumption $N^2/T \to 0$ is required. This explains the poor small-sample properties of the statistic t_{gls} if T is only slightly larger than N (cf. Breitung and Das, 2005).

5. CASE B: $f_t \sim I(1)$ AND $u_{it} \sim I(0)$

In this case we assume that the N time series share r common trends, whereas the idiosyncratic components are stationary. Accordingly, there exists an $N \times (N - r)$ matrix B such that $B' y_t$ is stationary. Following Banerjee et al. (2005) this situation is called *cross-unit cointegration*. Under the null hypothesis

$$H_0^{(b)}: A = I_r \quad \text{and} \quad |\theta| < 1$$

the process $\Delta y_{it} = \gamma_i' v_i + (1 - \theta L)^{-1}(1 - L)\varepsilon_{it}$ has an ARMA(1,1) representation with all roots outside the unit circle. Accordingly, the unit root test has to account for serial correlated errors by augmenting the equation with lagged differences (cf. Said and Dickey, 1984), by using the prewhitening approach proposed by Breitung and Das (2005), or by using the nonparametric approach suggested by Moon and Perron (2004). In the following theorem we nevertheless analyze the limiting behavior of the original test statistics introduced in Section 3. The effects of an adjustment for serial correlation are considered in Remark 5.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

THEOREM 2. *Assume that y_{it} is generated as in (1)–(3) with $A = I_r$ and $|\theta| < 1$.*

(i) If $(N, T \to \infty)$ and $N/T \to 0$, then

$$N^{-1/2} t_{ols} \Rightarrow \frac{\left[\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a)\right] - \frac{1-\theta}{1-\theta^2}\,\bar{\sigma}^2}{\bar{\omega}\sqrt{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da}},$$

where λ_i^*, $\bar{\sigma}^2$, and $\bar{\omega}^2$ are defined in Theorem 1,

$$t_{rob} \Rightarrow \frac{\left[\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a)\right] - \frac{1-\theta}{1-\theta^2}\,\bar{\sigma}^2}{\sqrt{\sum_{i=1}^{r} \lambda_i^{*2} \int_0^1 W_i(a)^2\, da}},$$

$$N^{-1/2} T^{-1/2} t_Q \xrightarrow{p} -\frac{(1-\theta)}{\sqrt{1-\theta^2}} \lim_{N\to\infty} \sqrt{\mathrm{tr}(Q\Sigma/N)}.$$

(ii) If in addition $N^2/T \to 0$, we have

$$N^{-1/2} t_{gls} \Rightarrow -\frac{\frac{1-\theta}{1-\theta^2} \lim_{N\to\infty} \mathrm{tr}(\Omega^{-1}\Sigma)/N}{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da}.$$

Remark 5. There are alternative ways to account for the serial correlation of the errors e_{it}. As shown by Chang (2004) the augmented Dickey–Fuller test, where lagged differences are included in the regression, is invalid in a panel data setup. Breitung and Das (2005) suggest an alternative approach where the test is applied to the prewhitened series $\tilde{y}_{it} = y_{it} - \hat{\alpha}_{i1} y_{i,t-1} - \cdots - \hat{\alpha}_{ip_i} y_{i,t-p_i}$ and the autoregressive coefficients are the least-squares estimators of the individual specific regressions

$$\Delta y_{it} = \alpha_{i1} \Delta y_{i,t-1} + \cdots + \alpha_{ip_i} \Delta y_{i,t-p_i} + e_{it}.$$

Under the null hypothesis the prewhitened series $\tilde{y}_{it}$ is a pure random walk series, and the test statistic suggested in Section 3 can be applied. Moon and Perron (2004) adopt a nonparametric approach that eliminates the nuisance parameters from the limiting distribution.

In a factor model, the prewhitening approach may fail if the short-run dynamics are individual specific. Under the null hypothesis $H_0^{(b)}$ the errors in (4) can be represented as $e_{it} = \gamma_i' v_i + w_{it}$, where $w_{it} = \theta w_{i,t-1} + \varepsilon_{it} - \varepsilon_{i,t-1}$. Because

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

the parameters of the ARMA(1,1) representation of e_{it} depend on γ_i, the ARMA representation is different across the panel units, in general. The prewhitening approach uses an autoregressive approximation to the ARMA(1,1) process yielding

$$\alpha_i(L)y_{it} = \gamma_i'[\alpha_i(L)f_t] + \alpha_i(L)u_{it}. \quad \textbf{(14)}$$

However, if $\alpha_i(L)$ is different across i the factors are no longer common to all units. As noted by Stock and Watson (2002) the dynamic factor model (14) can be represented as a factor model with $\tilde{f}_t = [f_t', f_{t-1}', \ldots, f_{t-p^*}']'$ as the vector of common factors, where $p^* = \max\{p_1, \ldots, p_N\}$.

Moreover, even if the number of factors is increased and the autoregressive representation is generous enough to yield sufficient approximations of the ARMA(1,1) processes, the GLS approach remains invalid. First, the prewhitening procedure removes the serial correlation of e_{it} but does not yield serially uncorrelated idiosyncratic errors u_{it} (which would eliminate the nuisance parameter θ). Second, if the idiosyncratic errors are stationary, the numerator of the GLS test statistic has a larger stochastic order than the denominator regardless of the serial correlation in e_{it}.

Remark 6. It is not surprising that the defactoring approach (i.e., the test statistic t_Q) is invalid in case B. Because the nonstationary factors are removed, the test applied to the defactored series tends to reject the null hypothesis of nonstationary time series.

6. CASE C: $f_t \sim I(0)$ AND $u_{it} \sim I(1)$

In the third case we assume that the common factors are stationary, whereas the idiosyncratic components are nonstationary. Accordingly, the null hypothesis is

$$H_0^{(c)}: \theta = 1 \quad \text{and} \quad A = \rho I_r, \quad \text{where } |\rho| < 1.$$

As in case B, this assumption implies that the errors e_{it} in regression (4) have an ARMA(1,1) representation, yet this time of the form $e_{it} = \gamma_i' f_t^* + \varepsilon_{it}$, where $f_t^* = (1 - \rho L)^{-1}(1 - L)f_t$. In the following theorem, the asymptotic properties of the test statistics are summarized.

THEOREM 3. *Assume that y_{it} is generated as given in (1)–(3) and $H_0^{(c)}$. As $(N, T \to \infty)$ and $N/T \to 0$, t_{ols} and t_{rob} diverge to $-\infty$, $t_Q \Rightarrow \mathcal{N}(0,1)$, and, if $N^2/T \to 0$, $t_{gls} \Rightarrow \mathcal{N}(0,1)$.*

It turns out that the GLS and the defactorization approach t_Q are valid in this case, whereas the OLS-based test statistics tend to indicate stationary time series as N and T tend to infinity.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

7. SMALL-SAMPLE PERFORMANCE

In this section, we present the results of some Monte Carlo experiments performed to investigate the finite-sample performance of the four tests analyzed in Sections 4–6, i.e., the pooled OLS t-statistic (t_{ols}), the robust test (t_{rob}), the GLS t-statistic (t_{gls}), and the test statistic t_Q defined in (12) using the transformation matrix Q_{MP} given in (13), where Ω is replaced by $\hat{\Omega}$. This test is denoted by $t_{\hat{Q}_{MP}}$. In addition we have included the GLS statistic that imposes the factor structure by using the (estimated) inverse covariance matrix given in (10). This test statistic is denoted by t^*_{gls}.

In the first set of our Monte Carlo simulations, the data are generated as in (1)–(3) with $r = 1$ (the single-factor model). All starting values are set equal to zero, and v_t is drawn from an $\mathcal{N}(0,1)$ distribution. The factor loadings γ_i are drawn from a $U(0,2)$ distribution, and the error vector $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt}]'$ is independently drawn from an $\mathcal{N}(0, I_N)$ distribution.[5]

Table 1 reports the actual sizes of the tests. In the top part of the table the results for case A are presented, where both components are generated as random walks (i.e., $\theta = 1$ and $\rho = 1$). Our theoretical findings summarized in Theorem 1 suggest that in this situation the OLS-based test statistics are biased, whereas the GLS test (t_{gls}), the modified GLS test (t^*_{gls}), and the $t_{\hat{Q}_{MP}}$ test are asymptotically valid. These results are confirmed by the simulation results presented in Table 1. According to Remark 2, we expect that the actual size of the test converges to 0.6827 as N and T tend to infinity. Indeed, this test statistic is severely biased, with an actual size exceeding 0.5 if $N = 50$ and $T = 100$.

Theorem 1(i) indicates that in a single-factor model the robust t-statistic is distributed as the Dickey–Fuller t-statistic. The 5% critical value of this distribution is -1.95 compared to the critical value of -1.645 for the standard normal distribution. Therefore, we expect a moderate positive size bias of the robust OLS t-statistic. Indeed, our Monte Carlo simulations suggest that this test statistic slightly overrejects as T becomes large.

As suggested by our Monte Carlo simulations, the actual sizes of the test statistics that impose the factor structure ($t_{\hat{Q}_{MP}}$ and t^*_{gls}) are close to the nominal size. In contrast, the general GLS t-statistic is severely biased if T is only slightly larger than N. Our results suggest that the nominal size is attained when T is at least 10 times larger than N. This result is in line with our theoretical results that require $N^2/T \to 0$ for the distribution to be asymptotically standard normal.

The second part of Table 1 presents the actual sizes of the test if the panel units are cross-cointegrated (case B). We set $\theta = 0.98$ to generate "near-integrated" errors, which seems plausible for many economic applications.[6] As stated in Theorem 2, all tests are biased in this case. The robust OLS test has the smallest size bias among all tests. From Theorem 2 it turns out that the GLS statistic t_{gls} and the $t_{\hat{Q}_{MP}}$ statistic tend to $-\infty$ and therefore the actual size tends to one. This is confirmed by our Monte Carlo simulations, where in large samples the rejection frequencies are close to unity.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

TABLE 1. Empirical sizes: One factor ($r = 1$)

N	T	t_{ols}	t_{rob}	$t_{\hat{Q}_{MP}}$	t_{gls}	t^*_{gls}
Case A: $f_t \sim I(1)$ and $u_{it} \sim I(1)$						
10	20	23.8	6.01	10.4	17.6	9.48
	50	23.6	7.05	8.58	9.49	8.39
	100	24.1	7.99	8.56	8.40	8.00
20	20	34.4	4.79	9.65	—	8.71
	50	35.5	7.48	7.28	18.2	7.18
	100	35.5	7.86	7.16	11.2	6.84
50	20	48.3	3.81	11.35	—	9.44
	50	49.5	6.42	7.18	—	7.05
	100	50.8	7.67	6.15	34.4	6.18
Case B: $f_t \sim I(1)$ and $u_{it} \sim I(0)$						
10	20	37.1	11.4	24.8	32.7	23.3
	50	49.4	20.8	56.4	50.8	53.0
	100	56.6	26.7	93.5	82.0	86.6
20	20	50.2	11.9	34.9	—	33.2
	50	59.1	22.9	80.9	70.9	78.2
	100	64.4	28.0	99.8	96.4	98.6
50	20	58.7	12.5	56.5	—	55.3
	50	66.8	22.4	98.2	—	97.9
	100	69.6	28.6	100	98.0	100
Case C: $f_t \sim I(0)$ and $u_{it} \sim I(1)$						
10	20	54.0	20.3	14.1	39.7	14.9
	50	63.4	43.0	11.0	24.4	14.1
	100	67.3	56.7	9.20	17.8	13.1
20	20	74.4	21.6	16.2	—	14.4
	50	84.9	54.0	10.4	53.6	11.9
	100	88.2	73.7	8.54	32.7	10.9
50	20	91.6	21.1	21.7	—	16.2
	50	98.3	67.1	11.1	—	11.1
	100	99.6	92.2	8.06	91.7	9.65

Note: t_{ols} is the pooled OLS t-statistic that ignores cross-section dependence, t_{rob} is the robust OLS t-statistic as defined in (7), $t_{\hat{Q}_{MP}}$ indicates the test statistic based on the approach suggested by Moon and Perron (2004), t_{gls} is the GLS t-statistic, and t^*_{gls} is the GLS statistic that imposes the factor structure on the covariance matrix. In case A the components are generated as pure random walks, in case B the stationary idiosyncratic component is generated with $\theta = 0.98$, and in case C the common factors are stationary with $\rho = 0.8$. The nominal size for all tests is 5%.

The theoretical considerations of Section 6 suggest that in case C the actual sizes of the OLS-based statistics tend to one, whereas the actual sizes of the $t_{\hat{Q}_{MP}}$ and t_{gls} statistics are expected to converge to the nominal ones as N and T tend to infinity. Setting $\rho = 0.8$ yields a stationary factor with long-run dynamics that are clearly different from the idiosyncratic components, and the theoretical findings should be relevant in fairly small samples also. The results of our Monte

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

TABLE 2. Empirical sizes: Two factors ($r = 2$)

N	T	t_{ols}	t_{rob}	$t_{\hat{Q}_{MP}}$	t_{gls}	t^*_{gls}
Case A: $f_t \sim I(1)$ and $u_{it} \sim I(1)$						
10	20	25.0	6.36	16.1	21.0	14.6
	50	25.8	7.72	10.9	11.0	10.1
	100	26.5	8.16	10.3	9.10	10.0
20	20	36.4	6.01	12.1	—	14.6
	50	38.3	7.28	8.78	23.4	9.92
	100	39.3	8.06	7.50	13.2	8.08
50	20	48.1	5.84	14.3	—	22.2
	50	50.0	7.66	6.68	—	9.52
	100	49.9	8.22	5.40	38.3	7.08
Case B: $f_t \sim I(1)$ and $u_{it} \sim I(0)$						
10	20	33.8	11.1	28.6	32.8	29.4
	50	41.3	14.9	55.2	45.4	51.8
	100	45.8	18.9	92.3	72.1	81.5
20	20	45.1	11.3	36.1	—	42.6
	50	51.2	15.8	78.2	65.7	75.6
	100	55.1	19.1	99.6	92.3	96.8
50	20	56.0	10.6	44.0	—	59.7
	50	61.5	16.8	98.9	—	97.9
	100	62.4	19.4	100	96.8	99.9
Case C: $f_t \sim I(0)$ and $u_{it} \sim I(1)$						
10	20	71.9	33.2	29.6	59.4	34.2
	50	88.7	69.1	17.2	43.9	25.0
	100	93.5	85.6	11.8	33.5	21.5
20	20	92.0	43.9	22.9	—	34.7
	50	98.1	80.9	10.4	86.4	21.9
	100	99.6	95.9	9.60	68.2	18.4
50	20	95.8	38.9	25.5	—	43.8
	50	99.8	90.0	6.86	—	21.2
	100	100	99.5	5.66	99.9	14.2

Note: t_{ols} is the pooled OLS t-statistic that ignores cross-section dependence, t_{rob} is the robust OLS t-statistic as defined in (7), $t_{\hat{Q}_{MP}}$ indicates the test statistic based on the approach suggested by Moon and Perron (2004), t_{gls} is the GLS t-statistic, and t^*_{gls} is the GLS statistic that imposes the factor structure on the covariance matrix. In case A the components are generated as pure random walks, in case B the stationary idiosyncratic component is generated with $\theta = 0.98$, and in case C the common factors are stationary with $\rho = 0.8$. The nominal size for all tests is 5%.

Carlo simulation presented in the third part of Table 1 show that this is indeed the case, as the GLS statistic suffers from size distortions in small samples. It generally has a larger size bias than the $t_{\hat{Q}_{MP}}$ statistic that disappears slowly as T increases.

Next we consider a factor model with $r = 2$, where the factors are generated as $f_{jt} = \rho f_{j,t-1} + v_{jt}$ $(j = 1, 2)$ and the idiosyncratic errors are generated as in the

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

single-factor model. The results of the Monte Carlo simulations for the three situations considered in Sections 4–6 are presented in Table 2. In general, the results are similar to the findings based on the single-factor model. A notable difference is that the tests that impose the two-factor structure ($t_{\hat{Q}_{MP}}$ and t^*_{gls}) perform substantially worse in small samples. Furthermore, it is important to note that these tests are computed assuming that the correct number of factors is known. If the number of factors is estimated, e.g., by using the information criteria suggested by Bai and Ng (2002), one can expect that the performance will be even worse.

8. CONCLUSIONS

In this paper we analyze the size properties of various unit root tests in panels where the cross-section dependence is due to common factors. Banerjee et al. (2005) demonstrate that first generation panel unit root tests may be severely biased. In this paper we also consider the asymptotic properties of second generation panel unit root tests that were suggested to cope with cross-section dependence. We analyze three different situations. First, the common factors and idiosyncratic components may both be nonstationary. In this case the GLS test and tests that apply a transformation for removing the common factors possess a standard normal limiting distribution, whereas the OLS-based test statistics are invalid. Second, if the common component is $I(1)$ and the idiosyncratic component is stationary (the case of cross-unit cointegration), then all tests are invalid. Finally, if the idiosyncratic components are $I(1)$ but the common factors are stationary, then the OLS-based test statistics are not applicable, whereas GLS test procedures that remove the cross-section dependence are asymptotically valid. These findings are confirmed by our Monte Carlo simulations.

NOTES

1. For a recent review of the literature see Breitung and Pesaran (2007).

2. There exist several other panel unit root statistics that we do not consider in our analysis. For example, we leave out the test suggested by Im, Pesaran, and Shin (2003), which was suggested only for the case of independent panel units. Furthermore, the test of Harris and Tzavalis (1999) is not considered as it assumes that T is fixed. In addition, there is no variant of this test that is applicable for cross-section-dependent units.

3. As pointed out by a referee, the assumption of identical short-run dynamics of the factors is not without loss of generality. In particular, if the vector of innovations of the factors has a singular covariance matrix (which is a reasonable assumption in dynamic factor models), then the covariance matrix of the factors has a reduced rank, which is excluded by Assumption 1.

4. Other transformations are available for the model with individual specific time trends (cf. Breitung, 2000).

5. We also investigate the small-sample properties of the tests in a model with individual specific constants, where the individual specific means are removed by subtracting the first observations (see Section 3). The results are quite similar to the case without a constant (see also Breitung and Das, 2005, where both cases are presented). The results for the model with individual specific constants are available from the authors on request.

6. We also tried other values of θ. For smaller values of θ the results are much more pronounced, and the asymptotic properties are apparent in smaller sample sizes also.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

REFERENCES

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

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

Banerjee, A., M. Marcellino, & C. Osbat (2005) Testing for PPP: Should we use panel methods? *Empirical Economics* 30, 77–91.

Breitung, J. (2000) The local power of some unit root tests for panel data. In B. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Advances in Econometrics, vol. 15, pp. 161–178. JAI.

Breitung, J. & S. Das (2005) Panel unit root tests under cross sectional dependence. *Statistica Neerlandica* 59, 1–20.

Breitung, J. & M.H. Pesaran (2007) Unit roots and cointegration in panels. In L. Mátyás & P. Sevestre (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Springer.

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

Chang, Y. (2004) Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics* 120, 263–293.

Choi, I. (2002) Combination Unit Root Tests for Cross-Sectionally Correlated Panels. Mimeo, Hong Kong University of Science and Technology.

Evans, G.B.A. & N.E. Savin (1981) Testing for unit roots, part 1. *Econometrica* 49, 753–779.

Gengenbach, C., F.C. Palm, & J.-P. Urbain (2005) Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling. METEOR Research Memorandum 04039, Maastricht University.

Harris, R.D.F. & E. Tzavalis (1999) Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics* 91, 201–226.

Harvey, A.C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Harvey, A. & D. Bates (2003) Multivariate Unit Root Tests, Stability and Convergence. DAE Working paper 301, University of Cambridge.

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

Jönsson, K. (2005) Cross-sectional dependency and size distortion in a small-sample homogeneous panel-data unit root test. *Oxford Bulletin of Economics and Statistics* 63, 369–392.

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

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

Pesaran, H.M. (2005) A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. Cambridge Working Papers in Economics 0346, revised version, University of Cambridge.

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

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

Said, S.E. & D.A. Dickey (1984) Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71, 599–607.

Stock, J.H. & M.W. Watson (2002) Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–1179.

Zellner, A. (1962) An efficient method of estimating seemingly unrelated regressions and tests of aggregation bids. *Journal of the American Statistical Association* 57, 348–368.

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

APPENDIX

Proof of Theorem 1. The matrix $\Omega = E(\Delta y_t \Delta y_t') = \Gamma\Gamma' + \Sigma$ is decomposed as $\Omega = V\Lambda V'$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_N)$, $\lambda_1 > \cdots > \lambda_N$ are the ordered eigenvalues and $V = [\nu_1, \ldots, \nu_N]$ is the matrix of orthonormal eigenvectors. Let $z_t = \Lambda^{-1/2} V' y_t$ denote an $N \times 1$ vector of random walks with unit covariance matrix. We have

$$a_{NT} = N^{-1}T^{-1} \sum_{t=1}^{T} y_{t-1}' \Delta y_t = N^{-1}T^{-1} \sum_{t=1}^{T} z_{t-1}' \Lambda \Delta z_t = N^{-1} \sum_{i=1}^{N} \lambda_i \xi_{iT},$$

where $\xi_{iT} = T^{-1} \sum_{t=1}^{T} z_{i,t-1} \Delta z_{it}$. Because ξ_{iT} is independent of ξ_{jT} for $i \neq j$ and other regularity conditions that are implied by Assumption 1, it follows from Theorem 1 of Phillips and Moon (1999) that

$$\frac{1}{N} \sum_{i=r+1}^{N} \lambda_i \xi_{iT} \xrightarrow{p} 0,$$

as $(N, T \to \infty)$ and $N/T \to 0$, where we have used the fact that λ_i is bounded for $i \geq r + 1$. It follows that

$$a_{NT} = \sum_{i=1}^{r} (\lambda_i / N) \xi_{iT} + o_p(1) \Rightarrow \sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a)$$

as $(N, T \to \infty)$.

In a similar manner it follows that

$$\begin{aligned} b_{NT} &= \frac{1}{NT^2} \sum_{t=1}^{T} y_{t-1}' y_{t-1} = \frac{1}{NT^2} \sum_{t=1}^{T} \left(\sum_{i=1}^{r} \lambda_i z_{i,t-1}^2 + \sum_{j=r+1}^{N} \lambda_j z_{j,t-1}^2 \right) \\ &= \sum_{i=1}^{r} (\lambda_i / N) S_{iT} + N^{-1} \sum_{j=r+1}^{N} \lambda_j S_{jT}, \quad \text{where } S_{iT} = T^{-2} \sum_{t=1}^{T} z_{i,t-1}^2 \\ &\Rightarrow \left(\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da \right) + \frac{1}{2} \bar{\sigma}^2, \end{aligned}$$

as $(N, T \to \infty)$, where $N^{-1} \lim_{N\to\infty} \sum_{i=r+1}^{N} \lambda_i = \lim_{N\to\infty} \text{tr}(\Sigma/N) \equiv \bar{\sigma}^2$.

Under the null hypothesis the least-squares estimator $\hat{\phi}$ is $O_p(T^{-1}N^{-1/2})$ (see, e.g., Moon and Perron, 2004), and therefore $\hat{e}_{it} = \gamma_i' v_t + u_{it} + O_p(T^{-1/2}N^{-1/2})$. It follows for the elements of the estimated covariance matrix $\hat{\Omega}$ that

$$\begin{aligned} \hat{\omega}_{ij} &= T^{-1} \sum_{t=1}^{T} \hat{e}_{it} \hat{e}_{jt} = T^{-1} \sum_{t=1}^{T} (\gamma_i' v_t + u_{it})(\gamma_j' v_t + u_{jt}) + O_p(T^{-1/2}N^{-1/2}) \\ &= \omega_{it} + O_p(T^{-1/2}) + O_p(T^{-1/2}N^{-1/2}), \end{aligned}$$

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

where ω_{ij} denotes the (i,j)th element of Ω. Let $\hat{\Omega} = \Omega + \Delta_{NT}$. We have

$$c_{NT} = N^{-2}T^{-2}\sum_{t=1}^{T} y'_{t-1}\hat{\Omega}y_{t-1} = N^{-2}T^{-2}\sum_{t=1}^{T} y'_{t-1}\Omega y_{t-1} + N^{-2}T^{-2}\sum_{t=1}^{T} y'_{t-1}\Delta_{NT}y_{t-1}$$

$$= N^{-2}\sum_{i=1}^{N}\lambda_i^2 S_{iT} + o_p(1).$$

Because λ_i is bounded for $i \geq r+1$ we have $N^{-1}\sum_{i=r+1}^{N}(\lambda_i^2/N)S_{iT} \xrightarrow{p} 0$, and therefore, as $(N,T \to \infty)$,

$$c_{NT} \Rightarrow \sum_{i=1}^{r}\lambda_i^{*2}\int_0^1 W_i(a)^2\,da.$$

Using

$$\hat{\sigma}^2 = N^{-1}T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{e}_{it}^2 \xrightarrow{p} \lim_{N\to\infty}\operatorname{tr}(\Omega/N) \equiv \bar{\omega}^2$$

it follows that

$$N^{-1/2}t_{ols} = \frac{a_{NT}}{\hat{\sigma}\sqrt{b_{NT}}} \Rightarrow \frac{\sum_{i=1}^{r}\lambda_i^*\int_0^1 W_i(a)\,dW_i(a)}{\bar{\omega}\sqrt{\frac{1}{2}\bar{\sigma}^2 + \sum_{i=1}^{r}\lambda_i^*\int_0^1 W_i(a)^2\,da}},$$

$$t_{rob} = \frac{a_{NT}}{\sqrt{c_{NT}}} \Rightarrow \frac{\sum_{i=1}^{r}\lambda_i^*\int_0^1 W_i(a)\,dW_i(a)}{\sqrt{\sum_{i=1}^{r}\lambda_i^{*2}\int_0^1 W_i(a)^2\,da}}.$$

Next we consider the GLS t-statistic. Let $\hat{\Omega} = \Omega + \Delta_{NT}$, where the elements of Δ_{NT} are $O_p(T^{-1/2})$. We make use of the following expansion (e.g., Zellner, 1962):

$$\hat{\Omega}^{-1} = \Omega^{-1} - \Omega^{-1}\Delta_{NT}\Omega^{-1} + \ldots,$$

where asymptotically negligible terms are suppressed. It follows that

$$d_{NT} = N^{-1/2}T^{-1}\sum_{t=1}^{T} y'_{t-1}\hat{\Omega}^{-1}\Delta y_t$$

$$= N^{-1/2}T^{-1}\sum_{t=1}^{T} y'_{t-1}\Omega^{-1}\Delta y_t - N^{-1/2}T^{-1}\sum_{t=1}^{T} y'_{t-1}\Omega^{-1}\Delta_{NT}\Omega^{-1}\Delta y_t + \ldots$$

$$= N^{-1/2}T^{-1}\sum_{t=1}^{T} y'_{t-1}\Omega^{-1}\Delta y_t + O_p(NT^{-1/2}),$$

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

where we have used the fact that $\|\Omega^{-1/2}y_{t-1}\| = O_p(N^{1/2})$, $\|\Omega^{-1/2}\Delta y_t\| = O_p(N^{1/2})$, and $\|\Omega^{-1/2}\Delta_{NT}\Omega^{-1/2}\| = O_p(N^{1/2}T^{-1/2})$. Thus, if $(N,T \to \infty)$ and $N^2/T \to 0$ we obtain

$$d_{NT} = N^{-1/2}\sum_{i=1}^{N}\xi_{iT} + o_p(1) \Rightarrow \mathcal{N}\left(0, \frac{1}{2}\right).$$

Similarly, we have for $(N,T \to \infty)$ and $N/T \to 0$

$$e_{NT} = \frac{1}{NT^2}\sum_{t=1}^{T} y'_{t-1}\hat{\Omega}^{-1}y_{t-1} = \frac{1}{NT^2}\sum_{t=1}^{T} y'_{t-1}\Omega^{-1}y_{t-1} + O_p\left(\sqrt{N}/\sqrt{T}\right) \xrightarrow{p} \frac{1}{2}.$$

It follows that $t_{gls} = d_{NT}/\sqrt{e_{NT}} \Rightarrow \mathcal{N}(0,1)$ as $(N,T \to \infty)$ and $N^2/T \to 0$.

Finally, as shown by Phillips and Sul (2003) and Moon and Perron (2004), the statistic t_Q has a standard normal limiting distribution under the conditions of case A. ■

Proof of Theorem 2. We define a_{NT} as in the proof of Theorem 1 and write

$$a_{NT} = N^{-1}T^{-1}\sum_{t=1}^{T}(f'_{t-1}\Gamma'\Gamma\Delta f_t + u'_{t-1}\Delta u_t + f'_{t-1}\Gamma'\Delta u_t + u'_{t-1}\Gamma\Delta f_t).$$

Because f_t and u_t are independent we obtain

$$N^{-1}T^{-1}\sum_{t=1}^{T} u'_{t-1}\Gamma\Delta f_t \xrightarrow{p} 0,$$

$$N^{-1}T^{-1}\sum_{t=1}^{T} f'_{t-1}\Gamma'\Delta u_t = -N^{-1}T^{-1}\sum_{t=1}^{T}\Delta f'_t\Gamma' u_t + O_p(T^{-1}) \xrightarrow{p} 0,$$

as $(N,T \to \infty)$. Let $\Delta z_t^{(r)} = [\Delta z_{1t},\ldots,\Delta z_{rt}]' = \Lambda_{(r)}^{-1/2}V'_{(r)}\Delta y_t$, where $\Lambda_{(r)}$ is a diagonal matrix with the first r eigenvalues of $E(\Delta y_t \Delta y'_t)$ on the leading diagonal and $V_{(r)}$ is an $N \times r$ matrix with the associated r eigenvectors. Because $N^{-1}\Lambda_{(r)}^{1/2}\Delta z_t^{(r)} = V'_{(r)}\Gamma\Delta f_t + O_p(N^{-1/2})$ it follows that $\sum_{i=1}^{r}\lambda_i z_{i,t-1}\Delta z_{it} = f'_{t-1}\Gamma'\Gamma\Delta f_t + O_p(N^{-1/2})$, and therefore

$$N^{-1}T^{-1}\sum_{t=1}^{T} f'_{t-1}\Gamma'\Gamma\Delta f_t \Rightarrow \sum_{i=1}^{r}\lambda_i^*\int W_i(a)\,dW(a),$$

as $(N,T \to \infty)$.

Because $E(\Delta u'_t u_{t-1}) = E\{[(\theta-1)u'_{t-1} + \varepsilon'_t]u_{t-1}\} = -(1-\theta)(1-\theta^2)^{-1}\mathrm{tr}(\Sigma)$ we have

$$N^{-1}T^{-1}\sum_{t=1}^{T} u'_{t-1}\Delta u_t \xrightarrow{p} -\frac{1-\theta}{1-\theta^2}\lim_{N\to\infty}\mathrm{tr}(\Sigma/N)$$

as $(N,T \to \infty)$. Therefore,

$$a_{NT} \Rightarrow \sum_{i=1}^{r}\lambda_i^*\int_0^1 W_i(a)\,dW_i(a) - \frac{1-\theta}{1-\theta^2}\bar{\sigma}^2.$$

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

As in Theorem 1 it can be shown that

$$b_{NT} = \sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da \quad \text{and} \quad c_{NT} \Rightarrow \sum_{i=1}^{r} \lambda_i^{*2} \int_0^1 W_i(a)^2\, da.$$

Thus,

$$N^{-1/2} t_{ols} = \frac{a_{NT}}{\hat{\sigma}\sqrt{b_{NT}}} \Rightarrow \frac{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a) - \frac{1-\theta}{1-\theta^2}\bar{\sigma}^2}{\bar{\omega}\sqrt{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da}},$$

$$t_{rob} = \frac{a_{NT}}{\sqrt{c_{NT}}} \Rightarrow \frac{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)\, dW_i(a) - \frac{1-\theta}{1-\theta^2}\bar{\sigma}^2}{\sqrt{\sum_{i=1}^{r} \lambda_i^{*2} \int_0^1 W_i(a)^2\, da}}.$$

Next we consider the GLS t-statistic. As $(N, T \to \infty)$ and $N^2/T \to 0$ we have

$$N^{-1/2} d_{NT} = N^{-1}T^{-1} \sum_{t=1}^{T} u_{t-1}' \Omega^{-1} \Delta u_t + o_p(1)$$

$$\xrightarrow{p} -\frac{1-\theta}{1-\theta^2} \lim_{N\to\infty} \operatorname{tr}(\Omega^{-1}\Sigma/N).$$

Because u_t is stationary we obtain as $N^2/T \to 0$

$$e_{NT} = N^{-1}T^{-2} \sum_{t=1}^{T} (f_{t-1}' \Gamma' \Gamma f_{t-1} + u_{t-1}' u_{t-1} + 2 f_{t-1}' \Gamma' u_{t-1}) + O_p(NT^{-1/2})$$

$$\Rightarrow \sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da.$$

It follows that as $(N, T \to \infty)$ and $N^2/T \to 0$

$$N^{-1/2} t_{gls} = \frac{N^{-1/2} d_{NT}}{\sqrt{e_{NT}}} \Rightarrow -\frac{\frac{1-\theta}{1-\theta^2} \lim_{N\to\infty} \operatorname{tr}(\Omega^{-1}\Sigma/N)}{\sum_{i=1}^{r} \lambda_i^* \int_0^1 W_i(a)^2\, da}.$$

Finally, we consider the defactorization approach proposed by Phillips and Sul (2003) and Moon and Perron (2004). Let $\zeta_t = Q y_t$, where $Q\Gamma = 0$. As for d_{NT} we obtain as $(N, T \to \infty)$

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

$$N^{-1/2}\tilde{d}_{NT} = N^{-1}T^{-1}\sum_{t=1}^{T} y'_{t-1}Q\Delta y_t = N^{-1}T^{-1}\sum_{t=1}^{T} u'_{t-1}Q\Delta u_t + o_p(1)$$

$$\xrightarrow{p} -\frac{1-\theta}{1-\theta^2}\lim_{N\to\infty}\operatorname{tr}(Q\Sigma/N)$$

and

$$T\tilde{e}_{NT} = N^{-1}T^{-1}\sum_{t=1}^{T} y'_{t-1}Qy_{t-1} \xrightarrow{p} \frac{1}{1-\theta^2}\lim_{N\to\infty}\operatorname{tr}(Q\Sigma/N).$$

This implies that

$$N^{-1/2}T^{-1/2}t_Q = \frac{N^{-1/2}\tilde{d}_{NT}}{\sqrt{T\tilde{e}_{NT}}} \xrightarrow{p} -\frac{(1-\theta)}{\sqrt{1-\theta^2}}\sqrt{\lim_{N\to\infty}\operatorname{tr}(Q\Sigma/N)}. \quad \blacksquare$$

Proof of Theorem 3. As in Theorems 1 and 2 we obtain as $(N, T \to \infty)$ and $N/T \to 0$

$$a_{NT} = N^{-1}T^{-1}\sum_{t=1}^{T} f'_{t-1}\Gamma'\Gamma\Delta f_t + O_p(N^{-1/2}) \xrightarrow{p} -\frac{1-\rho}{1-\rho^2}\sum_{i=1}^{r}\lambda_i^*$$

and

$$b_{NT} = \sum_{i=1}^{r}(\lambda_i/N)S_{iT} + N^{-1}\sum_{j=r+1}^{N}\lambda_j^2 S_{jT}$$

$$\Rightarrow \sum_{i=1}^{r}\lambda_i^*\int_0^1 W_i(a)\,dW_i(a) + \frac{1}{2}\lim_{N\to\infty}\operatorname{tr}(\Sigma^2/N).$$

Furthermore,

$$c_{NT} = \sum_{i=1}^{r}(\lambda_i^2/N^2)S_{iT} + N^{-2}\sum_{j=r+1}^{N}\lambda_j^2 S_{jT} + O_p(T^{-1/2}).$$

Because S_{iT} is $O_p(T^{-1})$ for $i = 1, \ldots, r$ and $E(S_{jT}) = \frac{1}{2}$ for $j = r+1, \ldots, N$ it follows that $c_{NT} = O_p(T^{-1/2}) + O(N^{-1})$. Therefore, $t_{ols} = N^{1/2}a_{NT}/\left(\hat{\sigma}\sqrt{b_{NT}}\right)$ and $t_{rob} = a_{NT}/\sqrt{c_{NT}}$ tend to $-\infty$ as N and T tend to infinity.

Using similar results as in Theorems 1 and 2 we obtain as $N^2/T \to 0$

$$d_{NT} = \left(N^{-1/2}T^{-1}\sum_{t=1}^{T} f'_{t-1}\Gamma'\Omega^{-1}\Gamma\Delta f_t\right) + \left(N^{-1/2}T^{-1}\sum_{t=1}^{T} u'_{t-1}\Omega^{-1}\Delta u_t\right) + o_p(1).$$

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press

The first term of the right-hand side is $O_p(N^{-1/2})$, whereas for the second term we have

$$\begin{aligned} N^{-1/2}T^{-1}\sum_{t=1}^{T} u_{t-1}'\Omega^{-1}\Delta u_t &= N^{-1/2}T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_i^{-1}\tilde{u}_{i,t-1}\Delta\tilde{u}_{it} \\ &= N^{-1/2}T^{-1}\sum_{i=r+1}^{N}\sum_{t=1}^{T}\lambda_i^{-1}\tilde{u}_{i,t-1}\Delta\tilde{u}_{it} + O_p(N^{-1/2}) \\ &= N^{-1/2}T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} u_{i,t-1}\Delta u_{it}/\sigma_i^2 + o_p(1), \end{aligned}$$

where $\tilde{u}_{it} = v_i' u_t$ and v_i is the ith eigenvector of Ω. Thus d_{NT} is asymptotically normally distributed with $E(d_{NT}) = 0$ and $\mathrm{var}(d_{NT}) \to \frac{1}{2}$.

In a similar manner we obtain as $(N, T \to \infty)$ and $N/T \to 0$

$$\begin{aligned} e_{NT} &= N^{-1}T^{-1}\sum_{t=1}^{T} u_{t-1}'\Omega^{-1}u_{t-1} + o_p(1) \\ &\xrightarrow{p} \frac{1}{2}. \end{aligned}$$

It follows that $t_{gls} = d_{NT}/\sqrt{e_{NT}} \Rightarrow \mathcal{N}(0,1)$.

Finally, because in the defactorization approach the factors cancel out, the test is applied to the (estimated) nonstationary idiosyncratic components. In this case Moon and Perron (2004) showed that the test statistic has a standard normal limiting distribution. ■

https://doi.org/10.1017/S0266466608080067 Published online by Cambridge University Press