Econometric Theory, **25**, 2009, 1851–1868. Printed in the United States of America.
doi:10.1017/S0266466609990351

A NOTE ON THE POOLING OF INDIVIDUAL PANIC UNIT ROOT TESTS

JOAKIM WESTERLUND
University of Gothenburg

ROLF LARSSON
Uppsala University

One of the most cited studies in recent years within the field of nonstationary panel data analysis is that of Bai and Ng (2004), in which the authors propose PANIC, a new framework for analyzing the nonstationarity of panels with idiosyncratic and common components. The problem is that the asymptotic validity of PANIC as a platform for constructing pooled panel unit root tests based on averaging is not fully proven. This paper provides the required results, whose usefulness is verified through simulations.

1. INTRODUCTION

Consider the observed variable X_{it}, where $t = 1, \ldots, T$, and $i = 1, \ldots, N$ indexes the time series and cross-sectional units, respectively. The starting point of PANIC is to decompose X_{it} in the following way:

$$X_{it} = c_i + \lambda_i' F_t + e_{it}, \tag{1}$$

where the common factor F_t and loading λ_i represents the common component of X_{it}, while e_{it} represents the idiosyncratic component, which is assumed to be generated as

$$(1 - \rho_i L)e_{it} = D_i(L)\epsilon_{it}, \tag{2}$$

Previous versions of this paper were presented at the 2007 Granger Centre Conference in Honour of Paul Newbold and at seminars at Lund University, the University of Maastricht, and the University of Cambridge. The authors would like to thank conference and seminar participants, and in particular Jushan Bai, David Edgerton, Graham Elliott, Maria Holmlund, Serena Ng, Hashem Pesaran, Peter Phillips, Vanessa Smith, Robert Taylor, Jean-Pierre Urbain Takashi Yamagata, and three anonymous referees for many valuable comments and suggestions. Westerlund would also like to thank the Maastricht Research School of Economics of Technology and Organizations for its hospitality during a visit at the Department of Quantitative Economics at the University of Maastricht, where a part of this paper was written. Financial support from the Jan Wallander and Tom Hedelius Foundation under research grant numbers P2005-0117:1 and W2006-0068:1 is gratefully acknowledged. Address correspondence to Joakim Westerlund, Department of Economics, University of Gothenburg, P.O. Box 640, SE-405 30 Gothenburg, Sweden; e-mail: joakim.westerlund@economics.gu.se.

© 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

where $\epsilon_{it} \sim \text{iid}(0, \sigma_{\epsilon i}^2)$ and $D_i(L) = \sum_{j=0}^{\infty} d_{ij} L^j$ is a polynomial in the lag operator L.

Although PANIC has many uses, in this paper we focus on the aspect concerned with the testing of the null hypothesis $H_0 : \rho_i = 1$ for all i. The basic idea is to first replace e_{it} by $\widehat{e}_{it}$, the accumulated principal components estimator of Δe_{it}, and then to replace (2) by the augmented regression

$$\Delta \widehat{e}_{it} = \delta_{i0} \widehat{e}_{it-1} + \sum_{j=1}^{p} \delta_{ij} \Delta \widehat{e}_{it-j} + \text{error}. \tag{3}$$

The testing of H_0 is implemented as a test of the restriction that $\delta_{0i} = 0$ for all i, which can be done by using the Im, Peseran, and Shin (2003) type statistic

$$Z_{\widehat{e}} = \frac{\sqrt{N}(\overline{ADF}_{\widehat{e}} - \text{E}(\mathcal{B}))}{\sqrt{\text{var}(\mathcal{B})}},$$

where $\overline{ADF}_{\widehat{e}}$ is the average of the individual t-tests for testing $\delta_{0i} = 0$, henceforth denoted $ADF_{\widehat{e}}(i)$, while $\mathcal{B}$ is the usual Dickey and Fuller (1979) test distribution.[1] The advantage of this is that since $ADF_{\widehat{e}}(i)$ is asymptotically equivalent to $ADF_e(i)$, the test based on e_{it}, and since Z_e is asymptotically standard normal under H_0, $Z_{\widehat{e}}$ should have a normal limit too. This makes $Z_{\widehat{e}}$ far more general than the original Im et al. (2003) test, Z_X in our notation, which is not normal unless X_{it} is independent across i.[2]

The current paper points out a weakness in this pooling aspect of PANIC, which seems to have been largely overlooked in the literature. In particular, it is shown that the results of Bai and Ng (2004) are not enough to ensure that $Z_{\widehat{e}}$ is asymptotically normal under H_0. This is because the order of the bias incurred when replacing e_{it} with $\widehat{e}_{it}$ is not sufficiently sharp to ensure that it vanishes as N increases. This paper provides additional results showing that $Z_{\widehat{e}}$ is in fact normal. However, the test is also shown to be biased in small samples, which leads naturally to the proposal of a bias-corrected test, whose small-sample properties are shown by simulations to be much better than those of the original test.

2. MAIN RESULTS

The data generating process is taken directly from Bai and Ng (2004). For simplicity, however, in this paper we also assume that ϵ_{it} is normally distributed and that $D_i(L)$ is of finite order and equal to $D(L)$ for all i, where $D(L) = \sum_{j=0}^{m} d_j L^j$ with $d_0 = 1$ and $m < \infty$.

The $ADF_{\widehat{e}}(i)$ statistic can be written in a general notation as

$$ADF_{\widehat{e}}(i) = \frac{\widetilde{U}_i}{\sqrt{\widetilde{V}_i}},$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

where $\widetilde{U}_i$ and $\widetilde{V}_i$ are the relevant sample moments obtained from the augmented regression in (3). The corresponding moments with no lags are henceforth denoted $\widehat{U}_i$ and $\widehat{V}_i$, respectively, while those based on e_{it} and no lags are denoted U_i and V_i.

The Appendix shows that under H_0,

$$ADF_{\widehat{e}}(i) = ADF_e(i) + O_p\left(\frac{1}{C_{NT}}\right) = ADF_e(i) + \mathcal{R}_i, \tag{4}$$

where $C_{NT} = \min\{\sqrt{T}, \sqrt{N}\}$ and $\mathcal{R}_i$ is a remainder term. Bai and Ng (2004) argue that since $ADF_{\widehat{e}}(i)$ is asymptotically equivalent to $ADF_e(i)$, which is independent across i, $Z_{\widehat{e}}$ must have a limiting standard normal distribution, and this should hold irrespectively of the relative expansion rate of N and T, as long as they both go to infinity. However, this is not correct, as seen by noting that as $N, T \to \infty$,

$$Z_{\widehat{e}} = Z_e + \frac{1}{\sqrt{\text{var}(\mathcal{B})}}\sqrt{N}\,\overline{\mathcal{R}}_N = N(0,1) + O_p\left(\frac{\sqrt{N}}{C_{NT}}\right), \tag{5}$$

where the asymptotic normality of Z_e follows from Theorem 3.1 of Im et al. (2003), and where $\overline{\mathcal{R}}_N = \frac{1}{N}\sum_{i=1}^{N}\mathcal{R}_i$. Thus, even if we assume that $N/T \to 0$ as $N, T \to \infty$, the order of the remainder is still $O_p(1)$. In other words, although valid at the level of the individual unit, based on the results provided by Bai and Ng (2004), it is not clear that $Z_{\widehat{e}}$ is valid in the sense that its asymptotic distribution need not be normal and free of nuisance parameters. In the following theorem we show that the $O_p(1)$ remainder actually is $o_p(1)$, and hence that $Z_{\widehat{e}}$ is valid. In so doing, it is useful to define $R_{NT} = \min\{\sqrt{T}, N\}$ and $a_\infty = \lim_{N\to\infty} a_N$, where $a_N = \frac{1}{N}\sum_{i=1}^{N} a_{ii}$ and $a_{ij} = \lambda_i'\left(\sum_{i=1}^{N}\lambda_i\lambda_i'\right)^{-1}\lambda_j$.

THEOREM 1. *Under H_0 and the assumptions given above, as $N, T \to \infty$ with $N/T \to 0$,*

$$\sqrt{N}\,\overline{\mathcal{R}}_N = \frac{a_\infty}{\sqrt{N}}\sum_{j=0}^{m} d_j w_j + O\left(\sqrt{N}||d||^2\right) + O_p\left(\frac{\sqrt{N}}{R_{NT}}\right),$$

where $O(||d||^2)$ is a second-order term in d_j, $w_0 = -1.67$, and, for $j > 0$,

$$w_j \approx 1.05 - 5.54\left(\frac{j}{T}\right) + 19.4\left(\frac{j}{T}\right)^2 - 52.38\left(\frac{j}{T}\right)^3 + 131.95\left(\frac{j}{T}\right)^4 - 317.7\left(\frac{j}{T}\right)^5.$$

Remarks.

(i) A detailed account of the remainder in (5) is provided in the Appendix. However, it is instructive to note that $\mathcal{R}_i$ can be written

$$\mathcal{R}_i = -\frac{1}{\sqrt{V_i}}\mathcal{R}_{1i} + \frac{U_i}{V_i^{3/2}}\mathcal{R}_{2i} + O_p\left(\frac{1}{R_{NT}}\right)$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

where $\mathcal{R}_{1i} = \frac{1}{T} e_{iT} A_{iT}$, $\mathcal{R}_{2i} = \sigma^2_{\epsilon i} \frac{1}{T^2} \sum_{t=2}^{T} e_{it-1} A_{it-1}$, and $A_{it} = \widehat{\lambda}_i' \widehat{F}_t - \lambda_i' F_t$, with $\widehat{\lambda}_i$ and $\widehat{F}_t$ being the principal component estimates of λ_i and F_t, respectively. This shows how the magnitude of the bias depends on A_{it}, and hence on the accuracy of the estimated common component. In the extreme case when λ_i and F_t are known, $A_{it} = 0$, and so there is no bias. At the other end of the scale, if λ_i and F_t are unknown, and in addition difficult to estimate, then the bias is expected to be much larger.

(ii) If there is no serial correlation, so that $\sum_{j=0}^{m} d_j w_j = -1.67$, since both $\text{var}(\mathcal{B})$ and a_∞ are nonnegative, the bias in $Z_{\widehat{e}}$ is negative for $N < \infty$. Being left-tailed, this is suggestive of an oversized test, a finding confirmed by our simulations. In view of this, we recommend bias-adjusting $Z_{\widehat{e}}$ in the following obvious manner:

$$Z^+_{\widehat{e}} = Z_{\widehat{e}} - \frac{1}{\sqrt{\text{var}(\mathcal{B})}} \frac{\widehat{a}_N}{\sqrt{N}} \sum_{j=0}^{m} \widehat{d}_j w_j,$$

where $\widehat{a}_N$ is a_N with $\widehat{\lambda}_i$ in place of λ_i, while $\widehat{d}_j$ with $\widehat{d}_0 = 1$ is any consistent estimator of d_j.

(iii) Observe that for $\frac{a_\infty}{\sqrt{N}} \sum_{j=0}^{m} d_j w_j$ to dominate the $O_p(\sqrt{N}/R_{NT})$ remainder, it must be the case that $T < N^2$ and so, in combination with the condition $N/T \to 0$, the bias correction is only expected to work when $N << T < N^2$.[3]

3. SIMULATIONS

A small-scale simulation was conducted to assess the impact of our asymptotic results in small samples. The data are generated according to (1), where ΔF_t, $\lambda_i \sim N(0, I_3)$ while (2) is replaced by

$$(1 - \rho L) e_{it} = \delta \Delta e_{it-1} + \epsilon_{it},$$

where $\epsilon_{it} \sim N(0, \sigma^2_{\epsilon i})$ and $\sigma_{\epsilon i} \sim U(0, \phi)$. Thus, in this data-generating process, $d_j = \delta^j$ and so, since $w_j \to 1.05$ as $T \to \infty$ for all $j > 0$, $\sum_{j=0}^{\infty} \delta^j w_j \to -1.67 + 1.05 \sum_{j=1}^{\infty} \delta^j$ with $\sum_{j=1}^{\infty} \delta^j = \delta/1 - \delta$. It follows that while $\delta < 0$ is expected to lead to an increase in bias when compared to the no-serial-correlation case, $\delta > 0$ is expected to have a compensating effect, making the bias less negative.

The bias-adjusted test is implemented with $m = 100$ and $\widehat{d}_j = \widehat{\delta}^j$ for $j > 0$, where $\widehat{\delta}$ is the least squares estimate of δ under H_0 with e_{it} replaced by $\widehat{e}_{it}$. Also, to better isolate the effect of pooling, we have assumed that the true number of factors and lags is known. All computations have been performed in GAUSS using 100,000 replications.

The results of the size at the 5% level are reported in Table 1. These may be summarized as follows. The first thing to notice is that $Z_{\widehat{e}}$ can be quite distorted.

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

TABLE 1. Simulated size at the 5% level

		$\delta = 0$						$\delta = 0.5$						$\delta = -0.5$					
N	T	$Z_{\widehat{e}}^{+}$	$Z_{\widehat{e}}^{+a}$	$Z_{\widehat{e}}$	$Z_{\widehat{e}}^{b}$	Z_e	$ADF_{\widehat{e}}$	$Z_{\widehat{e}}^{+}$	$Z_{\widehat{e}}^{+a}$	$Z_{\widehat{e}}$	$Z_{\widehat{e}}^{b}$	Z_e	$ADF_{\widehat{e}}$	$Z_{\widehat{e}}^{+}$	$Z_{\widehat{e}}^{+a}$	$Z_{\widehat{e}}$	$Z_{\widehat{e}}^{b}$	Z_e	$ADF_{\widehat{e}}$
								$\sigma_{\epsilon i} \sim U(0,5)$											
10	20	1.1	1.1	13.7	14.5	6.6	7.3	1.4	3.0	12.3	4.4	6.1	6.5	3.0	2.5	26.9	58.5	6.7	10.7
	40	0.8	0.8	13.0	19.4	5.6	6.5	0.4	1.3	6.2	1.1	5.2	4.4	2.3	2.1	26.6	81.9	5.7	10.2
	80	0.7	0.8	12.8	22.2	5.1	6.1	0.3	1.1	4.8	0.6	5.0	3.6	2.5	2.4	27.3	90.3	5.2	10.4
20	40	1.2	1.2	10.0	12.7	5.5	6.1	0.8	1.5	4.5	0.1	4.9	4.6	5.9	5.5	30.3	97.3	5.7	9.8
	80	1.0	1.0	9.1	13.4	5.1	5.5	0.4	0.8	2.8	0.0	4.8	3.8	4.2	4.1	25.8	99.1	5.2	8.9
	160	0.8	0.8	8.6	13.1	4.9	5.3	0.5	0.9	2.7	0.0	4.8	3.7	3.0	3.0	22.0	99.5	4.9	8.1
40	80	1.5	1.5	7.5	6.7	4.9	5.3	0.9	1.1	2.4	0.0	4.3	4.3	5.4	5.2	21.0	100.0	5.1	7.1
	160	1.5	1.5	7.3	7.6	4.9	5.1	1.5	1.7	3.6	0.0	4.6	4.3	2.6	2.6	13.8	100.0	5.0	6.1
	320	1.3	1.3	7.1	7.9	4.6	5.0	1.9	2.0	4.0	0.0	4.5	4.4	1.7	1.7	11.1	100.0	4.7	5.7
80	160	2.3	2.3	7.0	5.8	4.6	5.1	2.8	2.9	4.6	0.0	4.2	4.7	3.3	3.3	11.2	100.0	4.8	5.5
	320	2.2	2.2	6.9	6.6	4.5	5.0	3.2	3.3	5.0	0.0	4.3	4.7	2.5	2.5	9.3	100.0	4.6	5.3
	640	2.2	2.2	6.7	6.9	4.4	4.9	3.4	3.4	5.2	0.0	4.2	4.7	2.2	2.2	8.3	100.0	4.4	5.1
								$\sigma_{\epsilon i} \sim U(0,10)$											
10	20	1.1	1.2	14.2	18.8	6.6	7.5	2.5	5.3	17.7	10.6	6.1	7.7	2.8	2.2	25.8	54.5	6.7	10.6
	40	0.9	1.1	14.5	25.6	5.6	7.0	0.8	2.3	8.9	2.8	5.2	5.2	2.4	2.2	26.6	80.1	5.7	10.5
	80	1.1	1.2	15.4	30.6	5.1	7.1	0.6	1.8	6.5	1.6	5.0	4.3	3.2	3.0	29.8	90.2	5.2	11.3
20	40	1.5	1.6	12.3	25.2	5.5	6.8	2.0	3.9	9.4	1.3	4.9	5.6	7.2	6.7	33.9	97.3	5.7	10.8
	80	1.7	1.8	13.1	31.3	5.1	6.6	0.7	1.4	4.0	0.2	4.8	4.4	8.2	7.9	36.3	99.5	5.2	11.1
	160	1.8	2.0	13.7	34.9	4.9	6.7	0.5	1.0	3.0	0.1	4.8	3.9	9.8	9.6	38.8	99.8	4.9	11.5
40	80	2.6	2.6	11.1	27.3	4.9	6.1	0.9	1.5	3.1	0.0	4.3	4.4	21.0	20.6	48.3	100.0	5.1	10.7
	160	2.4	2.5	10.8	28.0	4.9	5.8	0.4	0.6	1.5	0.0	4.6	3.9	19.2	19.1	45.9	100.0	5.0	10.3
	320	2.1	2.2	9.8	26.0	4.6	5.6	0.4	0.6	1.3	0.0	4.5	3.6	16.6	16.5	42.0	100.0	4.7	9.8
80	160	2.6	2.6	7.6	12.3	4.6	5.2	0.4	0.5	0.9	0.0	4.2	3.9	25.3	25.1	45.3	100.0	4.8	8.6
	320	2.4	2.4	7.2	11.2	4.5	5.0	0.5	0.6	1.1	0.0	4.3	3.9	13.0	13.0	29.2	100.0	4.6	7.2
	640	2.3	2.3	7.0	10.7	4.4	5.0	0.9	1.0	1.7	0.0	4.2	4.1	7.2	7.2	19.4	100.0	4.4	6.4

[a]The test is adjusted using the true bias.

[b]The test is based on no lags.

Note: The data are generated according to (1) with $\Delta F_t, \lambda_i \sim N(0, I_3)$ and $\Delta e_{it} = \delta \Delta e_{it-1} + \epsilon_{it}$, where $\epsilon_{it} \sim N(0, \sigma_{\epsilon i}^2)$. $Z_{\widehat{e}}^{+}$ refers to the bias- adjusted panel test based on $\widehat{e}_{it}$; $Z_{\widehat{e}}$ and Z_e refer to the unadjusted tests based on $\widehat{e}_{it}$ and e_{it}, respectively. $ADF_{\widehat{e}}$ refers to the individual test based on $\widehat{e}_{it}$. All tests are based on the true number of factors and on the true lag length.

TABLE 2. Simulated size-adjusted power at the 5% level

N	T	$Z^{+}_{\widehat{e}}$	$Z^{+\mathrm{a}}_{\widehat{e}}$	$Z_{\widehat{e}}$	$Z^{\mathrm{b}}_{\widehat{e}}$	Z_e	$ADF_{\widehat{e}}$
10	20	6.5	6.5	6.5	6.5	8.3	5.3
	40	9.6	9.6	9.6	9.3	13.5	5.8
	80	18.0	18.1	18.0	16.6	29.4	7.0
20	40	15.6	15.5	15.6	14.8	20.9	5.9
	80	38.8	38.6	38.8	36.6	51.0	7.1
	160	84.3	84.2	84.3	82.5	94.6	10.0
40	80	68.3	68.1	68.3	69.4	79.7	7.2
	160	99.1	99.1	99.1	99.2	99.9	10.0
	320	100.0	100.0	100.0	100.0	100.0	18.1
80	160	100.0	100.0	100.0	100.0	100.0	10.1
	320	100.0	100.0	100.0	100.0	100.0	18.2
	640	100.0	100.0	100.0	100.0	100.0	43.2

Note: The equation for e_{it} is given by $e_{it} = 0.99e_{it-1} + \epsilon_{it}$, where $\epsilon_{it} \sim N(0, \sigma^2_{\epsilon i})$ and $\sigma_{\epsilon i} \sim U(0, 5)$. See Table 1 for an explanation of the remaining features.

In particular, as expected, the test is upwards biased when $\delta = 0$. We also see that the distortions get worse when $\delta = -0.5$, and that they move in the opposite direction when $\delta = 0.5$.

Second, it interesting to see how the size is affected by N and T on the one hand, and by $\sigma^2_{\epsilon i}$ on the other hand.[4] With the upper support ϕ of $\sigma_{\epsilon i}$ fixed, we see that while the effect of increasing T seems to be small, as expected, a larger N generally results in a substantial reduction in size distortion. By contrast, if we fix N and T, and instead let ϕ increase, then according to theory, the larger bias should tend to make $Z_{\widehat{e}}$ more distorted, and this is exactly what we see in Table 1.

Third, note that while the size of $Z_{\widehat{e}}$ is greatly affected by the parametrization of the data generating process, as expected, the size of the other tests is quite robust in this respect. Note especially how bias adjusting can greatly improve upon the accuracy of $Z_{\widehat{e}}$, and that there is little or no variation depending on whether the bias is treated as known or not.

Finally, as for the size-adjusted power reported in Table 2, we see that the better size accuracy obtained by bias adjusting has almost no cost in terms of power. As expected, Z_e is most powerful, while $ADF_{\widehat{e}}(i)$ is least powerful. The other tests are only slightly less powerful than Z_e, and the difference gets smaller as N and T grows.

4. CONCLUDING REMARKS

In this paper we point out a flaw in the theoretical results provided by Bai and Ng (2004) for their PANIC unit root methodology. The problem lies in the order of

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

the bias incurred when replacing the idiosyncratic component e_{it} by its estimated counterpart $\widehat{e}_{it}$, which is not sharp enough to ensure that PANIC can be used for pooling across i. The current paper provides more exact results establishing that PANIC can in fact be used for pooling purposes. They also lead naturally to the proposal of a bias-adjusted test whose small-sample properties are shown to be much better than those of the original test.

NOTES

1. Although we focus on $Z_{\widehat{e}}$ in our analysis, our results apply to all average type tests; see Breitung and Pesaran (2008) for a recent survey of the existing panel unit root literature.

2. Another advantage, even in comparison to other studies that also permit for common factors, is that in PANIC the factors need not be stationary. This makes $Z_{\widehat{e}}$ very general indeed, and is probably one of the main reasons why PANIC has become so popular in both applied and theoretical work; see, for example, Bai and Ng (2005), Gengenbach, Palm, and Urbain (2006), and Westerlund (2008).

3. The condition that $N/T \to 0$ as $N, T \to \infty$ is standard even when testing for unit roots in independent panels. The reason for this is the assumed heterogeneity, whose elimination induces an estimation error in T, which is then aggravated when pooling across N, see Im et al. (2003) for a discussion. By contrast, the condition that $T < N^2$ is new. Intuitively, if $T > N^2$ then $\sqrt{N}/R_{NT}$ is equal to $1/\sqrt{N}$, suggesting that there will be another remainder term that is of the same order as $\frac{a_\infty}{\sqrt{N}} \Sigma_{j=0}^{m} d_j w_j$. Thus, although the remainder still vanishes as $N \to \infty$, the bias adjustment will no longer be effective.

4. Note how N and T reflect the condition that $N << T < N^2$.

REFERENCES

Bai, J. (2003) Inferential theory for factor models of large dimensions. *Econometrica* 71, 135–171.

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

Bai, J. & S. Ng (2005) A new look at panel stationarity tests and the PPP hypothesis. In D.W.K. Andrews and J.H. Stock (eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*. Cambridge University Press.

Breitung, J. & M.H. Pesaran (2008) Unit roots and cointegration in panels. In L. Mátyás & P. Sevestre (eds.), *The Econometrics of Panel Data*, 3rd ed., pp. 279–322. Kluwer.

Dickey, D.A. & W.A. Fuller (1979) Distribution of the estimator for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.

Gengenbach, C., F.C. Palm, & J.-P. Urbain (2006) Panel cointegration testing in the presence of common factors. *Oxford Bulletin of Economics and Statistics* 68, 683–719.

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

Larsson, R. (1998) Bartlett corrections for unit root tests. *Journal of Time Series Analysis* 19, 425–438.

Magnus, J.R. (1978) The moments of products of quadratic forms in normal variables. *Statistica Neerlandica* 32, 201–210.

Westerlund, J. (2008) Panel cointegration tests of the Fisher effect. *Journal of Applied Econometrics* 23, 193–233.

Westerlund, J. & R. Larsson (2008) A Note on the Pooling of Individual PANIC Unit Root Tests: Supplement. Unpublished manuscript. Available at http://www.hgu.gu/item.aspx?id=17810.

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

APPENDIX: Mathematical Proofs

LEMMA A.1. *Under H_0, as $T \to \infty$,*

(a) $\widetilde{V}_i = \widehat{V}_i + O_p\left(\frac{1}{T}\right)$,

(b) $\widetilde{U}_i = \widehat{U}_i + O_p\left(\frac{1}{\sqrt{T}}\right)$.

Proof of Lemma A.1. We omit this proof in this paper. The required results are obtained by combining the rules for projections with the order results of Bai and Ng (2004). All details can be found in Westerlund and Larsson (2008). ■

LEMMA A.2. *Under H_0, as $N, T \to \infty$,*

(a) $\widehat{V}_i = V_i - 2\mathcal{R}_{2i} + O_p\left(\frac{1}{R_{NT}}\right)$,

(b) $\widehat{U}_i = U_i - \mathcal{R}_{1i} + O_p\left(\frac{1}{R_{NT}}\right)$.

Proof of Lemma A.2. We begin with (a). By definition,

$$\widehat{V}_i = \widehat{\sigma}_{\epsilon i}^2 \frac{1}{T^2} \sum_{t=2}^{T} \widehat{e}_{it-1}^2.$$

Consider $\widehat{\sigma}_{\epsilon i}^2$, which can be expanded as

$$\widehat{\sigma}_{\epsilon i}^2 = \frac{1}{T} \sum_{t=2}^{T} (\Delta \widehat{e}_{it} - \widehat{b}_i \widehat{e}_{it-1})^2$$

$$= \frac{1}{T} \sum_{t=2}^{T} (\Delta \widehat{e}_{it})^2 - 2\widehat{b}_i \frac{1}{T} \sum_{t=2}^{T} \widehat{e}_{it-1} \Delta \widehat{e}_{it} + T\widehat{b}_i^2 \frac{1}{T^2} \sum_{t=2}^{T} \widehat{e}_{it-1}^2, \tag{A1}$$

where $\widehat{b}_i$ is the least squares slope estimator when regressing $\Delta \widehat{e}_{it}$ onto $\widehat{e}_{it-1}$.

Note that

$$\Delta X_{it} = \lambda_i' \Delta F_t + \Delta e_{it} = \lambda_i' f_t + \Delta e_{it}.$$

Thus, if we denote by $\widehat{\lambda}_i$ and $\widehat{f}_t$ the principal component estimates of λ_i and f_t, respectively, then $\Delta \widehat{e}_{it}$ can be written as

$$\Delta \widehat{e}_{it} = \Delta X_{it} - \widehat{\lambda}_i' \widehat{f}_t = \Delta e_{it} - (\widehat{\lambda}_i' \widehat{f}_t - \lambda_i' f_t) = \Delta e_{it} - a_{it},$$

where

$$a_{it} = \lambda_i' H^{-1}(\widehat{f}_t - H f_t) + (\widehat{\lambda}_i - (H^{-1})' \lambda_i)' \widehat{f}_t = \lambda_i' H^{-1} v_t + d_i' \widehat{f}_t,$$

with $d_i = \widehat{\lambda}_i - (H^{-1})' \lambda_i$ and $v_t = \widehat{f}_t - H f_t$, and where H is a scaling matrix defined in Bai and Ng (2004). This implies that

$$\frac{1}{T} \sum_{t=2}^{T} (\Delta \widehat{e}_{it})^2 = \frac{1}{T} \sum_{t=2}^{T} (\Delta e_{it} - a_{it})^2 = \frac{1}{T} \sum_{t=2}^{T} (\Delta e_{it})^2 + \frac{1}{T} \sum_{t=2}^{T} a_{it}^2 - \frac{2}{T} \sum_{t=2}^{T} \Delta e_{it} a_{it}$$

$$= \frac{1}{T} \sum_{t=2}^{T} (\Delta e_{it})^2 + I - II. \tag{A2}$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

By using the same arguments as in Lemma B.1 of Bai and Ng (2004), $I = O_p(1/C_{NT}^2)$. Part II can be written as

$$II = 2\lambda_i' H^{-1} \frac{1}{T} \sum_{t=2}^{T} \Delta e_{it} v_t + 2d_i' \frac{1}{T} \sum_{t=2}^{T} \Delta e_{it} \widehat{f}_t = 2d_i' \frac{1}{T} \sum_{t=2}^{T} \Delta e_{it} \widehat{f}_t + O_p\left(\frac{1}{C_{NT}^2}\right)$$

$$= 2d_i' \frac{1}{T} \sum_{t=2}^{T} \Delta e_{it} (\widehat{f}_t - H f_t) + 2d_i' H \frac{1}{T} \sum_{t=2}^{T} \Delta e_{it} f_t + O_p\left(\frac{1}{C_{NT}^2}\right),$$

where the second equality follows by Lemma B.1 of Bai (2003). By applying $||AB|| \leq ||A||||B||$, where $||A|| = \sqrt{\text{tr}(A'A)}$ is the Euclidean norm, and the Cauchy-Schwarz and triangle inequalities, we get

$$|II| \leq 2||d_i|| \frac{1}{T} \sum_{t=2}^{T} ||\Delta e_{it} (\widehat{f}_t - H f_t)|| + 2||d_i||||H|| \frac{1}{T} \sum_{t=2}^{T} ||\Delta e_{it} f_t|| + O_p\left(\frac{1}{C_{NT}^2}\right)$$

$$= O_p\left(\frac{1}{R_{NT}}\right) O_p\left(\frac{1}{C_{NT}}\right) + O_p\left(\frac{1}{R_{NT}}\right) O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{C_{NT}^2}\right),$$

where $||H|| = O_p(1)$, $||d_i|| = O_p(1/R_{NT})$ by Lemma 1(c) of Bai and Ng (2004), and

$$\frac{1}{T} \sum_{t=2}^{T} ||\widehat{f}_t - H f_t||^2 = O_p\left(\frac{1}{C_{NT}^2}\right) \tag{A3}$$

by Lemma A.1 of Bai (2003). This implies that $II = O_p(1/C_{NT}^2)$, and so (A2) becomes

$$\frac{1}{T} \sum_{t=2}^{T} (\Delta \widehat{e}_{it})^2 = \frac{1}{T} \sum_{t=2}^{T} (\Delta e_{it})^2 + O_p\left(\frac{1}{C_{NT}^2}\right). \tag{A4}$$

Moreover, since $\widehat{e}_{i1} = 0$ by definition,

$$\frac{1}{T^2} \sum_{t=2}^{T} \widehat{e}_{it}^2 = \frac{1}{T^2} \sum_{t=2}^{T} (e_{it} - e_{i1} - A_{it})^2 = \frac{1}{T^2} \sum_{t=2}^{T} e_{it}^2 - \frac{2}{T^2} \sum_{t=2}^{T} e_{it} A_{it} + \frac{1}{T} e_{i1}^2 + \frac{1}{T^2} \sum_{t=2}^{T} A_{it}^2$$

$$-2e_{i1} \left(\frac{1}{T^2} \sum_{t=2}^{T} e_{it} - \frac{1}{T^2} \sum_{t=2}^{T} A_{it} \right)$$

$$= \frac{1}{T^2} \sum_{t=2}^{T} e_{it}^2 - \frac{2}{T^2} \sum_{t=2}^{T} e_{it} A_{it} + I + II - III. \tag{A5}$$

Part I is obviously $O_p(1/T)$, while II is $O_p(1/C_{NT}^2)$ by Lemma C.1(i) of Bai and Ng (2004).

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

Consider III. The first term within the parenthesis is $O_p(1/\sqrt{T})$. As for the second term, note that if V_t and $\widehat{F}_t$ are the cumulative sums of υ_t and $\widehat{f}_t$, respectively, then

$$A_{it} = \lambda_i' H^{-1} V_t + d_i' \widehat{F}_t.$$

This implies that

$$\begin{aligned}\frac{1}{T^2}\sum_{t=2}^{T} A_{it} &= \lambda_i' H^{-1} \frac{1}{T^2}\sum_{t=2}^{T} V_t + d_i' H \frac{1}{T^2}\sum_{t=2}^{T} F_t + d_i' \frac{1}{T^2}\sum_{t=2}^{T} (\widehat{F}_t - H F_t) \\ &= O_p\left(\frac{1}{\sqrt{T}C_{NT}}\right) + O_p\left(\frac{1}{R_{NT}}\right) O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{R_{NT}}\right) O_p\left(\frac{1}{\sqrt{T}C_{NT}}\right),\end{aligned}$$

where $V_t/\sqrt{T} = O_p(1/C_{NT})$ by equation (A3) of Bai and Ng (2004). Thus, by collecting all the terms, (A5) reduces to

$$\frac{1}{T^2}\sum_{t=2}^{T} \widehat{e}_{it}^2 = \frac{1}{T^2}\sum_{t=2}^{T} e_{it}^2 - \frac{2}{T^2}\sum_{t=2}^{T} e_{it} A_{it} + O_p\left(\frac{1}{R_{NT}}\right). \tag{A6}$$

By using (A4), (A6), part (b), and the fact that $T\widehat{b}_i = O_p(1)$ under H_0, (A1) becomes

$$\widehat{\sigma}_{\epsilon i}^2 = \frac{1}{T}\sum_{t=2}^{T} (\Delta e_{it})^2 + O_p\left(\frac{1}{C_{NT}^2}\right) = \sigma_{\epsilon i}^2 + O_p\left(\frac{1}{C_{NT}^2}\right). \tag{A7}$$

This result, together with (A6), implies part (a).

Next, consider (b). Again, by definition,

$$\widehat{U}_i = \frac{1}{T}\sum_{t=2}^{T} \widehat{e}_{it-1}\Delta\widehat{e}_{it}.$$

Note that $\widehat{e}_{it}^2 = (\widehat{e}_{it-1} + \Delta\widehat{e}_{it})^2 = \widehat{e}_{it-1}^2 + (\Delta\widehat{e}_{it})^2 + 2\widehat{e}_{it-1}\Delta\widehat{e}_{it}$, from which it follows that

$$\begin{aligned}\frac{1}{T}\sum_{t=2}^{T} \widehat{e}_{it-1}\Delta\widehat{e}_{it} &= \frac{1}{2T}\sum_{t=2}^{T} (\widehat{e}_{it}^2 - \widehat{e}_{it-1}^2 - (\Delta\widehat{e}_{it})^2) \\ &= \frac{1}{2T}\widehat{e}_{iT}^2 - \frac{1}{2T}\widehat{e}_{i1}^2 - \frac{1}{2T}\sum_{t=2}^{T} (\Delta\widehat{e}_{it})^2,\end{aligned} \tag{A8}$$

and by applying the same trick to e_{it}^2, we have

$$\frac{1}{T}\sum_{t=2}^{T} e_{it-1}\Delta e_{it} = \frac{1}{2T}e_{iT}^2 - \frac{1}{2T}e_{i1}^2 - \frac{1}{2T}\sum_{t=2}^{T} (\Delta e_{it})^2. \tag{A9}$$

Now, the terms in the middle of the right-hand side of (A8) and (A9) are clearly $O_p(1/T)$ as $\widehat{e}_{i1} = 0$, while $e_{i1} = O_p(1)$. Also, by using (A4), the difference between the third terms

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

is $O_p(1/C_{NT}^2)$. As for the first term, note that

$$\begin{aligned}\frac{1}{T}\widehat{e}_{iT}^2 &= \frac{1}{T}(e_{iT}-e_{i1}-A_{iT})^2 \\ &= \frac{1}{T}e_{iT}^2-\frac{2}{T}e_{iT}A_{iT}+\frac{1}{T}e_{i1}^2+\frac{1}{T}A_{iT}^2-\frac{2}{T}e_{i1}(e_{iT}-A_{iT}) \\ &= \frac{1}{T}e_{iT}^2-\frac{2}{T}e_{iT}A_{iT}+I+II-III. \end{aligned} \tag{A10}$$

Again, $I = O_p(1/T)$. Moreover, by using $(a+b)^2 \le 2(a^2+b^2)$, the triangle inequality, $||AB|| \le ||A||||B||$, and (A3), II can be written as

$$|II| = \frac{1}{T}(\lambda_i' H^{-1} V_T + d_i' \widehat{F}_T)^2 \;\le\; 2||\lambda_i' H^{-1}||^2 \frac{1}{T}\sum_{t=2}^{T}||v_t||^2 + 2||d_i||^2 \frac{1}{T}\sum_{t=2}^{T}||\widehat{f}_t||^2$$

$$= O_p\left(\frac{1}{C_{NT}^2}\right)+O_p\left(\frac{1}{R_{NT}^2}\right)O_p(1),$$

which is clearly $O_p(1/C_{NT}^2)$.

Part III is dominated by $2/T(e_{i1}e_{iT})$, which is $O_p(1/\sqrt{T})$. Therefore, by adding the terms, (A10) simplifies to

$$\frac{1}{T}\widehat{e}_{iT}^2 = \frac{1}{T}e_{iT}^2-\frac{2}{T}e_{iT}A_{iT}+O_p\left(\frac{1}{R_{NT}}\right),$$

which it turn implies (b), and thus the proof of Lemma A.2 is complete. ■

Proof of Theorem 1. By part (a) of Lemmas A.1 and A.2, and a first-order Taylor expansion of the inverse square root, we get

$$\frac{1}{\sqrt{\widetilde{V}_i}} = \frac{1}{\sqrt{\widehat{V}_i}}+O_p\left(\frac{1}{T}\right) = \frac{1}{\sqrt{V_i}}+\frac{1}{V_i^{3/2}}\mathcal{R}_{2i}+O_p\left(\frac{1}{R_{NT}}\right),$$

and by further application of part (b) of these lemmas,

$$ADF_{\widehat{e}}(i) = ADF_e(i)-\frac{1}{\sqrt{V_i}}\mathcal{R}_{1i}+\frac{U_i}{V_i^{3/2}}\mathcal{R}_{2i}+O_p\left(\frac{1}{R_{NT}}\right) = ADF_e(i)+\mathcal{R}_i.$$

It follows that

$$\begin{aligned}\sqrt{N}\,\overline{\mathcal{R}}_N &= -\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{\sqrt{V_i}}\mathcal{R}_{1i}+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{U_i}{V_i^{3/2}}\mathcal{R}_{2i}+O_p\left(\frac{\sqrt{N}}{R_{NT}}\right) \\ &= -I+II+O_p\left(\frac{\sqrt{N}}{R_{NT}}\right). \end{aligned} \tag{A11}$$

Consider I. From the proof of Theorem 3 in Bai (2003), we get

$$\frac{1}{\sigma_{\epsilon i}}A_{it} = \frac{1}{\sigma_{\epsilon i}}\sum_{j=1}^{N}\sigma_{\epsilon j}a_{ij}e_{jt}^{\circ}+\sum_{s=2}^{T}b_{st}\epsilon_{is}^{\circ}+O_p\left(\frac{\sqrt{T}}{\sqrt{N}C_{NT}^2}\right) = A_{it}^{\circ}+O_p\left(\frac{\sqrt{T}}{\sqrt{N}C_{NT}^2}\right),$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

where $b_{st} = \sum_{s=2}^{t} f_s' \left(\sum_{s=2}^{T} f_s f_s' \right)^{-1} f_s$, $\epsilon_{is}^{\circ} = \epsilon_{is}/\sigma_{\epsilon i}$, and $e_{is}^{\circ} = e_{is}/\sigma_{\epsilon i}$. A_{it}° can be written

$$A_{it}^{\circ} = \frac{1}{\sigma_{\epsilon i}} \sum_{j=1}^{N} \sigma_{\epsilon j} a_{ij} (e_{jt}^{\circ} - e_{j1}^{\circ}) + \sum_{s=2}^{T} b_{st} (e_{is}^{\circ} - e_{is-1}^{\circ})$$

$$= \frac{1}{\sigma_{\epsilon i}} \sum_{j=1}^{N} \sigma_{\epsilon j} a_{ij} (e_{jt}^{\circ} - e_{j1}^{\circ}) - b_{2t} e_{i1}^{\circ} - \sum_{s=3}^{T} (\Delta b_{st}) e_{is-1}^{\circ} + b_{Tt} e_{iT}^{\circ} = G_{it}' e^{\circ}, \tag{A12}$$

where $\Delta b_{st} = b_{st} - b_{(s-1)t}$ and e° is the NT-dimensional vector of stacked observations on e_{jt}°. $G_{it} = (g_{i1t}', \ldots, g_{iNt}')'$ stacks the T-dimensional vector g_{ijt}, which is such that if $i \neq j$, then $g_{ijt} = \sigma_{\epsilon j} a_{ij} (q_t - q_1)/\sigma_{\epsilon i}$, where q_t is a T-dimensional vector of zeroes except for a one at position t, whereas if $i = j$, then $g_{iit} = a_{ii} q_t - (a_{ii} + b_{2t}, \Delta b_{3t}, \ldots, \Delta b_{Tt}, -b_{Tt})'$. Hence, since $e_{iT}^{\circ} = (0, \ldots, 0, q_T', 0, \ldots, 0) e^{\circ}$, where q_T is at position i, it is clear that $\mathcal{R}_{1i}^{\circ} = e_{iT}^{\circ} A_{iT}^{\circ} = (e^{\circ})' G_i e^{\circ}$, where $G_i' = G_{iT}(0, \ldots, 0, q_T', 0, \ldots, 0)$. Similarly, $V_i^{\circ} = T^2 V_i / \sigma_{\epsilon i}^4 = (e^{\circ})' D_i e^{\circ}$, where D_i is a block diagonal $NT \times NT$ matrix with all diagonal blocks equal to zero except for the block at position i, which is given by

$$D_{ii} = D = \begin{bmatrix} I_{T-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

These results suggest that as $N, T \to \infty$,

$$I = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{\sqrt{V_i}} \mathcal{R}_{1i} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{\sqrt{V_i^{\circ}}} \mathcal{R}_{1i}^{\circ} + O_p\left(\frac{1}{C_{NT}^2} \right)$$

$$= \lim_{N \to \infty} \sqrt{N} \mathrm{E} \left(\frac{\mathcal{R}_{1i}^{\circ}}{\sqrt{V_i^{\circ}}} \right) + O_p\left(\frac{1}{C_{NT}^2} \right). \tag{A13}$$

Letting $M_j = L^j + (L')^j$, where L is now the $T \times T$ lag matrix, it holds that $\mathrm{var}(\Delta e^{\circ}) = I_N \otimes \Sigma$, where

$$\Sigma = I_T + \sum_{j=1}^{m} d_j M_j.$$

Thus, if $e^{\circ\circ} = (I_N \otimes \Sigma)^{-1/2} e^{\circ}$, the moment-generating function of $(\mathcal{R}_{1i}^{\circ}, V_i^{\circ})$ becomes

$$\varphi_i(r, -v) = \mathrm{E}\left(\exp(r \mathcal{R}_{1i}^{\circ} - v V_i^{\circ}) \right)$$

$$= \frac{1}{\left((2\pi)^T \det(\Sigma) \right)^{N/2}} \int \exp\left(r \mathcal{R}_{1i}^{\circ} - v V_i^{\circ} - \frac{1}{2} (\Delta e^{\circ})' (I_N \otimes \Sigma)^{-1} (\Delta e^{\circ}) \right) d(e^{\circ})$$

$$= \frac{1}{(2\pi)^{NT/2}} \int \exp\left(r \mathcal{R}_{1i}^{\circ\circ} - v V_i^{\circ\circ} - \frac{1}{2} (\Delta e^{\circ\circ})' (\Delta e^{\circ\circ}) \right) d(e^{\circ\circ})$$

$$= \frac{1}{(2\pi)^{NT/2}} \int \exp\left(-\frac{1}{2} (e^{\circ\circ})' P_i e^{\circ\circ} \right) d(e^{\circ\circ}) = \frac{1}{\sqrt{\det(P_i)}}, \tag{A14}$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

with $\mathcal{R}_{1i}^{\circ\circ} = (e^{\circ\circ})'(I_N \otimes \Sigma)^{1/2} G_i (I_N \otimes \Sigma)^{1/2} e^{\circ\circ}$ and where $V_i^{\circ\circ}$ is $\mathcal{R}_{1i}^{\circ\circ}$ with D_i in place of G_i. This implies that P_i can be written as $P_i = P_{0i} + r P_{1i} = P_{0i} + P_{mi}$, where $P_{1i} = -2(I_N \otimes \Sigma)^{1/2} G_i (I_N \otimes \Sigma)^{1/2}$ and

$$P_{0i} = \begin{bmatrix} P_{01i} & 0 & \cdots & 0 \\ 0 & P_{02i} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & P_{0Ni} \end{bmatrix}.$$

The diagonal matrices P_{0ij} are such that if $j \neq i$, then

$$P_{0ji} = P = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix},$$

whereas if $j = i$, then $P_{0ii} = P_0 = P - 2\upsilon \Sigma^{1/2} D \Sigma^{1/2}$.

Consider the identity

$$\frac{1}{(V_i^{\circ})^{n/2}} = \frac{1}{\Gamma(n/2)} \int_0^{\infty} \upsilon^{(n/2-1)} \exp(-\upsilon V_i^{\circ}) d\upsilon, \tag{A15}$$

where n is a positive integer, which, via the Fubini theorem, gives

$$\begin{aligned} \mathrm{E}\left(\frac{\mathcal{R}_{1i}^{\circ}}{\sqrt{V_i^{\circ}}} \right) &= \frac{1}{\Gamma(1/2)} \int_0^{\infty} \frac{1}{\sqrt{\upsilon}} \frac{\partial}{\partial r} \mathrm{E}\Big(\exp\left(r\mathcal{R}_{1i}^{\circ} - \upsilon V_i^{\circ} \right) \Big) d\upsilon \\ &= \frac{1}{\Gamma(1/2)} \int_0^{\infty} \frac{1}{\sqrt{\upsilon}} \frac{\partial}{\partial r} \varphi_i(r, -\upsilon) d\upsilon, \end{aligned} \tag{A16}$$

where the derivative is taken at $r = 0$. To find $\frac{\partial}{\partial r}\varphi_i(r, -\upsilon)$, we follow Larsson (1998) and use the Taylor expansion

$$\begin{aligned} \frac{1}{\sqrt{\det(P_i)}} &= \frac{1}{\sqrt{\det(P_{0i})}} \mathrm{E}\left(\exp\left(-\frac{1}{2} X' P_{mi} X \right) \right) \\ &= \frac{1}{\sqrt{\det(P_{0i})}} \left(1 - \frac{1}{2} \mathrm{E}\left(X' P_{mi} X \right) + \frac{1}{8} \mathrm{E}\left(\left(X' P_{mi} X \right)^2 \right) + \ldots \right), \end{aligned} \tag{A17}$$

where X is a generic NT-dimensional vector with normal distribution and covariance matrix P_{0i}^{-1}. Also, by using the results of Magnus (1978), we have that $\mathrm{E}(X' P_{mi} X) = \mathrm{tr}(P_{0i}^{-1} P_{mi})$, which together with (A14) and (A17), yields

$$\varphi_i(r, -\upsilon) = \frac{1}{\sqrt{\det(P_{0i})}} \left(1 - \frac{1}{2} \mathrm{tr}\Big(P_{0i}^{-1} P_{mi} \Big) + O(r^2) \right). \tag{A18}$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

This, in turn, implies that

$$\frac{\partial}{\partial r}\varphi_i(r,-\upsilon) = -\frac{\mathrm{tr}\left(P_{0i}^{-1}P_{1i}\right)}{2\sqrt{\det(P_{0i})}}. \tag{A19}$$

Lemma A.3 expresses the numerator of $\partial/\partial r(\varphi_i(r,-\upsilon))$ in terms of $x=\sqrt{2\upsilon}T$ and $y=j/T$, where $j=1,\ldots,m$.

LEMMA A.3. *We have*

$$\mathrm{tr}\left(P_{0i}^{-1}P_{1i}\right) = -\frac{2Ta_{ii}}{x\cosh(x)}\left(\sinh(x)+\sum_{j=1}^{m}d_j\delta(x,y)\right)+O_p(\sqrt{T}),$$

where $\delta(x,y)=\frac{x(1-y)}{\cosh(x)}\cosh(xy)-\frac{1}{2}\sinh(xy)$.

Proof of Lemma A.3. We omit this proof in this paper. All details can be found in Westerlund and Larsson (2008). ■

Let $D_0=P-2\upsilon D$, then via Taylor expansion,

$$P_0 = D_0-\upsilon\sum_{j=1}^{m}d_j\left(M_jD+DM_j\right)+O(||d||^2),$$

where $O(||d||^2)$ is a second-order term in d_j. By using this, and the fact that $\det(P_{0i})=\cosh(x)+O(1/T)$ as in Larsson (1998), we find that

$$\begin{aligned}\frac{1}{\sqrt{\det(P_{0i})}} &= \frac{1}{\sqrt{\det(P_0)}}\\ &= \frac{1}{\sqrt{\cosh(x)}}\left(1+\frac{\upsilon}{2}\sum_{j=1}^{m}d_j\mathrm{tr}\left(D_0^{-1}(M_jD+DM_j)\right)\right)+O\left(\frac{1}{T}\right)+O(||d||^2),\end{aligned}$$

where

$$\mathrm{tr}\left(D_0^{-1}(M_jD+DM_j)\right) = 2T^2\frac{\alpha(x,y)}{x\cosh(x)}+O(T)$$

with $\alpha(x,y)=(1-y)\sinh(x(1-y))$. This implies that

$$\frac{1}{\sqrt{\det(P_{0i})}} = \frac{1}{\sqrt{\cosh(x)}}\left(1-\frac{x}{2\cosh(x)}\sum_{j=1}^{m}d_j\alpha(x,y)\right)+O\left(\frac{1}{T}\right)+O(||d||^2),$$

which, together with (A19) and Lemma A.3, in turn suggests that

$$\begin{aligned}\frac{\partial}{\partial r}\varphi_i(r,-\upsilon) = \frac{Ta_{ii}}{x(\cosh(x))^{3/2}}&\left(\sinh(x)+\sum_{j=1}^{m}d_j\delta(x,y)\right)\left(1-\frac{x}{2\cosh(x)}\sum_{j=1}^{m}d_j\alpha(x,y)\right)\\ &\times\left(1+O\left(\frac{1}{T}\right)\right)+O_p(\sqrt{T})+O(||d||^2)\end{aligned}$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

$$= \frac{Ta_{ii}}{x(\cosh(x))^{3/2}} \left(\sinh(x) + \sum_{j=1}^{m} d_j \phi(x,y) \right) \left(1 + O\left(\frac{1}{T}\right) \right) + O_p(\sqrt{T}) + O(||d||^2), \tag{A20}$$

where $\phi(x,y) = -\frac{x}{2}\tanh(x)\alpha(x,y) + \delta(x,y)$.

Thus, by inserting (A20) into (A16), we get

$$\mathrm{E}\left(\frac{\mathcal{R}^{\circ}_{1i}}{\sqrt{V^{\circ}_i}} \right) = a_{ii} \sum_{j=0}^{m} d_j r_j + O_p\left(\frac{1}{\sqrt{T}}\right) + O(||d||^2),$$

where

$$r_0 = \frac{\sqrt{2}}{\Gamma(1/2)} \int_0^{\infty} \frac{\sinh(x)}{x(\cosh(x))^{3/2}}\, dx = 1.2459,$$

while if $j > 0$, we write $r(y) = r(j/T) = r_j$, where

$$r_j = \frac{\sqrt{2}}{\Gamma(1/2)} \int_0^{\infty} \frac{\phi(x,y)}{x(\cosh(x))^{3/2}}\, dx.$$

Making use of (A13), it follows that as $N,T \to \infty$,

$$I = \frac{a_{\infty}}{\sqrt{N}} \sum_{j=0}^{m} d_j r_j + O_p\left(\frac{\sqrt{N}}{\sqrt{T}}\right) + O\left(\sqrt{N}||d||^2\right) + O_p\left(\frac{1}{C^2_{NT}}\right). \tag{A21}$$

Next, consider II. Similar to (A13), letting $U^{\circ}_i = TU_i/\sigma^2_{\epsilon i}$ and

$$\mathcal{R}^{\circ}_{2i} = \sum_{t=2}^{T} e^{\circ}_{it-1} A^{\circ}_{it-1},$$

and then passing N, $T \to \infty$, we obtain

$$II = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{U^{\circ}_i}{(V^{\circ}_i)^{3/2}} \mathcal{R}^{\circ}_{2i} \right) + O_p\left(\frac{1}{C^2_{NT}}\right) = \lim_{N\to\infty} \sqrt{N}\mathrm{E}\left(\frac{U^{\circ}_i}{(V^{\circ}_i)^{3/2}} \mathcal{R}^{\circ}_{2i} \right) + O_p\left(\frac{1}{C^2_{NT}}\right).$$

Now, by using (A12), it is clear that $\mathcal{R}^{\circ}_{2i}$ can be written as

$$\mathcal{R}^{\circ}_{2i} = \sum_{t=2}^{T} e^{\circ}_{it-1} A^{\circ}_{it-1} = \sum_{t=1}^{T-1} e^{\circ}_{it} (G'_{it} e^{\circ}) = (e^{\circ})' D_i G' e^{\circ},$$

where G' is the $NT \times NT$ matrix stacking G'_{it}. Similarly, U°_i may be written as

$$U^{\circ}_i = \sum_{t=2}^{T} e^{\circ}_{it-1} \Delta e^{\circ}_{it} = -\frac{1}{2} (e^{\circ})' C_i e^{\circ},$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

where C_i is a block diagonal $NT \times NT$ matrix with diagonal $T \times T$ blocks that are equal to zero except for the block at position i, which is given by

$$C_{ii} = C = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & & 2 & -1 \\ 0 & \cdots & 0 & -1 & 0 \end{bmatrix}.$$

As before, the moment-generating function of $(U_i^\circ, \mathcal{R}_{2i}^\circ, V_i^\circ)$ can be written as

$$\varphi_i(u, r, -v) = \frac{1}{(2\pi)^{NT/2}} \int \exp\left(-\frac{1}{2}(e^{\circ\circ})' P_i e^{\circ\circ}\right) d\left(e^{\circ\circ}\right) = \frac{1}{\sqrt{\det(P_i)}},$$

where P_i is now given by $P_i = P_{0i} + uP_{1i} + rP_{2i} = P_{0i} + P_{mi}$, with P_{0i} unchanged but with $P_{1i} = (I_N \otimes \Sigma)^{1/2} C_i (I_N \otimes \Sigma)^{1/2}$ and where P_{2i} is P_{1i} with $-2D_i G'$ in place of C_i. Furthermore, (A15) implies that

$$\mathrm{E}\left(\frac{U_i^\circ \mathcal{R}_{2i}^\circ}{(V_i^\circ)^{3/2}}\right) = \frac{1}{\Gamma(3/2)} \int_0^\infty \sqrt{v} \frac{\partial^2}{\partial u \partial r} \varphi_i(u, r, -v)\, dv, \tag{A22}$$

where the derivative is evaluated at $r = u = 0$. Note that

$$\begin{aligned} \mathrm{E}\Big(\left(X' P_{mi} X\right)^2\Big) &= \Big(\mathrm{tr}\Big(P_{0i}^{-1} P_{mi}\Big)\Big)^2 + 2\mathrm{tr}\Big(\Big(P_{0i}^{-1} P_{mi}\Big)^2\Big) \\ &= 2ur\Big(\mathrm{tr}\big(P_{0i}^{-1} P_{1i}\big)\mathrm{tr}\big(P_{0i}^{-1} P_{2i}\big) + 2\mathrm{tr}\Big(P_{0i}^{-1} P_{1i} P_{0i}^{-1} P_{2i}\Big)\Big) + O(u^2) + O(r^2) \\ &= 2ur\, Q_i + O(u^2) + O(r^2), \end{aligned}$$

which implies that

$$\frac{\partial^2}{\partial u \partial r} \varphi_i(u, r, -v) = -\frac{Q_i}{2\sqrt{\det(P_{0i})}}, \tag{A23}$$

where Q_i is given in Lemma A.4.

LEMMA A.4. *We have*

$$Q_i = T^3 a_{ii} Q(x) + T^3 \sum_{j=1}^{m} d_j Q(x, y) + O(T^2) + O(||d||^2),$$

where

$$Q(x) = \frac{3x - x\cosh^2(x) + (x^2 - 2)\sinh(x)\cosh(x)}{2x^3 \cosh^2(x)}$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

and $Q(x,y) = a_{ii}(g_0(x) + g_1(x,y)\cosh(xy) + g_2(x,y)\sinh(xy))$ *with*

$$g_0(x) = -\frac{1}{2x^2}\left(1 - \frac{\sinh(x)}{x\cosh(x)}\right),$$

$$g_1(x,y) = \frac{1}{2x^3\cosh^3(x)}\left((1-x^2(1-y))yx\cosh^3(x) - (1-x^2+y^2x^2)\sinh(x)\cosh^2(x)\right.$$

$$\left. + x(1+x^2)(1-y)\cosh(x) - 6x^2(1-y)\sinh(x)\right),$$

and

$$g_2(x,y) = \frac{1}{2x^3\cosh^2(x)}\left(3x^2y(1-y) - 3 + (1-(1-y^2)x^2)\cosh^2(x)\right.$$

$$\left. -(1-(1-y)x^2)xy\sinh(x)\cosh(x) + 2\cosh(x)\right).$$

Proof of Lemma A.4. The proof is omitted and can be found in Westerlund and Larsson (2008). ■

By inserting the result of Lemma A.4 into (A23), we get

$$\begin{aligned}\frac{\partial^2}{\partial u\partial r}\varphi_i(u,r,-v) &= -\frac{Q_i}{2\sqrt{\det(P_{0i})}}\\ &= -\frac{T^3}{2\sqrt{\cosh(x)}}\left(1 - \frac{x}{2\cosh(x)}\sum_{j=1}^{n} d_j\alpha(x,y)\right)\\ &\quad\times\left(a_{ii}Q(x) + \sum_{j=1}^{m} d_j Q(x,y)\right) + O(T^2) + O(||d||^2)\\ &= -\frac{T^3}{2\sqrt{\cosh(x)}}\left(a_{ii}Q(x) + \sum_{j=1}^{m} d_j R(x,y)\right) + O(T^2) + O(||d||^2),\end{aligned}$$

where

$$\begin{aligned}R(x,y) &= Q(x,y) - a_{ii}Q(x)\frac{\alpha(x,y)x}{2\cosh(x)}\\ &= a_{ii}(g_0(x) + h_1(x,y)\cosh(xy) + h_2(x,y)\sinh(xy))\end{aligned}$$

with

$$h_1(x,y) = g_1(x,y) - Q(x)\frac{x(1-y)\sinh(x)}{2\cosh(x)}$$

and $h_2(x,y) = g_2(x,y) + Q(x)x(1-y)/2$.

Hence, (A22) becomes

$$\begin{aligned}\mathrm{E}\left(\frac{U_i^{\circ}\mathcal{R}_{2i}^{\circ}}{(V_i^{\circ})^{3/2}}\right) &= -\frac{1}{2^{3/2}\Gamma(3/2)}\int_0^{\infty}\frac{x^2}{\sqrt{\cosh(x)}}\left(a_{ii}Q(x) + \sum_{j=1}^{m} d_j R_j\right)dx\\ &= a_{ii}\sum_{j=0}^{m} d_j s_j + O_p\left(\frac{1}{\sqrt{T}}\right) + O(||d||^2),\end{aligned}$$

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press

where

$$s_0 = -\frac{1}{2^{3/2}\Gamma(3/2)} \int_0^\infty \frac{x^2}{\sqrt{\cosh(x)}} \left(\left(1 - \frac{\sinh(x)}{x\cosh(x)} \right) \frac{\sinh(x)}{2(x\cosh(x))} - \frac{\sinh(2x) - 2x}{2x^3(\cosh(x))^2} \right) dx,$$

which equals -0.4231, while if $j > 0$, we again write $s(y) = s(j/T) = s_j$, where

$$s_j = -\frac{1}{2^{3/2}\Gamma(3/2)} \int_0^\infty \frac{x^2}{\sqrt{\cosh(x)}} \big(h(x) + h_1(x, y)\cosh(xy) + h_2(x, y)\sinh(xy) \big) dx.$$

It follows that as $N, T \to \infty$,

$$II = \frac{a_\infty}{\sqrt{N}} \sum_{j=0}^{m} d_j s_j + O_p\left(\frac{\sqrt{N}}{\sqrt{T}} \right) + O\big(\sqrt{N} ||d||^2 \big) + O_p\left(\frac{1}{C_{NT}^2} \right),$$

which, in conjunction with (A11) and the results obtained for I, yields

$$\sqrt{N}\,\overline{\mathcal{R}}_N = -I + II + O_p\left(\frac{\sqrt{N}}{R_{NT}} \right) = \frac{a_\infty}{\sqrt{N}} \sum_{j=0}^{m} d_j w_j + O\big(\sqrt{N} ||d||^2 \big) + O_p\left(\frac{\sqrt{N}}{R_{NT}} \right),$$

where $w_j = (s_j - r_j)$ with $w_0 = -1.669$ and, via Taylor expansion of w_j in y, and then integration over x, we get $w_j = 1.04605 - 5.54336y + 19.4018y^2 - 52.384y^3 + 131.95y^4 - 317.697y^5 + O(y^6)$. This completes the proof. ■

https://doi.org/10.1017/S0266466609990351 Published online by Cambridge University Press