On the Breitung Test for Panel Unit Roots and Local Asymptotic Power*

H.R. Moon, B. Perron, and P.C.B. Phillips

January 2006

Abstract

This note analyzes the local asymptotic power properties of a test proposed by Breitung (2000). We demonstrate that the Breitung test, like many other tests (including point optimal tests) for panel unit roots in the presence of incidental trends, has non-trivial power in neighborhoods that shrink towards the null hypothesis at the rate of $n^{-1/4}T^{-1}$ where n and T are the cross-section and time-series dimensions respectively. This rate is slower than the $n^{-1/2}T^{-1}$ rate claimed by Breitung. Simulation evidence documents the usefulness of the asymptotic approximations given here.

JEL Classification: C22 & C23

Keywords and Phrases: Asymptotic power, Breitung test, heterogeneous alternatives, incidental trends, local to unity, panel unit root test.

1 Motivation

In the past decade or so, there has been much interest in testing for the presence of a unit root in panel data. Many researchers have proposed statistics to test the hypothesis of a common unit autoregressive root. Recent surveys by Baltagi and Kao (2000), Choi (2004), Hurlin and Mignon (2004), and Breitung and Pesaran (2005) provide an overview of these developments.

In this context, Breitung (2000) proposed a $t-$ ratio type test statistic for testing a panel unit root. Through numerical analysis, he claimed that his test has 'nice' power properties within a certain local neighborhood of unity. The present paper investigates analytically the asymptotic power properties of Breitung's test and clarifies some of the analytic results in Breitung (2000). Specifically, we show that the limiting distribution of the Breitung test is the same under the null and the $O\left(n^{-1/2}T^{-1}\right)$ local alternatives considered by Breitung,

*The authors thank Paolo Paruolo and a referee for comments on an earlier version of the paper. Phillips acknowledges partial support from a Kelly Fellowship and the NSF under Grant SES 04-142254. Perron acknowledges financial support from FQRSC, SSHRC, and MITACS.

so that the test has trivial power in such narrow neighborhoods. We provide expressions for the local asymptotic power of this test for wider $n^{-1/4}T^{-1}$ local departures from the null, discuss comparative results for other procedures such as point optimal tests, and study the accuracy of the asymptotic approximations in finite samples.

2 Breitung's Test Statistic and Claimed Power

Suppose that panel data y_{it} is generated by the following simple components model

$$y_{it} = \mu_i + \beta_i t + x_{it},$$

where the unobserved error term x_{it} follows

$$x_{it} = \rho_i x_{it-1} + \varepsilon_{it}.$$

Our main interest is in testing the presence of a unit root in all cross-sectional units, viz.,

$$H_0 : \rho_i = 1 \text{ for all } i. \tag{1}$$

For this, we assume that $\varepsilon_{it} \sim iid\left(0, \sigma^2\right)$ with $E\left(\varepsilon_{it}^4\right) < \infty$, the initial observations x_{i0} are iid across i with $E\left(x_{i0}^4\right) < \infty$ and independent of ε_{it} all $t \geq 1$ and i.[1]

Notice that this testing problem is invariant to the following linear transformation: $y_{it}^* = y_{it} + \mu_i^* + \beta_i^* t$. To construct a test that is invariant to the transformation, Breitung (2000) suggested the use of the following transformed data:

$$(\Delta y_{it})^* = s_t \left[\Delta y_{it} - \frac{1}{T-t}\left(\Delta y_{it+1} + ... + \Delta y_{iT}\right)\right],$$

for $t = 1, ..., T-1$, where $s_t^2 = (T-t) / (T-t+1)$, and

$$y_{it-1}^* = y_{it-1} - y_{i0} - \frac{t-1}{T}\left(y_{iT} - y_{i0}\right),$$

for $t = 2, ..., T$.[2] The panel unit root test for the null hypothesis (1) proposed by Breitung (2000) is to reject the null for the small values of the following statistic:

$$\begin{aligned} \mathcal{B}_{nT} &= \left(\frac{\hat{\sigma}^2}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T-1}\left(y_{it-1}^*\right)^2\right)^{-1/2} \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T-1}\left(\Delta y_{it}\right)^* y_{it-1}^* \\ &= \left(\mathcal{B}_{2nT}\right)^{-1/2}\mathcal{B}_{1nT}, \end{aligned}$$

[1] These restrictions are made for simplicity in the following analysis and can be relaxed to cover more general cases.

[2] Notice that $(\Delta y_{it})^*$ and y_{it-1}^* correspond to the terms in equations (16) and (17), respectively of Breitung (2000).

where $\hat{\sigma}^2$ is a consistent estimator of σ^2. This is a $t-$ ratio statistic from a pooled regression of the transformed data $(\Delta y_{it})^*$ on y^*_{it-1}. Define $x^*_{it} = \sum_{s=1}^{t} \varepsilon_{is}$. Under the null hypothesis, a direct calculation (given in the appendix) shows that for $2 \leq t \leq T-1$

$$
\begin{aligned}
& E\left[(\Delta y_{it})^* \, y^*_{it-1}\right] \\
= \;& s_t E\left[\left(\Delta x_{it} - \frac{1}{T-t}\left(\Delta x_{it+1} + ... + \Delta x_{iT}\right)\right)\left(x_{it-1} - x_{i0} - \frac{t-1}{T}\left(x_{iT} - x_{i0}\right)\right)\right] \\
= \;& s_t E\left[\left\{\varepsilon_{it} - \frac{1}{T-t}\left(x^*_{iT} - x^*_{it}\right)\right\}\left\{x^*_{it-1} - \frac{t-1}{T} x^*_{iT}\right\}\right] = 0,
\end{aligned}
$$

and

$$
\lim_{n,T} \frac{1}{n} \sum_{i=1}^{n} Var\left(\frac{1}{T} \sum_{t=2}^{T-1} (\Delta y_{it})^* \, y^*_{it-1}\right) = \frac{1}{6}\sigma^4.
$$

Hence, by central limit theory (e.g., see Phillips and Moon (1999)), it is possible to show that $\mathcal{B}_{1nT} \Rightarrow N\left(0, \frac{1}{6}\sigma^4\right)$ as $n, T \to \infty$. On the other hand,

$$
\lim_{n,T} \frac{1}{n} \sum_{i=1}^{n} E\left(\frac{1}{T^2} \sum_{t=2}^{T-1} \left(y^*_{it-1}\right)^2\right) = \frac{1}{6}\sigma^2,
$$

so that $\mathcal{B}_{2nT} \to_p \frac{1}{6}\sigma^4$ as $n, T \to \infty$. Therefore, under the null hypothesis, it follows that as $n, T \to \infty$,

$$
\mathcal{B}_{nT} \Rightarrow N(0,1).
$$

Breitung's test for H_0 with size α rejects H_0 if $\mathcal{B}_{nT} < -z_\alpha$, where z_α is the $(1-\alpha)-$quantile of the $N(0,1)$ distribution.

To analyze the local power of the test based on $\mathcal{B}_{nT}$, Breitung (2000) considered the following local parameterization

$$
\rho_i = 1 - \frac{c}{n^{1/2}T}, \tag{2}
$$

and he claimed (see Theorem 5 of Breitung (2000)) that the test B_{nT} has asymptotically significant local power against the local alternative

$$
H_1 : c < 0.
$$

To verify the claim, since $\Delta x_{it} = -\frac{c}{\sqrt{n}T} x_{it-1} + \varepsilon_{it}$ under the local alternative (2), we can express the transformed variables as follows:

$$
\begin{aligned}
& (\Delta y_{it})^* \\
= \;& s_t \left[\varepsilon_{it} - \frac{1}{T-t}\left(\varepsilon_{it+1} + ... + \varepsilon_{iT}\right)\right] - \frac{c}{\sqrt{n}T} s_t \left[x_{it-1} - \frac{1}{T-t}\left(x_{it} + ... + x_{iT-1}\right)\right] \\
= \;& A_{it} - \frac{c}{\sqrt{n}T} B_{it}, \text{ say;}
\end{aligned}
$$

and

$$
\begin{aligned}
& y_{it-1}^{*} \\
= & \left[(\varepsilon_{i1} + ... + \varepsilon_{it-1}) - \frac{t-1}{T} (\varepsilon_{i1} + ... + \varepsilon_{iT}) \right] \\
& - \frac{c}{\sqrt{n}T} \left[(x_{i0} + ... + x_{it-2}) - \frac{t-1}{T} (x_{i0} + ... + x_{iT-1}) \right] \\
= & C_{it-1} - \frac{c}{\sqrt{n}T} D_{it-1}, \text{ say.}
\end{aligned}
$$

Note that A_{it} and C_{it} do not depend on c. Then, the above formulae lead to the following decomposition of $\mathcal{B}_{1nT}$

$$
\begin{aligned}
& \mathcal{B}_{1nT} \\
= & \frac{1}{\sqrt{n}T} \sum_{i=1}^{n} \sum_{t=2}^{T-1} A_{it} C_{it-1} - \frac{c}{nT^2} \sum_{i=1}^{n} \sum_{t=2}^{T-1} (A_{it} D_{it-1} + B_{it} C_{it-1}) \\
& + \frac{c^2}{n^{3/2} T^3} \sum_{i=1}^{n} \sum_{t=2}^{T-1} B_{it} D_{it-1}. \qquad (3)
\end{aligned}
$$

Lemma 1 *The following hold under the local alternative* $\rho_i = 1 - \frac{c}{n^{1/2}T}$ *as* $n, T \to \infty$.

(a) $\frac{1}{\sqrt{n}T} \sum_{i=1}^{n} \sum_{t=2}^{T-1} A_{it} C_{it-1} \Rightarrow N\left(0, \frac{1}{6}\sigma^4\right).$

(b) $\frac{1}{nT^2} \sum_{i=1}^{n} \sum_{t=2}^{T-1} (A_{it} D_{it-1} + B_{it} C_{it-1}) \to_p 0.$

(c) $\frac{1}{n.T^3} \sum_{i=1}^{n} \sum_{t=2}^{T-1} B_{it} D_{it-1} = O_p(1).$

(d) $\mathcal{B}_{2nT} \to_p \frac{1}{6}\sigma^4.$

It follows from Lemma 1 that $\mathcal{B}_{1nT} \Rightarrow N\left(0, \frac{1}{6}\sigma^4\right)$. Together with the limit of $\mathcal{B}_{2nT}$, under the local alternative in (2), we therefore obtain $\mathcal{B}_{nT} \Rightarrow N(0,1)$ and have the following result.

Theorem 2 *Under the local alternative* $\rho_i = 1 - \frac{c}{n^{1/2}T}$ *assumed in Breitung (2000), the Breitung test statistic* $\mathcal{B}_{nT}$ *has asymptotic power equal to the size of the test.*

The above result does not imply a mistake in Breitung (2000). Specifically, Breitung stated that the power of his test depends on a quantity that in our notation is $E\left[\frac{1}{T}\sum_{t=2}^{T-1} y_{it}^{*} (\Delta y_{it})^{*}\right]$. However, the above results demonstrate that this expectation has zero limit. Instead of an analytical evaluation of the limit, Breitung relied on a numerical approximation to the expectation which suggested non-negligible power against alternatives of the form (2). Our results indicate that the numerical approximation is poor and that consequently the implications regarding power in neighborhoods of the form (2) are misleading.

3 Asymptotic Local Power of Breitung's Test

The main problem with Theorem 2 is that the local neighborhood in (2) shrinks too fast to unity and so the asymptotic power is trivial. To correct the result, this section considers local neighborhoods that shrink to unity more slowly and derives an asymptotic local power function for the Breitung test.

First, consider the local to unity region defined by

$$\rho_i = 1 - \frac{c_i}{n^{1/4}T}, \tag{4}$$

where the c_i may be defined as the realizations of a sequence of iid random variables whose support lies in an interval of the form $[-M_{lc}, M_{uc}]$ for some $0 < M_{lc} < \infty$ and $0 \leq M_{uc} < \infty$. Let c_i be independent of ε_{js} for all i, j, s. Under these conditions, the numerator of Breitung's test statistic can be decomposed as

$$\begin{aligned} &\mathcal{B}_{1nT} \\ = \; & \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T-1} A_{it}C_{it-1} - \frac{1}{n^{3/4}T^2}\sum_{i=1}^{n} c_i \sum_{t=2}^{T-1}\left(A_{it}D_{it-1} + B_{it}C_{it-1}\right) \\ & + \frac{1}{nT^3}\sum_{i=1}^{n} c_i^2 \sum_{t=2}^{T-1} B_{it}D_{it-1}, \end{aligned}$$

where the components A_{it}, B_{it}, C_{it-1}, and D_{it-1} are defined in the previous section.

Lemma 3 *Under the local alternative* (4), *the following hold as* $n, T \to \infty$.

(a) $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T-1} A_{it}C_{it-1} \Rightarrow N\left(0, \frac{1}{6}\sigma^4\right)$.

(b) $\frac{1}{n^{3/4}T^2}\sum_{i=1}^{n} c_i \sum_{t=2}^{T-1}\left(A_{it}D_{it-1} + B_{it}C_{it-1}\right) \to_p -\frac{E\left(c_i^2\right)\sigma^2}{18}$.

(c) $\frac{1}{nT^3}\sum_{i=1}^{n} c_i^2 \sum_{t=2}^{T-1} B_{it}D_{it-1} \to_p \frac{E\left(c_i^2\right)\sigma^2}{36}$.

(d) $\mathcal{B}_{2nT} \to_p \frac{1}{6}\sigma^4$.

Using Lemma 3, the limit distribution under the local alternative hypothesis (4) is

$$\mathcal{B}_{nT} \Rightarrow N\left(-\frac{E\left(c_i^2\right)}{6\sqrt{6}}, 1\right),$$

from which we deduce the following theorem.

Theorem 4 *Against the local alternative in* (4), *the asymptotic local power of Breitung's test is* $\Phi\left(\frac{E\left(c_i^2\right)}{6\sqrt{6}} - z_\alpha\right)$.

Remarks

1. Contrary to Breitung's (2000) claim, non-trivial local power is defined in neighborhoods that shrink towards the null hypothesis at the rate $\frac{1}{n^{1/4}T}$. This is the same rate worked out by Moon, Perron, and Phillips (2005) in defining the power envelope for panel unit root testing in the context of incidental trends for models in this form.

2. In contrast to Breitung (2000), the result above is obtained against heterogeneous alternatives. The test therefore has significant power against this type of hypothesis despite pooling.

3. The power of Breitung's test depends on the second moments of the local-to-unity parameters. Thus, for a given mean autoregressive parameter, the more heterogeneous the alternatives are, the easier they are to detect.

4. Moon, Perron, and Phillips (2005) derive the power envelope for the above testing problem, suggest a common point optimal (CPO) test and discuss the local asymptotic power of other tests such as those proposed by Levin et al. (2002), Ploberger and Phillips (2002), and Moon and Phillips (2004). According to those results, the test based on B_{nT} is more powerful than the Levin et al. and Moon and Phillips tests, but less powerful than the Ploberger-Phillips and CPO tests.

A small scale simulation was conducted to assess the accuracy of these asymptotic results in finite samples. We use the following data generating process:

$$\begin{aligned} z_{it} &= b_{0i} + b_{1i}t + y_{it}, \\ y_{it} &= \rho_i y_{it-1} + u_{it}, \\ y_{i,-1} &= 0, \quad u_{it} \sim iid\ N(0,1). \end{aligned}$$

The heterogeneous trend coefficients are taken to be $iid\ N(0,1)$. We assume that the error term is independent in both time and cross-section dimensions with a Gaussian distribution and identical variances. We consider four values for n (10, 25, 100, and 250) and three values for T (50, 100, and 250). All tests are carried out at the 5% significance level, and the number of replications is set at 10,000.

The autoregressive parameters are generated according to (4). We consider the following nine distributions for the local-to-unity parameters:

(0) $c_i = 0 \quad \forall i$ (size),

(1) $c_i \sim iidU[0,2]$,

(2) $c_i \sim iidU[0,4]$,

(3) $c_i \sim iidU[0,8]$,

(4) $c_i \sim iid\chi^2(1)$,

(5) $c_i \sim iid\chi^2(2)$,

(6) $c_i \sim iid\chi^2(4)$,

(7) $c_i = 1\ \forall i$,

(8) $c_i = 2\ \forall i$.

These distributions enable us to examine performance of the tests as the mass of the distribution of the localizing parameters moves away from the null hypothesis. We can also look at the effect of homogeneous versus heterogeneous alternatives (cases (1) and (4) versus (7), and cases (2) and (5) versus (8)) together with the role of the higher-order moments of the distribution. For instance, case (7) has the same mean as cases (1) and (4) but smaller higher-order moments than the other two cases. The same situation arises for cases (2), (5), and (8), and cases (3) and (6). Note that the alternatives with χ^2 distributions do not fit our asymptotic framework since they have unbounded support.

Table 1 presents the results. The second column provides the size and power predicted by our asymptotic theory using the moments of c_i. The other columns in the table report the size and size-adjusted power of the tests for the various combinations of n and T. If asymptotic theory were a reliable guide to finite-sample behavior, all columns in the table would be very close.

Table 1. Size and size-adjusted power of Breitung's test

		$T=50$				$T=100$				$T=250$			
	Theory	$n=10$	25	100	250	$n=10$	25	100	250	$n=10$	25	100	250
$c_i = 0$ (size)	5.0	6.6	5.1	3.8	2.7	6.1	5.5	4.8	3.8	6.4	5.8	4.7	4.3
$c_i \sim U[0,2]$	6.0	5.0	5.6	6.0	5.4	5.3	5.4	5.2	5.6	5.6	5.2	6.2	6.7
$c_i \sim U[0,4]$	10.0	6.9	8.2	8.7	8.6	7.3	7.8	8.3	8.8	8.3	7.9	9.6	9.9
$c_i \sim U[0,8]$	42.3	13.3	17.6	22.0	24.4	15.6	18.5	22.9	27.0	16.2	19.3	24.3	30.0
$c_i \sim \chi^2(1)$	7.5	5.5	6.4	6.2	6.8	6.3	5.8	5.9	6.7	6.4	6.2	7.7	7.0
$c_i \sim \chi^2(2)$	13.6	7.1	8.4	9.1	10.0	7.7	8.7	8.8	10.0	8.4	8.6	10.4	11.0
$c_i \sim \chi^2(4)$	49.5	13.8	17.3	21.8	24.8	15.4	17.1	21.1	26.1	16.0	18.2	24.8	29.4
$c_i = 1$	5.7	4.4	5.5	6.1	5.3	5.8	5.1	5.2	5.6	5.8	5.4	5.5	6.1
$c_i = 2$	8.5	5.8	7.4	7.7	7.8	7.0	7.4	7.2	8.1	7.0	7.5	9.1	8.6

Note: The second column reports the rejection frequency of the panel unit root hypothesis based on a one-sided test according to our theory. The remaining columns report the actual rejection frequencies using ether the asymptotic critical value (for size) or the empirical critical value (for size-adjusted power) based on 10,000 replications.

Overall, we see that the test performs much more closely to the asymptotic theory as both n and T increase. The test underrejects for large n relative to T.

Power is usually below what is predicted by asymptotic theory, especially for the more distant alternatives, but the discrepancy diminishes with increases in either n or T or both. Finally, experiments where the local-to-unity parameters have a fatter tailed distribution tend to have higher power as predicted. Thus, for a given mean autoregressive parameter, more heterogeneous alternatives are easier to detect (despite the pooling approach used in the test). There also seems no sign that the unboundedness of the χ^2 distributions affects the validity of the asymptotic theory.

References

[1] Baltagi, B.H. and C. Kao (2000): Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey, in Baltagi, B.H., T B. Fomby and R.C. Hill, eds., *Advances in Econometrics volume 15: Nonstationary Panels, Cointegration in Panels and Dynamic Panels*, Greenwich, CT: JAI Press, 7-51.

[2] 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,* JAI: Amsterdam, 161-178.

[3] Breitung, J. and M. H. Pesaran (2005): Unit Roots and Cointegration in Panels, forthcoming in Matyas, L. and P. Sevestre (ed.), *The Econometrics of Panel Data (Third Edition)*, Kluwer Academic Publishers. Available at http://www.econ.cam.ac.uk/faculty/pesaran/panelUnitCoin05November05.pdf.

[4] Choi, I. (2004): Nonstationary Panels, *Palgrave Handbooks of Econometrics, Vol. 1*, forthcoming. Available at http://ihome.ust.hk/~inchoi/InChoiRev3.pdf.

[5] Hurlin, C. and V. Mignon (2004): Second Generation Panel Unit Root Tests, THEMA-CNRS, Université de Paris X, *Mimeo.* Avaliable at http://thema.u-paris10.fr/papers/2004-21Mignon.pdf.

[6] Levin, A., F. Lin, and C. Chu (2002): Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties, *Journal of Econometrics*, 108, 1–24.

[7] Moon, H.R., B. Perron, and P.C.B. Phillips (2005): Incidental Trends and the Power of Panel Unit Root Tests, *Mimeo.* Avaliable at http://mapageweb.umontreal.ca/perrob/power.pdf.

[8] Moon, H.R., B. Perron, and P.C.B. Phillips (2006): On the Britung Test for Panel Unit Roots and Local Asymptotic Power - Further Calculations, *Mimeo.* Avaliable at http://mapageweb.umontreal.ca/perrob/breitung_appendix.pdf.

[9] Moon, H.R. and P.C.B. Phillips (2004): GMM Estimation of Autoregressive Roots Near Unity with Panel Data, *Econometrica*, 72, 467-522.

[10] Phillips, P.C.B. and H.R. Moon (1999): Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, 67, 1057–1111.

[11] Ploberger, W. and P.C.B. Phillips (2002): Optimal Testing for Unit Roots in Panel Data, *Mimeo*.

4 Appendix: Proofs

Proof of Lemma 1
Part (a): Part (a) follows because the distribution of $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=2}^{T-1} A_{it}C_{it-1}$ is identical to that of $\mathcal{B}_{1nT}$ under the null since neither A_{it} nor C_{it-1} depends on c. ■

Part (b): Under the given assumptions, by a WLLN,

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=2}^{T-1}\left(A_{it}D_{it-1}+B_{it}C_{it-1}\right)\to_p \lim_{T\to\infty}\frac{1}{T^2}\sum_{t=2}^{T-1}E\left[A_{it}D_{it-1}+B_{it}C_{it-1}\right].$$

Recall the definition $x_{it}^{*}=\sum_{s=1}^{t}\varepsilon_{is}$. For $2\le t\le T-1$, we have

$$\begin{aligned}
& E\left[A_{it}D_{it-1}\right] \\
= \;& E\left[s_t\left\{\varepsilon_{it}-\frac{1}{T-t}\left(\varepsilon_{it+1}+...+\varepsilon_{iT}\right)\right\}\left\{\left(x_{i0}+...+x_{it-2}\right)-\frac{t-1}{T}\left(x_{i0}+...+x_{iT-1}\right)\right\}\right] \\
= \;& E\left[s_t\left\{\varepsilon_{it}-\frac{1}{T-t}\left(x_{iT}^{*}-x_{it}^{*}\right)\right\}\left\{\left(x_{i1}^{*}+...+x_{it-2}^{*}\right)-\frac{t-1}{T}\left(x_{i1}^{*}+...+x_{iT-1}^{*}\right)\right\}\right]+o\left(1\right) \\
= \;& E\left[s_t\varepsilon_{it}\left(x_{i1}^{*}+...+x_{it-2}^{*}\right)\right]-\frac{1}{T-t}E\left[s_t\left(x_{iT}^{*}-x_{it}^{*}\right)\left(x_{i1}^{*}+...+x_{it-2}^{*}\right)\right] \\
& -\frac{t-1}{T}E\left[s_t\varepsilon_{it}\left(x_{i1}^{*}+...+x_{iT-1}^{*}\right)\right]+\frac{1}{T-t}\frac{t-1}{T}E\left[s_t\left(\varepsilon_{it+1}+...+\varepsilon_{iT}\right)\left(x_{i1}^{*}+...+x_{iT-1}^{*}\right)\right] \\
= \;& s_t\left(-\frac{t-1}{T}\left(T-t\right)+\frac{t-1}{T}\frac{T-t-1}{2}\right).
\end{aligned}$$

Also,

$$\begin{aligned}
& E\left[B_{it}C_{it-1}\right] \\
= \;& E\left[s_t\left\{x_{it-1}-\frac{1}{T-t}\left(x_{it}+...+x_{iT-1}\right)\right\}\left\{\left(\varepsilon_{i1}+...+\varepsilon_{it-1}\right)-\frac{t-1}{T}\left(\varepsilon_{i1}+...+\varepsilon_{iT}\right)\right\}\right] \\
= \;& E\left[s_t\left\{x^*_{it-1}-\frac{1}{T-t}\left(x^*_{it}+...+x^*_{iT-1}\right)\right\}\left\{x^*_{it-1}-\frac{t-1}{T}x^*_{iT}\right\}\right]+o\left(1\right) \\
= \;& E\left[s_t x^*_{it-1}x^*_{it-1}\right]-\frac{1}{T-t}E\left[s_t\left(x^*_{it}+...+x^*_{iT-1}\right)x^*_{it-1}\right] \\
& -\frac{t-1}{T}E\left[s_t x^*_{it-1}x^*_{iT}\right]+\frac{1}{T-t}\frac{t-1}{T}E\left[s_t\left(x^*_{it}+...+x^*_{iT-1}\right)x^*_{iT}\right] \\
= \;& s_t\left[-\frac{t-1}{T}\left(t-1\right)+\frac{t-1}{T}\frac{\left(T+t-1\right)}{2}\right].
\end{aligned}$$

Combining these, we have the required result for Part (b) since

$$\frac{1}{T^2}\sum_{t=2}^{T-1}E\left[A_{it}D_{it-1}+B_{it}C_{it-1}\right]\to\int_0^1\left(-r\left(1-r\right)+\frac{1}{2}r\left(1-r\right)-r^2+\frac{1}{2}r\left(1+r\right)\right)dr=0. \blacksquare$$

Parts (c) and (d): These follow by a WLLN. ■

Proof of Lemma 3

In this proof we use the notation $x^*_{it}=\sum_{q=1}^{t}\varepsilon_{iq}+x^*_{i0}$, where $x^*_{i0}=x_{i0}$. Also, for notational convenience, we write $\varepsilon_{i0}=x_{i0}=x^*_{i0}$. Then, by definition, for $t\geq 1$,

$$x_{it}-x^*_{it}=\sum_{p=0}^{t-1}\left(\rho_i^{t-p}-1\right)\varepsilon_{ip}=\sum_{p=0}^{t-1}\sum_{l=1}^{t-p}\binom{t-p}{l}\left(-\frac{c_i}{n^{1/4}T}\right)^l\varepsilon_{ip}. \qquad (5)$$

Parts (a) and (d): They follow in the same fashion as Parts (a) and (d) in Lemma 1. ■

Part (b): Using (5), we can approximate the quantity of interest as:

$$\begin{aligned}
& \frac{1}{n^{3/4}T^2}\sum_{i=1}^{n}c_i\sum_{t=2}^{T-1}\left(A_{it}D_{it-1}+B_{it}C_{it-1}\right) \\
= \;& \frac{1}{n^{3/4}T^2}\sum_{i=1}^{n}c_i\sum_{t=2}^{T-1}\left(A_{it}D_{1,it-1}+B_{1,it}C_{it-1}\right)-\frac{1}{nT^2}\sum_{i=1}^{n}c_i^2\sum_{t=2}^{T-1}\left(A_{it}D_{2,it-1}+B_{2,it}C_{it-1}\right)+o_p\left(1\right),
\end{aligned}$$

where

$$
\begin{aligned}
A_{it} &= s_t \left[\varepsilon_{it} - \frac{1}{T-t} \left(\sum_{q=t+1}^{T} \varepsilon_{iq} \right) \right], \; C_{it-1} = \left[\left(\sum_{q=1}^{t-1} \varepsilon_{iq} \right) - \frac{t-1}{T} \left(\sum_{q=1}^{T} \varepsilon_{iq} \right) \right] \\
B_{1,it} &= s_t \left[x^*_{it-1} - \frac{1}{T-t} \left(\sum_{q=t}^{T-1} x^*_{iq} \right) \right], \\
B_{2,it} &= s_t \left[\sum_{p=0}^{t-2} \left(\frac{t-p-1}{T} \right) \varepsilon_{ip} - \frac{1}{T-t} \left(\sum_{q=t}^{T-1} \sum_{p=0}^{q-1} \left(\frac{q-p}{T} \right) \varepsilon_{ip} \right) \right], \\
D_{1,it-1} &= \sum_{q=0}^{t-2} x^*_{iq} - \frac{t-1}{T} \sum_{q=0}^{T-1} x^*_{iq}, \\
D_{2,it-1} &= \left[\sum_{q=1}^{t-2} \sum_{p=0}^{q-1} \left(\frac{q-p}{T} \right) \varepsilon_{ip} - \frac{t-1}{T} \sum_{q=1}^{T-1} \sum_{p=0}^{q-1} \left(\frac{q-p}{T} \right) \varepsilon_{ip} \right].
\end{aligned}
$$

Note that $E\left[\frac{1}{T^2}\sum_{t=2}^{T-1}\left(A_{it}D_{1,it-1} + B_{1,it}C_{it-1}\right)\right] \leq \frac{\bar{M}}{T}$ and

$$
Var\left[\frac{1}{T^2}\sum_{t=2}^{T-1}\left(A_{it}D_{1,it-1} + B_{1,it}C_{it-1}\right)\right] \leq \bar{M},
$$

for some constant $\bar{M} > 0$. By Chebychev's inequality, we then have

$$
\frac{1}{n^{3/4}T^2}\sum_{i=1}^{n} c_i \sum_{t=2}^{T-1}\left(A_{it}D_{1,it-1} + B_{1,it}C_{it-1}\right) = O_p\left(\frac{1}{n^{1/4}}\right) = o_p(1).
$$

Next, we apply a WLLN (e.g., see Phillips and Moon, 1999) to deduce that

$$
\begin{aligned}
\frac{1}{nT^2}\sum_{i=1}^{n} c_i^2 \sum_{t=2}^{T-1}\left(A_{it}D_{2,it-1} + B_{2,it}C_{it-1}\right) &\rightarrow {}_p E\left(c_i^2\right) \lim_{T\to\infty} E\left[\frac{1}{T^2}\sum_{t=2}^{T-1}\left(A_{it}D_{2,it-1} + B_{2,it}C_{it-1}\right)\right] \\
&= -\frac{E\left(c_i^2\right)\sigma^2}{18}
\end{aligned}
$$

as shown in Moon, Perron, and Phillips (2006). ■

Part (c): By the WLLN, we have

$$
\begin{aligned}
\frac{1}{nT^3}\sum_{i=1}^{n} c_i^2 \sum_{t=2}^{T-1} B_{it}D_{it-1} &= \frac{1}{nT^3}\sum_{i=1}^{n} c_i^2 \sum_{t=2}^{T-1} B_{1,it}D_{1,it-1} + o_p(1) \\
&\rightarrow_p E\left(c_i^2\right) \lim_{T\to\infty} E\left[\frac{1}{T^3}\sum_{t=2}^{T-1} E\left(B_{1,it}D_{1,it-1}\right)\right] = \frac{E\left(c_i^2\right)\sigma^2}{36}
\end{aligned}
$$

as shown in Moon, Perron, and Phillips (2006). ■