Supplementary Appendix

to

Unit Root Tests and Heavy-Tailed Innovations

by

I. Georgiev, P.M.M. Rodrigues and A.M.R. Taylor

Date: January 12, 2017

S.1 Introduction

This supplement contains additional Monte Carlo results and proof details for our paper "Unit Root Tests and Heavy-Tailed Innovations". Equation references (S.n) for $n \geq 1$ refer to equations in this supplement and other equation references are to the main paper.

The supplement is organised as follows. Additional Monte Carlo results relating to $T = 500$ are reported in section S.2. Proofs of Lemma 4.2, as well as selected details of Theorem 4.1 and Proposition 4.2, are given in section S.3. All additional references are included at the end of the supplement.

S.2 Additional Monte Carlo Results

Tables S.1-S.3 and Figures S.1-S.5 report for the case of $T = 500$ complementary finite sample local power results to those given for $T = 200$ in Tables 1-4 and Figures 1-5, respectively, in the main text. The Monte Carlo DGP and set-up of these experiments were otherwise exactly as detailed in sections 5.2 and 5.3.

S.3 Additional Proof Details

Let $\mathcal{P}_t := \sum_{j=0}^{t-1} \phi_T^{t-j} \varepsilon_{t-j}$ and $\mathcal{P}_{i,t} := \sum_{j=0}^{t-1} \phi_T^{t-j} \varepsilon_{i,t-j}$ $(i = 1, 2)$, so that $\mathcal{P}_t = \mathcal{P}_{1,t} + \gamma a_T^{-1} T^{1/2} \mathcal{P}_{2,t}$. In view of Lemma 4.1, summation by parts can be used to justify the standard joint convergence $(T^{-1/2}\mathcal{P}_{1,[Tr]}, a_T^{-1}\mathcal{P}_{2,[Tr]})' \Rightarrow (\sigma_1 J_c(r), \mathfrak{J}_{c,\alpha}(r))'$ in $D_2[0, 1]$, which by the continuity of $J_c(r)$ implies that $T^{-1/2}\mathcal{P}_{[Tr]} \Rightarrow \sigma_1 H_{c,\nu}(r)$ on $D[0, 1]$. As a direct result, for $\xi_t := \psi(1)\mathcal{P}_t$ it holds on $D[0, 1]$ that

$$T^{-1/2}\xi_{[Tr]} \Rightarrow \sigma_1 \psi(1) H_{c,\nu}(r). \tag{S.1}$$

Recall further that $u_t = \psi(L)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ with $\psi(L) := \sum_{j=0}^{\infty} \psi_j L^j$ and $\varepsilon_t := \varepsilon_{1t} + \gamma a_T^{-1} T^{1/2} \varepsilon_{2t}$. Thus, considering a Beveridge-Nelson decomposition of u_t we obtain,

$$u_t = \psi(1)\varepsilon_t + \widetilde{\varepsilon}_{t-1} - \widetilde{\varepsilon}_t \tag{S.2}$$

where $\widetilde{\varepsilon}_t := \widetilde{\psi}(L)\varepsilon_t = \sum_{j=0}^{\infty} \widetilde{\psi}_j \varepsilon_{t-j}$, with $\widetilde{\psi}_j := \sum_{k=j+1}^{\infty} \psi_k$. Alternatively, considering u_t as given in (3), we can write (S.2) as,

$$u_t = [\psi(1)\varepsilon_{1t} + \widetilde{\varepsilon}_{1,t-1} - \widetilde{\varepsilon}_{1t}] + \gamma a_T^{-1} T^{1/2} [\psi(1)\varepsilon_{2t} + \widetilde{\varepsilon}_{2,t-1} - \widetilde{\varepsilon}_{2t}].$$

Here the series for $\widetilde{\varepsilon}_{it}$, $i = 1, 2$, are well-defined $a.s.$ given that $\sum_{j=0}^{\infty} |\widetilde{\psi}_j|^\delta < \infty$ for δ of Assumption $\mathcal{A}.5$ ($\sum_{j=0}^{\infty} |\widetilde{\psi}_j|^\delta < \sum_{k=0}^{\infty} k|\psi_k|^\delta < \infty$; cf. Phillips and Solo (1992, pp.976,984)), and $\widetilde{\varepsilon}_{2t}$ belongs to the normal domain of attraction of a stable law with characteristic exponent α.

Finally, for x_t of (1) we find that

$$\begin{aligned} x_t &= \sum_{j=0}^{t} \phi_T^j u_{t-j} + \phi_T^t x_0 = \psi(1)\mathcal{P}_t - \widetilde{\varepsilon}_t + (1 - \phi_T) \sum_{j=1}^{t-1} \phi_T^{j-1} \widetilde{\varepsilon}_{t-j} + \widetilde{\varepsilon}_0 + \phi_T^t x_0 \\ &= \xi_t - \widetilde{\varepsilon}_t + \zeta_t, \end{aligned} \tag{S.3}$$

where the equality defines ζ_t. From $\max_{t=1,\dots,T} |\sum_{j=1}^{t-1} \phi_T^{j-1} \widetilde{\varepsilon}_{t-j}| \leq \sum_{t=1}^{T} |\widetilde{\varepsilon}_{1t}| + \gamma a_T^{-1} T^{1/2} \sum_{t=1}^{T} |\widetilde{\varepsilon}_{2t}|$, Markov's inequality and, for $\alpha = 1$, Karamata's theorem, it follows that $\max_{t=1,\dots,T} |\zeta_t| = O_p(1)$.

Proof of Lemma 4.2

Without loss of generality under our assumption that $x_0 = O_p(1)$, we may set $x_0 = 0$ in what follows.

i) Let $\Psi^2 := \sum_{j=0}^{\infty} \psi_j^2$. Consider first the sample variance of u_t, that is,

$$T^{-1}\sum_{t=1}^{T} u_t^2 = T^{-1}\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]^2 + \frac{\gamma^2}{a_T^2}\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{2t}\right]^2 + \frac{2\gamma}{T^{1/2}a_T}\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]\left[\psi(L)\varepsilon_{2t}\right].$$

Here $T^{-1}\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]^2 \xrightarrow{p} Var(\psi(L)\varepsilon_{11}) = \Psi^2\sigma_1^2$ by a law of large numbers [LLN], $a_T^{-2}\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{2t}\right]^2 \Rightarrow \Psi^2[\mathcal{U}_\alpha]_1$ by Theorem 4.2 of Davis and Resnick (1985), and $\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]\left[\psi(L)\varepsilon_{2t}\right] = o_p(T^{1/2}a_T)$ by Markov's inequality. In fact,

$$E\left|\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]\left[\psi(L)\varepsilon_{2t}\right]\right|^{\eta} \leq TE|\varepsilon_{1t}|^{\eta}E|\varepsilon_{2t}|^{\eta}(\sum_{j=1}^{\infty}|\psi_j|^{\eta})^2 = O(T),$$

where $\eta = 1$ if $\alpha > 1$ and $\eta \in [\delta/2, 1)$ is arbitrary if $\alpha = 1$, so $\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]\left[\psi(L)\varepsilon_{2t}\right] = O_p(T) = o_p(T^{1/2}a_T)$ if $\alpha > 1$ and $\sum_{t=1}^{T}\left[\psi(L)\varepsilon_{1t}\right]\left[\psi(L)\varepsilon_{2t}\right] = O(T^{1+\epsilon})$ for all $\epsilon > 0$ if $\alpha = 1$, with $O(T^{1+\epsilon}) = o_p(T^{1/2}a_T)$ for $\epsilon \in (0, \frac{1}{2})$ in the latter case. By collecting these facts, we establish that,

$$T^{-1}\sum_{t=1}^{T} u_t^2 \Rightarrow \Psi^2\sigma_1^2 + \gamma^2\Psi^2[\mathcal{U}_\alpha]_1 = \Psi^2\sigma_1^2\left(1 + \nu^2[\mathcal{U}_\alpha]_1\right).$$

ii) Using (S.3) and the uniform evaluation of ξ_t there, we find that

$$T^{-3/2}\sum_{t=1}^{T} x_t = T^{-3/2}\sum_{t=1}^{T}\xi_t - T^{-3/2}\sum_{t=1}^{T}\widetilde{\varepsilon}_t + o_p(1),$$

where further $\sum_{t=1}^{T}\widetilde{\varepsilon}_t = O_p(T)$ by the same argument as for the remainder in (S.3). Hence, by (S.1) and the Continuous mapping theorem [CMT], $T^{-3/2}\sum_{t=1}^{T} x_t \Rightarrow \psi(1)\sigma_1\int_0^1 H_{c,\nu}dr$.

iii) Again by (S.3) with a uniformly $O_p(1)$ remainder ζ_t,

$$\begin{aligned}
\frac{1}{4}\left|T^{-1}\sum_{t=1}^{T}(x_{t-1}^2 - \xi_{t-1}^2)\right| &\leq T^{-1}\sum_{t=1}^{T}|\xi_{t-1}(\widetilde{\varepsilon}_{t-1} - \zeta_{t-1})| + \frac{T^{-1}}{4}\sum_{t=1}^{T}(\widetilde{\varepsilon}_{t-1} - \zeta_{t-1})^2 \\
&\leq \max_{t=1,\ldots,T}|\xi_{t-1}|(\frac{1}{T}\sum_{t=1}^{T}|\widetilde{\varepsilon}_{2,t-1}| + \frac{1}{T^{1/2}a_T}\sum_{t=1}^{T}|\widetilde{\varepsilon}_{2,t-1}| + O_p(1)) \\
&\quad + T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{1,t-1}^2 + \frac{1}{a_T^2}\sum_{t=1}^{T}\widetilde{\varepsilon}_{2,t-1}^2 + o_p(T) = o_p(T)
\end{aligned}$$

because $\max_{t=1,\ldots,T}|T^{-1/2}\xi_{t-1}| \Rightarrow \sigma_1|\psi(1)|\sup_{[0,1]}|H_{c,\nu}| < \infty$ a.s. by (S.1) and the CMT, $\sum_{t=1}^{T}|\widetilde{\varepsilon}_{2,t-1}^i| = O_p(T)$, $i = 1, 2$, by an LLN, $\sum_{t=1}^{T}|\widetilde{\varepsilon}_{2,t-1}| = O_p(T)$ for $\alpha > 1$ by an LLN, $\sum_{t=1}^{T}|\widetilde{\varepsilon}_{2,t-1}| = O_p(Tl_T)$ with a slowly varying l_T for $\alpha = 1$ by Markov's inequality, and $\sum_{t=1}^{T}\widetilde{\varepsilon}_{2,t-1}^2 = O_p(a_T^2)$ by Theorem 4.2 of Davis and Resnick (1985). Therefore, $T^{-2}\sum_{t=1}^{T} x_{t-1}^2 = \sum_{t=1}^{T}\xi_{t-1}^2 + o_p(1) \Rightarrow \{\psi(1)\}^2\sigma_1^2\int_0^1 H_{c,\nu}^2 dr$ by (S.1) and the CMT.

iv) Regarding $T^{-1}\sum_{t=1}^{T} x_{t-1}u_t$, following Phillips (1988, 1990) we observe that:

$$\sum_{t=1}^{T} x_t^2 = \sum_{t=1}^{T}(\phi_T x_{t-1} + u_t)^2 = \sum_{t=1}^{T}(\phi_T^2 x_{t-1}^2 + 2\phi_T x_{t-1}u_t + u_t^2).$$

Since $\phi_T^2 = (1 - c/T)^2 = 1 - 2c/T + c^2/T^2$, it follows that

$$\sum_{t=1}^{T} x_t^2 = \sum_{t=1}^{T} x_{t-1}^2 - \frac{2c}{T}\sum_{t=1}^{T} x_{t-1}^2 + \frac{c^2}{T^2}\sum_{t=1}^{T} x_{t-1}^2 + 2\sum_{t=1}^{T} x_{t-1}u_t - \frac{2c}{T}\sum_{t=1}^{T} x_{t-1}u_t + \sum_{t=1}^{T} u_t^2.$$

Hence,

$$x_T^2 = \sum_{t=1}^{T}\left(x_t^2 - x_{t-1}^2\right) = -\frac{2c}{T}\sum_{t=1}^{T} x_{t-1}^2 + 2\sum_{t=1}^{T} x_{t-1}u_t + \sum_{t=1}^{T} u_t^2 + o_p(T).$$

Thus, we establish that

$$T^{-1}\sum_{t=1}^{T} x_{t-1}u_t = \frac{1}{2}\left(T^{-1}x_T^2 + T^{-2}2c\sum_{t=1}^{T} x_{t-1}^2 - T^{-1}\sum_{t=1}^{T} u_t^2\right) + o_p(1) \tag{S.4}$$

From (S.3), $x_T^2 = \xi_T^2 - 2\xi_T(\widetilde{\varepsilon}_T - \zeta_T) + (\widetilde{\varepsilon}_T - \zeta_T)^2$, where $T^{-1/2}\xi_T \Rightarrow \sigma_1\psi(1)\,H_{c,\nu}(1)$ by (S.1) and the CMT, and $\widetilde{\varepsilon}_T = \widetilde{\varepsilon}_{1T} + T^{1/2}a_T^{-1}\widetilde{\varepsilon}_{2T} = \widetilde{\varepsilon}_{1T} + o_p(1) = O_p(1)$ because $\{\widetilde{\varepsilon}_{1t}\}$ and $\{\widetilde{\varepsilon}_{2t}\}$ are stationary with a.s. finite terms. Thus, $x_T^2 \Rightarrow \{\psi(1)\}^2\sigma_1^2 H_{c,\nu}^2(1)$. Considering also Lemma 4.2(i, iii), we establish that,

$$T^{-1}\sum_{t=1}^{T} x_{t-1}u_t \Rightarrow \frac{1}{2}\left\{\{\psi(1)\}^2\sigma_1^2 H_{c,\nu}^2(1) + 2c\sigma_1^2\{\psi(1)\}^2\int_0^1 H_{c,\nu}^2(r)dr - \Psi^2\sigma_1^2[H_{0,\nu}]_1\right\}.$$

Finally, we obtain the limit in Lemma 4.2(iv) by straightforward manipulations and using the identity

$$H_{c,\nu}^2(1) \equiv [H_{0,\nu}]_1 - 2c\int_0^1 H_{c,\nu}^2(r)dr + 2\int_0^1 H_{c,\nu}(r)dH_{0,\nu}(r). \tag{S.5}$$

v) The convergence of $T^{-1}\sum_{t=1}^{T} x_{t-1}\varepsilon_t$ can be deduced from part (iv) and the identities

$$\begin{aligned}
\sum_{t=1}^{T} x_{t-1}\varepsilon_t &= \{\psi(1)\}^{-1}\sum_{t=1}^{T} x_{t-1}u_t + \sum_{t=1}^{T} x_{t-1}\Delta\widetilde{\varepsilon}_t \\
&= \{\psi(1)\}^{-1}\sum_{t=1}^{T} x_{t-1}u_t - \sum_{t=1}^{T} \Delta x_t\widetilde{\varepsilon}_t + x_T\widetilde{\varepsilon}_T \\
&= \{\psi(1)\}^{-1}\sum_{t=1}^{T} x_{t-1}u_t - \sum_{t=1}^{T} u_t\widetilde{\varepsilon}_t + T^{-1}c\sum_{t=1}^{T} x_{t-1}\widetilde{\varepsilon}_t + x_T\widetilde{\varepsilon}_T.
\end{aligned}$$

Handling mixed products as in the proof of part (i), we find that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} u_t\widetilde{\varepsilon}_t &= \frac{1}{T}\sum_{t=1}^{T} u_{1t}\widetilde{\varepsilon}_{1t} + \frac{\gamma^2}{a_T^2}\sum_{t=1}^{T} u_{2t}\widetilde{\varepsilon}_{2t} + o_p(1) \\
&\Rightarrow Cov(u_{11}, \widetilde{\varepsilon}_{11}) + \gamma^2[\mathcal{U}_\alpha]_1\sum_{i=0}^{\infty}\psi_i\widetilde{\psi}_i = \sigma_1^2[H_{0,\nu}]_1\sum_{i=0}^{\infty}\psi_i\widetilde{\psi}_i
\end{aligned}$$

by an LLN and Theorem 4.2 of Davis and Resnick (1985). As $|\sum_{t=1}^{T} x_{t-1}\widetilde{\varepsilon}_t| \le \max_{t=1,\ldots,T}|x_t|\sum_{t=1}^{T}|\widetilde{\varepsilon}_t| = O_p(T^{3/2})$, see the derivation of (S.3), and $x_T\widetilde{\varepsilon}_T = O_p(T^{1/2})$, it remains to apply part (iv) to $\sum_{t=1}^{T} x_{t-1}u_t$ and to observe that $\sum_{i=0}^{\infty}\psi_i\widetilde{\psi}_i = \frac{1}{2}(\{\psi(1)\}^2 - \Psi^2)$.

vi) First, $T^{-2}\sum_{t=1}^{T} x_{t-1}^2\varepsilon_t^2 = T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2\varepsilon_t^2 + o_p(1)$ since, using (S.3),

$$\begin{aligned}\sum_{t=1}^{T}\left|x_{t-1}^2-\xi_{t-1}^2\right|\varepsilon_t^2 &= \sum_{t=1}^{T}\left|2\xi_{t-1}(\widetilde{\varepsilon}_t-\zeta_t)+(\widetilde{\varepsilon}_t-\zeta_t)^2\right|\varepsilon_t^2\\ &\leq 2\max_{t=1,\dots,T}|\xi_t|\,(\sum_{t=1}^{T}|\widetilde{\varepsilon}_{t-1}|\varepsilon_t^2+\max_{t=1,\dots,T}|\zeta_t|\sum_{t=1}^{T}\varepsilon_t^2)\\ &\quad +2\sum_{t=1}^{T}\widetilde{\varepsilon}_{t-1}^2\varepsilon_t^2+2\max_{t=1,\dots,T}\zeta_t^2\sum_{t=1}^{T}\varepsilon_t^2\end{aligned}$$

with (i) $\max|T^{-1/2}\xi_t| = O_p(1)$, as a consequence of the fact that it converges weakly, (ii),

$$\sum_{t=1}^{T}|\widetilde{\varepsilon}_{t-1}|\varepsilon_t^2 \leq 2\sum_{i,j=1}^{2}(\gamma a_T^{-1}T^{1/2})^{i+2j-3}\sum_{t=1}^{T}|\widetilde{\varepsilon}_{i,t-1}|\varepsilon_{jt}^2 = o_p(T^{3/2})$$

by LLN for $j=1$ and by Markov's inequality for $j=2$:

$$E(\sum_{t=1}^{T}|\widetilde{\varepsilon}_{i,t-1}|\varepsilon_{2t}^2)^{\eta/2} \leq \sum_{t=1}^{T}E|\widetilde{\varepsilon}_{i,t-1}|^{\eta/2}E|\varepsilon_{2t}|^{\eta} = O\left(T\right),$$

so $\sum|\widetilde{\varepsilon}_{i,t-1}|\varepsilon_{2t}^2 = O\left(T^{2/\eta}\right)$ for all $\eta\in(0,\alpha)$, (iii), $T^{-1}\sum\varepsilon_t^2 = O_p(1)$, again because it converges weakly, and (iv),

$$\sum_{t=1}^{T}\widetilde{\varepsilon}_{t-1}^2\varepsilon_t^2 \leq 4\sum_{i,j=1}^{2}(\gamma^2a_T^{-2}T)^{i+j-2}\sum_{t=1}^{T}\widetilde{\varepsilon}_{i,t-1}^2\varepsilon_{jt}^2 = o_p(T^2)$$

by LLN for $i=j=1$ and by Markov's inequality applied to the $\eta/2$ powers otherwise.

Second, we turn to $T^{-2}\sum\xi_{t-1}^2\varepsilon_t^2$. It holds that $(T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_t, T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_t^2)\Rightarrow(\sigma_1H_{\nu,0}(r),\sigma_1^2[H_{\nu,0}]_r)$ in $D_2[0,1]$ because, (i), $(T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{1t}, a_T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{2t}, a_T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{2t}^2)\Rightarrow(\sigma_1W(r),\mathcal{U}_\alpha(r),[\mathcal{U}_\alpha]_r)$ in $D_3[0,1]$ by Theorem 4 of Resnick and Greenwood (1979) and the independence of $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$, and (ii),

$$\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_t^2 = \frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{1t}^2+\frac{\gamma^2}{a_T^2}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{2t}^2+\frac{2\gamma}{T^{1/2}a_T}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{1t}\varepsilon_{2t}\Rightarrow\sigma_1^2r+\gamma^2[\mathcal{U}_\alpha]_r=\sigma_1^2[H_{\nu,0}]_r \quad \text{(S.6)}$$

because $\{\varepsilon_{1t}\varepsilon_{2t}\}$ is IID with tail index α, so $\max_{r\in[0,1]}|T^{-1/2}a_T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_{1t}\varepsilon_{2t}| \xrightarrow{p} 0$. By Theorem 2.7 of Kurtz and Protter (1991), it follows that $T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2\varepsilon_t^2 = [\psi(1)]^2T^{-2}\sum_{t=1}^{T}(\sum_{s=1}^{t-1}\varepsilon_s)^2\varepsilon_t^2 \Rightarrow [\psi(1)]^2\sigma_1^4\int H_{\nu,0}^2d[H_{\nu,0}]$, where condition C2.7 of the theorem can be checked as on pp.784-786 of Paulauskas and Rachev (1998). Recalling the previous paragraph, we conclude that $T^{-2}\sum_{t=1}^{T}x_{t-1}^2\varepsilon_t^2$ converges weakly to the same limit as that of $T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2\varepsilon_t^2$. ■

Proof of Theorem 4.1 (complements).

Evaluation of S_{00}. Upon splitting the observations and the product moments into the contributions of the finite and the infinite variance components, with notation corresponding to decomposition in (4), we argue in steps that $\|S_{00}-S_{00}^{(1)}-S_{00}^{(2)}\|_* = o_p(T)$, where $\|\cdot\|_*$ denotes the spectral matrix norm, $S_{00}^{(i)} := \sum_{t=1}^{T}\Delta X_{i,t-1}\Delta X_{i,t-1}'$ $(i=1,2)$ and the norming sequence $\gamma a_T^{-1}T^{1/2}$ is incorporated into ΔX_{2t}. Thus, defining $C_T := \{c_{ij}\}_{i,j=1}^{p_T}$ with $c_{ij} := \sum_{t=0}^{T-1}\Delta x_{1t}\Delta x_{2,t-|i-j|}$, we find that

$$\begin{aligned}\frac{1}{2}\|S_{00}-S_{00}^{(1)}-S_{00}^{(2)}\|_* &\leq \|\sum_{t=1}^{T}\Delta X_{1,t-1}\Delta X_{2,t-1}'\|_* \leq \|C_T\|_*+p_T^2\max_{i=1,\dots,p_T}\{(\Delta x_{-i})^2+(\Delta x_{T-i})^2\}\\ &= \|C_T\|_*+p_T^2\{o_p(p_T)+a_T^{-2}TO_p(a_{p_T}^2)\} = \|C_T\|_*+o_p(T) \quad \text{(S.7)}\end{aligned}$$

under $p_T^3/T \to 0$. Further, given the Toeplitz structure of C_T,

$$\frac{1}{2}\|C_T\|_* \leq \sum_{i=1}^{p_T} |c_{1i}| \leq \sum_{i=1}^{p_T} |\sum_{t=0}^{T-1} u_{1t}u_{2,t-i+1}| + O_p(p_T T^{1/2+\epsilon}), \tag{S.8}$$

where $O_p(p_T T^{1/2+\epsilon})$, with $\epsilon > 0$ arbitrary, stands for

$$\sum_{i=1}^{p_T} \{(\frac{c}{T})^2 \sum_{t=0}^{T-1} x_{1,t-1}x_{2,t-i} + \frac{c}{T} \max_{t=1,\ldots,T}(|x_{1,t-1}| + |x_{2,t-i}|) \sum_{t=0}^{T-1}(|u_{1t}| + |u_{2,t-i+1}|)\},$$

given that $\max_{t=-p_T,\ldots,T} |x_{1t}| = O_p(T^{1/2})$ and $\sum_{i=1}^{p_T} \sum_{t=0}^{T-1}(|u_{1,t}| + |u_{2,t-i+1}|) \leq p_T \sum_{t=0}^{T-1} |u_{1,t}| + p_T \sum_{t=-p_T}^{T-1} |u_{2,t}| = O_p(p_T l_T T)$ with a slowly varying l_T (constant except for $\alpha = 1$). Regarding $\sum_{t=0}^{T-1} u_{1,t}u_{2,t-i+1} = a_T^{-1}T^{1/2}\gamma(\chi_i^{\leq} + \chi_i^{>})$, with

$$\chi_i^R := \sum_{t=0}^{T-1} \sum_{u,v=0}^{\infty} \psi_u\psi_v\varepsilon_{1,t-u}\varepsilon_{2,t-v-i+1}\mathbb{I}_{|\varepsilon_{2,t-v-i+1}|Ra_T},\ R \in \{\leq, >\},$$

it holds that (i), $E\sum_{i=1}^{p_T} |\chi_i^{\leq}| \leq \sum_{i=1}^{p_T}\{E(\chi_i^{\leq})^2\}^{1/2}$ by Jensen's inequality, where, using Karamata's theorem, we find that

$$E(\chi_i^{\leq})^2 \leq TE\varepsilon_{11}^2 E(\varepsilon_{21}^2\mathbb{I}_{|\varepsilon_{21}|\leq a_T})(\sum_{u=0}^{\infty} |\psi_u|)^4 = O(a_T^2),$$

where $\mathbb{I}$ denotes the usual indicator function, because $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ are independent, $E\varepsilon_{1,t-u} = 0$, $E(\varepsilon_{2,t-v-i+1}\mathbb{I}_{|\varepsilon_{2,t-v-i+1}|\leq a_T}) = 0$ by symmetry, and $E(\varepsilon_{21}^2\mathbb{I}_{|\varepsilon_{21}|\leq a_T}) = E(\varepsilon_{2,t-v-i+1}^2\mathbb{I}_{|\varepsilon_{2,t-v-i+1}|\leq a_T}) = O(T^{-1}a_T^2)$, and (ii), $E(\sum_{i=1}^{p_T} |\chi_i^{>}|)^\eta \leq \sum_{i=1}^{p_T} E|\chi_i^{>}|^\eta$, where $\eta = 1$ for $\alpha > 1$, $\eta \in [\delta, 1)$ is arbitrary for $\alpha = 1$, and

$$E|\chi_i^{>}|^\eta \leq \sum_{t=0}^{T-1} \sum_{u,v=0}^{\infty} |\psi_u|^\eta|\psi_v|^\eta E|\varepsilon_{11}|^\eta E(|\varepsilon_{21}|^\eta\mathbb{I}_{|\varepsilon_{21}|>a_T}) = O(a_T^\eta)(\sum_{u=0}^{\infty} |\psi_u|^\eta)^2 = O(a_T^\eta)$$

using Karamata's theorem again, so eventually, by Markov's inequality, $\|S_{00} - S_{00}^{(1)} - S_{00}^{(2)}\|_* \leq O_p(p_T T^{1/2+\epsilon}) + o_p(T) = o_p(T)$, because $p_T^3/T \to 0$ as $T \to \infty$, where $\epsilon > 0$ is arbitrary. Let $\Sigma_p := \{r_{|i-j|}\}_{i,j=1}^{p_T}$ with $r_i := \sum_{j=0}^{\infty} \psi_i\psi_{j+i}$; then the eigenvalues of Σ_p are bounded and bounded away from zero under Assumptions $\mathcal{A}$.1 and $\mathcal{A}$.5. As additionally, under $p_T^3/T \to 0$, $\|S_{00}^{(1)} - T\Sigma_p\sigma_1^2\|_* = o_p(T)$ by Lemma 3 of Berk (1974) and $\|S_{00}^{(2)} - a_T^{-2}T\gamma^2\Sigma_p\sum_{t=1}^{T}\varepsilon_{2t}^2\|_* = o_p(T)$ by Lemma 2 of Cavaliere *et al.* (2016), by combining the previous results it follows that

$$\|T^{-1}S_{00} - \Sigma_p(\sigma_1^2 + \gamma^2 a_T^{-2}\sum_{t=1}^{T}\varepsilon_{2t}^2)\|_* = o_p(1), \tag{S.9}$$

and using inequality (2.15) of Berk (1974), also (A.3) holds.

Evaluation of $S_{00}^{-1}S_{0\varepsilon}$. The vector $S_{00}^{-1}S_{0\varepsilon}$ can be decomposed as

$$(S_{00}^{-1}S_{00}^{(1)})(S_{00}^{(1)})^{-1}S_{0\varepsilon}^{(1)} + (S_{00}^{-1}S_{00}^{(2)})(S_{00}^{(2)})^{-1}S_{0\varepsilon}^{(2)} + S_{00}^{-1}\sum_{t=1}^{T}(\Delta X_{1,t-1}\varepsilon_{2t,p_T} + \Delta X_{2,t-1}\varepsilon_{1t,p_T}),$$

where $\|(S_{00}^{(1)})^{-1}S_{0\varepsilon}^{(1)}\| = o_p(p_T^{-2})$ as in Lemma 3.2 of Chang and Park (2002), $(S_{00}^{(2)})^{-1}S_{0\varepsilon}^{(2)} = O_p(a_{p_T}a_T^{\epsilon-1} + \sum_{i=p_T+1}^{\infty}|\beta_i|)$ for all $\epsilon > 0$ as in Equation (7.1) of Cavaliere *et al.* (2016), both under the condition that $p_T^2/T + 1/p_T \to 0$ as $T \to \infty$, and $\| \sum_{t=1}^T (\Delta X_{1,t-1}\varepsilon_{2t,p_T} + \Delta X_{2,t-1}\varepsilon_{1t,p_T})\| = O_p(p_T T^{1/2})$ by Markov's inequality and Karamata's theorem as, e.g., for the first kind of summands,

$$\begin{aligned}
\| \sum_{t=1}^T \Delta X_{1,t-1}\varepsilon_{2t,p_T}\|^2 &= \sum_{i=1}^{p_T}(\sum_{t=1}^T \Delta x_{1,t-i}\varepsilon_{2t,p_T})^2 \\
&\le 2\sum_{i=1}^{p_T}\{\sum_{t=1}^T \Delta x_{1,t-i}(a_T^{-1}T^{1/2}\varepsilon_{2t} + \sum_{i=p_T+1}^{\infty} \beta_i u_{2,t-i})\}^2 \\
&\quad +2c^2 p_T T^{-2} \max_{t=-p_T,\dots,T} x_{2t}^2 (\sum_{t=1-p_T}^{T-1} |\Delta x_{1t}|)^2 (\sum_{i=1}^{\infty} |\beta_i|)^2
\end{aligned}$$

with $\max_{t=-p_T,\dots,T} |x_{2t}| = O_p(T^{1/2})$, $\sum_{t=1-p_T}^{T-1} |\Delta x_{1t}| = O_p(T)$,

$$\begin{aligned}
E\sum_{i=1}^{p_T}\{\sum_{t=1}^T \Delta x_{1,t-i}\varepsilon_{2t}\mathbb{I}_{|\varepsilon_{2t}|\le a_T}\}^2 &= p_T T E(\varepsilon_{21}^2\mathbb{I}_{|\varepsilon_{21}|\le a_T})[E(\Delta x_{11})^2] = O(p_T a_T^2), \\
E[\sum_{i=1}^{p_T}\{\sum_{t=1}^T \Delta x_{1,t-i}\varepsilon_{2t}\mathbb{I}_{|\varepsilon_{2t}|> a_T}\}^2]^{\eta/2} &\le p_T T E(|\varepsilon_{21}|^{\eta}\mathbb{I}_{|\varepsilon_{21}|> a_T}) E|\Delta x_{11}|^{\eta} = O(p_T a_T^{\eta})
\end{aligned}$$

for $\eta = 1$ if $\alpha > 1$ and $\eta \in [\delta, 1)$ arbitrary if $\alpha = 1$, and similarly for the terms involving u_{2t}:

$$\begin{aligned}
&E\sum_{m=1}^{p_T}\{\sum_{t=1}^T \Delta x_{1,t-m} \sum_{i=p_T+1}^{\infty} \beta_i \sum_{j=0}^{\infty} \psi_j \varepsilon_{2,t-i-j}\mathbb{I}_{|\varepsilon_{2,t-i-j}|\le a_T}\}^2 \\
\le\ & p_T T E(\varepsilon_{21}^2\mathbb{I}_{|\varepsilon_{21}|\le a_T})[E(\Delta x_{11})^2](\sum_{i=p_T+1}^{\infty} |\beta_i| \sum_{j=0}^{\infty} |\psi_j|)^2 = o(a_T^2),
\end{aligned}$$

$$\begin{aligned}
&E[\sum_{m=1}^{p_T}\{\sum_{t=1}^T \Delta x_{1,t-m} \sum_{i=p_T+1}^{\infty} \beta_i \sum_{j=0}^{\infty} \psi_j \varepsilon_{2,t-i-j}\mathbb{I}_{|\varepsilon_{2,t-i-j}|> a_T}\}^2]^{\eta/2} \\
\le\ & p_T T E(|\varepsilon_{21}|^{\eta}\mathbb{I}_{|\varepsilon_{21}|> a_T}) E|\Delta x_{11}|^{\eta} \sum_{i=p_T+1}^{\infty} |\beta_i|^{\eta} \sum_{j=0}^{\infty} |\psi_j|^{\eta} = o(p_T a_T^{\eta})
\end{aligned}$$

since $p_T \sum_{i=p_T+1}^{\infty} |\beta_i| \to 0$, $\sum_{i=p_T+1}^{\infty} |\beta_i|^{\eta} \to 0$ as $p_T \to \infty$. Accounting also for (A.3), which implies that $\|TS_{00}^{-1}\|_* = O_p(1)$, it follows that for all $\epsilon > 0$, (A.4) holds.

Evaluation of S_{10}. It holds that

$$\|S_{10} - \mathbf{1}_{p_T}' \sum_{t=1}^T x_{t-1}\Delta x_t - Tr\{(S_{00}^{(1)} + S_{00}^{(2)} + 2C_T)\Upsilon_{p_T}\}\| = o_p(T),$$

where $\mathbf{1}_{p_T}$ is a p_T-vector of ones, Υ_{p_T} is an upper triangular matrix with ones on and above the main diagonal, and the difference is due to presample and end-of-sample contributions as in (S.7). Thus, further,

$$\begin{aligned}
\|S_{10}\| &\le p_T^{1/2}|\sum_{t=1}^T x_{t-1}\Delta x_t| + \|Tr(S_{00}^{(1)}\Upsilon_{p_T})\| + \|Tr(S_{00}^{(2)}\Upsilon_{p_T})\| + 2\|Tr(C_T\Upsilon_{p_T})\| + o_p(T) \\
&\le O_p(p_T^{1/2}T) + p_T^{1/2}\sum_{i=1}^{p_T} |c_{1i}| = O_p(p_T^{1/2}T)
\end{aligned}$$

since $\sum_{t=1}^T x_{t-1}\Delta x_t = \sum_{t=1}^T x_{t-1}u_t - (c/T)\sum_{t=1}^T x_{t-1}^2 = O_p(T)$ by Lemma 4.2(iii,iv), $\|Tr(S_{00}^{(i)}\Upsilon_{p_T})\| = O_p(p_T^{1/2}T)$ $(i=1,2)$ is shown in the proof of Lemma 3.2(b) of Chang and Park (2002) and Lemma A.1(d) of Cavaliere *et al.* (2017), and $\sum_{i=1}^{p_T}|c_{1i}| = O_p(p_T T^{1/2+\epsilon})$ for all $\epsilon > 0$ by the argument following (S.8) and $p_T^3/T \to 0$. ■

Proof of Proposition 4.2 (complements).

Convergence of $a_T^{-2}\sum_{t=1}^{[Tr]} Z_t$ and $a_T^{-2}\sum_{t=1}^{[Tr]} Z_t^*$. Possibly upon an expansion of the probability space, take $\{\boldsymbol{\varepsilon}_t^*\}$ distributed as before and independent of $\{\varepsilon_{1t}, \varepsilon_{2t}\}$. For a fixed $\delta > 0$, let I_t be the indicator of the event that $\{|a_T^{-1}\varepsilon_{2t}| > \delta$ and $|a_T^{-1}\varepsilon_{2,t-2}| \le \delta$ and $\{|a_T^{-1}\varepsilon_{2,t+2}| \le \delta\}$. Then $\tilde{Z}_t = \iota_{t\delta}Z_t + (1-\iota_{t\delta})Z_t^*$, $t \in \mathbb{N}$, defines an IID sequence independent of $\{\varepsilon_{2t}\}$, and hence, $a_T^{-2}\sum_{t=1}^{[Tr]}\tilde{Z}_t \Rightarrow L(r)$ in $D_5[0,1]$. On the other hand, for every $\lambda > 0$,

$$\lim_{\delta\to 0}\limsup_{T\to\infty} P\left(\max_{s=1,\dots,T}\left\|a_T^{-2}\sum_{t=1}^{s}(\tilde{Z}_t - Z_t)\right\| \ge \lambda\right) = 0,$$

which by Theorem 4.2 of Billingsley (1968) implies that also $a_T^{-2}\sum_{t=1}^{[Tr]} Z_t \Rightarrow L(r)$ in $D_5[0,1]$. In fact, let

$$e_t = \{Var(\varepsilon_{2t}^2\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2,t-2}|\le\delta\}})\}^{-1/2}\{\varepsilon_{2t}^2\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2,t-2}|\le\delta\}} - E(\varepsilon_{2t}^2\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2,t-2}|\le\delta\}})\};$$

since $\sum_{t=1}^T \iota_{t\delta} = \sum_{t=1}^T \mathbb{I}_{\{|a_T^{-1}\varepsilon_{2t}|>\delta\}}$ with probability approaching one as $T\to\infty$, with the same probability it holds that

$$\begin{aligned}
\max_{s=1,\dots,T}\left\|a_T^{-2}\sum_{t=1}^{s}(\tilde{Z}_t - Z_t)\right\| &= \max_{s=1,\dots,T}\left\|a_T^{-2}\sum_{t=1}^{[Tr]}\varepsilon_{2t}^2\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2t}|<\delta\}}(\boldsymbol{\varepsilon}_t^* - \boldsymbol{\varepsilon}_t)\right\| \\
&\le T^{1/2}E(a_T^{-2}\varepsilon_{2t}^2\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2t}|\le\delta\}})\max_{r\in[0,1]}\left\|T^{-1/2}\sum_{t=1}^{[Tr]}(\boldsymbol{\varepsilon}_t^* - \boldsymbol{\varepsilon}_t)\right\| \\
&\quad + \{TE(a_T^{-4}\varepsilon_{2t}^4\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2,t-2}|\le\delta\}})\}^{1/2}\max_{s=1,\dots,T}\left\|T^{-1/2}\sum_{t=1}^{[Tr]}e_t(\boldsymbol{\varepsilon}_t^* - \boldsymbol{\varepsilon}_t)\right\|,
\end{aligned}$$

where $T^{1/2}E(a_T^{-2}\varepsilon_{2t}^2\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2,t}|\le\delta\}}) \to 0$ as $T\to\infty$ and $TE(a_T^{-4}\varepsilon_{2t}^4\mathbb{I}_{\{|a_T^{-1}\varepsilon_{2,t}|\le\delta\}}) \to \delta^{4-\alpha}\alpha/(4-\alpha) \to 0$ as $T\to\infty$ followed by $\delta\to 0$, both by Karamata's theorem, whereas the maximum over r does not depend on δ and converges weakly as $T\to\infty$ to the maximum on $[0,1]$ of a Wiener processes, while the maximum over s is $O_P(1)$ as $T\to\infty$, uniformly in δ, by Kolmogorov's maximal inequality.

Derivation of eq. (A.11). Consider additionally an IID sequence $\{\varepsilon_{1t}^{**}\}$ independent of the random elements introduced so far and with ε_{1t}^{**} distributed like ε_{11}. Next, in W_T replace ε_{1t} by ε_{1t}^{**} whenever ε_{1t} was retained in $\{\tilde{Z}_t\}$:

$$W_{T,\delta}(r) = W_T(r) + T^{-1/2}\sum_{t=1}^{[Tr]}\iota_{t\delta}(\varepsilon_{t-1}^{**} + \varepsilon_{t+1}^{**} - \varepsilon_{t-1} - \varepsilon_{t+1}).$$

Then $W_{T,\delta}$ is distributed like W_T, so $W_{T,\delta}(r) \Rightarrow W(r)$. Since $W_{T,\delta}(r)$ and $a_T^{-2}\sum_{t=1}^{[Tr]}\tilde{Z}_t$ are independent, their convergence is joint and to independent limits. On the other hand, since

$$\begin{aligned}
\max_{r\in[0,1]}\|W_{T,\delta}(r) - W_T(r)\| &\le (E\iota_{t\delta})^{1/2}\max_{r\in[0,1]}\left\|T^{-1/2}\sum_{t=1}^{[Tr]}\frac{\iota_{t\delta} - E\iota_{t\delta}}{\{Var(\iota_{t\delta})\}^{1/2}}(\varepsilon_{t-1}^{**} + \varepsilon_{t+1}^{**} - \varepsilon_{t-1} - \varepsilon_{t-2})\right\| \\
&\quad + E\iota_{t\delta}\max_{r\in[0,1]}\left\|T^{-1/2}\sum_{t=1}^{[Tr]}(\varepsilon_{t-1}^{**} + \varepsilon_{t+1}^{**} - \varepsilon_{t-1} - \varepsilon_{t-2})\right\| \xrightarrow{P} 0
\end{aligned}$$

as $T \to \infty$, because $E\iota_{t\delta} \to 0$ and the maxima over r converge weakly to maxima of Wiener processes with variances independent of δ, we can conclude that $a_T^{-2} \sum_{t=1}^{[Tr]} Z_t$ and $W_T(r)$ converge like $a_T^{-2} \sum_{t=1}^{[Tr]} \tilde{Z}_t$ and $W_{T,\delta}(r)$, as stated in (A.11). ∎

S.4 Additional references

Davis, R.A. and S. Resnick (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities, The Annals of Probability 13(1), 179-195.

Resnick, S.I. and P. Greenwood (1979) A bivariate stable characterization and domains of attraction, Journal of Multivariate Analysis 9, 206-221.

Table S.1. Empirical size of unit root tests under OLS and local GLS de-meaning. The DGP is (25) and (26) with $T = 500$ and $\gamma = 0.1$.

			OLS de-meaning							Local GLS de-meaning						
φ	θ	α_2	$t_{\widehat{\rho}}$	$\mathcal{VRT}$	$\mathcal{MSB}$	$\mathcal{MZ}_{\varphi}$	$\mathcal{MZ}_t$	$t^W_{\widehat{\rho},1}$	$t^{W,\widehat{\alpha}_u}_{\widehat{\rho},1}$	$t_{\widehat{\rho}}$	$\mathcal{VRT}$	$\mathcal{MSB}$	$\mathcal{MZ}_{\varphi}$	$\mathcal{MZ}_t$	$t^W_{\widehat{\rho},1}$	$t^{W,\widehat{\alpha}_u}_{\widehat{\rho},1}$
0	0	1.75	0.044	0.053	0.045	0.045	0.042	0.050	0.050	0.045	0.054	0.045	0.046	0.045	0.051	0.050
		1.5	0.044	0.050	0.043	0.044	0.041	0.046	0.049	0.054	0.051	0.053	0.052	0.053	0.056	0.048
		1.25	0.042	0.054	0.043	0.046	0.042	0.050	0.044	0.049	0.051	0.051	0.049	0.049	0.054	0.051
		1	0.043	0.054	0.047	0.046	0.042	0.044	0.040	0.048	0.056	0.050	0.047	0.047	0.051	0.063
0.5	0	1.75	0.045	0.051	0.052	0.052	0.046	0.051	0.054	0.045	0.053	0.047	0.046	0.046	0.053	0.050
		1.5	0.043	0.047	0.051	0.050	0.043	0.045	0.051	0.052	0.050	0.055	0.053	0.053	0.054	0.048
		1.25	0.043	0.051	0.051	0.052	0.043	0.048	0.044	0.049	0.050	0.054	0.052	0.052	0.054	0.052
		1	0.044	0.051	0.052	0.051	0.045	0.044	0.040	0.049	0.055	0.051	0.052	0.051	0.053	0.062
-0.5	0	1.75	0.043	0.056	0.042	0.042	0.040	0.057	0.057	0.047	0.055	0.045	0.045	0.046	0.055	0.055
		1.5	0.042	0.053	0.041	0.041	0.038	0.049	0.056	0.051	0.053	0.051	0.048	0.048	0.059	0.051
		1.25	0.041	0.057	0.039	0.042	0.038	0.054	0.047	0.049	0.052	0.049	0.046	0.046	0.057	0.055
		1	0.044	0.056	0.044	0.044	0.040	0.048	0.046	0.047	0.057	0.047	0.045	0.046	0.053	0.063
0	0.5	1.75	0.043	0.052	0.051	0.054	0.048	0.056	0.056	0.048	0.053	0.052	0.049	0.049	0.055	0.053
		1.5	0.040	0.049	0.052	0.054	0.047	0.049	0.052	0.049	0.051	0.054	0.052	0.052	0.056	0.049
		1.25	0.041	0.053	0.050	0.051	0.046	0.053	0.044	0.047	0.047	0.051	0.051	0.052	0.053	0.052
		1	0.036	0.053	0.053	0.050	0.045	0.045	0.045	0.046	0.056	0.054	0.049	0.048	0.054	0.062
0	-0.5	1.75	0.050	0.062	0.057	0.056	0.049	0.077	0.078	0.058	0.054	0.056	0.055	0.054	0.073	0.074
		1.5	0.049	0.060	0.056	0.055	0.047	0.075	0.075	0.059	0.057	0.056	0.054	0.056	0.074	0.068
		1.25	0.047	0.061	0.058	0.059	0.049	0.073	0.068	0.058	0.054	0.057	0.052	0.055	0.073	0.072
		1	0.048	0.061	0.057	0.057	0.044	0.067	0.065	0.057	0.060	0.052	0.052	0.055	0.074	0.081

TABLE S.2. Empirical size of unit root tests under OLS and local GLS de-meaning. The DGP is (25) and (26) with $T = 500$ and $\gamma = 1$.

			OLS de-meaning							Local GLS de-meaning						
φ	θ	α_2	$t_{\widehat{\rho}}$	$\mathcal{VRT}$	$\mathcal{MSB}$	$\mathcal{MZ}_{\varphi}$	$\mathcal{MZ}_t$	$t^{W}_{\widehat{\rho},1}$	$t^{W,\widehat{\alpha}_u}_{\widehat{\rho},1}$	$t_{\widehat{\rho}}$	$\mathcal{VRT}$	$\mathcal{MSB}$	$\mathcal{MZ}_{\varphi}$	$\mathcal{MZ}_t$	$t^{W}_{\widehat{\rho},1}$	$t^{W,\widehat{\alpha}_u}_{\widehat{\rho},1}$
0	0	1.75	0.049	0.050	0.046	0.047	0.047	0.041	0.051	0.046	0.054	0.051	0.047	0.047	0.051	0.058
		1.5	0.052	0.050	0.046	0.048	0.053	0.028	0.040	0.042	0.063	0.051	0.044	0.043	0.048	0.046
		1.25	0.050	0.053	0.046	0.047	0.048	0.022	0.029	0.038	0.064	0.048	0.039	0.037	0.039	0.048
		1	0.052	0.043	0.043	0.044	0.051	0.021	0.024	0.033	0.074	0.044	0.034	0.031	0.040	0.042
0.5	0	1.75	0.049	0.047	0.048	0.047	0.049	0.041	0.054	0.046	0.052	0.052	0.049	0.047	0.048	0.062
		1.5	0.053	0.046	0.049	0.053	0.056	0.028	0.046	0.044	0.063	0.051	0.045	0.047	0.049	0.050
		1.25	0.048	0.049	0.052	0.054	0.050	0.025	0.032	0.039	0.061	0.051	0.043	0.042	0.039	0.049
		1	0.052	0.041	0.048	0.046	0.055	0.023	0.026	0.032	0.073	0.045	0.037	0.036	0.041	0.043
-0.5	0	1.75	0.048	0.053	0.045	0.045	0.047	0.045	0.049	0.046	0.056	0.047	0.044	0.045	0.053	0.059
		1.5	0.052	0.054	0.044	0.045	0.050	0.031	0.042	0.042	0.064	0.049	0.041	0.041	0.051	0.046
		1.25	0.049	0.055	0.046	0.046	0.047	0.027	0.037	0.039	0.065	0.048	0.040	0.039	0.041	0.049
		1	0.051	0.044	0.042	0.042	0.051	0.025	0.024	0.032	0.075	0.042	0.033	0.031	0.044	0.043
0	0.5	1.75	0.044	0.049	0.052	0.052	0.051	0.043	0.048	0.043	0.056	0.051	0.046	0.046	0.050	0.064
		1.5	0.049	0.049	0.051	0.054	0.055	0.030	0.044	0.041	0.062	0.052	0.044	0.044	0.049	0.048
		1.25	0.047	0.051	0.051	0.054	0.050	0.025	0.030	0.038	0.062	0.055	0.045	0.043	0.043	0.049
		1	0.049	0.042	0.047	0.049	0.055	0.024	0.025	0.032	0.072	0.045	0.036	0.035	0.042	0.046
0	-0.5	1.75	0.053	0.059	0.052	0.056	0.055	0.067	0.074	0.053	0.055	0.054	0.050	0.051	0.070	0.085
		1.5	0.058	0.060	0.055	0.060	0.059	0.050	0.064	0.054	0.065	0.056	0.051	0.050	0.072	0.073
		1.25	0.055	0.062	0.061	0.062	0.055	0.044	0.049	0.047	0.067	0.060	0.049	0.046	0.074	0.079
		1	0.058	0.050	0.055	0.054	0.057	0.042	0.042	0.041	0.081	0.050	0.042	0.039	0.080	0.082

Table S.3. Empirical size of unit root tests under OLS and local GLS de-meaning. The DGP is (25) and (26) with $T = 500$ and $\gamma = 10$.

			OLS de-meaning							Local GLS de-meaning						
φ	θ	α_2	$t_{\widehat{\rho}}$	$\mathcal{VRT}$	$\mathcal{MSB}$	$\mathcal{MZ}_{\varphi}$	$\mathcal{MZ}_t$	$t^{W}_{\widehat{\rho},1}$	$t^{W,\widehat{\alpha}_u}_{\widehat{\rho},1}$	$t_{\widehat{\rho}}$	$\mathcal{VRT}$	$\mathcal{MSB}$	$\mathcal{MZ}_{\varphi}$	$\mathcal{MZ}_t$	$t^{W}_{\widehat{\rho},1}$	$t^{W,\widehat{\alpha}_u}_{\widehat{\rho},1}$
0	0	1.75	0.045	0.048	0.041	0.044	0.044	0.029	0.056	0.044	0.050	0.048	0.043	0.043	0.044	0.048
		1.5	0.053	0.054	0.044	0.047	0.054	0.017	0.057	0.037	0.067	0.047	0.040	0.038	0.041	0.055
		1.25	0.058	0.053	0.046	0.049	0.059	0.007	0.027	0.028	0.079	0.042	0.030	0.028	0.028	0.047
		1	0.062	0.048	0.045	0.048	0.064	0.002	0.007	0.025	0.095	0.040	0.025	0.024	0.022	0.025
0.5	0	1.75	0.046	0.044	0.048	0.047	0.045	0.028	0.054	0.045	0.048	0.049	0.048	0.048	0.045	0.048
		1.5	0.055	0.050	0.047	0.049	0.058	0.017	0.058	0.038	0.066	0.051	0.043	0.041	0.042	0.057
		1.25	0.057	0.050	0.050	0.051	0.062	0.007	0.027	0.029	0.078	0.045	0.033	0.032	0.031	0.050
		1	0.063	0.044	0.048	0.050	0.066	0.003	0.007	0.026	0.092	0.045	0.032	0.029	0.022	0.023
-0.5	0	1.75	0.045	0.050	0.039	0.041	0.042	0.034	0.060	0.044	0.051	0.047	0.043	0.042	0.049	0.050
		1.5	0.054	0.057	0.043	0.044	0.052	0.019	0.062	0.037	0.068	0.044	0.038	0.037	0.043	0.059
		1.25	0.057	0.055	0.045	0.048	0.059	0.010	0.031	0.027	0.080	0.041	0.030	0.028	0.031	0.052
		1	0.062	0.052	0.043	0.049	0.063	0.005	0.010	0.025	0.096	0.041	0.025	0.025	0.026	0.027
0	0.5	1.75	0.042	0.047	0.044	0.046	0.048	0.031	0.056	0.042	0.052	0.049	0.044	0.044	0.046	0.052
		1.5	0.054	0.052	0.049	0.051	0.057	0.019	0.063	0.037	0.066	0.050	0.043	0.041	0.044	0.058
		1.25	0.057	0.053	0.052	0.055	0.063	0.007	0.032	0.029	0.078	0.047	0.035	0.034	0.034	0.050
		1	0.061	0.047	0.051	0.054	0.068	0.003	0.010	0.028	0.094	0.048	0.032	0.032	0.023	0.028
0	-0.5	1.75	0.052	0.056	0.050	0.053	0.050	0.052	0.088	0.056	0.053	0.056	0.053	0.052	0.071	0.076
		1.5	0.060	0.062	0.054	0.057	0.061	0.037	0.093	0.048	0.073	0.055	0.047	0.048	0.070	0.093
		1.25	0.064	0.060	0.055	0.061	0.066	0.022	0.058	0.039	0.083	0.052	0.040	0.040	0.067	0.097
		1	0.069	0.057	0.058	0.060	0.073	0.014	0.025	0.038	0.099	0.053	0.039	0.037	0.074	0.085

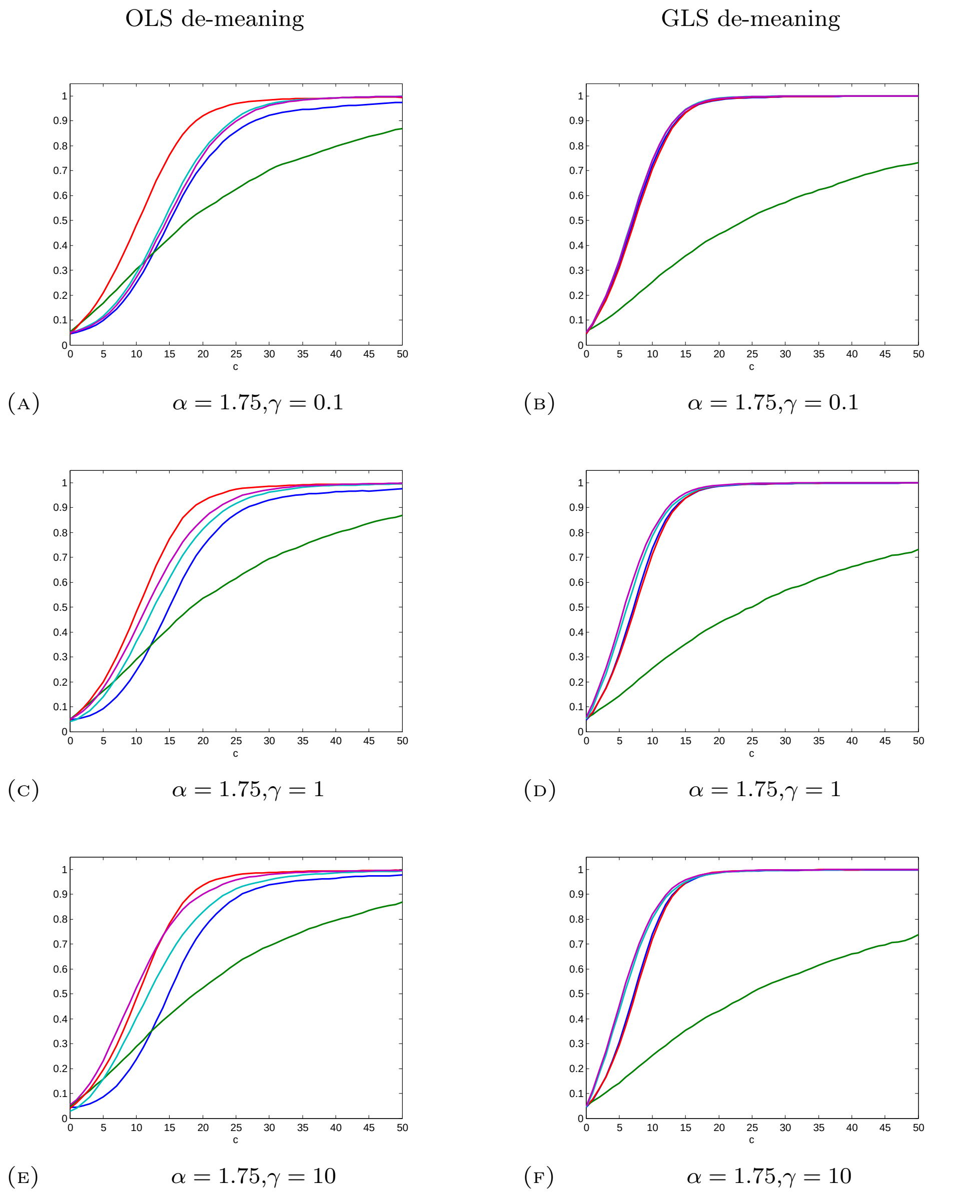
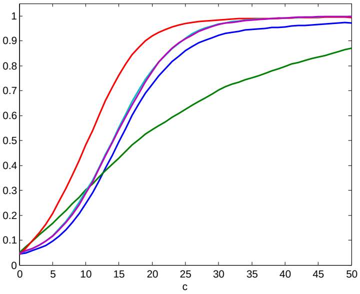
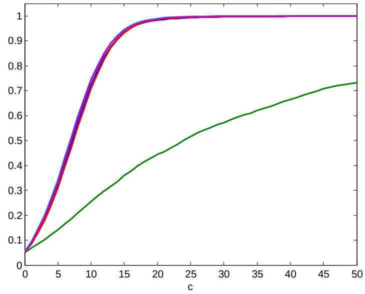


FIGURE S.1: Local power of unit root tests under OLS and local GLS de-meaning when $T = 500$. The DGP is (25) and (26) with $c \in 0, 1, 2, ..., 50$ and $\varphi = \theta = 0$. Key: $t_{\hat{\rho}}$ (blue); $\mathcal{VRT}$ (green); $\mathcal{MSB}$ (red); $t^{W}_{\hat{\rho},1}$ (cyan); $t^{W,\alpha_u}_{\hat{\rho},1}$ (magenta).

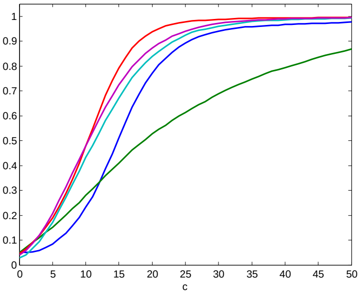
OLS de-meaning | GLS de-meaning



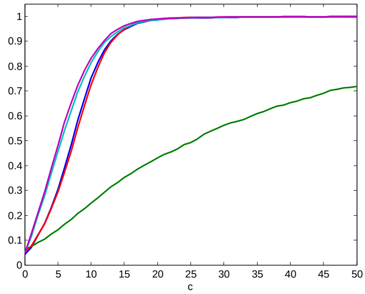
(A) $\alpha = 1.5, \gamma = 0.1$



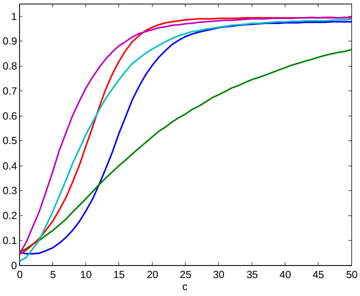
(B) $\alpha = 1.5, \gamma = 0.1$



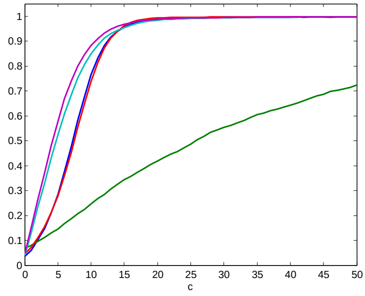
(C) $\alpha = 1.5, \gamma = 1$



(D) $\alpha = 1.5, \gamma = 1$



(E) $\alpha = 1.5, \gamma = 10$

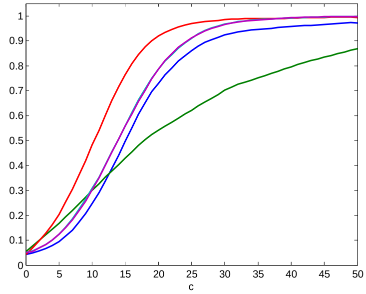


(F) $\alpha = 1.5, \gamma = 10$

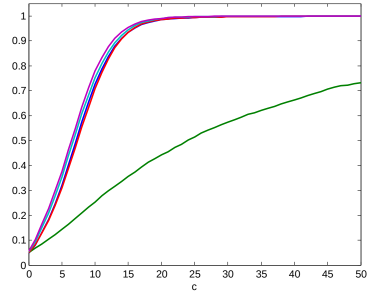
FIGURE S.2: Local power of unit root tests under OLS and local GLS de-meaning when $T = 500$. The DGP is (25) and (26) with $c \in 0, 1, 2, ..., 50$ and $\varphi = \theta = 0$. Key: $t_{\hat{\rho}}$ (blue); $\mathcal{VRT}$ (green); $\mathcal{MSB}$ (red); $t^{W}_{\hat{\rho},1}$ (cyan); $t^{W,\alpha_u}_{\hat{\rho},1}$ (magenta).

OLS de-meaning

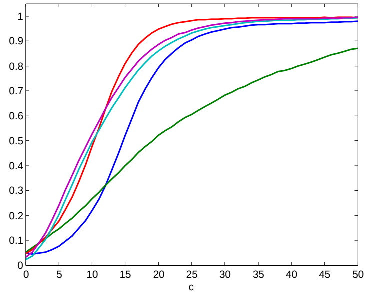
GLS de-meaning



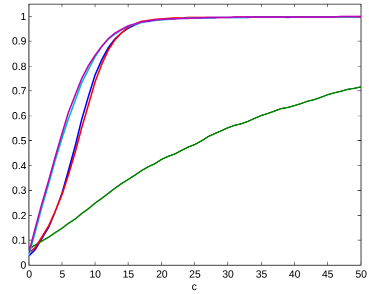
(A) $\alpha = 1.25, \gamma = 0.1$



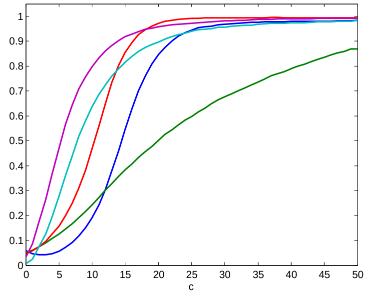
(B) $\alpha = 1.25, \gamma = 0.1$



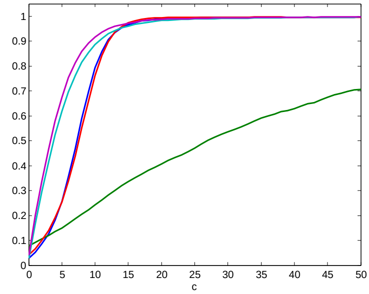
(C) $\alpha = 1.25, \gamma = 1$



(D) $\alpha = 1.25, \gamma = 1$



(E) $\alpha = 1.25, \gamma = 10$

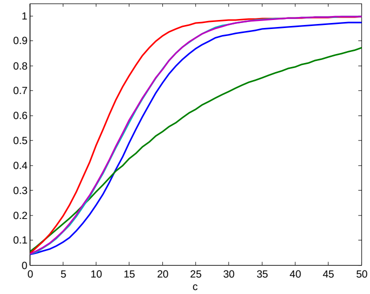


(F) $\alpha = 1.25, \gamma = 10$

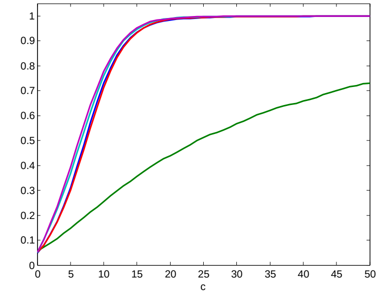
FIGURE S.3: Local power of unit root tests under OLS and local GLS de-meaning when $T = 500$. The DGP is (25) and (26) with $c \in 0, 1, 2, ..., 50$ and $\varphi = \theta = 0$. Key: $t_{\hat{\rho}}$ blue; $\mathcal{VRT}$ green; $\mathcal{MSB}$ red; $t^{W}_{\hat{\rho},1}$ cyan; $t^{W,\alpha_u}_{\hat{\rho},1}$ magenta.

OLS de-meaning

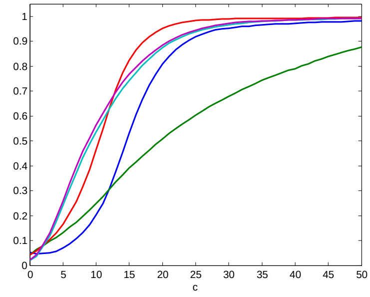
GLS de-meaning



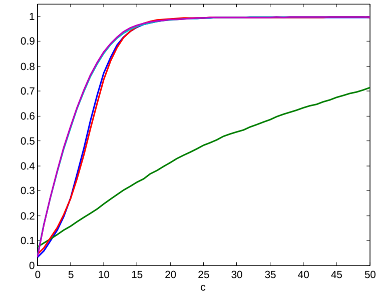
(A) $\alpha = 1.0, \gamma = 0.1$



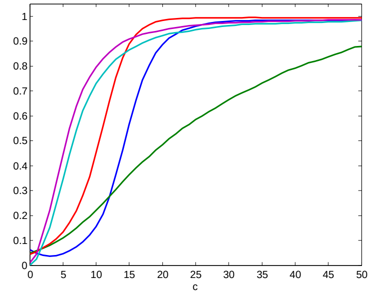
(B) $\alpha = 1.0, \gamma = 0.1$



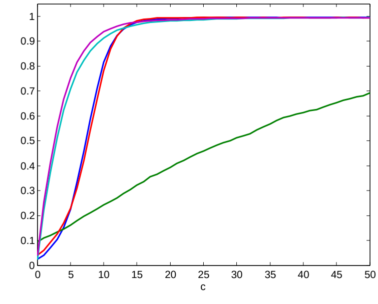
(C) $\alpha = 1.0, \gamma = 1$



(D) $\alpha = 1.0, \gamma = 1$



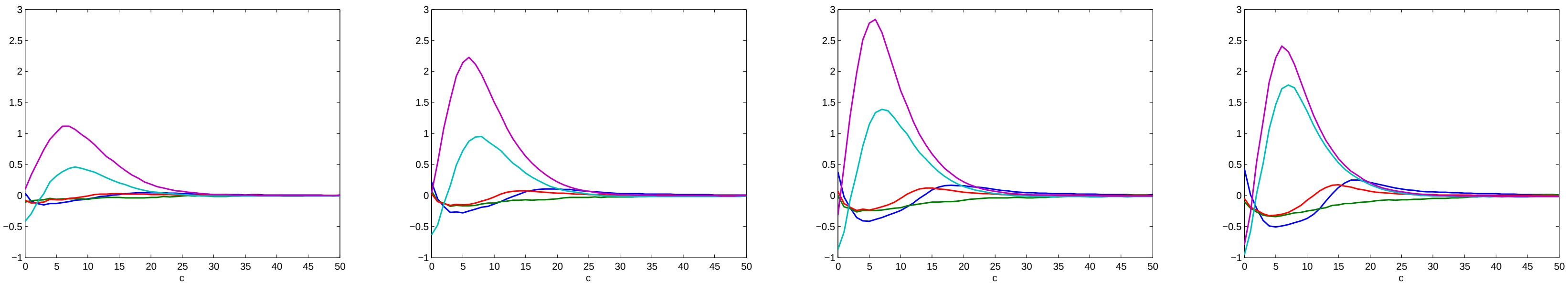
(E) $\alpha = 1.0, \gamma = 10$



(F) $\alpha = 1.0, \gamma = 10$

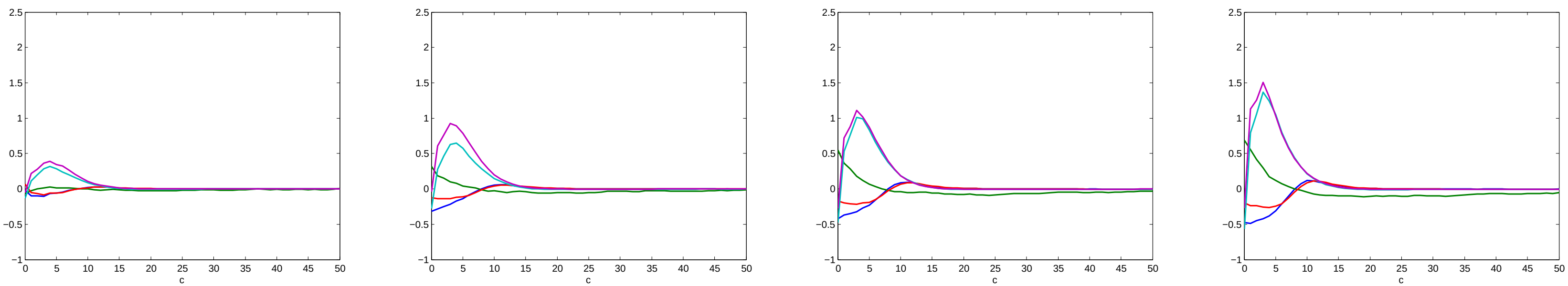
FIGURE S.4: Local power of unit root tests under OLS and local GLS de-meaning when $T = 500$. The DGP is (25) and (26) with $c \in 0, 1, 2, ..., 50$ and $\varphi = \theta = 0$. Key: $t_{\hat{\rho}}$ (blue); $\mathcal{VRT}$ (green); $\mathcal{MSB}$ (red); $t^{W}_{\hat{\rho},1}$ (cyan); $t^{W,\alpha_u}_{\hat{\rho},1}$ (magenta).

OLS de-meaning



(A) $\alpha = 1.75$ (B) $\alpha = 1.50$ (C) $\alpha = 1.25$ (D) $\alpha = 1.00$

GLS de-meaning



(E) $\alpha = 1.75$ (F) $\alpha = 1.50$ (G) $\alpha = 1.25$ (H) $\alpha = 1.00$

FIGURE S.5: Local power variation of the unit root tests. The DGP is (25) and (26) with $c \in 0, 1, 2, ..., 50$, $\varphi = \theta = 0$ and $T = 500$.

Key: $t_{\hat{\rho}}$ ——; $\mathcal{VRT}$ ——; $\mathcal{MSB}$ ——; $t^{W}_{\hat{\rho},1}$ ——; $t^{W,\alpha_u}_{\hat{\rho},1}$ ——.