Testing for common trends in non-stationary large datasets

SUPPLEMENTARY APPENDIX

Matteo Barigozzi
Department of Economics, Università di Bologna
and
Lorenzo Trapani
School of Economics, University of Nottingham

Contents

A Technical Appendix

A.1 Preliminary lemmas

Henceforth, $\nu^{(p)}(A)$ represent the eigenvalues, sorted in decreasing order, for a matrix A; we occasionally employ the notation $\nu^{(\min)}(A)$ to denote the smallest eigenvalue of A. Also, "$\overset{D}{=}$" denotes equality in distribution. We also use the following matrix notation

$$\begin{aligned} X_t &= \Lambda^{(1)} f_t^{(1)} + u_t^{(1)} \\ &= \Lambda^{(1)} f_t^{(1)} + \Lambda^{(2)} f_t^{(2)} + u_t^{(2)} \\ &= \Lambda^{(1)} f_t^{(1)} + \Lambda^{(2)} f_t^{(2)} + \Lambda^{(3)} f_t^{(3)} + u_t. \end{aligned}$$

As far as the notation is concerned, $\Lambda^{(1)}$ is $N \times r_1$; $\Lambda^{(2)}$ is $N \times r_2$; and, finally, $\Lambda^{(3)}$ is $N \times r_3$.

We begin with the following lemma, which is useful to derive almost sure rates.

Lemma A1. *Consider a multi-index random variable $U_{i_1,...,i_h}$, with $1 \leq i_1 \leq S_1$, $1 \leq i_2 \leq S_2$, etc... Assume that*

$$\sum_{S_1} \cdot \cdot \sum_{S_h} \frac{1}{S_1 \cdot ... \cdot S_h} P\left(\max_{1 \leq i_1 \leq S_1, ..., 1 \leq i_h \leq S_h} |U_{i_1,...,i_h}| > \epsilon L_{S_1,...,S_h} \right) < \infty, \tag{A1}$$

for some $\epsilon > 0$ and a sequence $L_{S_1,...,S_h}$ defined as

$$L_{S_1,...,S_h} = S_1^{d_1} \cdot ... \cdot S_h^{d_h} l_1(S_1) \cdot ... l_h(S_h),$$

where d_1, d_2, etc. are non-negative numbers and $l_1(\cdot)$, $l_2(\cdot)$, etc. are slowly varying functions in the sense of Karamata. Then it holds that

$$\lim \sup_{(S_1,...,S_h) \to \infty} \frac{|U_{S_1,...,S_h}|}{L_{S_1,...,S_h}} = 0 \ a.s. \tag{A2}$$

Proof. The proof follows similar arguments as the proof of Lemma 2 in Trapani (2018) - see also Cai (2006). We begin by noting that, for every h-tuple $(S_1, ..., S_h)$, there is a h-tuple of integers $(k_1, ..., k_h)$ such that $2^{k_1} \leq S_1 < 2^{k_1+1}$, $2^{k_2} \leq S_2 < 2^{k_2+1}$, etc. Similarly, there is a h-tuple of real numbers defined over $[0, 1)$, say $(\rho_1, ..., \rho_h)$, such that $2^{k_1+\rho_1} = S_1$, $2^{k_2+\rho_2} = S_2$, etc. Consider now the short-hand notation

$$L_{k_1,...,k_h} = \left(2^{k_1+1}\right)^{d_1} \cdot ... \cdot \left(2^{k_h+1}\right)^{d_h} l_1(S_1) \cdot ... l_h(S_h),$$

$$P\left(\max_{1 \leq i_1 \leq S_1, ..., 1 \leq i_h \leq S_h} |U_{i_1,...,i_h}| > \epsilon L_{k_1,...,k_h} \right) = P_{k_1,...,k_h};$$

by (A1), we have

$$\sum_{k_1=0}^{\infty} \cdot ... \cdot \sum_{k_h=0}^{\infty} \frac{2^{k_1+1} \cdot ... \cdot 2^{k_h+1}}{(2^{k_1+1}-1) \cdot ... \cdot (2^{k_h+1}-1)} P_{k_1,...,k_h} < \infty.$$

This, in turn, entails that

$$\sum_{k_1=0}^{\infty} \cdot ... \cdot \sum_{k_h=0}^{\infty} P_{k_1,...,k_h} \leq \sum_{k_1=0}^{\infty} \cdot ... \cdot \sum_{k_h=0}^{\infty} \frac{2^{k_1+1} \cdot ... \cdot 2^{k_h+1}}{(2^{k_1+1}-1) \cdot ... \cdot (2^{k_h+1}-1)} P_{k_1,...,k_h} < \infty;$$

thus, by the Borel-Cantelli Lemma

$$\frac{\max_{1 \leq i_1 \leq S_1,...,1 \leq i_h \leq S_h} |U_{i_1,...,i_h}|}{L_{S_1,...,S_h}} \to 0 \text{ a.s.}$$

Therefore we have

$$\begin{aligned}
\frac{|U_{S_1,...,S_h}|}{L_{S_1,...,S_h}} &\leq \frac{\max_{1 \leq i_1 \leq S_1,...,1 \leq i_h \leq S_h} |U_{i_1,...,i_h}|}{L_{k_1,...,k_h}} \frac{L_{k_1,...,k_h}}{L_{S_1,...,S_h}} \\
&\leq C \frac{\max_{1 \leq i_1 \leq S_1,...,1 \leq i_h \leq S_h} |U_{i_1,...,i_h}|}{L_{k_1,...,k_h}} \to 0 \text{ a.s.,}
\end{aligned}$$

which, finally, implies (A2). □

Let now $\gamma^{(p)}$ and $\omega^{(p)}$ denote the p-th largest eigenvalues of $\Lambda T^{-1} \sum_{t=1}^{T} E\left(\Delta f_t \Delta f_t'\right) \Lambda'$ and $T^{-1} \sum_{t=1}^{T} E\left(\Delta u_t \Delta u_t'\right)$ respectively. By Assumption 6, it can be easily verified using the arguments in the proof of Lemma 1 in Trapani (2018) that $\gamma^{(p)} = C_p N$ for $1 \leq p \leq r$; $\omega^{(1)} \leq C_1$; and $\liminf_{N \to \infty} \omega^{(N)} > 0$.

We will often need the following lemma, shown in Trapani (2018) (see Lemma A1), which we report here for convenience.

Lemma A2. *Under Assumption 6, it holds that, as* $\min(N, T) \to \infty$

$$\begin{aligned}
\limsup_{N \to \infty} \overline{\nu}_{3,p}(k) &= \overline{\nu}_{3,p}^{U}(k) < \infty, \\
\liminf_{N \to \infty} \overline{\nu}_{3,p}(k) &= \overline{\nu}_{3,p}^{L}(k) > 0,
\end{aligned}$$

for every p *and* k*, where* $\overline{\nu}_{3,p}(k)$ *is defined in equation (18).*

Proof. We begin by showing that

$$\limsup_{N \to \infty} \frac{1}{N-k+1} \sum_{h=k}^{N} \nu_3^{(h)} = \overline{\nu}_{3,p}^{U}(k) < \infty, \tag{A3}$$

$$\liminf_{N \to \infty} \frac{1}{N-k+1} \sum_{h=k}^{N} \nu_3^{(h)} = \overline{\nu}_{3,p}^{L}(k) > 0. \tag{A4}$$

Letting

$$\overline{\nu}_{3,p}(k) = \frac{1}{N-k+1} \sum_{h=k}^{N} \nu_3^{(h)},$$

note that, by Weyl's inequalities, we have $\gamma^{(h)} + \omega^{(N)} \leq \nu_3^{(h)} \leq \gamma^{(h)} + \omega^{(1)}$. Thus

$$\omega^{(N)} + \frac{1}{N-k+1} \sum_{h=k}^{N} \gamma^{(h)} \leq \overline{\nu}_{3,p}(k) \leq \omega^{(1)} + \frac{1}{N-k+1} \sum_{h=k}^{N} \gamma^{(h)}. \tag{A5}$$

Assumption 6 implies that

$$0 \leq \frac{1}{N-k+1} \sum_{h=k}^{N} \gamma^{(h)} \leq C_{k+1} < \infty,$$

so that (A5) becomes

$$\omega^{(N)} \leq \overline{\nu}_{3,p}(k) \leq C_0 + C_{k+1},$$

whence (A3) and (A4) follow for each k. Hereafter, the proof is exactly the same as that of Lemma A1 in Trapani (2018) and thus omitted. □

Lemma A3. *Under Assumption 2, it holds that*

$$\lim\inf_{T\to\infty} \frac{\ln \ln T}{T^2} \sum_{t=1}^{T} f_t^* f_t^{*\prime} = D \quad a.s.,$$

where D is a positive definite matrix of dimension $[r_2 + r_1(1-d_1)d_2] \times [r_2 + r_1(1-d_1)d_2]$.

Proof. We have

$$\begin{aligned} f_t^* f_t^{*\prime} &= \left(f_t^* \pm \Sigma_{\Delta f^*}^{1/2} W(t)\right)\left(f_t^* \pm \Sigma_{\Delta f^*}^{1/2} W(t)\right)' \\ &= \Sigma_{\Delta f^*}^{1/2} W(t) W(t)' \Sigma_{\Delta f^*}^{1/2} + \Sigma_{\Delta f^*}^{1/2} W(t) \left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right)' + \left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right) W(t)' \Sigma_{\Delta f^*}^{1/2} \\ &\quad + \left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right)\left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right)'. \\ &= I + II + III + IV. \end{aligned}$$

Let b be a nonzero vector of dimension $r_1 + r_2$, such that $\|b\| < \infty$. We will prove that

$$\lim\inf_{T\to\infty} \frac{\ln \ln T}{T^2} \sum_{t=1}^{T} b' f_t^* f_t^{*\prime} b > 0 \quad a.s.,$$

for every b, thus proving the lemma. Clearly

$$\frac{\ln \ln T}{T^2} \sum_{t=1}^{T} b' \left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right)\left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right)' b \leq C_0 T^{-2\epsilon} \ln \ln T = o_{a.s.}(1),$$

by Assumption 2*(iv)*. This entails that IV is dominated. Consider now II and III. By the Law of the Iterated Logarithm (henceforth, LIL), we have that there exists a random t_0 such that, for all $t \geq t_0$, there exists a positive finite constant C_0 such that $\|W(t)\|^2 \leq C_0 t^{1/2}(\ln\ln t)^{1/2}$. Thus, using Assumption 2*(iv)*

$$\begin{aligned}
& \frac{\ln\ln T}{T^2}\sum_{t=1}^{T} b'\Sigma_{\Delta f^*}^{1/2} W(t)\left(f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right)' b \\
\leq\ & C_0 \frac{\ln\ln T}{T^2}\sum_{t=1}^{T}\|W(t)\|\left\|f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right\| \\
\leq\ & C_0 \frac{\ln\ln T}{T^2}\left(\sup_{1\leq t\leq T}\left\|f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right\|\right)\sum_{t=1}^{T}\|W(t)\| \\
\leq\ & C_0 \frac{\ln\ln T}{T^2} T^{1/2-\epsilon}\sum_{t=1}^{T} t^{1/2}(\ln\ln t)^{1/2} = o_{a.s.}(1).
\end{aligned}$$

Finally it holds that

$$\lim\inf_{T\to\infty}\frac{\ln\ln T}{T^2}\sum_{t=1}^{T} b'\Sigma_{\Delta f^*}^{1/2} W(t) W(t)' \Sigma_{\Delta f^*}^{1/2} b = \frac{1}{4}\left(b'\Sigma_{\Delta f^*} b\right) > 0,$$

by noting that $b'\Sigma_{\Delta f^*}^{1/2} W(t) \overset{D}{=} (b'\Sigma_{\Delta f^*} b)^{1/2} B(t)$ with $B(t)$ a scalar, standard Wiener process and by applying equation (4.6) in Donsker and Varadhan (1977), and by the positive definitness of $\Sigma_{\Delta f^*}$. Since this holds for all b, the Lemma follows. □

We will now make extensive use of the notation $\widetilde{f}_t^{(1)} = d_1 t + d_2 f_t^{(1)\dagger}$.

Lemma A4. *Let* $f_t^{(1,2)} = \left[\widetilde{f}_t^{(1)}, f_t^{(2)\prime}\right]'$. *Under Assumptions 2 and 4-6, it holds that*

$$\nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) \geq C_0 T \text{ if } d_1 = 1, \tag{A6}$$

$$\nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) \geq \frac{C_0}{\ln\ln T}, \text{ for } d_1 + 1 \leq p \leq r_2 + \max\{d_1, d_2\}, \tag{A7}$$

$$\nu^{(r_2+1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) \leq \frac{C_0}{T}(\ln T)^{3/2+\epsilon} \text{ if } d_1 = d_2 = 0, \tag{A8}$$

for N, T *large enough.*

Proof. Let $\widetilde{d}_1 = [d_1, 0, ..., 0]'$ be an (r_2+1)-dimensional vector. We have

$$\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime} = \frac{1}{T^2}\sum_{t=1}^{T} t^2 \widetilde{d}_1 \widetilde{d}_1' + \frac{1}{T^2}\widetilde{d}_1 \sum_{t=1}^{T} t f_t^{*\prime} + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* t \widetilde{d}_1' + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}.$$

In the proof, we make repeated use of the lower bound entailed by Weyl's inequality (see Horn and Johnson, 2012, p.181)

$$\nu^{(p)}(A+B) \geq \nu^{(p)}(A) + \nu^{(\min)}(B),$$

for two symmetric matrices A and B. Clearly

$$\nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) \geq \nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} t^2 \widetilde{d}_1 \widetilde{d}_1{}'\right) + \nu^{(\min)}(B), \tag{A9}$$

with

$$B = \frac{1}{T^2}\widetilde{d}_1 \sum_{t=1}^{T} t f_t^{*\prime} + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* t \widetilde{d}_1{}' + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}.$$

Simple algebra yields

$$\nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} t^2 \widetilde{d}_1 \widetilde{d}_1{}'\right) = \frac{d_1^2}{3} T.$$

Also, we have that $\left|\nu^{(\min)}(B)\right| = O_{a.s.}(\ln\ln T)$; indeed

$$\nu^{(\min)}(B) \leq \nu^{(\min)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right) + \nu^{(1)}\left(\frac{1}{T^2}\widetilde{d}_1 \sum_{t=1}^{T} t f_t^{*\prime} + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* t \widetilde{d}_1{}'\right).$$

By Donsker and Varadhan (1977, Example 2), it holds that

$$\nu^{(\min)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right) \leq C_0 \ln\ln T.$$

Also, the matrix

$$B_2 = \frac{1}{T^2}\widetilde{d}_1 \sum_{t=1}^{T} t f_t^{*\prime} + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* t \widetilde{d}_1{}',$$

is symmetric, and after some algebra it can be shown that

$$\nu^{(1)}(B_2) \leq C_0 \left(\sum_{i=1}^{r_2+1} \left|\frac{1}{T^2}\sum_{t=1}^{T} t f_{i,t}^*\right|^2\right)^{1/2}.$$

Given that

$$E\left|\frac{1}{T^2}\sum_{t=1}^{T} t f_{i,t}^*\right|^2 \leq C_0 \frac{1}{T^3}\sum_{t=1}^{T} t^2 E\left(f_{i,t}^*\right)^2 \leq C_1 T,$$

by Lemma A1 it holds that $\nu^{(1)}(B_2) = O_{a.s.}\left(T^{1/2}(\ln T)^{1+\epsilon}\right)$ for every $\epsilon > 0$. Thus, finally, by (A9) it holds that there exists a random T_0 such that for $T \geq T_0$

$$\nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) \geq C_0 T,$$

which proves (A6). Turning to (A7), for each $p > 1$

$$\begin{aligned}
\nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) &\geq \nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T} t^2 \widetilde{d}_1 \widetilde{d}_1{}' + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right) \\
&\quad + \nu^{(\min)}\left(\frac{1}{T^2}\widetilde{d}_1 \sum_{t=1}^{T} t f_t^{*\prime} + \frac{1}{T^2}\sum_{t=1}^{T} f_t^* t \widetilde{d}_1{}'\right) \\
&\geq \nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right) + \nu^{(\min)}\left(\frac{1}{T^2}\sum_{t=1}^{T} t^2 \widetilde{d}_1 \widetilde{d}_1{}'\right) = \nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right),
\end{aligned}$$

so that the desired result follows immediately from Lemma A3. Finally, consider (A8). Let $\widetilde{g}_t = [g_t, 0, ..., 0]'$ and $\widetilde{f}_t = \left[0, f_t^{(2)\prime}\right]'$ be two (r_2+1)-dimensional vectors; in this case we have

$$\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime} = \frac{1}{T^2}\sum_{t=1}^{T} \widetilde{g}_t \widetilde{g}_t' + \frac{1}{T^2}\sum_{t=1}^{T} \widetilde{f}_t \widetilde{f}_t';$$

thus

$$\nu^{(\min)}\left(\frac{1}{T^2}\sum_{t=1}^{T} f_t^{(1,2)} f_t^{(1,2)\prime}\right) \leq \nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T} \widetilde{g}_t \widetilde{g}_t'\right) + \nu^{(\min)}\left(\frac{1}{T^2}\sum_{t=1}^{T} \widetilde{f}_t \widetilde{f}_t'\right) \leq \frac{1}{T^2}\sum_{t=1}^{T} g_t^2.$$

Assumption 3*(i)* and equation (2.3) in Serfling (1970) imply that

$$E \max_{1 \leq \widetilde{t} \leq T} \left\| \sum_{t=1}^{\widetilde{t}} g_t^2 \right\|^2 \leq C_0 (\ln T)^2 T,$$

which, through Lemma A1, yields the desired result. □

Lemma A5. *Under Assumptions 2-4*

$$\max_{1 \leq p \leq N} \left| \nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T} u_t^{(1)} u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T} \Lambda^{(1)} \widetilde{f}_t^{(1)} u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T} u_t^{(1)} \widetilde{f}_t^{(1)} \Lambda^{(1)\prime}\right) \right| = O_{a.s.}\left(\frac{N}{\sqrt{T}} l_{N,T}\right).$$

Proof. We show the lemma for the case $d_1 = d_2 = 1$; when either dummy is zero, calculations become easier and the result can be readily shown. Let

$$\max_{1 \leq p \leq N} \left| \nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T} u_t^{(1)} u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T} \Lambda^{(1)} \widetilde{f}_t^{(1)} u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T} u_t^{(1)} \widetilde{f}_t^{(1)} \Lambda^{(1)\prime}\right) \right| = \nu^{(\max)},$$

for short. It holds that

$$\frac{1}{3}\nu^{(\max)} \leq \frac{1}{3}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}u_{i,t}^{(1)}u_{j,t}^{(1)}+\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_i^{(1)}\widetilde{f}_t^{(1)}u_{j,t}^{(1)}+\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_j^{(1)}\widetilde{f}_t^{(1)}u_{i,t}^{(1)}\right|^2\right)^{1/2}. \quad \text{(A10)}$$

$$\leq \left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}u_{i,t}^{(1)}u_{j,t}^{(1)}\right|^2\right)^{1/2}+\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r}\Lambda_i^{(1)}\widetilde{f}_t^{(1)}u_{j,t}^{(1)}\right|^2\right)^{1/2}$$

$$+\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_j^{(1)}\widetilde{f}_t^{(1)}u_{i,t}^{(1)}\right|^2\right)^{1/2},$$

where the first passage is the usual spectral norm inequality, and the last passage follows from applying (twice) the C_r-inequality (Davidson, 1994, p. 140).
Let now

$$u_{i,t}^{(2)}=\lambda^{(1)}g_t+\lambda^{(3)\prime}f_t^{(3)}+u_{i,t} \quad \text{(A11)}$$

and note that

$$\frac{1}{3}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}u_{i,t}^{(1)}u_{j,t}^{(1)}\right|^2$$

$$\leq \sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}u_{i,t}^{(2)}u_{j,t}^{(2)}\right|^2+\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r_2}\Lambda_{i,k}^{(2)}f_{k,t}^{(2)}f_{k,t}^{(2)\prime}\Lambda_{i,k}^{(2)\prime}\right|^2$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r_2}\Lambda_{i,k}^{(2)}f_{k,t}^{(2)}u_{j,t}^{(2)}\right|^2+\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r_2}\Lambda_{j,k}^{(2)}f_{k,t}^{(2)}u_{i,t}^{(2)}\right|^2.$$

We have

$$E\max_{h_1,h_2,\widetilde{t}}\sum_{i=1}^{h_1}\sum_{j=1}^{h_2}\left|\frac{1}{T^3}\sum_{t=1}^{\widetilde{t}}u_{i,t}^{(2)}u_{j,t}^{(2)}\right|^2 \leq \sum_{i=1}^{N}\sum_{j=1}^{N}E\max_{\widetilde{t}}\left|\frac{1}{T^3}\sum_{t=1}^{\widetilde{t}}u_{i,t}^{(2)}u_{j,t}^{(2)}\right|^2$$

$$\leq C_0T\frac{1}{T^6}\sum_{i=1}^{N}\sum_{j=1}^{N}E\sum_{t=1}^{T}\left|u_{i,t}^{(2)}\right|^2\left|u_{j,t}^{(2)}\right|^2 \leq C_0N^2T^{-4}\max_{1\leq i\leq N}E\left|u_{i,t}^{(2)}\right|^4$$

$$\leq C_0N^2T^{-4}\left(\max_{1\leq i\leq N}E\left|u_{i,t}\right|^4+\max_{1\leq i\leq N}\left\|\lambda_i^{(3)}\right\|^4E\left\|f_t^{(3)}\right\|^4++\max_{1\leq i\leq N}\left\|\lambda_i^{(1)}\right\|^4E\left\|g_t\right\|^4\right)$$

$$\leq C_0N^2T^{-4},$$

so that, by Lemma A1

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}u_{i,t}^{(2)}u_{j,t}^{(2)}\right|^2=O_{a.s.}\left(\frac{N^2}{T^4}\ln^{2+\epsilon}N\ln^{1+\epsilon}T\right). \quad \text{(A12)}$$

Also

$$\begin{aligned}&\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r_2}\Lambda_{i,k}^{(2)}f_{k,t}^{(2)}f_{k,t}^{(2)\prime}\Lambda_{i,k}^{(2)}\right|^2\\ \leq\ & T^{-6}N^2\left(\max_i\left\|\Lambda_{i,k}^{(2)}\right\|\right)^2\left\|\sum_{t=1}^{T}f_t^{(2)}f_t^{(2)\prime}\right\|^2;\end{aligned}$$

on account of Assumption 2*(iv)*, it holds that

$$\left\|\frac{\sum_{t=1}^{T}f_t^{(2)}f_t^{(2)\prime}}{T^2\ln\ln T}\right\|^2=\left\|\Sigma_{\Delta f^*}^{1/2}\frac{\sum_{t=1}^{T}W(t)\,W(t)'}{T^2\ln\ln T}\Sigma_{\Delta f^*}^{1/2}\right\|^2+o_{a.s.}(1)=O_{a.s.}(1);$$

the final result follows from Donsker and Varadhan (1977, Example 2)). Thus

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r_2}\Lambda_{i,k}^{(2)}f_{k,t}^{(2)}f_{k,t}^{(2)\prime}\Lambda_{i,k}^{(2)}\right|^2=O_{a.s.}\left(\frac{N^2}{T^2}(\ln\ln T)^2\right). \quad \text{(A13)}$$

Finally, consider

$$\begin{aligned}&E\max_{h_1,h_2,\widetilde{t}}\sum_{i=1}^{h_1}\sum_{j=1}^{h_2}\left|\sum_{t=1}^{\widetilde{t}}\Lambda_i^{(2)\prime}f_t^{(2)}u_{j,t}^{(2)}\right|^2\leq\sum_{i=1}^{N}\sum_{j=1}^{N}E\max_{\widetilde{t}}\left|\sum_{t=1}^{\widetilde{t}}\Lambda_i^{(2)\prime}f_t^{(2)}u_{j,t}^{(2)}\right|^2\\ \leq\ & C_0(\ln T)^2\sum_{i=1}^{N}\sum_{j=1}^{N}E\left|\sum_{t=1}^{T}\Lambda_i^{(2)\prime}f_t^{(2)}u_{j,t}^{(2)}\right|^2\leq C_0(\ln T)^2\sum_{i=1}^{N}\sum_{j=1}^{N}\Lambda_i^{(2)\prime}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(f_t^{(2)}u_{j,t}^{(2)}f_s^{(2)\prime}u_{j,s}^{(2)}\right)\Lambda_i^{(2)}\\ \leq\ & C_1\left(\max_i\left\|\Lambda_i^{(2)\prime}\right\|\right)^2(\ln T)^2\sum_{i=1}^{N}\sum_{j=1}^{N}E\left\|\sum_{t=1}^{T}f_t^{(2)}u_{j,t}^{(2)}\right\|^2\leq C_2N^2T^2(\ln T)^2,\end{aligned}$$

having used Assumption 3*(ii)*, so that

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\sum_{k=1}^{r_2}\Lambda_{i,k}^{(2)}f_{k,t}^{(2)}u_{j,t}^{(2)}\right|^2=O_{a.s.}\left(\frac{N^2}{T^4}\ln^{2+\epsilon}N\ln^{3+\epsilon}T\right).$$

Putting all together, we have

$$\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}u_{i,t}^{(1)}u_{j,t}^{(1)}\right|^{2}\right)^{1/2}=O_{a.s.}\left(\frac{N}{T}\ln\ln T\right).$$

Consider now

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}\widetilde{f}_{t}^{(1)}u_{j,t}^{(1)}\right|^{2}\\
\leq\ &\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}tu_{j,t}^{(1)}\right|^{2}+\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}f_{t}^{(1)\dagger}u_{j,t}^{(1)}\right|^{2}.
\end{aligned}\tag{A14}$$

We have

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}tu_{j,t}^{(1)}\right|^{2}\\
\leq\ &\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}tf_{t}^{(2)\prime}\Lambda_{j}^{(2)\prime}\right|^{2}+\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}tu_{j,t}^{(2)}\right|^{2}.
\end{aligned}$$

Note that

$$\begin{aligned}
&E\max_{h_1,h_2,\widetilde{t}}\sum_{i=1}^{h_1}\sum_{j=1}^{h_2}\left|\frac{1}{T^{3}}\sum_{t=1}^{\widetilde{t}}\Lambda_{i}^{(1)}tf_{t}^{(2)\prime}\Lambda_{j}^{(2)\prime}\right|^{2}\leq T^{-6}\sum_{i=1}^{N}\sum_{j=1}^{N}E\max_{\widetilde{t}}\left|\sum_{t=1}^{\widetilde{t}}\Lambda_{i}^{(1)}tf_{t}^{(2)\prime}\Lambda_{j}^{(2)\prime}\right|^{2}\\
\leq\ &C_{0}\left(\ln T\right)^{2}T^{-6}\sum_{i=1}^{N}\sum_{j=1}^{N}E\left|\sum_{t=1}^{T}\Lambda_{i}^{(1)}tf_{t}^{(2)\prime}\Lambda_{j}^{(2)\prime}\right|^{2}\leq C_{0}\left(\ln T\right)^{2}T^{-6}\sum_{i=1}^{N}\sum_{j=1}^{N}\Lambda_{i}^{(1)}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(tf_{s}^{(2)\prime}\right)\Lambda_{j}^{(2)}\\
\leq\ &C_{1}N^{2}T^{-6}\left(\max_{i}\left\|\Lambda_{i}^{(1)}\right\|\right)^{2}\left(\max_{i}\left\|\Lambda_{i}^{(2)}\right\|\right)^{2}\left(\ln T\right)^{2}E\left\|\sum_{t=1}^{T}tf_{t}^{(2)}\right\|^{2}\leq C_{2}N^{2}T^{-1}\left(\ln T\right)^{2},
\end{aligned}$$

having used Assumption 3*(iii)*; Lemma A1 entails that

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}tf_{t}^{(2)\prime}\Lambda_{j}^{(2)\prime}\right|^{2}=O_{a.s.}\left(\frac{N^{2}}{T}\ln^{2+\epsilon}N\ln^{3+\epsilon}T\right).$$

Similar passages yield

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^{3}}\sum_{t=1}^{T}\Lambda_{i}^{(1)}tu_{j,t}^{(2)}\right|^{2}=O_{a.s.}\left(\frac{N^{2}}{T^{3}}\ln^{2+\epsilon}N\ln^{3+\epsilon}T\right).$$

Thus, finally

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_i^{(1)}tu_{j,t}^{(1)}\right|^2 = O_{a.s.}\left(\frac{N^2}{T}\ln^{2+\epsilon}N\ln^{3+\epsilon}T\right).$$

We now consider the next term in equation (A14). We have

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_i^{(1)}f_t^{(1)\dagger}u_{j,t}^{(1)}\right|^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_i^{(1)}f_t^{(1)\dagger}f_t^{(2)'}\Lambda_j^{(2)\prime}\right|^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_i^{(1)}f_t^{(1)\dagger}u_{j,t}^{(2)}\right|^2.$$

Similar passages as above yield

$$\begin{aligned}
& E\max_{h_1,h_2,\widetilde{t}}\sum_{i=1}^{h_1}\sum_{j=1}^{h_2}\left|\frac{1}{T^3}\sum_{t=1}^{\widetilde{t}}\Lambda_i^{(1)}f_t^{(1)\dagger}f_t^{(2)'}\Lambda_j^{(2)\prime}\right|^2 \\
\leq\ & C_0N^2T^{-6}\left(\max_i\left\|\Lambda_i^{(1)}\right\|\right)^2\left(\max_i\left\|\Lambda_i^{(2)}\right\|\right)^2(\ln T)^2\,E\left\|\sum_{t=1}^{T}f_t^{(1)\dagger}f_t^{(2)'}\right\|^2 \leq C_1N^2T^{-6}(\ln T)^2T^4,
\end{aligned}$$

having used Assumption 2*(vi)*. Similarly

$$\begin{aligned}
& E\max_{h_1,h_2,\widetilde{t}}\sum_{i=1}^{h_1}\sum_{j=1}^{h_2}\left|\frac{1}{T^3}\sum_{t=1}^{\widetilde{t}}\Lambda_i^{(1)}f_t^{(1)\dagger}u_{j,t}^{(2)}\right|^2 \\
\leq\ & C_0N^2T^{-6}\left(\max_i\left\|\Lambda_i^{(1)}\right\|\right)^2(\ln T)^2\,E\left\|\sum_{t=1}^{T}f_t^{(1)\dagger}u_{j,t}^{(2)}\right\|^2 \leq C_1N^2T^{-6}(\ln T)^2T^2,
\end{aligned}$$

having used Assumption 3*(ii)*. Thus, using Lemma A1

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^3}\sum_{t=1}^{T}\Lambda_i^{(1)}f_t^{(1)\dagger}u_{j,t}^{(1)}\right|^2 = O_{a.s.}\left(\frac{N}{T}\ln^{1+\epsilon}N\ln^{\frac{3}{2}+\epsilon}T\right).$$

Using (A10) and putting all together, the desired result obtains. □

Lemma A6. *Under Assumptions 2-4*

$$\max_{1\leq p\leq N}\left|\nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right)\right| = O_{a.s.}\left(\frac{N}{\sqrt{T}}l_{N,T}\right), \tag{A15}$$

where $u_t^{(2)}$ is defined in (A11).

Proof. Let

$$\max_{1\leq p\leq N}\left|\nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right)\right|=\nu^{(\max)},$$

for short. As before

$$\begin{aligned}
& \frac{1}{3}\nu^{(\max)} \qquad\qquad \text{(A16)}\\
\leq\; & \frac{1}{3}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^2}\sum_{t=1}^{T}u_{i,t}^{(2)}u_{j,t}^{(2)}+\frac{1}{T^2}\sum_{t=1}^{T}\sum_{k=1}^{r_2+r_1}\Lambda_{i,k}^{(1,2)}f_{k,t}^{(1,2)}u_{j,t}^{(2)}+\frac{1}{T^2}\sum_{t=1}^{T}\sum_{k=1}^{r_2+r_1}\Lambda_{j,k}^{(1,2)}f_{k,t}^{(1,2)}u_{i,t}^{(2)}\right|^2\right)^{1/2}.\\
\leq\; & \left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^2}\sum_{t=1}^{T}u_{i,t}^{(2)}u_{j,t}^{(2)}\right|^2\right)^{1/2}+\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{k=1}^{r_2+r_1}\Lambda_{i,k}^{(1,2)}f_{k,t}^{(1,2)}u_{j,t}^{(2)}\right|^2\right)^{1/2}\\
& +\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^2}\sum_{t=1}^{T}\sum_{k=1}^{r_2+r_1}\Lambda_{j,k}^{(1,2)}f_{k,t}^{(1,2)}u_{i,t}^{(2)}\right|^2\right)^{1/2}.
\end{aligned}$$

Consider the first term; by (A12),

$$\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{1}{T^2}\sum_{t=1}^{T}u_{i,t}^{(2)}u_{j,t}^{(2)}\right|^2\right)^{1/2}=O_{a.s.}\left(\frac{N}{T}\left(\ln N\right)^{1+\epsilon}\left(\ln T\right)^{(1+\epsilon)/2}\right).$$

Similarly, considering the second term in (A16) we have

$$\begin{aligned}
& E\max_{h_1,h_2,\widetilde{t}}\sum_{i=1}^{h_1}\sum_{j=1}^{h_2}\left|\frac{1}{T^2}\sum_{t=1}^{\widetilde{t}}\sum_{k=1}^{r_2+r_1}\Lambda_{i,k}^{(1,2)}f_{k,t}^{(1,2)}u_{j,t}^{(2)}\right|^2\leq T^{-4}\sum_{i=1}^{N}\sum_{j=1}^{N}E\max_{\widetilde{t}}\left|\sum_{t=1}^{\widetilde{t}}\sum_{k=1}^{r_2+r_1}\Lambda_{i,k}^{(1,2)}f_{k,t}^{(1,2)}u_{j,t}^{(2)}\right|^2\\
\leq\; & C_0T^{-4}\left(\ln T\right)^2\sum_{i=1}^{N}\sum_{j=1}^{N}E\left|\sum_{t=1}^{T}\sum_{k=1}^{r_2+r_1}\Lambda_{i,k}^{(1,2)}f_{k,t}^{(1,2)}u_{j,t}^{(2)}\right|^2\\
\leq\; & C_0T^{-4}\left(\ln T\right)^2\left(\max_{1\leq i\leq N}\left\|\Lambda_i^{(1,2)}\right\|\right)^2\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(f_t^{(1,2)}f_s^{(1,2)\prime}u_{j,t}^{(2)}u_{j,s}^{(2)}\right)\\
\leq\; & C_0T^{-4}\left(\ln T\right)^2\sum_{i=1}^{N}\sum_{j=1}^{N}E\left\|\sum_{t=1}^{T}f_t^{(1,2)}u_{j,t}^{(2)}\right\|^2\leq C_0N^2T^{-1}\left(\ln T\right)^2,
\end{aligned}$$

having used equation (2.3) in Serfling (1970), Assumption 4*(i)* and Assumption 3*(ii)*. From here henceforth, the proof is the same as for the first tem in (A16); also, the proof for the third term in (A16) is exactly the same, and it is therefore omitted. Putting everything together, the lemma follows. □

A.2 Proofs of main results

Proof of Lemma 1. When $d_1 = 0$, the lemma follows immediately from B having full rank. When $d_1 = 1$, the proof follows the arguments in Maciejowska (2010). Let

$$\mathcal{F}_t = (a|B)\begin{pmatrix} t \\ \psi_t \end{pmatrix} = C\begin{pmatrix} t \\ \psi_t \end{pmatrix};$$

by Assumption 1*(ii)*, C has full rank. It is therefore possible to re-write the expression above as

$$\mathcal{F}_t = P\left(D_1|D_2\right)\begin{pmatrix} t \\ \psi_t \end{pmatrix},$$

where $D_1 = [1, 0, ..., 0]'$ is $r \times 1$, and P and D_2 are $r \times r$ and have full rank. Among the possible matrices that satisfy this representation one can consider $(D_1|D_2) = (I_r|E)$, where $E = [E_1, ..., E_r]$ is a nonzero vector. The desired result follows immediately after computing

$$P^{-1}\mathcal{F}_t = \begin{pmatrix} t + E_1\psi_{r,t} \\ \psi_{1,t} + E_2\psi_{r,t} \\ \psi_{2,t} + E_3\psi_{r,t} \\ . \\ \psi_{r-1,t} + E_r\psi_{r,t} \end{pmatrix}.$$

□

Proof of Theorem 1. We start with (11)-(12). Weyl's inequality entails that, for $0 \leq p \leq r_1$

$$\begin{aligned} \nu_1^{(p)} \geq\ & \nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right) \\ & +\nu^{(N)}\left(\frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right). \end{aligned}$$

We already know that, by Lemma A5

$$\nu^{(N)}\left(\frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}u_t^{(1)\prime} + \frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right) = O_{a.s.}\left(\frac{N}{\sqrt{T}}l_{N,T}\right).$$

Consider now

$$
\begin{aligned}
& \nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right) \\
\geq\ & d_1^2\nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T}t^2\Lambda^{(1)\prime}\Lambda^{(1)}\right) \\
& +\nu^{(N)}\left(2\frac{d_1}{T^3}\sum_{t=1}^{T}tf_t^{(1)\dagger}\Lambda^{(1)}\Lambda^{(1)\prime}+\frac{1}{T^3}\Lambda^{(1)}\sum_{t=1}^{T}f_t^{(1)\dagger}f_t^{(1)\dagger\prime}\Lambda^{(1)\prime}\right)
\end{aligned}
$$

We have

$$
\nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T}t^2\Lambda^{(1)\prime}\Lambda^{(1)}\right)=\left(\frac{1}{T^3}\sum_{t=1}^{T}t^2\right)\nu^{(p)}\left(\Lambda^{(1)\prime}\Lambda^{(1)}\right)\geq C_0N,
$$

in view of Assumption 4*(ii)*. Consider now

$$
\nu^{(N)}\left(2\frac{d_1}{T^3}\sum_{t=1}^{T}tf_t^{(1)\dagger}\Lambda^{(1)}\Lambda^{(1)\prime}+\frac{1}{T^3}\sum_{t=1}^{T}\left(f_t^{(1)\dagger}\right)^2\Lambda^{(1)}\Lambda^{(1)\prime}\right);
$$

by Donsker and Varadhan (1977, Example 2) we have

$$
\frac{1}{T^3}\sum_{t=1}^{T}\left(f_t^{(1)\dagger}\right)^2=O_{a.s.}\left(\frac{\ln\ln T}{T}\right).
$$

Also, by Assumption 3*(iii)* and equation (2.3) in Serfling (1970) we have

$$
E\max_{1\leq t\leq T}\left|\sum_{j=1}^{t}jf_j^{(1)\dagger}\right|^2=C_0T^5\left(\ln T\right)^2,
$$

so that by Lemma A1 we have

$$
\frac{1}{T^3}\left|\sum_{t=1}^{T}tf_t^{(1)\dagger}\right|=O_{a.s.}\left(T^{-1/2}\left(\ln T\right)^{3/2+\epsilon}\right).
$$

The same steps as in the proofs of Lemmas A5 and A6 entail

$$
\nu^{(N)}\left(2\frac{d_1}{T^3}\sum_{t=1}^{T}tf_t^{(1)\dagger}\Lambda^{(1)}\Lambda^{(1)\prime}+\frac{1}{T^3}\sum_{t=1}^{T}\left(f_t^{(1)\dagger}\right)^2\Lambda^{(1)}\Lambda^{(1)\prime}\right)=O_{a.s.}\left(\frac{N}{\sqrt{T}}l_{N,T}\right).
$$

Putting everything together, the desired result follows. When $p > r_1$

$$\begin{aligned}
\nu_1^{(p)} &\leq \nu^{(p)}\left(\frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right) \\
&+\nu^{(1)}\left(\frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right) \\
&\leq \nu^{(1)}\left(\frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^3}\sum_{t=1}^{T}\Lambda^{(1)}\widetilde{f}_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^3}\sum_{t=1}^{T}u_t^{(1)}\widetilde{f}_t^{(1)\prime}\Lambda^{(1)\prime}\right);
\end{aligned}$$

Lemma A5 immediately yields the desired result.

The proof of (13)-(14) is very similar. Whenever $1 \leq p \leq r_1 + r_2 + (1 - r_1)\, d_2$, we have

$$\begin{aligned}
\nu_2^{(p)} &\geq \nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right) \\
&+\nu^{(N)}\left(\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right).
\end{aligned}$$

By Lemma A6 we have

$$\nu^{(N)}\left(\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}u_t^{(2)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right)=O_{a.s.}\left(\frac{N}{\sqrt{T}}l_{N,T}\right).$$

Also, using Theorem 7 in Merikoski and Kumar (2004) and by Assumption 4*(ii)*

$$\begin{aligned}
\nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right) &\geq \nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T}f_t^{(1,2)}f_t^{(1,2)\prime}\right)\nu^{(\min)}\left(\Lambda^{(1,2)\prime}\Lambda^{(1,2)}\right) \\
&\geq C_0\frac{N}{\ln\ln T},
\end{aligned}$$

where the last passage follows from equation (A7) in Lemma A4. Equation (13) now follows readily. Turning to (14), whenever $p > r_1 + r_2 + (1 - r_1)\, d_2$,

$$\begin{aligned}
\nu_2^{(p)} &\leq \nu^{(p)}\left(\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1,2)}f_t^{(1,2)}f_t^{(1,2)\prime}\Lambda^{(1,2)\prime}\right) \\
&+\nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1)}f_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(1)}f_t^{(1)\prime}\Lambda^{(1)\prime}\right) \\
&= \nu^{(1)}\left(\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}\Lambda^{(1)}f_t^{(1)}u_t^{(1)\prime}+\frac{1}{T^2}\sum_{t=1}^{T}u_t^{(1)}f_t^{(1)\prime}\Lambda^{(1)\prime}\right),
\end{aligned}$$

and Lemma A6 immediately yields the desired result. □

Proof of Theorem 2. The proof is similar to that of related results in other papers - see e.g. Trapani (2018). We begin with (22). Note that, under $H_{0,1}^{(p)}$, (11) and Lemma A2 entail that

$$P\left\{\omega : \lim_{\min(N,T)\to\infty} \phi_1^{(p)} \exp\left\{-N^{1-\delta-\varepsilon}\right\} = \infty\right\} = 1,$$

for every $\varepsilon > 0$, and therefore we can henceforth assume that $\lim_{\min(N,T)\to\infty} \phi_1^{(p)} = \infty$ and

$$\left(\phi_1^{(p)}\right)^{-1} = O\left(\exp\left\{-N^{1-\delta}\right\}\right). \tag{A17}$$

Let E^* and V^* denote, respectively, expectation and variance conditional on P^*; we have, for $1 \le j \le R_1$

$$E^*\left(\zeta_{1,j}^{(p)}(u)\right) = G_1(0) \text{ and } V^*\left(\zeta_{1,j}^{(p)}(u)\right) = G_1(0)(1 - G_1(0)).$$

Also

$$\begin{aligned}
& \frac{1}{\sqrt{R_1}} \sum_{j=1}^{R_1} \left(\zeta_{1,j}^{(p)}(u) - G_1(0)\right) \\
= & \frac{1}{\sqrt{R_1}} \sum_{j=1}^{R_1} \left(I\left(\xi_{1,j}^{(p)} \le 0\right) - G_1(0)\right) + \frac{1}{\sqrt{R_1}} d_u \sum_{j=1}^{R_1} \left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right) \\
& + \frac{1}{\sqrt{R_1}} \sum_{j=1}^{R_1} \left[I\left(0 \le \left|\xi_{1,j}^{(p)}\right| \le u/\phi_1^{(p)}\right) - \left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right) d_u\right],
\end{aligned}$$

with $d_u = 1$ for $u \ge 0$ and -1 otherwise. Letting m_{G_1} denote the upper bound for the density of G_1, we have

$$R_1^{-1} \int_{-\infty}^{\infty} \left(\sum_{j=1}^{R_1} \left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right)\right)^2 dF_1(u) \le m_{G_1}^2 \frac{R_1}{\left(\phi_1^{(p)}\right)^2} \int_{-\infty}^{\infty} u^2 dF_1(u),$$

which drifts to zero under (21) by (A17) and Assumption 7. Also, consider

$$\begin{aligned}
& E^* \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{R_1}} \sum_{j=1}^{R_1} I\left(0 \le \left|\xi_{1,j}^{(p)}\right| \le u/\phi_1^{(p)}\right) - \left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right) d_u\right)^2 dF_1(u) \\
= & E^* \int_{-\infty}^{\infty} \left(I\left(0 \le \left|\xi_{1,1}^{(p)}\right| \le u/\phi_1^{(p)}\right) - \left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right) d_u\right)^2 dF_1(u) \\
= & \int_{-\infty}^{\infty} V^*\{I\left(0 \le \left|\xi_{1,1}^{(p)}\right| \le u/\phi_1^{(p)}\right)\} dF_1(u)
\end{aligned}$$

by the independence of the $\xi_{1,j}^{(p)}$. Elementary arguments yield

$$V^*\{I\left(0 \leq \xi_{1,1}^{(p)} \leq u/\phi_1^{(p)}\right)\} = \left|G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right|\left(1 - \left|G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right|\right)$$
$$\leq \left|G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right| \leq m_{G_1}\frac{|u|}{\phi_1^{(p)}},$$

so that

$$\int_{-\infty}^{\infty} V^*\{I\left(0 \leq \left|\xi_{1,1}^{(p)}\right| \leq u/\phi_1^{(p)}\right)\}dF_1(u) \to 0,$$

as $\phi_1^{(p)} \to \infty$. Thus, by Markov inequality, under (21)

$$\Theta_1^{(p)} = \int_{-\infty}^{\infty}\left(\frac{\sum_{j=1}^{R_1}\left(\zeta_{1,j}^{(p)}(u) - G_1(0)\right)}{\sqrt{R_1}\sqrt{G_1(0)(1-G_1(0))}}\right)^2 dF_1(u)$$
$$= \int_{-\infty}^{\infty}\left(\frac{\sum_{j=1}^{R_1}\left(I\left(\xi_{1,j}^{(p)} \leq 0\right) - G_1(0)\right)}{\sqrt{R_1}\sqrt{G_1(0)(1-G_1(0))}}\right)^2 dF_1(u) + o_{P^*}(1) \stackrel{D^*}{\to} \chi_1^2,$$

with the last passage following from the CLT for Bernoulli random variables and continuity. This proves (22).

We now turn to (23). By (12) and Lemma A2, we have that, under $H_{A,1}^{(p)}$

$$P\left\{\omega : \lim_{\min(N,T)\to\infty} \phi_1^{(p)} = 1\right\} = 1,$$

and therefore we can henceforth assume that

$$\lim_{\min(N,T)\to\infty} \phi_1^{(p)} = 1. \tag{A18}$$

We can write

$$\zeta_{1,j}^{(p)}(u) - G_1(0) = \zeta_{1,j}^{(p)}(u) - G_1(0) \pm G_1\left(u/\phi_1^{(p)}\right),$$

so that

$$E^*\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{R_1}}\sum_{j=1}^{R_1}\zeta_{1,j}^{(p)}(u) - G_1(0)\right)^2 dF_1(u)$$
$$= E^*\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{R_1}}\sum_{j=1}^{R_1}\zeta_{1,j}^{(p)}(u) - G_1\left(u/\phi_1^{(p)}\right)\right)^2 dF_1(u) + R_1\int_{-\infty}^{\infty}\left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right)^2 dF_1(u)$$
$$= \int_{-\infty}^{\infty} V^*\left(\zeta_{1,j}^{(p)}(u)\right) dF_1(u) + R_1\int_{-\infty}^{\infty}\left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right)^2 dF_1(u),$$

having used again the independence of the $\zeta_{1,j}^{(p)}(u)$. Clearly, $V^*\left(\zeta_{1,j}^{(p)}(u)\right) < \infty$; also, as $\min(N,T) \to \infty$, (A18) yields

$$\int_{-\infty}^{\infty} \left(G_1\left(u/\phi_1^{(p)}\right) - G_1(0)\right)^2 dF_1(u) = \int_{-\infty}^{\infty} \left(G_1(u) - G_1(0)\right)^2 dF_1(u),$$

so that, finally

$$\frac{1}{R_1}\Theta_1^{(p)} = \frac{1}{R_1}\int_{-\infty}^{\infty} \left(\frac{\sum_{j=1}^{R_1}\left(\zeta_{1,j}^{(p)}(u) - G_1(0)\right)}{\sqrt{R_1}\sqrt{G_1(0)(1-G_1(0))}}\right)^2 dF_1(u) = \frac{\int_{-\infty}^{\infty}\left(G_1(u) - G_1(0)\right)^2 dF_1(u)}{G_1(0)(1-G_1(0))} + o(1).$$

□

Proof of Lemma 2. Let Z be a $N(0,1)$ random variable. By (22), using Bernstein concentration inequality we have that

$$P^*\left(\Theta_1^{(p)} > c_{\alpha,1}\right) = P^*\left(Z^2 > c_{\alpha,1}\right) + o_{P^*}(1) \leq 2\exp\left(-\frac{1}{2}c_{\alpha,1}\right) + o_{P^*}(1), \tag{A19}$$

which implies that $P^*\left(\Theta_1^{(p)} > c_{\alpha,1}\right)$ drifts to zero as long as $c_{\alpha,1} \to \infty$. Therefore, under $H_{0,1}^{(1)}$, there is zero probability of a Type I error. Under $H_{A,1}^{(1)}$, by (23) we have

$$\begin{aligned} P^*\left(\Theta_1^{(p)} \leq c_{\alpha,1}\right) &= P^*\left[\left(Z + C_0\sqrt{R_1}\right)^2 \leq c_{\alpha,1}\right] + o_{P^*}(1) \\ &\leq P^*\left(|Z| \leq |c_{\alpha,1}|^{1/2} - C_0\sqrt{R_1}\right) + o_{P^*}(1) \\ &\to P^*\left(|Z| \leq -\infty\right) = 0, \end{aligned}$$

since $c_{\alpha,1} = o(R_1)$. Thus, under the alternative there is zero probability of a Type II error. This proves the desired result. □

Proof of Theorem 3. The proof is exactly the same as the proof of Theorem 2. □

Proof of Lemma 3. The proof is exactly the same as the proof of Theorem 3 in Trapani (2018). □

A.3 Discussion of the main assumptions

In this section, we shed further light on Assumptions 2 and 3. We begin by spelling out some easier-to-verify sufficient conditions for the assumptions to hold (Section A.3.1). We then verify such sufficient conditions under various dependence assumptions which are typically employed in the literature (Section A.3.2). Finally, we present several examples of DGPs for which the assumptions are satisfied (Section A.3.3).

Recall that the vector of zero-mean, $I(1)$ common factors f_t^* has dimension $r_2 + d_2$; henceforth, we use the short-hand notation $d = r_2 + d_2$.

A.3.1 Sufficient conditions

We present a set of sufficient conditions which imply Assumption 3 and are easier to verify. In all our subsequent arguments, we will check under which assumptions our sufficient conditions hold, and prove the validity of Assumption 3 by showing them.

Let

$$\begin{aligned}\gamma_{i,ts} &= E\left(u_{i,t}u_{i,s}\right),\\ \gamma_{ts}^{(3)} &= E\left(f_t^{(3)\prime}f_t^{(3)}\right),\\ \gamma_{ts}^{(g)} &= E\left(g_t^{\prime}g_s\right),\end{aligned}$$

and let $|Z|_p$ denote the L_p-norm of an n-dimensional vector Z, viz. $|Z|_p = \left(E\|Z\|^p\right)^{1/p}$, where we recall that $\|\cdot\|$ denotes the Euclidean norm. Consider the following relations

$$\max_{1\leq i\leq N}\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{i,ts}\right| \leq c_0 T; \tag{A20}$$

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{ts}^{(3)}\right| \leq c_0 T; \tag{A21}$$

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{ts}^{(g)}\right| \leq c_0 T; \tag{A22}$$

$$\left|f_t^*\right|_2^2 \leq c_0 t; \tag{A23}$$

$$\left|f_t^*\right|_4^4 \leq c_0 t^2, \tag{A24}$$

where the last two equations hold for all $1 \leq t \leq T$.

Lemma A7. *It holds that*

(i) under (A20) and (A23), Assumption 3(ii)(a) holds, viz.

$$\max_{1\leq i\leq N}\left|\sum_{t=1}^{T} f_t^* u_{i,t}\right|_2^2 \leq c_0 T^2; \tag{A25}$$

(ii) under (A21) and (A23), Assumption 3(ii)(b) holds, viz.

$$E\left\|\sum_{t=1}^{T} f_t^* f_t^{(3)\prime}\right\|^2 \leq c_0 T^2; \tag{A26}$$

(iii) under (A22) and (A23), Assumption 3(ii)(c) holds, viz.

$$E\left\|\sum_{t=1}^{T} f_t^* g_t'\right\|^2 \leq c_0 T^2; \tag{A27}$$

(iv) under (A23), Assumption 3(iii) holds, viz.

$$\left|\sum_{t=1}^{T} t f_t^*\right|_2^2 \leq c_0 T^5; \tag{A28}$$

(v) under (A20), (A21), (A22) Assumptions 3(iv)(a)-(b)-(c) hold respectively, viz.

$$\max_{1\leq i\leq N} E\left|\sum_{t=1}^{T} t u_{i,t}\right|^2 \leq c_0 T; \tag{A29}$$

$$\left|\sum_{t=1}^{T} t f_t^*\right|_2^2 \leq c_0 T; \tag{A30}$$

$$\left|\sum_{t=1}^{T} t g_t\right|_2^2 \leq c_0 T; \tag{A31}$$

(vi) under (A24), Assumption 3(v) holds, viz.

$$E\left\|\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right\|^2 \leq c_0 T^4. \tag{A32}$$

Note that (A24) also immediately implies Assumption 2*(v)*.

A.3.2 Various dependence assumptions

We now prove the validity of Assumptions 2*(iv)* and 3 under several dependence assumptions which are often employed in the literature. Recall that

$$f_t^* = \sum_{j=1}^{t} e_j.$$

The *i.i.d.* case We begin by discussing the *i.i.d.* case as a benchmark, for the sake of completeness. It holds that

Corollary 1. *We assume that* $\{e_t\}$, $\{u_{i,t}\}$, $\{g_t\}$ *and* $\left\{f_t^{(3)}\right\}$ *are i.i.d. for all* $1 \leq i \leq N$*, with* $|e_0|_4 < \infty$, $\max_{1\leq i\leq N} Eu_{i,0}^2 < \infty$, $|g_0|_4 < \infty$ *and* $\left|f_0^{(3)}\right|_4 < \infty$*. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

Linear processes Assume that, for all $1 \leq i \leq N$

$$e_t = \sum_{j=0}^{\infty} c_j^e \varepsilon_{t-j}, \tag{A33}$$

$$u_{i,t} = \sum_{j=0}^{\infty} c_{i,j}^u v_{i,t-j}, \tag{A34}$$

$$f_t^{(3)} = \sum_{j=0}^{\infty} c_j^{(3)} \varepsilon_{t-j}^{(3)}, \tag{A35}$$

$$g_t = \sum_{j=0}^{\infty} c_j^g \varepsilon_{t-j}^g. \tag{A36}$$

Corollary 2. *We assume that the innovations* ε_t, $v_{i,t}$, $\varepsilon_t^{(3)}$ *and* ε_t^g *are i.i.d. with mean zero with* $|\varepsilon_0|_{4+\epsilon} < \infty$ *for some* $\epsilon > 0$, $|v_{i,0}|_2 < \infty$, $\left|\varepsilon_0^{(3)}\right|_2 < \infty$ *and* $|\varepsilon_0^g|_2 < \infty$*, with*

$$\sum_{j=0}^{\infty} j \left\|c_j^e\right\| < \infty \tag{A37}$$

$$\sum_{j=0}^{\infty} j \left|c_{i,j}^u\right| < \infty, \tag{A38}$$

and similarly $\sum_{j=0}^{\infty} j \left\|c_j^{(3)}\right\| < \infty$ *and* $\sum_{j=0}^{\infty} j \left\|c_j^g\right\| < \infty$*. Also,* $c^e(1) = \sum_{j=0}^{\infty} c_j^e$ *has full rank with* $\|c^e(1)\| < \infty$*, and* $E(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon$ *such that* $\|\Sigma_\varepsilon\| < \infty$ *is a positive definite matrix. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

A general result for mixingales Henceforth, for a generic process x_t, we let $\mathcal{F}_{x,t-m}^{t+m} = \sigma(x_{t+m}, ..., x_{t-m})$; further, we write $\mathcal{F}_{x,t-m} = \sigma(x_{t-m}, x_{t-m-1}, ..., x_{t-\infty})$ for short.

We consider the following assumptions

Assumption 1. *The innovations* e_t *are* $L_{4+\epsilon}$*-bounded (for some* $\epsilon > 0$*);* $u_{i,t}$ *(for* $1 \leq i \leq N$*),* $f_t^{(3)}$ *and* g_t *are all* L_4*-bounded; all sequences are zero mean, weak stationary, strong mixing with mixing numbers* $\alpha_m = O(\rho^m)$*, where* $0 < \rho < 1$*.*

Assumption 2. *The innovations e_t, $u_{i,t}$ (for $1 \leq i \leq N$), $f_t^{(3)}$ and g_t are all L_4-bounded, zero mean, weak stationary, uniformly mixing, with mixing numbers $\phi_m = O\left(m^{-\varphi_\phi}\right)$ with $\varphi_\phi > -\left(\frac{4}{3} + \epsilon\right)$ for some $\epsilon > 0$.*

Assumption 3. *It holds that $Var\left(\sum_{j=1}^{t} e_{i,j}\right) \geq r_i(t)$, for $1 \leq i \leq d$, where $r_i(t) \rightarrow \infty$ as $t \rightarrow \infty$.*

Remarks

We note that the exponential rates of strong mixing in Assumption 1 could be replaced, as typical in this literature (see the book by Davidson, 1994), with higher order moment conditions. Assumption 3 is the same as in Corollary 2.2 in Corradi (1999), and it requires that the variance of the partial sums of the $e_{i,t}$s diverge, even at a very slow rate.

Corollary 3. *We assume that e_t, $u_{i,t}$ (for $1 \leq i \leq N$), $f_t^{(3)}$ and g_t satisfy either Assumption 1 or Assumption 2, and that Assumption 3 holds. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

Results under Near Epoch Dependence Consider the following - possibly vector valued - sequences: $\{v_t^e\}_{t=-\infty}^{\infty}$, $\{v_t^{u_i}\}_{t=-\infty}^{\infty}$ (with $1 \leq i \leq N$), $\{v_t^g\}_{t=-\infty}^{\infty}$, $\left\{v_t^3\right\}_{t=-\infty}^{\infty}$, which form four mutually independent groups. We assume that the DGP of e_t, $u_{i,t}$, $f_t^{(3)}$ and g_t are given by

$$e_t = g_e\left(v_{\infty}^e, ..., v_0^e, ..., v_{-\infty}^e\right), \tag{A39}$$

$$u_{i,t} = g_{u_i}\left(v_{\infty}^{u_i}, ..., v_0^{u_i}, ..., v_{-\infty}^{u_i}\right), \tag{A40}$$

$$g_t = g_g\left(v_{\infty}^g, ..., v_0^g, ..., v_{-\infty}^g\right), \tag{A41}$$

$$f_t^{(3)} = g_3\left(v_{\infty}^3, ..., v_0^3, ..., v_{-\infty}^3\right), \tag{A42}$$

for all $1 \leq i \leq N$, where the functions $g_j(\cdot)$ are measurable for all $j \in \{e, u_1, ..., u_N, g, 3\}$.

Assumption 4. *We assume that $\left\{v_t^j\right\}_{t=-\infty}^{\infty}$, for $j \in \{e, u_1, ..., u_N, g, 3\}$, are $L_{4+\epsilon}$-bounded (for some $\epsilon > 0$), zero mean, stationary, strong mixing of mixing size $-(4+\epsilon)$ for some $\epsilon > 0$.*

Assumption 5. *We assume that $\left\{v_t^j\right\}_{t=-\infty}^{\infty}$, for $j \in \{e, u_1, ..., u_N, g, 3\}$, are $L_{4+\epsilon}$-bounded (for some $\epsilon > 0$), zero mean, stationary, uniformly mixing of mixing size $-\left(\frac{4}{3} + \epsilon\right)$ for some $\epsilon > 0$.*

Assumption 6. *We assume that, for $p = 4$*

$$\left|e_t - E\left(e_t | \mathcal{F}_{v^e, t-m}^{t+m}\right)\right|_p \leq c_{e,t} v_{p,m}^e,$$

$$\left|u_{i,t} - E\left(u_{i,t} | \mathcal{F}_{v^{u_i}, t-m}^{t+m}\right)\right|_p \leq c_{u_i,t} v_{p,m}^{u_i},$$

$$\left|g_t - E\left(g_t | \mathcal{F}_{v^g, t-m}^{t+m}\right)\right|_p \leq c_{g,t} v_{p,m}^g,$$

$$\left|f_t^{(3)} - E\left(f_t^{(3)} | \mathcal{F}_{v^3, t-m}^{t+m}\right)\right|_p \leq c_{3,t} v_{p,m}^3,$$

for all $1 \leq i \leq N$, with $v_m^j = o\left(m^{-1}\right)$ for for $j \in \{e, u_1, ..., u_N, g, 3\}$.

Assumption 6 entails that all the sequences are L_p-NED of size -1 on the relevant mixing basis on which they are defined.

Corollary 4. *We assume that (A39)-(A42) hold and that Assumptions 3 and 4-6 are satisfied. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

Results for causal processes We consider the following DGPs

$$e_t = f_e\left(\varepsilon_t, \varepsilon_{t-1}, ...\varepsilon_{-\infty}\right), \tag{A43}$$

$$u_{i,t} = f_{u_i}\left(v_{i,t}, v_{i,t-1}, ...v_{i,-\infty}\right), \tag{A44}$$

$$g_t = f_g\left(\varepsilon_t^g, \varepsilon_{t-1}^g, ...\varepsilon_{-\infty}^g\right), \tag{A45}$$

$$f_t^{(3)} = f_3\left(\varepsilon_t^3, \varepsilon_{t-1}^3, ...\varepsilon_{-\infty}^3\right), \tag{A46}$$

where the shocks $\{\varepsilon_t\}$, $\{v_{i,t}\}$ (for $1 \leq i \leq N$), $\{\varepsilon_t^g\}$, $\left\{\varepsilon_t^3\right\}$ are mutually independent groups of *i.i.d.* variables and the functions $f_j(\cdot)$ are all measurable.

We define the *functional measures of dependence* (Wu, 2005) as

$$\delta_{t,p}^e = \left|f_e\left(\varepsilon_t, ..., \varepsilon_0, ...\varepsilon_{-\infty}\right) - f_e\left(\varepsilon_t, ..., \varepsilon_0', ...\varepsilon_{-\infty}\right)\right|_p$$
$$\delta_{t,p}^{u_i} = \left|f_{u_i}\left(v_{i,t}, ..., v_{i,0}, ...v_{i,-\infty}\right) - f_{u_i}\left(v_{i,t}, ..., v_{i,0}', ...v_{i,-\infty}\right)\right|_p,$$

where ε_0' is a copy of ε_0 such that $\varepsilon_0' \overset{D}{=} \varepsilon_0$ and ε_0' is independent of $\{\varepsilon_t\}$, and $v_{i,0}'$ is defined similarly. We also define $\delta_{t,2}^g$ and $\delta_{t,2}^3$ in the same way.

Assumption 7. *We assume that $\left|u_{i,0}\right|_2 < \infty$, $\left|g_0\right|_2 < \infty$ and $\left|f_0^{(3)}\right|_2 < \infty$, with*

$$\sum_{m=1}^{T} \sum_{j=m}^{\infty} \delta_{j,2}^{u_i} \leq c_0 T \tag{A47}$$

holds for $1 \leq i \leq N$. Similarly

$$\sum_{m=1}^{T} \sum_{j=m}^{\infty} \delta_{j,2}^{g} \leq c_0 T, \tag{A48}$$

$$\sum_{m=1}^{T} \sum_{j=m}^{\infty} \delta_{j,2}^{3} \leq c_0 T. \tag{A49}$$

Assumption 8. *We assume that $\left|e_0\right|_4 < \infty$ and $\sum_{j=m}^{\infty} \delta_{j,p}^e = O\left(m^{-\eta}\right)$, for $p \leq 4$ and for some $\eta > 0$.*

Remarks

We refer to Section A.3.3 for several examples of DGPs that satisfy these assumptions.

Corollary 5. *We assume that (A43)-(A46) hold and that Assumptions 7-8 are satisfied. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

A.3.3 Results for various DGPs

In this section, we build on the results in Section A.3.2 to study various DGPs for which Corollary 5 can be applied.

Transformations of linear processes The set-up in Section A.3.2 lends itself (similarly to the NED set-up, see Chapter 17 in Davidson, 1994) to studying nonlinear tranformations of causal processes. Consider in particular

$$\begin{aligned} e_t^* &= h_e(e_t), \\ u_{i,t}^* &= h_{u,i}(u_{i,t}), \\ g_t^* &= h_g(g_t), \\ f_t^{(3)*} &= h_3\left(f_t^{(3)}\right), \end{aligned}$$

where e_t, $u_{i,t}$, g_t and $f_t^{(3)}$ are the linear processes defined in (A33)-(A36)

$$\|h_e(x) - h_e(x)\| \leq c_0 \|x - y\|, \tag{A50}$$

and the same for $h_{u,i}(\cdot)$, $h_g(\cdot)$ and $h_3(\cdot)$.

Corollary 6. *We assume that e_t^*, $u_{i,t}^*$, g_t^* and $f_t^{(3)^*}$ satisfy the assumptions of Corollary 2. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

Nonlinear autoregression models Consider the nonlinear autoregressions

$$\begin{aligned} e_t &= f^e(e_{t-1}) + \varepsilon_t, \\ u_{i,t} &= f^{u_i}(u_{i,t-1}) + v_{i,t}, \end{aligned}$$

with the DGPs for $f_t^{(3)}$ and g_t defined similarly. The functions $f^e(\cdot)$ and $f^{u_i}(\cdot)$ are assumed to be *contracting maps*, i.e. $\|f^{u_i}(x) - f^{u_i}(y)\| \leq c_0 \|x - y\|$ with $0 \leq c_0 < 1$. Note that a possible example could be the Threshold AutoRegression model, defined as

$$u_{i,t} = \rho_1 \max\{u_{i,t-1}, 0\} + \rho_2 \min\{u_{i,t-1}, 0\} + v_{i,t},$$

where $\max\{|\rho_1|, |\rho_2|\} < 1$.

Corollary 7. *We assume that ε_t, $v_{i,t}$, $\varepsilon_t^{(3)}$ and ε_t^g satisfy the assumptions of Corollary 2. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

Random coefficient autoregressive models We consider

$$e_t = (\Phi_e + b_t^e) e_{t-1} + \varepsilon_t, \tag{A51}$$

$$u_{i,t} = \left(\phi_{u_i} + b_{i,t}^u\right) u_{i,t-1} + v_{i,t}, \tag{A52}$$

with the DGPs for $f_t^{(3)}$ and g_t defined similarly.

Corollary 8. *We assume that* e_t*, and* $u_{i,t}$ *are generated as in (A51)-(A52), and similarly* $f_t^{(3)}$ *and* g_t*. We assume that* b_t^e *and* ε_t *are two independent groups, each of which is is i.i.d. with* $|\varepsilon_0|_4 < \infty$ *and* $|b_0|_4 < \infty$ *and* $0 \leq E \|\Phi_e + b_0^e\|^p < 1$ *for all* $2 \leq p \leq 4$*. We assume also that* $b_{i,t}^u$ *and* $v_{i,t}$ *are two independent groups, each of which is is i.i.d. with* $|v_{i,0}|_2 < \infty$ *and* $|b_0|_2 < \infty$ *and* $\phi_{u_i}^2 + E\left(b_{i,0}^u\right)^2 < 1$*. Then, (A20)-(A22); (A23)-(A24); and Assumption 2(iv) hold.*

GARCH-type models These models lend themselves to being defined as causal processes similar to the ones discussed in Section A.3.2. Recall that a d-dimensional causal process is defined as

$$X_t = g\left(\varepsilon_t, ..., \varepsilon_0, \varepsilon_{-1}, ...\right),$$

where ε_t is a d'-dimensional *i.i.d.* sequance and $g : R^{d' \times \infty} \to R^d$ is a measurable function; consider the construction

$$X_t' = g\left(\varepsilon_t, ..., \varepsilon_0', \varepsilon_{-1}, ...\right),$$

where ε_0' is an independent copy of ε_0. Recalling the functional dependence measure for X_t, viz. $\delta_{t,p}^X = |X_t - X_t'|_p$, the typical result when GARCH-type DGPs are considered is that $\delta_{t,p}^X$ declines *exponentially* with t, similarly to the RCA case considered in Section A.3.3, viz. $\delta_{t,p}^X = O\left(\rho^t\right)$ for some $0 < \rho < 1$. Upon inspecting the proof of Corollary 5, this automatically yields all the results required for Lemma A7.

We report some examples of GARCH-type models which could be considered for the innovations e_t, $u_{i,t}$, $f_t^{(3)}$ and g_t. The key difference is between univariate and multivariate GARCH-type models, since in the latter case there are fewer specifications usually considered in the literature.

We begin by considering a large family of univariate GARCH models for $u_{i,t}$

$$u_{i,t} = h_{i,t} v_{i,t}, \tag{A53}$$

$$\Lambda\left(h_{i,t}^2\right) = c\left(v_{i,t-1}\right) \Lambda\left(h_{i,t-1}^2\right) + d\left(v_{i,t-1}\right), \tag{A54}$$

for all $1 \leq i \leq N$, where $v_{i,t}$ is an *i.i.d.* sequence, $h_{i,t}$ is a nonnegative random variable and the functions $\Lambda(\cdot)$, $c(\cdot)$ and $d(\cdot)$ are all real valued. Model (A53)-(A54) is known as the *augmented GARCH* model (see Duan, 1997). We assume that $\Lambda^{-1}(\cdot)$ and $\Lambda'(\cdot)$ exist (with $\Lambda'(x) \geq 0$ for all x), and

$$E \ln |c\left(v_{i,0}\right)| < \infty, \tag{A55}$$

$$E \ln^+ |c\left(v_{i,0}\right)| < \infty, \tag{A56}$$

$$E \ln^+ |d\left(v_{i,0}\right)| < \infty, \tag{A57}$$

$$\Lambda'\left(h_{i,0}^2\right) \geq \omega_i > 0 \text{ for all } x, \tag{A58}$$

$$\Lambda'(x) \geq 0 \text{ for all } x, \tag{A59}$$

$$\left|\frac{1}{\Lambda'\left(\Lambda^{1-}(x)\right)}\right| \leq C x^{\gamma} \text{ for all } x \geq \omega_i. \tag{A60}$$

Corollary 9. *Assume that (A53)-(A54), (A55)-(A60) and* $|v_{i,0}|_2 < \infty$ *hold, with*

$$E\left|c\left(v_{i,0}\right)\right| < \infty, \tag{A61}$$

$$E\left|d\left(v_{i,0}\right)\right| < \infty. \tag{A62}$$

Then (A20) holds.

Remarks

The DGP (A53)-(A54), and the conditions (A55)-(A62), encompass many GARCH-type specifications. Aue, Berkes, Horváth, et al. (2006) provide a wide variety of examples, which can be grouped in two categories:

1. polynomial GARCH models, where $\Lambda(x) = x^p$, for some $p > 0$. These include the standard GARCH(1,1) model, and its variants such as the GJR model (where $p = 1$), and the Power-ARCH model of Carrasco and Chen (2002);
2. exponential GARCH models, where $\Lambda(x) = \ln x$, which includes the EGARCH as a leading example.

We now provide examples of multivariate GARCH models which have an exponential rate of decay for the functional dependence measure coefficients. While we refer only to e_t in the following, the same results would apply to g_t and $f_t^{(3)}$.

We consider the following specifications (see also Aue, Hörmann, Horváth, and Reimherr (2009)), where $\odot$ denotes the Hadamard product

1. the CCC-GARCH (Bollerslev, 1990)

$$e_t = h_t \odot \varepsilon_t, \tag{A63}$$

$$h_t \odot h_t = \omega + \sum_{l=1}^{p} \alpha_l \odot h_{t-l} \odot h_{t-l} + \sum_{l=1}^{q} \beta_l \odot e_{t-l} \odot e_{t-l}, \tag{A64}$$

where the vectors ω is coordinate-wise strictly positive and the vectors $\{\alpha_l\}_{l=1}^{p}$ and $\{\beta_l\}_{l=1}^{q}$ are coordinate-wise nonnegative;

2. the CCC-GARCH variant of Jeantheau (1998)

$$e_t = h_t \odot \varepsilon_t, \tag{A65}$$

$$h_t \odot h_t = \omega + \sum_{l=1}^{p} A_l \left(h_{t-l} \odot h_{t-l}\right) + \sum_{l=1}^{q} B_l \left(e_{t-l} \odot e_{t-l}\right), \tag{A66}$$

where $\{A_l\}_{l=1}^{p}$ and $\{B_l\}_{l=1}^{q}$ are nonnegative definite matrices;

3. the multivariate exponential GARCH of Kawakatsu (2006)

$$e_t = H_t^{1/2} \varepsilon_t, \tag{A67}$$

$$vech\left(\ln H_t - C\right) = Avech\left(\ln H_{t-1} - C\right) + F\left(\varepsilon_{t-1}, ..., \varepsilon_{t-q}\right), \tag{A68}$$

where $F(\cdot): R^d \to R^{\frac{d(d+1)}{2}}$ is a measurable function and C is a symmetric $d \times d$ matrix; we require that

$$E \ln |F(\varepsilon_{-1}, ..., \varepsilon_{-q})| < \infty, \tag{A69}$$

$$E(t \exp(F(\varepsilon_{t-1}, ..., \varepsilon_{t-q}))) < \infty, \tag{A70}$$

for some $t > \sqrt{8}q$.

Corollary 10. *We assume that, in (A63), (A65) and (A67), ε_t is i.i.d. with mean zero and $|\varepsilon_0|_4 < \infty$. Further, we assume that*

$$\gamma_C = \max_{1 \le j \le d} \sum_{l=1}^{\max\{p,q\}} \left|\alpha_{j,l} + \beta_{j,l}\varepsilon_{j,0}^2\right|_2 < 1,$$

in (A63);

$$\gamma_J = \sum_{l=1}^{\max\{p,q\}} |A_l + B_l E_0|_2 < 1,$$

in (A65), where E_0 is a diagonal matrix whose main diagonal is given by $\varepsilon_0 \odot \varepsilon_0$; and, in (A67), that (A69)-(A70) hold, with $\|A\| < 1$. In all the preceding displays, extra coefficients are set to zero. Then, (A23)-(A24) and Assumption 2(iv) hold for models (A63)-(A64), (A65)-(A66) and (A67)-(A68).

A.3.4 Proofs

Proof of Lemma A7. Consider (A25), and note that, for $1 \le i \le N$

$$\begin{aligned} E\left\|\sum_{t=1}^T f_t^* u_{i,t}\right\|^2 &= E\sum_{t=1}^T\sum_{s=1}^T f_t^{*\prime} f_s^* u_{i,t}u_{i,s} = \sum_{t=1}^T\sum_{s=1}^T E\left(f_t^{*\prime} f_s^*\right) E\left(u_{i,t}u_{i,s}\right) \\ &\le \sum_{t=1}^T\sum_{s=1}^T \left|E\left(f_t^{*\prime} f_s^*\right)\right| |\gamma_{i,ts}| \le \sum_{t=1}^T\sum_{s=1}^T |f_t^*|_2 |f_s^*|_2 |\gamma_{i,ts}| \\ &\le c_0 \sum_{t=1}^T\sum_{s=1}^T t^{1/2}s^{1/2} |\gamma_{i,ts}| \le c_1 T \sum_{t=1}^T\sum_{s=1}^T |\gamma_{i,ts}| \\ &\le c_0 T^2, \end{aligned}$$

having used the Cauchy-Schwartz inequality in the fourth passage, (A23) in the fifth passage, and (A20) in the last one. Consider (A27), and note that

$$
\begin{aligned}
E\left\|\sum_{t=1}^{T} f_t^* f_t^{(3)\prime}\right\|^2 &= E\, tr\left(\sum_{t=1}^{T} f_t^* f_t^{(3)\prime} \sum_{t=1}^{T} f_t^{(3)} f_t^{*\prime}\right) = E\, tr\left(\sum_{t=1}^{T}\sum_{s=1}^{T} f_t^* f_t^{(3)\prime} f_s^{(3)} f_s^{*\prime}\right) \\
&= E\sum_{t=1}^{T}\sum_{s=1}^{T} tr\left(f_t^* f_t^{(3)\prime} f_s^{(3)} f_s^{*\prime}\right) = E\sum_{t=1}^{T}\sum_{s=1}^{T} f_t^{(3)\prime} f_s^{(3)} f_s^{*\prime} f_t^* \\
&= \sum_{t=1}^{T}\sum_{s=1}^{T} E\left(f_s^{*\prime} f_t^*\right) E\left(f_t^{(3)\prime} f_s^{(3)}\right);
\end{aligned}
$$

henceforth, the proof is the same as for (A25), and it is therefore omitted. Similarly, the proof of (A26) follows exactly the same passages and is therefore omitted. Turning to (A28), we have

$$
\begin{aligned}
\left|\sum_{t=1}^{T} t f_t^*\right|_2^2 &= E\left(\sum_{t=1}^{T} t f_t^*\right)'\left(\sum_{t=1}^{T} t f_t^*\right) = E\sum_{t=1}^{T}\sum_{s=1}^{T} ts f_t^{*\prime} f_s^* \\
&= \sum_{t=1}^{T}\sum_{s=1}^{T} ts E\left(f_t^{*\prime} f_s^*\right) \le \sum_{t=1}^{T}\sum_{s=1}^{T} ts\left|E\left(f_t^{*\prime} f_s^*\right)\right| \\
&\le \sum_{t=1}^{T}\sum_{s=1}^{T} ts\left|f_t^*\right|_2\left|f_s^*\right|_2 \le c_0 T^2 \sum_{t=1}^{T}\sum_{s=1}^{T}\left|f_t^*\right|_2\left|f_s^*\right|_2 \\
&= c_0 T^2\left(\sum_{t=1}^{T}\left|f_t^*\right|_2\right)^2 \le c_1 T^2\left(\sum_{t=1}^{T} t^{1/2}\right)^2 \le c_2 T^5,
\end{aligned}
$$

having used the Cauchy-Schwartz inequality in the fifth passage and (A23) in the seventh passage. We now consider (A29)-(A31), and provide a full proof only for the first result (the other two follow from the same passages); note that, for all $1 \le i \le N$

$$
E\left|\sum_{t=1}^{T} t u_{i,t}\right|^2 = \sum_{t=1}^{T}\sum_{s=1}^{T} ts E\left(u_{i,t} u_{i,s}\right) \le c_0 T^2 \sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{i,ts}\right| \le c_1 T^3.
$$

We finally consider (A32)

$$
\begin{aligned}
E\left\|\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right\|^2 &= Etr\left(\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right)\left(\sum_{t=1}^{T} f_t^* f_t^{*\prime}\right) = E\sum_{t=1}^{T}\sum_{s=1}^{T} tr\left(f_t^* f_t^{*\prime} f_s^* f_s^{*\prime}\right) \\
&= E\sum_{t=1}^{T}\sum_{s=1}^{T}\left(f_t^{*\prime} f_s^*\right)^2 \leq E\sum_{t=1}^{T}\sum_{s=1}^{T}\left|f_t^*\right|_2^2\left|f_s^*\right|_2^2 \\
&\leq \sum_{t=1}^{T}\sum_{s=1}^{T}\left|f_t^*\right|_4^2\left|f_s^*\right|_4^2 = \left(\sum_{t=1}^{T}\left|f_t^*\right|_4^2\right)^2 \leq c_0\left(\sum_{t=1}^{T} t\right)^2 \\
&\leq c_1 T^4,
\end{aligned}
$$

having used the Cauchy-Schwartz inequality in the fifth and sixth passages, and (A24) in the eight passage. □

Proof of Corollary 1. It is immediate, by direct calculation, to verify (A20)-(A22). Similarly, (A23) also follows by direct calculation. Finally, note that

$$
\left|f_t^*\right|_4^4 = \left|\sum_{j=1}^{t} e_j\right|_4^4 \leq c_0 E\left(\sum_{j=1}^{t} e_j^2\right)^2,
$$

from the Burkholder inequality. Using Holder's inequality

$$
E\left(\sum_{j=1}^{t} e_j^2\right)^2 \leq c_0 t\sum_{j=1}^{t} Ee_j^4 \leq c_1 t^2,
$$

recalling that e_j is i.i.d. with finite fourth moment. Thus, Lemma A7 follows and therefore Assumption 3 is proven. As far as Assumption 2*(iv)* is concerned, it follows from Theorem 1 in Berkes and Philipp (1979). □

Proof of Corollary 2. We begin by showing that (A20) holds; as before, the proofs of (A21) and (A22) follow from the same arguments and we therefore omit them to save space. Let $c_i^u(L) = \sum_{j=0}^{\infty} c_{i,j}^u L^j$; using the Beveridge-Nelson decomposition, we can write

$$
c_i^u(L) = c_i^u(1) - (1-L)\widetilde{c}_i^u(L),
$$

where

$$
\begin{aligned}
c_i^u(1) &= \sum_{j=0}^{\infty} c_{i,j}^u, \\
\widetilde{c}_i^u(L) &= \sum_{j=0}^{\infty} \widetilde{c}_{i,j}^u L^j,
\end{aligned}
$$

with $\widetilde{c}^u_{i,j} = \sum_{k=j+1}^{\infty} c^u_{i,k}$. Note that

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{i,ts}\right| = TEu_{i,0}^2 + \sum_{h=1}^{T-1}(T-h)\left|\gamma_{i,0h}\right|.$$

Also, noting that (A38) entails that $\sum_{j=0}^{\infty}\left(c^u_{i,j}\right)^2 < \infty$, it follows that $Eu_{i,0}^2 = E\left(v_{i,0}^2\right)\sum_{j=0}^{\infty}\left(c^u_{i,j}\right)^2 < \infty$. Equation (28) in Phillips and Solo (1992) yields $\gamma_{i,0h} = \sum_{j=0}^{\infty} c^u_{i,j}c^u_{i,j+h}E\left(v_{i,0}^2\right)$. Hence we can write

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\gamma_{i,ts}\right| \le c_0 T + c_0 T\sum_{h=1}^{T-1}\left|\sum_{j=0}^{\infty} c^u_{i,j}c^u_{i,j+h}\right| \le c_1 T\sum_{h=0}^{\infty}\left|\sum_{j=0}^{\infty} c^u_{i,j}c^u_{i,j+h}\right|.$$

We now show that $\sum_{h=0}^{\infty}\left|\sum_{j=0}^{\infty} c^u_{i,j}c^u_{i,j+h}\right| < \infty$. Indeed, using the same logic in the proof of Lemma 3.6 in Phillips and Solo (1992)

$$\begin{aligned}\sum_{h=0}^{\infty}\left|\sum_{j=0}^{\infty} c^u_{i,j}c^u_{i,j+h}\right| &\le \sum_{h=0}^{\infty}\sum_{j=0}^{\infty}\left|c^u_{i,j}c^u_{i,j+h}\right| \\ &= \sum_{h=0}^{\infty}\sum_{j=1}^{h}\left|c^u_{i,j}c^u_{i,j+h}\right| + \sum_{h=0}^{\infty}\sum_{j=h+1}^{\infty}\left|c^u_{i,j}c^u_{i,j+h}\right| + \sum_{h=0}^{\infty}\left|c^u_{i,j}c^u_{i,h}\right| \\ &= I + II + III.\end{aligned}$$

Hence

$$\begin{aligned}I = \sum_{h=0}^{\infty}\sum_{j=1}^{h}\left|c^u_{i,j}c^u_{i,j+h}\right| &\le \sum_{h=0}^{\infty}\left(\sum_{j=1}^{h}\left(c^u_{i,j}\right)^2\right)^{1/2}\left(\sum_{j=1}^{h}\left(c^u_{i,j+h}\right)^2\right)^{1/2} \\ &\le \sum_{h=0}^{\infty}\left(\sum_{j=1}^{h}\left|c^u_{i,j}\right|\right)\left(\sum_{j=1}^{h}\left|c^u_{i,j+h}\right|\right) \le c_0\sum_{h=0}^{\infty}\sum_{j=1}^{h}\left|c^u_{i,j+h}\right| \\ &\le c_0\sum_{j=1}^{\infty} j\left|c^u_{i,j+h}\right| \le c_0\sum_{j=1}^{\infty}(j+h)\left|c^u_{i,j+h}\right| < \infty,\end{aligned}$$

having used (A38) repeatedly, and the fact that it implies $\sum_{j=1}^{h}\left|c_{i,j}^{u}\right|<\infty$ for all h. Similarly

$$
\begin{aligned}
II &\leq \sum_{h=0}^{\infty}\left(\sum_{j=h+1}^{\infty}\left(c_{i,j}^{u}\right)^{2}\right)^{1/2}\left(\sum_{j=h+1}^{\infty}\left(c_{i,j+h}^{u}\right)^{2}\right)^{1/2} \\
&\leq c_{0}\sum_{h=0}^{\infty}\left(\sum_{j=h+1}^{\infty}\left|c_{i,j}^{u}\right|\right)\left(\sum_{j=h+1}^{\infty}\left|c_{i,j+h}^{u}\right|\right)\leq c_{0}\sum_{h=0}^{\infty}\sum_{j=h+1}^{\infty}\left|c_{i,j+h}^{u}\right| \\
&\leq c_{0}\sum_{j=1}^{\infty}j\left|c_{i,j+h}^{u}\right|\leq c_{0}\sum_{j=1}^{\infty}(j+h)\left|c_{i,j+h}^{u}\right|<\infty,
\end{aligned}
$$

again by (A38). Finally, it is easy to see that $III\leq\left|c_{i,j}^{u}\right|\sum_{h=0}^{\infty}\left|c_{i,h}^{u}\right|<\infty$. Putting all together, the desired result follows. The proof of (A22) and (A21) follows from the same logic.

We now estimate $|f_t^*|_2^2$. Let $c^e(1)=\sum_{j=0}^{\infty}c_j^e$, and note

$$
f_t^* = c^e(1)\sum_{j=1}^{t}\varepsilon_j+\widetilde{\varepsilon}_0-\widetilde{\varepsilon}_t,
$$

where $\widetilde{\varepsilon}_t=\sum_{j=0}^{\infty}\widetilde{c}_j^e L^j\varepsilon_{t-j}$, with $\widetilde{c}_j^e=\sum_{k=j+1}^{\infty}\widetilde{c}_k^e$. We have

$$
|f_t^*|_2^2\leq c_0\left(\left|c^e(1)\sum_{j=1}^{t}\varepsilon_j\right|_2^2+2\left|\widetilde{\varepsilon}_0\right|_2^2\right). \tag{A71}
$$

Also

$$
\begin{aligned}
\left|c^e(1)\sum_{j=1}^{t}\varepsilon_j\right|_2^2 &= E\left(\sum_{j=1}^{t}\varepsilon_j\right)'(c^e(1))'(c^e(1))\left(\sum_{j=1}^{t}\varepsilon_j\right) \\
&= Etr\left((c^e(1))'(c^e(1))\left(\sum_{j=1}^{t}\varepsilon_j\right)\left(\sum_{j=1}^{t}\varepsilon_j\right)'\right) \\
&= tr\left((c^e(1))'(c^e(1))\,E\left(\sum_{j=1}^{t}\varepsilon_j\right)\left(\sum_{j=1}^{t}\varepsilon_j\right)'\right) \\
&= tr\left((c^e(1))'c^e(1)\,t\Sigma_\varepsilon\right) \\
&= t\left(tr\left((c^e(1))'c^e(1)\,\Sigma_\varepsilon\right)\right)\leq c_0 t,
\end{aligned}
$$

using the fact that

$$\sum_{i=1}^{d} \lambda_i \left((c^e(1))' c^e(1) \right) \lambda_{d-i}(\Sigma_\varepsilon) \leq tr\left((c^e(1))' c^e(1) \Sigma_\varepsilon \right) \leq \sum_{i=1}^{d} \lambda_i \left((c^e(1))' c^e(1) \right) \lambda_i(\Sigma_\varepsilon),$$

where $\lambda_i(A)$ denotes the i-th smallest eigenvalue of a matrix A (see Bushell and Trustrum, 1990), and recalling that $(c^e(1))'(c^e(1))$ and Σ_ε are both positive definite. Further, by the proof of Theorem 3.2 in Phillips and Solo (1992), it follows that, under our assumptions, $|\widetilde{\varepsilon}_0|_4 < \infty$, whence (A24) follows from (A71). Similarly, note that

$$|f_t^*|_4^4 \leq c_0 \left(\left| c^e(1) \sum_{j=1}^{t} \varepsilon_j \right|_4^4 + 2 |\widetilde{\varepsilon}_0|_4^4 \right), \tag{A72}$$

and

$$\left| c^e(1) \sum_{j=1}^{t} \varepsilon_j \right|_4^4 \leq \| c^e(1) \| \left| \sum_{j=1}^{t} \varepsilon_j \right|_4^4 \leq c_0 t \sum_{j=1}^{t} |\varepsilon_j|_4^4 = c_0 |\varepsilon_0|_4^4 t^2,$$

which yields the desired result.

Finally, we prove that Assumption 2*(iv)* holds. Note that (A37) entails that $\sum_{j=m}^{\infty} \| c_j^e \| = O(m^{-1})$; thus, equation (3.20) in Liu and Lin (2009) holds with e.g. $p = 3$. This entails that we can use Corollary 3.7 in Liu and Lin (2009), whence it follows that, on a richer probability space, there exists a standard, real-valued, d-dimensional Brownian motion $W(t)$ such that

$$\sup_{1 \leq t \leq T} \left\| f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t) \right\| = o_{a.s.}\left(T^{1/3} \ln \ln T \right), \tag{A73}$$

which ensures that Assumption 2*(iv)* holds. □

Proof of Corollary 3. We show (A20); (A21) and (A22) follow from exactly the same logic. Recall first that $u_{i,t}$ is weakly stationary and L_4-bounded. Then it holds that (see Davidson, 1994, p. 212)

$$|E(u_{i,0}, u_{i,m})| \leq c_0 \alpha_m^{1/2} |u_{i,0}|_2 |u_{i,m}|_2,$$
$$|E(u_{i,0}, u_{i,m})| \leq c_0 \phi_m^{3/4} |u_{i,0}|_{4/3} |u_{i,m}|_4,$$

under Assumptions 1 and 2 respectively. This entails that the $|E(u_{i,0}, u_{i,m})|$ are summable across m under both Assumptions 1 and 2. Since

$$\sum_{t=1}^{T} \sum_{s=1}^{T} |E(u_{i,t}, u_{i,s})| \leq T \left| E\left(u_{i,0}^2\right) \right| + T \sum_{m=1}^{T} |E(u_{i,0}, u_{i,m})|,$$

(A20) follows immediately. We now turn to showing (A23)-(A24). Note that Assumptions 1 and 2 entail (see Davidson, 1994, p. 248) that, for $2 \leq p \leq 4$, $\left|E\left(e_t|\mathcal{F}_{e,t-m}\right)\right|_p \leq c_t \xi^e_{p,m}$ with $\sup_t c_t < \infty$

$$\xi^e_{p,m} = \alpha_m^{1/p-1/4}, \tag{A74}$$

$$\xi^e_{p,m} = \phi_m^{1-1/4}, \tag{A75}$$

according as Assumption 1 or 2 holds. Recall that e_t is (weakly) stationary; also

$$\left|E\left(\sum_{j=1}^t e_j|\mathcal{F}_{e,0}\right)\right|_p \leq c_0 \sum_{j=1}^t \left|E\left(e_j|\mathcal{F}_{e,0}\right)\right|_p .$$

Thus, using (A74) and (A75), it is easy to see that

$$\sum_{j=1}^{\infty} j^{-3/2} \sum_{k=1}^{j} \xi^e_{p,k} < \infty.$$

Thus, by Theorem 1 in Peligrad, Utev, and Wu (2007), we have

$$\left|\sum_{j=1}^t e_j\right|_p^p \leq E \max_{1\leq k\leq t}\left|\sum_{j=1}^k e_j\right|_p^p \leq c_0 t^{p/2}\left(|e_1|_p^p + \sum_{j=1}^{\infty} j^{-3/2}\sum_{k=1}^{j}\xi^e_{p,k}\right) \leq c_1 t^{p/2},$$

which yields (A23) and (A24) for $p = 2$ and $p = 4$ respectively.

We now verify that Assumption 2*(iv)* holds. Write $f_t^m = \sum_{j=m+1}^{m+t} e_t$. Equations (A74) and (A75) entail that the $\xi^e_{2,m}$s are summable across m; hence, by the definition of mixingale

$$\left|E\left(f_t^m|\mathcal{F}_{e,m}\right)\right|_2 \leq \sum_{j=m+1}^{m+t}\left|E\left(e_t|\mathcal{F}_{e,m}\right)\right|_2 \leq \sup_t c_t \sum_{j=1}^t \xi^e_{2,j} \leq c_0,$$

for all m and t. Thus, equation (1.3) in Eberlein (1986) is satisfied; this, and Assumption 3 (see Corollary 2.2 in Corradi (1999)) entail that all the assumptions in Theorem 2 in Eberlein (1986) are satisfied, which in turn entails that there exists a $\kappa > 0$ such that, on a richer probability space, there exists a standard, real-valued, d-dimensional Brownian motion $W(t)$ such that

$$\sup_{1\leq t\leq T}\left\|f_t^* - \Sigma_{\Delta f^*}^{1/2} W(t)\right\| = O_{a.s.}\left(T^{1/2-\kappa}\right),$$

which ensures that Assumption 2*(iv)* holds. □

Proof of Corollary 4. We start by showing (A20) (note that the proofs of (A21) and (A22) are, as usual, the same). Since $u_{i,t}$ is an L_4-NED sequence of size -1, Theorem 17.7 in Davidson (1994) stipulates that $\sum_{m=1}^T \left|E\left(u_{i,0}u_{i,m}\right)\right| < \infty$, which immediately yields (A20). We now turn to showing that (A23)-(A24) and Assumption 2*(iv)* hold. By Theorem 17.5 in Davidson (1994), our assumptions imply that e_t is an L_4-mixingale of size -1. In turn, this immediately entails that (A23)-(A24) hold by Corollary 3. Further, from Corollary 3, it also holds that $\left|E\left(f_t^m|\mathcal{F}_{e,m}\right)\right|_2 \leq c_0$ for all m and t. Also, by the measurability of $g_e(\cdot)$, e_t is a stationary sequence; thus, Corollary 2.2 in Corradi (1999) entails that all the assumptions in Theorem 2 in Eberlein (1986) hold, which yields Assumption 2*(iv)*. □

Proof of Corollary 5. We start with (A20). Note that, by the measurability of f_{u_i}, it follows that $u_{i,t}$ is a stationary sequence. Thus

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\left|E\left(u_{i,t}u_{i,s}\right)\right| = TE\left(u_{i,0}^2\right) + 2\sum_{m=1}^{T}\left(T-m\right)\left|E\left(u_{i,0}u_{i,m}\right)\right|.$$

Note now that

$$\left|E\left(u_{i,0}u_{i,m}\right)\right| = \left|E\left(u_{i,m}\left(u_{i,0} - E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right) + E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right)\right)\right)\right|,$$

where $0 < a < 1$. Thus

$$\begin{aligned}
&\left|E\left(u_{i,0}u_{i,m}\right)\right| \\
&\leq \left|E\left(u_{i,m}\left(u_{i,0} - E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right)\right)\right)\right| \\
&+ \left|E\left(u_{i,m}E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right)\right)\right| \\
&= I + II.
\end{aligned}$$

Also

$$I \leq \left|u_{i,m}\right|_2\left|u_{i,0} - E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right)\right|_2 \leq c_0\left|u_{i,0} - E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right)\right|_2.$$

Now, define the projection $Q_jX_0 = E\left(X_0|\mathcal{F}^0_{X,-j}\right)$ for a generic sequence X_t; clearly, $X_0 = \lim_{j\to\infty} Q_jX_0$. Define also $\widetilde{Q}_jX_0 = Q_jX_0 - Q_{j-1}X_0$. With this notation,

$$u_{i,0} - E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right) = \sum_{j=\lfloor am\rfloor}^{\infty}\widetilde{Q}_ju_{i,0}.$$

Wu (2005) shows that $\left|\widetilde{Q}_ju_{i,0}\right|_p \leq \delta_{j,p}^{u_i}$ for $p \geq 1$. Hence

$$\left|u_{i,0} - E\left(u_{i,0}|\mathcal{F}^0_{v_{i,t},-\lfloor am\rfloor}\right)\right|_2 \leq \sum_{j=\lfloor am\rfloor}^{\infty}\left|\widetilde{Q}_ju_{i,0}\right|_2 \leq \sum_{j=\lfloor am\rfloor}^{\infty}\delta_{j,2}^{u_i}. \tag{A76}$$

Similarly

$$
\begin{aligned}
&\left|E\left(u_{i,m}E\left(u_{i,0}|\mathcal{F}^{0}_{v_{i,t},-\lfloor am\rfloor}\right)\right)\right| \\
&= \left|E\left(E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\lfloor am\rfloor}\right)E\left(u_{i,0}|\mathcal{F}^{0}_{v_{i,t},-\lfloor am\rfloor}\right)\right)\right| \\
&\leq \left|E\left(u_{i,0}|\mathcal{F}^{0}_{v_{i,t},-\lfloor am\rfloor}\right)\right|_2 \left|E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\lfloor am\rfloor}\right)\right|_2 \\
&\leq \left|u_{i,0}\right|_2 \left|E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\infty}\right)\right|_2 \\
&\leq c_0 \left|E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\infty}\right)\right|_2 .
\end{aligned}
$$

Consider now the projection $P_0X_j = E\left(X_j|\mathcal{F}^{0}_{X,-\infty}\right) - E\left(X_j|\mathcal{F}^{-1}_{X,-\infty}\right)$. It holds that

$$
E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\infty}\right) = \sum_{l=0}^{\infty} P_{m-l}u_{i,m} - \sum_{l=0}^{m-\lfloor am\rfloor-1} P_{m-l}u_{i,m} = \sum_{l=m-\lfloor am\rfloor}^{\infty} P_{m-l}u_{i,m},
$$

so that

$$
\left|E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\infty}\right)\right|_2 \leq \sum_{l=m-\lfloor am\rfloor}^{\infty} \left|P_{m-l}u_{i,m}\right|_2 = \sum_{l=m-\lfloor am\rfloor}^{\infty} \left|P_0u_{i,l}\right|_2 ,
$$

by stationarity. As shown in Wu (2005), $\left|P_0u_{i,l}\right|_2 \leq \delta^{u_i}_{l,2}$, so that ultimately

$$
\left|E\left(u_{i,m}|\mathcal{F}^{\lfloor am\rfloor}_{v_{i,t},-\infty}\right)\right|_2 \leq \sum_{j=m-\lfloor am\rfloor}^{\infty} \delta^{u_i}_{j,2}. \tag{A77}
$$

Putting together (A76) and (A77), and recalling (A47), (A20) follows. The same applies to (A21) and (A22). We now turn to showing (A23) and (A24). Assumption 8 entails that, by Lemma A.2 in Liu and Lin (2009),

$$
\left|\sum_{j=1}^{t} \varepsilon_j\right|_p^p \leq c_0 t^{p/2},
$$

for all $p \leq 4$, thus providing the desired result. Note that the result only requires $\sum_{j=0}^{\infty} \delta^{e}_{j,p} < \infty$. Finally, consider Assumption 2*(iv)*. All the assumptions of Theorem 2.2 in Liu and Lin (2009) are satisfied, and therefore it follows that, on a richer probability space, there exists a standard, real-valued, d-dimensional Brownian motion $W(t)$ such that

$$
\sup_{1\leq t\leq T} \left\| f^*_t - \Sigma^{1/2}_{\Delta f^*} W(t) \right\| = o_{a.s.}\left(T^{-\frac{\eta}{1+4\eta}+\epsilon}\right),
$$

for any $\epsilon > 0$, which ensures that Assumption 2*(iv)* holds. □

Proof of Corollary 6. The proof hinges on the fact that e^*_t, $u^*_{i,t}$, g^*_t and $f^{(3)^*}_t$ can be represented as causal processes, i.e.

$$\begin{aligned}
e_t^* &= f_e^* \left(\varepsilon_t, ..., \varepsilon_0, ...\varepsilon_{-\infty}\right), \\
u_{i,t}^* &= f_{u,i}^* \left(v_{i,t}, ..., v_{i,0}, ..., v_{i,-\infty}\right), \\
g_t^* &= f_g^* \left(\varepsilon_t^g, ..., \varepsilon_0^g, ...\varepsilon_{-\infty}^g\right), \\
f_t^{(3)*} &= f_3^* \left(\varepsilon_t^{(3)}, ..., \varepsilon_0^{(3)}, ...\varepsilon_{-\infty}^{(3)}\right),
\end{aligned}$$

where the *i.i.d.* innovations are defined in (A33)-(A36). We have

$$\begin{aligned}
\delta_{t,2}^{u_i*} &= \left| f_{u,i}^* \left(v_{i,t}, ..., v_{i,0}, ..., v_{i,-\infty}\right) - f_{u,i}^* \left(v_{i,t}, ..., v_{i,0}', ..., v_{i,-\infty}\right)\right|_2 \\
&\leq c_0 \left| c_{i,t}^u \left(v_{i,0} - v_{i,0}'\right)\right|_2 \leq c_1 \left| c_{i,t}^u \right|,
\end{aligned}$$

where $v_{i,0}'$ is an independent copy of $v_{i,0}$. Assumption (A38) therefore entails that $\sum_{m=1}^T \sum_{t=m}^\infty \delta_{t,2}^{u_i*} < \infty$; thus, using the same passages as in the proof of Corollary 5, the desired result obtains. We now turn to (A23)-(A24). It holds that

$$\begin{aligned}
\delta_{t,p}^{e*} &= \left| f_e^* \left(\varepsilon_t, ..., \varepsilon_0, ...\varepsilon_{-\infty}\right) - f_e^* \left(\varepsilon_t, ..., \varepsilon_0', ...\varepsilon_{-\infty}\right)\right|_p \\
&\leq c_0 \left| c_t^e \left(\varepsilon_0 - \varepsilon_0'\right)\right|_p \leq c_1 \left\| c_t^e \right\|,
\end{aligned}$$

and by (A37) we have $\sum_{t=0}^\infty \delta_{t,p}^{e*} < \infty$ for all $p \leq 4$; hence, Lemma A.2 in Liu and Lin (2009) yields (A23)-(A24). Finally, we show the validity of Assumption 2*(iv)*. Note that $|e_t^*|_4 \leq c_0 |e_t|_4 < \infty$. Further, (A37) entails that $\sum_{t=m}^\infty \delta_{t,p}^{e*} = O\left(m^{-1}\right)$, so that equation (3.20) in Liu and Lin (2009) holds with e.g. $p = 3$. This yields (A73). □

Proof of Corollary 7. Note that, under the assumptions on $f^e(\cdot)$, it is possible to write all sequences as causal processes, viz.

$$e_t = g_e \left(\varepsilon_t, \varepsilon_{t-1}, ...\right),$$

for a measurable function $g_e(\cdot)$ such that, for all $p \leq 4$, $\delta_{t,p}^e = O\left(\rho^t\right)$, with $0 < \rho < 1$, and the same holds for all the other sequences considered. The proof then follows the same arguments as the proof of Corollary 5. □

Proof of Corollary 8. We begin with (A20), noting that, as usual, (A21) and (A22) follow exactly the same logic. Under the assumptions of the corollary, it is well known (see e.g. Aue, Horvath, and Steinebach, 2006) that (A52) converges exponentially fast, for all initial values $u_{i,0}$, to a unique stationary solution defined as

$$\overline{u}_{i,t} = \sum_{s=-\infty}^{t} v_{i,s} \prod_{z=s+1}^{t} \left(\phi_{u_i} + b_{i,z}^u\right). \tag{A78}$$

We estimate

$$
\begin{aligned}
E\left(u_{i,t}u_{i,s}\right) &= E\left(\overline{u}_{i,t}+\left(u_{i,t}-\overline{u}_{i,t}\right)\right)\left(\overline{u}_{i,s}+\left(u_{i,s}-\overline{u}_{i,s}\right)\right) \\
&= E\left(\overline{u}_{i,t}\overline{u}_{i,s}\right)+E\left(\overline{u}_{i,t}\left(u_{i,s}-\overline{u}_{i,s}\right)\right) \\
&\quad +E\left(\overline{u}_{i,s}\left(u_{i,t}-\overline{u}_{i,t}\right)\right)+E\left(\left(u_{i,s}-\overline{u}_{i,s}\right)\left(u_{i,t}-\overline{u}_{i,t}\right)\right) \\
&= I+II+III+IV.
\end{aligned}
$$

Note that

$$
II \leq \left|\overline{u}_{i,t}\right|_2\left|u_{i,s}-\overline{u}_{i,s}\right|_2,
$$

and that $\left|\overline{u}_{i,t}\right|_2=\left|\overline{u}_{i,0}\right|_2<\infty$ because

$$
E\overline{u}_{i,0}^2=Ev_{i,0}^2\sum_{s=-\infty}^{0}\prod_{z=s+1}^{0}E\left(\phi_{u_i}+b_{i,z}^u\right)^2=\rho\left(Ev_{i,0}^2\right)\sum_{s=0}^{\infty}\rho_{u_i}^{-s},
$$

where $\rho_{u_i}=\phi_{u_i}^2+E\left(b_{i,0}^u\right)^2<1$, whence $E\overline{u}_{i,0}^2<\infty$ follows immediately. Also, using equation (A10) in Horváth and Trapani (2019), $\left|u_{i,s}-\overline{u}_{i,s}\right|_2=O\left(\exp\left(-c_1 s\right)\right)$ for $c_1>0$. Thus, $II=O\left(\exp\left(-c_1 s\right)\right)$; similarly, $III=O\left(\exp\left(-c_2 t\right)\right)$, and using the Cauchy-Schwartz inequality it also follows that $IV=O\left(\exp\left(-c_1 s-c_2 t\right)\right)$. Finally, standard algebra yields

$$
E\left(\overline{u}_{i,0}\overline{u}_{i,m}\right)\leq c_0\rho_{u_i}^m.
$$

Thus, putting everything together

$$
\sum_{t=1}^{T}\sum_{s=1}^{T}\left|E\left(u_{i,t}u_{i,s}\right)\right|\leq c_0+c_1T+T\sum_{m=1}^{T}E\left(\overline{u}_{i,0}\overline{u}_{i,m}\right)\leq c_2T.
$$

We now turn to (A23)-(A24) and Assumption 2*(iv)*. Solving (A51) recursively, we have

$$
e_t=\varepsilon_t+\sum_{k=1}^{\infty}\left(\prod_{j=0}^{k-1}\left(\Phi_e+b_{t-j}^e\right)\right)\varepsilon_{t-k}; \tag{A79}
$$

defining now e_t' as in (A79), but with ε_0 replaced by an independent copy ε_0', it follows that, for all $p\leq 4$

$$
\begin{aligned}
\delta_{p,t}^e=\left|e_t-e_t'\right|_p &= \left|\prod_{j=0}^{t-1}\left(\Phi_e+b_{t-j}^e\right)\left(\varepsilon_0-\varepsilon_0'\right)\right|_p \\
&\leq E\left\|\prod_{j=0}^{t-1}\left(\Phi_e+b_{t-j}^e\right)\right\|^p\left|\varepsilon_0-\varepsilon_0'\right|_p \\
&\leq c_0\prod_{j=0}^{t-1}E\left\|\left(\Phi_e+b_{t-j}^e\right)\right\|^p=c_1\rho^t,
\end{aligned}
$$

where $\rho = E \left\| (\Phi_e + b_0^e) \right\|^p < 1$ by assumption. Henceforth, the proof follows exactly the same arguments as in the proof of Corollary 7. $\square$

Proof of Corollary 9. Under (A53)-(A54), if (A55)-(A60) hold and $|v_{i,0}|_2 < \infty$, Theorem 2.3 in Aue et al. (2006) states that there exists a unique stationary solution to (A54). In turn, this entails that there exists a measurable function $g_{u_i}(\cdot) : R^\infty \to R$ such that $u_{i,t} = g_{u_i}(v_{i,t}, ..., v_{i,0}, ...)$. Given the coupling $u'_{i,t} = g_{u_i}\left(v_{i,t}, ..., v'_{i,0}, ...\right)$, Liu and Lin (2009) prove (see the proof of their Corollary 3.3) there exists a $0 < \rho < 1$ such that $\delta_{t,2}^{u_i} = \left|u_{i,t} - u'_{i,t}\right|_2 = O(\rho^t)$. Using Corollary 5, (A20) obtains immediately. $\square$

Proof of Corollary 10. For all three models, it is possible (see e.g. Aue et al., 2009) to write $y_t = g(\varepsilon_t, \varepsilon_{t-1}, ...)$; thus, we show that the functional dependence measure $\delta_{t,4} = |y_t - y'_t|_4$ - where, as usual, $y'_t = g(\varepsilon_t, ..., \varepsilon'_0, ...)$ with ε'_0 an independent copy of ε_0 - is such that $\delta_{t,4} = O(\rho^t)$ for some $0 < \rho < 1$. Hence, Corollary 5 affords the desired result. We begin by noting that our assumptions entail that, for all the three models (A63)-(A67), $|e_t|_4 < \infty$ (see Aue et al., 2009 where this result is shown). We begin with (A63); the use of the Hadamard product entails that the functional dependence measure can be computed coordinate-wise, i.e. for each $1 \le j \le d$. Standard arguments (see also Aue et al., 2006) yield the unique stationary solution

$$h_{j,t}^2 = \omega_j + \omega_j \sum_{m=1}^{\infty} \sum_{1 \le l_1, ..., l_m \le \max\{p,q\}} \prod_{i=1}^{m} \left(\alpha_{j,l_i} + \beta_{j,l_i} \varepsilon_{j,t-(l_1+...+l_i)}^2\right), \tag{A80}$$

for each $1 \le j \le d$. Write now $t = \max\{p,q\} n + s$, where $n \ge 0$ and $0 \le s \le \max\{p,q\}$ are two integer numbers. Using the fact that $\left|e_{j,t} - e'_{j,t}\right|^4 \le \left|(e_{j,t})^2 - (e'_{j,t})^2\right|^2$, it follows that

$$\begin{aligned}
\left|e_{j,t} - e'_{j,t}\right|_4 &\le \left(E\left|(e_{j,t})^2 - (e'_{j,t})^2\right|^2\right)^{1/4} \\
&\le \left(E\left|\varepsilon_{j,t}^2\left(h_{j,t}^2 - (h'_{j,t})^2\right)\right|^2\right)^{1/4} \\
&= |\varepsilon_{j,0}|_4 \left|h_{j,t}^2 - (h'_{j,t})^2\right|_2^{1/2} \\
&\le c_0 \omega_j^{1/2} |\varepsilon_{j,0}|_4 \left(E\left|\sum_{m=n}^{\infty} \sum_{1 \le l_1, ..., l_m \le \max\{p,q\}} \prod_{i=1}^{m} \left(\alpha_{j,l_i} + \beta_{j,l_i} \varepsilon_{j,t-(l_1+...+l_i)}^2\right)\right|^2\right)^{1/4} \\
&\le c_0 \omega_j^{1/2} |\varepsilon_{j,0}|_4 \left(\sum_{m=n}^{\infty} \sum_{1 \le l_1, ..., l_m \le \max\{p,q\}} \prod_{i=1}^{m} \left|\alpha_{j,l_i} + \beta_{j,l_i} \varepsilon_{j,t-(l_1+...+l_i)}^2\right|_2\right)^{1/2} \\
&= c_0 \omega_j^{1/2} |\varepsilon_{j,0}|_4 \left(\sum_{m=n}^{\infty} \left(\sum_{l=1}^{\max\{p,q\}} \left|\alpha_{j,l} + \beta_{j,l} \varepsilon_{j,0}^2\right|_2\right)^m\right)^{1/2} \\
&\le c_0 \omega_j^{1/2} |\varepsilon_{j,0}|_4 \left(\sum_{m=n}^{\infty} \gamma_C^m\right)^{1/2} \le c_1 \gamma_C^{n/2},
\end{aligned}$$

having repeatedly used Minkowski's inequality. Recalling the definition of n, this ultimately entails that $\delta_{t,4} = O\left(\rho^t\right)$ for some $0 < \rho < 1$, which proves the desired result.

The arguments for (A65) are very similar, except the model needs to be studied as a whole as opposed to component-wise. Indeed, we have the stationary solution

$$h_t \odot h_t = \omega + \left[\sum_{m=1}^{\infty} \sum_{1 \leq l_1,...,l_m \leq \max\{p,q\}} \prod_{i=1}^{m} \left(A_{l_i} + B_{l_i} E_{t-(l_1+...+l_i)}\right)\right] \omega.$$

Letting as before $t = \max\{p, q\}\, n + s$

$$h_t \odot h_t - h_t' \odot h_t' = \sum_{m=n+1}^{t} \sum_{1 \leq l_1,...,l_m \leq \max\{p,q\}} \prod_{i=1}^{m} \left(A_{l_i} + B_{l_i} E_{t-(l_1+...+l_i)}\right) \omega.$$

We now have

$$\left|e_t - e_t'\right|_4 \leq \left|\varepsilon_0\right|_4 \left|h_t - h_t'\right|_4 \leq \sqrt{d}\left|\varepsilon_0\right|_4 \left|h_t \odot h_t - h_t' \odot h_t'\right|_2^{1/2},$$

having used equation (B.7) in Aue et al. (2009). Also

$$\begin{aligned}
\left|h_t \odot h_t - h_t' \odot h_t'\right|_2 &\leq c_0 \|\omega\| \sum_{m=n}^{\infty} \sum_{1 \leq l_1,...,l_m \leq \max\{p,q\}} \prod_{i=1}^{m} \left|A_{l_i} + B_{l_i} E_0\right|_2 \\
&= c_0 \|\omega\| \sum_{m=n}^{\infty} \left(\sum_{l=1}^{\max\{p,q\}} \left|A_l + B_l E_0\right|_2\right)^m \leq c_1 \gamma_J^n.
\end{aligned}$$

Since $\gamma_J < 1$, putting all together, we obtain that $\left|e_t - e_t'\right|_4 = O\left(\rho^t\right)$ for some $0 < \rho < 1$.

Finally, we consider (A67)-(A68); by Theorem 4.6 in Aue et al. (2009), (A69) and $\|A\| < 1$ entail that the unique, nonanticipative, stationary and ergodic solution of (A68) is

$$vech\left(\ln H_t - C\right) = \sum_{k=0}^{\infty} A^k F\left(\varepsilon_{t-(k+1)}, ..., \varepsilon_{t-(k+q)}\right).$$

Thus, defining the coupling H_t' in the usual way with ε_0', we have

$$vech\left(\ln H_t - C\right) - vech\left(\ln H_t' - C\right) = \sum_{j=1}^{q} A^{t-j} F\left(\varepsilon_{j-1}, ..., \varepsilon_{j-q}\right).$$

Letting $math(\cdot)$ be the inverse of $vech(\cdot)$, we can write

$$\begin{aligned}
H_t &= \exp\left(S_t\right), \\
H_t' &= \exp\left(S_t'\right),
\end{aligned}$$

where

$$S_t = C + math\left(\sum_{k=0}^{\infty} A^k F\left(\varepsilon_{t-(k+1)}, ..., \varepsilon_{t-(k+q)}\right)\right),$$
$$S_t' = C + math\left(\sum_{k=0}^{\infty} A^k F\left(\widetilde{\varepsilon}_{t-(k+1)}, ..., \widetilde{\varepsilon}_{t-(k+q)}\right)\right),$$

where $\widetilde{\varepsilon}_i = \varepsilon_i$ whenever $i \neq 0$ and $\widetilde{\varepsilon}_i = \varepsilon_i'$ for $i = 0$. Thus, for $\alpha, \beta > 1$ such that $\alpha^{-1} + \beta^{-1} = 1$, adapting equation (B.9) in Aue et al. (2009)we have

$$|e_t - e_t'|_4 \leq |\varepsilon_0|_4 \, |\|S_t - S_t'\||_{8\beta} \left|\exp\left(\frac{1}{2}\|S_t'\|\right)\right|_{4\alpha} \left|\exp\left(\frac{1}{2}\|S_t - S_t'\|\right)\right|_{8\beta}.$$

We know by assumption that $|\varepsilon_0|_4 < \infty$. Lemma B.2 in Aue et al. (2009), Minkowski's inequality and (A70) entail

$$\begin{aligned}
&|\|S_t - S_t'\||_{8\beta} \\
&\leq 2^{1/2} \left|vech\left(S_t - S_t'\right)\right|_{8\beta} \\
&\leq 2^{1/2} \left|\sum_{j=1}^{q} \|A\|^{t-j} \|F\left(\varepsilon_{j-1}, ..., \varepsilon_{j-q}\right)\|\right|_{8\beta} \\
&\leq 2^{1/2} \sum_{j=1}^{q} \|A\|^{t-j} \left|F\left(\varepsilon_{j-1}, ..., \varepsilon_{j-q}\right)\right|_{8\beta} \\
&\leq c_0 \|A\|^t .
\end{aligned}$$

Also, by the same arguments as in the proof of Theorem 4.7 in Aue et al. (2009), it holds that $\left|\exp\left(\frac{1}{2}\|S_t'\|\right)\right|_{4\alpha} <$

∞. Finally

$$
\begin{aligned}
& \left| \exp\left(\frac{1}{2}\left\| S_t - S_t' \right\|\right) \right|_{8\beta} \\
& \leq \left| \exp\left(2^{-1/2}\left\| vech\left(S_t - S_t'\right) \right\|\right) \right|_{8\beta} \\
& \leq \left(E \prod_{j=1}^{q} \exp\left(\frac{8\beta}{2^{1/2}} \left\| A \right\|^{t-j} \left\| F\left(\varepsilon_{j-1}, ..., \varepsilon_{j-q}\right) \right\| \right) \right)^{1/8\beta} \\
& \leq \prod_{j=1}^{q} \left(E \exp\left(\frac{8q\beta}{2^{1/2}} \left\| A \right\|^{t-j} \left\| F\left(\varepsilon_{j-1}, ..., \varepsilon_{j-q}\right) \right\| \right) \right)^{1/8q\beta} \\
& \leq \prod_{j=1}^{q} \left(1 + c_0 \left\| A \right\|^{t-j} E \left\| F\left(\varepsilon_{-1}, ..., \varepsilon_{-q}\right) \right\| \right)^{1/8q\beta} \\
& \leq \exp\left(c_0 \sum_{j=1}^{q} \left\| A \right\|^{t-j} E \left\| F\left(\varepsilon_{-1}, ..., \varepsilon_{-q}\right) \right\| \right) < \infty
\end{aligned}
$$

having used: Lemma B.2 in Aue et al. (2009), Holder's inequality and Lemma B.4 in Aue et al. (2009). Note that these calculations assume t "large enough"; otherwise we can set $H_t' = 0$. Putting all together, we obtain

$$
\left| e_t - e_t' \right|_4 \leq c_0 \left\| A \right\|^t ,
$$

which entails that the functional dependence measure $\delta_{t,4}$ is, even in this case, $O\left(\rho^t\right)$, $0 < \rho < 1$. □

A.4 Testing for trend stationarity versus unit root

When r_1 is found to be equal to 1, the question arises as to whether the corresponding factor is a trend-stationary series, or an $I(1)$ process with drift. Indeed, our methodology can only determine whether $r_1 = 0$ or 1, but it does not provide an answer to this question: the sum of the squares of the common factor, in both cases, diverge at a rate which is exactly $O\left(T^3\right)$.

If the common factor were observable, this issue could be tackled via a battery of "classical" tests (see e.g. Bierens, 1997). In our context, however, we can only estimate the space spanned by the common factors. This entails that the first common factor would be a weighted average of all the latent ones, thus containing by construction an $I(1)$ component even though the unobservable first common factor is a genuinely trend stationary series.

In order to propose a solution to this issue, suppose that we have found that one factor has a trend (i.e. $r_1 = 1$), and define the vector $f_t = \left(f_t^{(1)}, f_t^{(2)\prime}, f_t^{(3)\prime}\right)'$, where $f_t^{(1)}$, $f_t^{(2)}$ and $f_t^{(3)}$ are defined in Lemma 1 in the main paper. By construction, f_t is of dimension $k = 1 + r_2 + r_3$, where recall that r_2 is the number of $I(1)$ common factors with zero mean, and r_3 is the number of $I(0)$ common factors. We can write

$$
f_t = \mu t + \phi_t,
$$

where the vector μ is nonzero in its first element and zero elsewhere, and $\phi_t = \left(f_t^{(1)\dagger}, f_t^{(2)\prime}, f_t^{(3)\prime} \right)'$. By construction, ϕ_t has r_3 $I(0)$ components. There are now two alternatives

1. ϕ_t has only r_3 $I(0)$ components: this means that the first component of f_t is trend stationary;
2. ϕ_t has $r_3 + 1$ $I(0)$ components: this means that the first component of f_t is random walk with drift.

In order to find out which of the two alternatives we are in, if f_t is observable we could determine the rank of the cointegrated system f_t (say R) from the hypothesis testing problem

$$\begin{cases} H_0 : R = r_3 \\ H_A : R = r_3 + 1 \end{cases},$$

which is customarily carried out using the Likelihood Ratio tests developed in Johansen (1991). In our case, f_t is not observable; however, a possible approach would be to apply Johansen's procedure to $\widehat{f}_t$.

Another possible approach which is more in line with this paper (and could be viewed as related to variance ratio tests, see e.g. Cai and Shintani, 2006) is based on the following arguments. Note first that a rotation of μ can be estimated using

$$\widehat{\mu} = \left(\sum_{t=1}^{T} t^2 \right)^{-2} \sum_{t=1}^{T} t \widehat{f}_t.$$

Lemma A8. *We assume that Assumptions 2 and 4 hold, and that Assumptions A-D in Maciejowska (2010) are satisfied. Then it holds that*

$$E \left\| \widehat{\mu} - H'\mu \right\|^2 \leq c_0 T^{-1},$$

for a nonsingular matrix H.

Let now

$$\Delta\phi_t = \Delta f_t - \mu,$$

and assume that $\Delta\phi_t$ admits an $MA(\infty)$ representation, viz.

$$\Delta\phi_t = C(L) u_t, \tag{A81}$$

where $C(L) = \sum_{m=0}^{\infty} C_m L^m$ with $\sum_{m=0}^{\infty} m \|C_m\| < \infty$, and u_t is *i.i.d.*, with finite fourth moments and covariance matrix Σ_u.

By definition, the long-run variance matrix $\Sigma = C(1) \Sigma_u C(1)'$, has reduced rank. We can propose the following estimator for a rotation of Σ

$$\widehat{\Sigma} = \widehat{\gamma}_0 + \sum_{k=1}^{B} \left(1 - \frac{k}{B+1} \right) \left(\widehat{\gamma}_k + \widehat{\gamma}_k' \right), \tag{A82}$$

where

$$\widehat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^{T} x_t x_{t-k}',$$

having defined $x_t = \Delta \widehat{f}_t - \widehat{\mu}$. The following result is very similar to Theorem 1 in Ipatova and Trapani (2013), and states the consistency of $\widehat{\Sigma}$.

Lemma A9. *Under the assumptions of this paper and of Maciejowska (2010), it holds that*

$$\widehat{\Sigma} - H'\Sigma H = O\left(\frac{1}{B}\right) + o_{a.s.}\left(\frac{B}{N^{1/2}}(\ln N)^{1+\epsilon}\right) + o_{a.s.}\left(\frac{B}{T^{1/2}}(\ln T)^{1+\epsilon}\right),$$

for all $\epsilon > 0$.

We are now ready to discuss a methodology to test for

$$\begin{cases} H_0 : R = r_3 \\ H_A : R = r_3 + 1 \end{cases}.$$

Let $\lambda_{r_3+1}(H'\Sigma H)$ denote the (r_3+1)-th largest eigenvalue of Σ. On account of H being full rank, we have

$$\begin{aligned} \lambda_{r_3+1}(H'\Sigma H) &> 0 \text{ under } H_0, \\ \lambda_{r_3+1}(H'\Sigma H) &= 0 \text{ under } H_A. \end{aligned}$$

Upon choosing $B = \min\{T^{1/4}, N^{1/4}\}$, it holds that

$$\lambda_{r_3+1}\left(\widehat{\Sigma}\right) = \lambda_{r_3+1}(H'\Sigma H) + o_{a.s.}\left(\frac{\left(\ln\left(\max\{T^{1/4}, N^{1/4}\}\right)\right)^{1+\epsilon}}{\min\{T^{1/4}, N^{1/4}\}}\right), \tag{A83}$$

for all $\epsilon > 0$. Hence, it is possible to define

$$c_{N,T} = \left(\min\left\{T^{1/4}, N^{1/4}\right\}\right)^{1-\delta} \lambda_{r_3+1}\left(\widehat{\Sigma}\right),$$

where $0 < \delta < 1$ is user defined. Equation (A83) entails that

$$\begin{aligned} P\left\{\omega : \lim_{\min(N,T)\to\infty} c_{N,T} = \infty\right\} &= 1 \text{ under } H_0, \\ P\left\{\omega : \lim_{\min(N,T)\to\infty} c_{N,T} = 0\right\} &= 1 \text{ under } H_A. \end{aligned}$$

Henceforth, it is possible to construct a randomised test based on $c_{N,T}$, along exactly the same lines as in the main paper.

A.4.1 Proofs

Proof of Lemma A8. It holds that

$$\begin{aligned}\widehat{\mu} &= \left(\sum_{t=1}^{T} t^2\right)^{-2} \sum_{t=1}^{T} t\left(H'f_t + \widehat{f}_t - H'f_t\right) \\ &= H'\mu + \left(\sum_{t=1}^{T} t^2\right)^{-2} H' \sum_{t=1}^{T} t\phi_t + \left(\sum_{t=1}^{T} t^2\right)^{-2} \sum_{t=1}^{T} t\left(\widehat{f}_t - H'f_t\right).\end{aligned}$$

Assumption 4*(iii)* entails that $E\left\|\sum_{t=1}^{T} t\phi_t\right\|^2 \leq c_0 T^5$. Also

$$E\left\|\sum_{t=1}^{T} t\left(\widehat{f}_t - H'f_t\right)\right\|^2 \leq \left(\sum_{t=1}^{T} t^2\right) E\sum_{t=1}^{T}\left\|\widehat{f}_t - H'f_t\right\|^2 \leq c_0 T^4 \min\left\{T^{-1}, N^{-1}\right\}$$

where the last line follows from Proposition 3 in Maciejowska (2010). Putting all together, the desired result obtains. □

Proof of Lemma A9. We start by showing that

$$E\left\|\widehat{\gamma}_k - H'\gamma_k H\right\| \leq c_0, \tag{A84}$$

for all $0 \leq k \leq B$. It holds that

$$\begin{aligned}\widehat{\gamma}_k &= \frac{1}{T}\sum_{t=k+1}^{T} x_t x_{t-k}' \\ &= \frac{1}{T}\sum_{t=k+1}^{T} H'\Delta\phi_t\Delta\phi_{t-k}' H + \frac{1}{T}\sum_{t=k+1}^{T} H'\Delta\phi_t\left(\left(\Delta\widehat{f}_{t-k} - H'\Delta f_{t-k}\right) - \left(\widehat{\mu} - H'\mu\right)\right)' \\ &\quad + \frac{1}{T}\sum_{t=k+1}^{T}\left(\left(\Delta\widehat{f}_t - H'\Delta f_t\right) - \left(\widehat{\mu} - H'\mu\right)\right)\Delta\phi_{t-k}' H \\ &\quad + \frac{1}{T}\sum_{t=k+1}^{T}\left(\left(\Delta\widehat{f}_t - H'\Delta f_t\right) - \left(\widehat{\mu} - H'\mu\right)\right)\left(\left(\Delta\widehat{f}_{t-k} - H'\Delta f_{t-k}\right) - \left(\widehat{\mu} - H'\mu\right)\right)' \\ &= \frac{1}{T}\sum_{t=k+1}^{T} H'\Delta\phi_t\Delta\phi_{t-k}' H + I + II + III.\end{aligned}$$

Now

$$E\left\|I\right\| \leq E\left\|\frac{1}{T}\sum_{t=k+1}^{T} H'\Delta\psi_t\left(\Delta\widehat{f}_{t-k} - H'\Delta f_{t-k}\right)'\right\| + E\left\|\frac{1}{T}\sum_{t=k+1}^{T} H'\Delta\psi_t\left(\widehat{\mu} - H'\mu\right)'\right\| = I_a + I_b.$$

We have

$$\begin{aligned}
I_a &\leq \frac{1}{T}E\left\|\sum_{t=k+1}^{T} H'\Delta\psi_t\left(\Delta\widehat{f}_{t-k}-H'\Delta f_{t-k}\right)'\right\| \\
&\leq E\left(\frac{1}{T}\sum_{t=k+1}^{T}\left\|H'\Delta\psi_t\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=k+1}^{T}\left\|\Delta\widehat{f}_{t-k}-H'\Delta f'_{t-k}\right\|^2\right)^{1/2} \\
&\leq \left(\frac{1}{T}\sum_{t=k+1}^{T}E\left\|H'\Delta\psi_t\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=k+1}^{T}E\left\|\Delta\widehat{f}_{t-k}-H'\Delta f'_{t-k}\right\|^2\right)^{1/2} \\
&\leq c_0\min\left\{T^{-1/2},N^{-1/2}\right\},
\end{aligned}$$

having used Proposition 3 in Maciejowska (2010). Also

$$I_b \leq \left(\frac{1}{T}\sum_{t=k+1}^{T}E\left\|H'\Delta\psi_t\right\|^2\right)^{1/2}\left(E\left\|\widehat{\mu}-H'\mu\right\|^2\right)^{1/2}=O_p\left(T^{-1}\right),$$

after Lemma A8. Hence, $E\|I\|\leq c_0\min\left\{T^{-1/2},N^{-1/2}\right\}$. The same result holds for $E\|II\|$. Finally

$$\begin{aligned}
III &= \frac{1}{T}\sum_{t=k+1}^{T}\left(\Delta\widehat{f}_t-H'\Delta f_t\right)\left(\Delta\widehat{f}_{t-k}-H'\Delta f_{t-k}\right)'+\frac{1}{T}\sum_{t=k+1}^{T}\left(\widehat{\mu}-H'\mu\right)\left(\widehat{\mu}-H'\mu\right)' \\
&\quad +\frac{1}{T}\sum_{t=k+1}^{T}\left(\Delta\widehat{f}_t-H'\Delta f_t\right)\left(\widehat{\mu}-H'\mu\right)'+\frac{1}{T}\sum_{t=k+1}^{T}\left(\widehat{\mu}-H'\mu\right)\left(\Delta\widehat{f}_{t-k}-H'\Delta f_{t-k}\right)' \\
&= III_a+III_b+III_c+III_d.
\end{aligned}$$

We have

$$E\left\|III_a\right\|\leq\left(\frac{1}{T}\sum_{t=k+1}^{T}E\left\|\Delta\widehat{f}_t-H'\Delta f_t\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=k+1}^{T}E\left\|\Delta\widehat{f}_{t-k}-H'\Delta f_{t-k}\right\|^2\right)^{1/2}\leq c_0\min\left\{T^{-1},N^{-1}\right\}.$$

Lemma A8 immediately entails that and, by construction $E\|III_b\|\leq c_0T^{-1}$. Similarly

$$E\left\|III_c\right\|=\left(\frac{1}{T}\sum_{t=k+1}^{T}E\left\|\Delta\widehat{f}_t-H'\Delta f_t\right\|^2\right)^{1/2}\left(E\left\|\widehat{\mu}-H'\mu\right\|^2\right)^{1/2}\leq c_0T^{-1/2}\min\left\{T^{-1/2},N^{-1/2}\right\},$$

and the same holds for $E\|III_d\|$. Finally, it is easy to see that

$$E\left\|H'\left(\frac{1}{T}\sum_{t=k+1}^{T}\Delta\psi_t\Delta\psi_{t-k}-\gamma_k\right)H\right\|\leq c_0T^{-1/2}.$$

Putting all together, it follows that

$$E\left\|\widehat{\gamma}_k - H'\gamma_k H\right\| \leq c_0 T^{-1/2}\min\left\{T^{-1/2}, N^{-1/2}\right\}.$$

Now the final result obtains from the proof of Theorem 1 in Ipatova and Trapani (2013). □

A.5 Weak factors: discussion

By Assumption 4, all the common factors are assumed to be strongly pervasive. This is a direct consequence of having $\|\Lambda\|^2 = O(N)$. It is however possible to imagine a situation in which some of the common factors are "weak", or "less pervasive": this can arise from e.g. having genuinely weak factors, or from having strong factors which impact only on a small number of units - see, for example, Onatski (2012) and the references therein.

In this section, we report some heuristic arguments (similar to Trapani, 2018), on the ability of our procedure to determine weak factors. For the sake of a concise discussion, but with no loss of generality, we consider the case where all r factors are zero-mean $I(1)$, and $\Lambda'\Lambda$ is diagonal, with diagonal elements $c_p(N)$ given by

$$c_p(N) = \left\{\begin{array}{ll} N & \text{for } 1 \leq p \leq p' \\ N^{1-\kappa_p} & \text{for } p' < p \leq r \end{array}\right. .$$

Allowing for $\kappa_p \in (0,1)$ corresponds to the case of having weak factors, and the larger κ_p the weaker the corresponding factor. Suppose that the researcher is using Σ_2 and its eigenvalues $\nu_2^{(p)}$ in order to determine r. Repeating exactly the same arguments in the proof of Theorem 1, it can be shown that

$$\nu_2^{(p)} \geq C_0 \frac{c_p(N)}{\ln\ln T}. \tag{A85}$$

Equation (A85) entails that, whenever $p' < p \leq r$,

$$\nu_2^{(p)} \geq C_0 \frac{N^{1-\kappa_p}}{\ln\ln T}. \tag{A86}$$

Recall that, our procedure, essentially, is based on testing whether, as $\min(N,T) \to \infty$

$$\left\{\begin{array}{l} H_{0,2}^{(p)} : (\ln\ln T)N^{-\delta}\nu_2^{(p)} \to \infty \\ H_{A,2}^{(p)} : (\ln\ln T)N^{-\delta}\nu_2^{(p)} \to 0 \end{array}\right. ,$$

with δ selected as per (35). Thus, based on (A86), weak factors can be determined if

$$\lim_{\min(N,T)\to\infty} N^{1-\kappa_p-\delta} \to \infty,$$

which requires

$$\kappa_p < 1 - \delta. \tag{A87}$$

On the grounds of (35), the constraint in (A87) explains up to which extent weak factors can be detected. When $\beta \leq \frac{1}{2}$, that is $\frac{N}{\sqrt{T}} = O(1)$, then $\delta = 0$, and we need $\kappa_p < 1$. This entails that, when N is much smaller than T, our procedure is able to detect even very weak factors. Conversely, when $\beta > \frac{1}{2}$, that is $\frac{\sqrt{T}}{N} = o(1)$, it is required that $\kappa_p < 1 - \frac{1}{2\beta}$: as β increases, i.e. N increases, the test is less and less able to detect weak factors. Note that when N and T have the same order of magnitude, and thus $\beta = 1$, weak factors can be detected as long as $\kappa_p < \frac{1}{2}$ - that is, when the eigenvalues associated with that factor diverge to infinity a bit faster than $\sqrt{N}$.

References

Aue, A., I. Berkes, L. Horváth, et al. (2006). Strong approximation for the sums of squares of augmented garch sequences. *Bernoulli 12*(4), 583–608.

Aue, A., S. Hörmann, L. Horváth, and M. Reimherr (2009). Break detection in the covariance structure of multivariate time series models. *Annals of Statistics 37*, 4046–4087.

Aue, A., L. Horvath, and J. Steinebach (2006). Estimation in random coefficient autoregressive models. *Journal of Time Series Analysis 27*(1), 61–76.

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics 122*, 137–183.

Berkes, I. and W. Philipp (1979). Approximation thorems for independent and weakly dependent random vectors. *The Annals of Probability*, 29–54.

Bierens, H. J. (1997). Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the us price level and interest rate. *Journal of Econometrics 81*(1), 29–64.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, 498–505.

Bushell, P. and G. Trustrum (1990). Trace inequalities for positive definite matrix power products. *Linear Algebra and its Applications 132*, 173–178.

Cai, G. H. (2006). Chover-type laws of the iterated logarithm for weighted sums of ρ-mixing sequences. *International Journal of Stochastic Analysis 2006*.

Cai, Y. and M. Shintani (2006). On the alternative long-run variance ratio test for a unit root. *Econometric Theory*, 347–372.

Carrasco, M. and X. Chen (2002). Mixing and moment properties of various garch and stochastic volatility models. *Econometric Theory*, 17–39.

Corradi, V. (1999). Deciding between i (0) and i (1) via flil-based bounds. *Econometric Theory*, 643–663.

Davidson, J. (1994). *Stochastic limit theory: An introduction for econometricians*. OUP Oxford.

Donsker, M. and S. Varadhan (1977). On laws of the iterated logarithm for local times. *Communications on Pure and Applied Mathematics 30*(6), 707–753.

Duan, J.-C. (1997). Augmented garch (p, q) process and its diffusion limit. *Journal of Econometrics 79*(1), 97–127.

Eberlein, E. (1986). On strong invariance principles under dependence assumptions. *The Annals of Probability*, 260–270.

Horn, R. A. and C. R. Johnson (2012). *Matrix analysis*. Cambridge university press.

Horváth, L. and L. Trapani (2019). Testing for randomness in a random coefficient autoregression model. *Journal of Econometrics 209*(2), 338–352.

Ipatova, E. and L. Trapani (2013). First-differenced inference for panel factor series. *Economics Letters 118*(2), 364–366.

Jeantheau, T. (1998). Strong consistency of estimators for multivariate arch models. *Econometric Theory*, 70–86.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 1551–1580.

Kawakatsu, H. (2006). Matrix exponential garch. *Journal of Econometrics 134*(1), 95–128.

Liu, W. and Z. Lin (2009). Strong approximation for a class of stationary processes. *Stochastic Processes and their Applications 119*(1), 249–280.

Maciejowska, K. (2010). Common factors in nonstationary panel data with a deterministic trend–estimation and distribution theory.

Merikoski, J. K. and R. Kumar (2004). Inequalities for spreads of matrix sums and products. *Applied Mathematics E-Notes 4*, 150–159.

Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics 168*, 244–258.

Peligrad, M., S. Utev, and W. Wu (2007). A maximal l_p-inequality for stationary sequences and its applications. *Proceedings of the American Mathematical Society 135*(2), 541–550.

Phillips, P. C. and V. Solo (1992). Asymptotics for linear processes. *The Annals of Statistics*, 971–1001.

Serfling, R. J. (1970). Moment inequalities for the maximum cumulative sum. *The Annals of Mathematical Statistics*, 1227–1234.

Trapani, L. (2018). A randomized sequential procedure to determine the number of factors. *Journal of the American Statistical Association 113*(523), 1341–1349.

Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences 102*(40), 14150–14154.

B Additional numerical results

We report further Monte Carlo evidence under different specifications of the tests. In all cases, when estimating r_2, we use r_1; results using $\widehat{r}_1$ are are similar and available upon request.

B.1 Bai Information Criterion for estimating r_2 when r_1 is known

As a complement to Tables 2 and 4 in the main paper, we report there the result obtained with the Information Criteria by Bai (2004), denoted as IC - this corresponds to $IC3$ in the original paper; we note that the other criteria, known as IC_1 and IC_2, deliver a similar (or worse) performance and are therefore not reported.

Table B1: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$

		$N = 50, T = 100$			$N = 100, T = 100$		
r_2	r_3	average	std. dev.	% correct	average	std. dev.	% correct
0	0	1.00	0.00	0.00	1.00	0.00	0.00
0	1	1.00	0.00	0.00	1.00	0.00	0.00
0	2	1.00	0.00	0.00	1.00	0.00	0.00
1	0	1.00	0.00	1.00	1.00	0.00	1.00
1	1	1.00	0.00	1.00	1.00	0.00	1.00
1	2	1.00	0.00	1.00	1.00	0.00	1.00
2	0	2.00	0.00	1.00	2.00	0.00	1.00
2	1	2.00	0.00	1.00	2.00	0.06	1.00
2	2	2.00	0.00	1.00	2.00	0.00	1.00
		$N = 200, T = 100$			$N = 100, T = 200$		
r_2	r_3	average	std. dev.	% correct	average	std. dev.	% correct
0	0	1.00	0.00	0.00	1.00	0.00	0.00
0	1	1.00	0.00	0.00	1.00	0.00	0.00
0	2	1.00	0.00	0.00	1.00	0.00	0.00
1	0	1.00	0.00	1.00	1.00	0.00	1.00
1	1	1.00	0.00	1.00	1.00	0.00	1.00
1	2	1.00	0.00	1.00	1.00	0.00	1.00
2	0	2.00	0.00	1.00	2.00	0.00	1.00
2	1	2.00	0.00	1.00	2.00	0.04	1.00
2	2	2.00	0.00	1.00	2.00	0.00	1.00
		$N = 200, T = 200$			$N = 200, T = 500$		
r_2	r_3	average	std. dev.	% correct	average	std. dev.	% correct
0	0	1.00	0.00	0.00	1.00	0.00	0.00
0	1	1.00	0.00	0.00	1.00	0.00	0.00
0	2	1.00	0.00	0.00	1.00	0.00	0.00
1	0	1.00	0.00	1.00	1.00	0.00	1.00
1	1	1.00	0.00	1.00	1.00	0.00	1.00
1	2	1.00	0.00	1.00	1.00	0.00	1.00
2	0	2.00	0.00	1.00	2.00	0.00	1.00
2	1	2.00	0.00	1.00	2.00	0.00	1.00
2	2	2.00	0.00	1.00	2.00	0.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$. Same configuration as tables in main text.

Table B2: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$

		$N=50$, $T=100$			$N=100$, $T=100$		
r_2	r_3	average	std. dev.	% correct	average	std. dev.	% correct
0	0	0.00	0.00	1.00	0.00	0.00	1.00
0	1	0.00	0.00	1.00	0.00	0.00	1.00
0	2	0.00	0.00	1.00	0.00	0.00	1.00
1	0	1.00	0.00	1.00	1.00	0.00	1.00
1	1	1.00	0.00	1.00	1.00	0.00	1.00
1	2	1.00	0.00	1.00	1.00	0.00	1.00
2	0	1.97	0.16	0.97	1.99	0.11	0.99
2	1	1.85	0.35	0.85	1.94	0.24	0.94
2	2	1.81	0.39	0.81	1.89	0.32	0.89
		$N=200$, $T=100$			$N=100$, $T=200$		
r_2	r_3	average	std. dev.	% correct	average	std. dev.	% correct
0	0	0.00	0.00	1.00	0.00	0.00	1.00
0	1	0.00	0.00	1.00	0.00	0.00	1.00
0	2	0.00	0.00	1.00	0.00	0.00	1.00
1	0	1.00	0.00	1.00	1.00	0.00	1.00
1	1	1.00	0.00	1.00	1.00	0.00	1.00
1	2	1.00	0.00	1.00	1.00	0.00	1.00
2	0	1.99	0.09	0.99	2.00	0.06	1.00
2	1	1.95	0.21	0.95	1.98	0.14	0.98
2	2	1.95	0.23	0.95	1.96	0.21	0.96
		$N=200$, $T=200$			$N=200$, $T=500$		
r_2	r_3	average	std. dev.	% correct	average	std. dev.	% correct
0	0	0.00	0.00	1.00	0.00	0.00	1.00
0	1	0.00	0.00	1.00	0.00	0.00	1.00
0	2	0.00	0.00	1.00	0.00	0.00	1.00
1	0	1.00	0.00	1.00	1.00	0.00	1.00
1	1	1.00	0.00	1.00	1.00	0.00	1.00
1	2	1.00	0.00	1.00	1.00	0.00	1.00
2	0	2.00	0.00	1.00	2.00	0.00	1.00
2	1	1.99	0.08	0.99	2.00	0.04	1.00
2	2	2.00	0.06	1.00	2.00	0.04	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$. Same DGP as for the tables in the main text.

B.2 Cases $\bar{\rho} = 0$ and $\bar{\rho} = 0.8$

We report results obtained when, in (32), $\rho_j \sim U[\bar{\rho}, 0.8]$, with $\bar{\rho} = 0$ and $\bar{\rho} = 0.8$. As an overall comment on the impact of $\bar{\rho}$, results are in general unaffected but for two cases. The first case is when $r_1 = 0$ and $r_2 = 1$ and we compute $\widehat{r}_1$; in this case, we find that lower values of $\bar{\rho}$ improve the results. Conversely, higher values of $\bar{\rho}$ make the innovations of the zero-mean $I(1)$ factors more persistent, thus making the associated eigenvalues larger: in this case, we are therefore more likely to falsely detect trends. The second case arises when $r_2 = 2$, and we compute $\widehat{r}_2$. In this case, we find the exact opposite. This can be explained upon noting that, for lower values of $\bar{\rho}$, the two $I(1)$ factors become closer to two pure random walks which are highly collinear, thus making the second eigenvalue $\nu_2^{(2)}$ much smaller than the first one $\nu_2^{(1)}$: thus, in this case, we are less likely to detect the second factor. For the same reason, higher values of $\bar{\rho}$ make the two factors less collinear, so that then the second factor is detected more easily.

B.2.1 Case $\bar{\rho} = 0$

Table B3: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.10	0.09	0.30	0.29	0.90	0.91	0.03	0.04	0.18	0.19	0.97	0.96
0	2	0.01	0.00	0.08	0.04	0.99	1.00	0.00	0.00	0.04	0.04	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.06	0.00	1.00	1.00

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.01	0.19	0.09	0.39	0.99	0.81	0.01	0.01	0.10	0.12	0.99	0.99
0	2	0.00	0.01	0.00	0.09	1.00	0.99	0.00	0.00	0.00	0.04	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.09	0.04	0.29	1.00	0.91	0.00	0.05	0.00	0.21	1.00	0.95
0	2	0.00	0.00	0.00	0.06	1.00	1.00	0.00	0.00	0.00	0.04	1.00	1.00
1	0	1.00	1.00	0.06	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B4: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	0.99	0.06	0.09	1.00	0.99
2	0	1.98	1.97	0.16	0.16	0.98	0.97	1.94	1.94	0.25	0.23	0.94	0.94
2	1	1.89	1.87	0.32	0.34	0.89	0.88	1.74	1.73	0.44	0.44	0.74	0.73
2	2	1.91	1.91	0.31	0.30	0.91	0.91	1.72	1.74	0.46	0.44	0.73	0.74

		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.06	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.99	1.00	0.11	0.04	0.99	1.00	1.00	1.00	0.04	0.00	1.00	1.00
2	0	1.97	1.98	0.18	0.13	0.97	0.98	1.97	1.97	0.18	0.16	0.97	0.97
2	1	1.81	1.94	0.41	0.25	0.82	0.94	1.92	1.93	0.27	0.26	0.93	0.93
2	2	1.76	1.89	0.43	0.31	0.76	0.89	1.94	1.93	0.24	0.26	0.94	0.93

		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	1.99	2.00	0.08	0.06	0.99	1.00	2.00	2.00	0.04	0.00	1.00	1.00
2	1	1.95	1.99	0.21	0.09	0.95	0.99	1.98	1.99	0.14	0.09	0.98	0.99
2	2	1.94	1.98	0.24	0.15	0.94	0.98	1.99	2.00	0.09	0.04	0.99	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B5: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.08	0.12	0.99	0.99	0.02	0.03	0.17	0.17	0.98	0.97
0	1	0.02	0.02	0.14	0.13	0.98	0.98	0.04	0.05	0.22	0.22	0.96	0.95
0	2	0.01	0.01	0.08	0.10	0.99	0.99	0.03	0.04	0.18	0.20	0.97	0.96
1	0	1.00	0.98	0.10	0.14	0.99	0.99	1.00	1.01	0.15	0.14	0.98	0.98
1	1	0.98	0.98	0.14	0.22	0.98	0.96	1.03	1.01	0.22	0.22	0.95	0.96
1	2	0.99	1.00	0.15	0.17	0.98	0.97	1.00	1.02	0.24	0.19	0.95	0.96
2	0	1.67	1.65	0.50	0.52	0.68	0.65	1.79	1.79	0.44	0.47	0.77	0.76
2	1	1.48	1.46	0.57	0.58	0.49	0.48	1.60	1.59	0.55	0.59	0.60	0.58
2	2	1.43	1.42	0.59	0.57	0.46	0.45	1.60	1.57	0.56	0.56	0.59	0.56
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.04	1.00	1.00	0.02	0.01	0.17	0.11	0.98	0.99
0	1	0.00	0.07	0.06	0.25	1.00	0.93	0.04	0.03	0.21	0.18	0.96	0.97
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.02	0.04	0.16	0.21	0.97	0.96
1	0	1.00	1.00	0.08	0.04	0.99	1.00	1.04	1.03	0.21	0.20	0.95	0.96
1	1	1.00	1.02	0.08	0.15	0.99	0.98	1.03	1.03	0.19	0.19	0.97	0.97
1	2	0.99	1.00	0.12	0.09	0.99	1.00	1.02	1.03	0.19	0.16	0.97	0.97
2	0	1.97	1.99	0.18	0.11	0.97	0.99	1.93	1.93	0.31	0.33	0.90	0.90
2	1	1.83	1.93	0.40	0.29	0.81	0.91	1.86	1.86	0.44	0.41	0.80	0.81
2	2	1.77	1.89	0.44	0.32	0.78	0.89	1.84	1.82	0.41	0.49	0.82	0.76
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.01	0.06	0.08	0.24	0.99	0.94	0.00	0.02	0.06	0.15	1.00	0.98
0	2	0.00	0.01	0.06	0.10	1.00	0.99	0.00	0.00	0.04	0.00	1.00	1.00
1	0	1.00	1.00	0.10	0.09	1.00	1.00	1.00	1.00	0.00	0.06	1.00	1.00
1	1	1.01	1.02	0.13	0.13	0.99	0.98	1.00	1.01	0.00	0.08	1.00	0.99
1	2	1.00	1.00	0.09	0.00	1.00	1.00	1.00	1.00	0.00	0.06	1.00	1.00
2	0	2.00	2.00	0.06	0.08	1.00	0.99	2.00	2.00	0.09	0.06	1.00	1.00
2	1	1.97	2.01	0.22	0.18	0.95	0.97	1.99	2.00	0.09	0.08	0.99	0.99
2	2	1.96	1.98	0.21	0.14	0.96	0.98	1.99	1.99	0.15	0.14	0.99	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.2.2 Case $\bar{\rho} = 0.8$

Table B6: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.44	0.41	0.50	0.49	0.56	0.59	0.30	0.30	0.46	0.46	0.70	0.70
0	2	0.26	0.25	0.44	0.43	0.74	0.75	0.13	0.13	0.34	0.33	0.87	0.87
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.06	0.00	1.00	1.00

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.21	0.59	0.41	0.49	0.79	0.41	0.24	0.24	0.43	0.43	0.76	0.76
0	2	0.08	0.31	0.26	0.46	0.92	0.69	0.07	0.07	0.26	0.25	0.93	0.93
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.06	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.13	0.44	0.33	0.50	0.87	0.56	0.07	0.36	0.25	0.48	0.93	0.64
0	2	0.03	0.19	0.17	0.40	0.97	0.81	0.01	0.07	0.10	0.26	0.99	0.93
1	0	1.00	1.00	0.06	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B7: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.09	0.04	1.00	1.00	2.00	2.00	0.09	0.00	1.00	1.00
2	1	2.00	1.99	0.06	0.10	1.00	1.00	1.99	1.99	0.09	0.11	0.99	0.99
2	2	1.99	1.99	0.10	0.10	1.00	1.00	1.99	1.99	0.13	0.10	0.99	0.99
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.04	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.04	0.00	1.00	1.00
2	0	2.00	2.00	0.09	0.04	1.00	1.00	1.99	2.00	0.11	0.06	0.99	1.00
2	1	1.99	1.99	0.12	0.12	0.99	0.99	1.99	2.00	0.10	0.04	1.00	1.00
2	2	1.99	2.00	0.08	0.06	0.99	1.00	2.00	2.00	0.00	0.00	1.00	1.00
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.00	0.04	1.00	1.00	2.00	2.00	0.04	0.00	1.00	1.00
2	1	2.00	2.00	0.04	0.00	1.00	1.00	2.00	2.00	0.00	0.00	1.00	1.00
2	2	2.00	2.00	0.00	0.00	1.00	1.00	2.00	2.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B8: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.08	0.12	0.99	0.99	0.02	0.03	0.17	0.17	0.98	0.97
0	1	0.02	0.02	0.14	0.13	0.98	0.98	0.04	0.05	0.22	0.22	0.96	0.95
0	2	0.01	0.01	0.08	0.10	0.99	0.99	0.03	0.04	0.18	0.20	0.97	0.96
1	0	1.00	0.99	0.08	0.10	0.99	1.00	1.01	1.01	0.12	0.12	0.99	0.99
1	1	1.00	1.01	0.06	0.16	1.00	0.98	1.04	1.03	0.20	0.18	0.96	0.97
1	2	1.01	1.01	0.09	0.12	0.99	0.99	1.02	1.03	0.20	0.17	0.97	0.97
2	0	1.95	1.96	0.25	0.27	0.94	0.94	1.97	1.98	0.23	0.25	0.94	0.94
2	1	1.88	1.90	0.36	0.34	0.86	0.88	1.95	1.95	0.32	0.33	0.90	0.91
2	2	1.89	1.87	0.36	0.36	0.88	0.87	1.92	1.92	0.33	0.34	0.89	0.88
		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.04	1.00	1.00	0.02	0.01	0.17	0.11	0.98	0.99
0	1	0.00	0.07	0.06	0.25	1.00	0.93	0.04	0.03	0.21	0.18	0.96	0.97
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.02	0.04	0.16	0.21	0.97	0.96
1	0	1.00	1.00	0.08	0.04	0.99	1.00	1.04	1.03	0.20	0.20	0.96	0.96
1	1	1.00	1.02	0.06	0.15	1.00	0.98	1.03	1.03	0.19	0.19	0.97	0.97
1	2	1.00	1.00	0.10	0.09	1.00	1.00	1.02	1.03	0.19	0.16	0.97	0.97
2	0	2.00	2.00	0.06	0.06	1.00	1.00	2.01	2.01	0.15	0.19	0.98	0.98
2	1	1.99	2.00	0.17	0.13	0.97	0.98	2.03	2.03	0.21	0.19	0.96	0.96
2	2	1.96	1.99	0.23	0.13	0.97	0.98	2.00	2.02	0.17	0.28	0.97	0.95
		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.01	0.06	0.08	0.24	0.99	0.94	0.00	0.02	0.06	0.15	1.00	0.98
0	2	0.00	0.01	0.06	0.10	1.00	0.99	0.00	0.00	0.04	0.00	1.00	1.00
1	0	1.00	1.00	0.10	0.09	1.00	1.00	1.00	1.00	0.00	0.06	1.00	1.00
1	1	1.01	1.02	0.13	0.13	0.99	0.98	1.00	1.01	0.00	0.08	1.00	0.99
1	2	1.00	1.00	0.09	0.00	1.00	1.00	1.00	1.00	0.00	0.06	1.00	1.00
2	0	2.00	2.00	0.00	0.06	1.00	1.00	2.00	2.00	0.09	0.06	1.00	1.00
2	1	2.01	2.02	0.09	0.15	0.99	0.98	2.00	2.00	0.00	0.04	1.00	1.00
2	2	2.00	2.00	0.11	0.04	0.99	1.00	1.99	1.99	0.13	0.13	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.3 Sensitivity to $F_1(u)$, $F_2(u)$, $G_1(u)$ and $G_2(u)$

We consider different, alternative choices of $F_1(u)$, $F_2(u)$, $G_1(u)$ and $G_2(u)$. In particular, as far as $F_1(u)$ and $F_2(u)$ are concerned, we use $u = \sqrt{(a)}$ and $u = -\sqrt{(a)}$ with equal weights, using $a = 2$ and $a = 25$. In the latter case, the theory predicts that tests will have higher power (at the expense of size distortion), thus yielding understatements of r_1 and r_2 for small N and T.

B.3.1 $F_1(u)$, $F_2(u)$ uniform on $\{-\sqrt{2}, \sqrt{2}\}$

Table B9: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.12	0.11	0.33	0.32	0.88	0.89	0.07	0.07	0.26	0.26	0.93	0.93
0	2	0.01	0.00	0.11	0.06	0.99	1.00	0.00	0.01	0.06	0.08	1.00	0.99
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.06	0.04	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00

		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.03	0.27	0.16	0.44	0.97	0.73	0.03	0.03	0.18	0.17	0.97	0.97
0	2	0.00	0.03	0.04	0.16	1.00	0.97	0.01	0.01	0.08	0.09	0.99	0.99
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.06	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.99	1.00	0.11	0.00	0.99	1.00	1.00	1.00	0.00	0.04	1.00	1.00

		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.02	0.15	0.13	0.36	0.98	0.85	0.00	0.10	0.00	0.30	1.00	0.90
0	2	0.00	0.01	0.06	0.09	1.00	0.99	0.00	0.01	0.00	0.09	1.00	0.99
1	0	1.00	1.00	0.06	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B10: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00
2	0	1.98	1.98	0.15	0.14	0.98	0.98	1.96	1.96	0.21	0.19	0.96	0.96
2	1	1.91	1.91	0.30	0.29	0.91	0.91	1.86	1.86	0.35	0.35	0.86	0.86
2	2	1.93	1.93	0.26	0.26	0.93	0.94	1.86	1.86	0.35	0.34	0.87	0.86

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.06	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.04	0.00	1.00	1.00
2	0	1.98	1.99	0.14	0.12	0.99	0.99	1.98	1.99	0.15	0.12	0.98	0.99
2	1	1.90	1.97	0.31	0.19	0.90	0.97	1.97	1.98	0.17	0.15	0.98	0.98
2	2	1.88	1.96	0.32	0.21	0.88	0.96	1.97	1.97	0.17	0.17	0.97	0.97

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.04	0.06	1.00	1.00	2.00	2.00	0.04	0.00	1.00	1.00
2	1	1.99	2.00	0.12	0.06	0.99	1.00	1.99	2.00	0.11	0.04	0.99	1.00
2	2	1.99	2.00	0.11	0.04	0.99	1.00	2.00	2.00	0.04	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B11: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.01	0.04	0.08	1.00	0.99	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.06	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
1	0	0.99	0.98	0.11	0.15	0.99	0.98	0.98	0.99	0.13	0.12	0.98	0.99
1	1	0.96	0.96	0.19	0.21	0.96	0.96	0.97	0.98	0.18	0.15	0.97	0.98
1	2	0.97	0.95	0.17	0.21	0.97	0.95	0.95	0.98	0.24	0.14	0.95	0.98
2	0	1.49	1.46	0.53	0.54	0.51	0.48	1.55	1.54	0.51	0.51	0.56	0.55
2	1	1.26	1.25	0.54	0.55	0.31	0.31	1.31	1.30	0.52	0.54	0.34	0.34
2	2	1.21	1.22	0.52	0.54	0.26	0.28	1.32	1.27	0.54	0.53	0.36	0.31
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.04	0.04	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	0.99	0.99	0.08	0.11	0.99	0.99
1	1	1.00	1.00	0.04	0.06	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.99	1.00	0.11	0.09	0.99	1.00	1.00	1.00	0.09	0.00	1.00	1.00
2	0	1.96	1.98	0.19	0.13	0.96	0.98	1.84	1.83	0.37	0.39	0.84	0.83
2	1	1.75	1.89	0.43	0.31	0.75	0.89	1.73	1.70	0.45	0.46	0.73	0.70
2	2	1.70	1.84	0.47	0.36	0.71	0.84	1.65	1.68	0.48	0.48	0.65	0.68
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.09	0.09	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	1	1.00	1.00	0.09	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.09	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
2	0	2.00	1.99	0.06	0.08	1.00	0.99	2.00	2.00	0.09	0.00	1.00	1.00
2	1	1.96	1.99	0.19	0.11	0.96	0.99	2.00	2.00	0.06	0.04	1.00	1.00
2	2	1.95	1.97	0.23	0.16	0.95	0.97	1.99	1.99	0.15	0.14	0.99	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.3.2 $F_1(u)$, $F_2(u)$ uniform on $\{-5, 5\}$

Table B12: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.02	0.02	0.13	0.13	0.98	0.98	0.01	0.01	0.09	0.09	0.99	0.99
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.04	0.04	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.04	0.00	1.00	1.00
1	1	0.93	0.92	0.26	0.27	0.93	0.92	0.82	0.81	0.38	0.39	0.82	0.81
1	2	0.90	0.91	0.29	0.29	0.90	0.91	0.84	0.84	0.36	0.37	0.84	0.84

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.15	0.06	0.35	1.00	0.85	0.00	0.00	0.06	0.06	1.00	1.00
0	2	0.00	0.01	0.00	0.08	1.00	0.99	0.00	0.00	0.00	0.00	1.00	1.00
1	0	0.99	1.00	0.09	0.00	0.99	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.69	1.00	0.46	0.04	0.69	1.00	1.00	1.00	0.06	0.06	1.00	1.00
1	2	0.66	1.00	0.48	0.06	0.66	1.00	1.00	1.00	0.04	0.06	1.00	1.00

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.08	0.00	0.26	1.00	0.92	0.00	0.02	0.00	0.14	1.00	0.98
0	2	0.00	0.00	0.00	0.06	1.00	1.00	0.00	0.00	0.00	0.04	1.00	1.00
1	0	1.00	1.00	0.06	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.99	1.00	0.08	0.00	0.99	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B13: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.06	0.04	1.00	1.00	0.98	0.98	0.13	0.13	0.98	0.98
1	2	0.99	0.99	0.12	0.12	0.99	0.99	0.99	0.99	0.12	0.12	0.99	0.99
2	0	1.90	1.89	0.31	0.31	0.90	0.89	1.83	1.82	0.39	0.39	0.83	0.82
2	1	1.55	1.54	0.51	0.51	0.56	0.55	1.52	1.51	0.52	0.52	0.53	0.52
2	2	1.61	1.61	0.50	0.50	0.61	0.61	1.49	1.49	0.52	0.52	0.50	0.50
		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.97	1.00	0.16	0.04	0.97	1.00	1.00	1.00	0.04	0.00	1.00	1.00
1	2	0.99	1.00	0.12	0.04	0.99	1.00	1.00	1.00	0.06	0.06	1.00	1.00
2	0	1.95	1.98	0.22	0.15	0.95	0.98	1.92	1.93	0.27	0.26	0.93	0.93
2	1	1.71	1.88	0.47	0.34	0.72	0.88	1.81	1.82	0.41	0.39	0.81	0.82
2	2	1.66	1.79	0.49	0.41	0.67	0.79	1.80	1.81	0.40	0.40	0.80	0.81
		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	1.99	1.99	0.09	0.10	0.99	0.99	2.00	2.00	0.04	0.00	1.00	1.00
2	1	1.90	1.98	0.31	0.13	0.90	0.98	1.98	1.99	0.15	0.11	0.98	0.99
2	2	1.87	1.96	0.34	0.20	0.87	0.96	1.98	1.99	0.14	0.09	0.98	0.99

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B14: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.01	0.04	0.08	1.00	0.99	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.06	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
1	0	0.81	0.79	0.39	0.41	0.81	0.79	0.84	0.84	0.37	0.37	0.84	0.84
1	1	0.69	0.69	0.46	0.47	0.69	0.69	0.71	0.71	0.45	0.45	0.71	0.71
1	2	0.70	0.69	0.46	0.46	0.70	0.69	0.67	0.68	0.48	0.47	0.67	0.68
2	0	0.81	0.78	0.49	0.50	0.04	0.04	0.82	0.81	0.49	0.48	0.04	0.04
2	1	0.59	0.58	0.50	0.50	0.01	0.01	0.58	0.58	0.51	0.51	0.01	0.01
2	2	0.54	0.53	0.52	0.51	0.01	0.01	0.60	0.57	0.50	0.51	0.01	0.01

		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.04	0.04	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
1	0	0.98	1.00	0.15	0.06	0.98	1.00	0.96	0.95	0.19	0.23	0.96	0.95
1	1	0.86	0.98	0.34	0.14	0.86	0.98	0.90	0.89	0.31	0.32	0.90	0.89
1	2	0.87	0.95	0.35	0.22	0.87	0.95	0.90	0.90	0.31	0.31	0.90	0.90
2	0	1.63	1.82	0.52	0.38	0.65	0.82	1.20	1.17	0.53	0.54	0.26	0.25
2	1	1.05	1.38	0.62	0.55	0.22	0.42	1.00	0.97	0.52	0.50	0.13	0.11
2	2	1.04	1.24	0.59	0.58	0.19	0.32	0.95	0.95	0.50	0.50	0.10	0.10

		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	0.99	0.99	0.10	0.10	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	1	0.98	1.00	0.14	0.00	0.99	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.97	0.99	0.17	0.08	0.98	0.99	1.00	1.00	0.00	0.04	1.00	1.00
2	0	1.90	1.95	0.30	0.21	0.90	0.95	1.99	2.00	0.11	0.04	0.99	1.00
2	1	1.53	1.79	0.57	0.42	0.57	0.80	1.90	1.96	0.30	0.20	0.90	0.96
2	2	1.50	1.67	0.56	0.49	0.53	0.68	1.83	1.90	0.39	0.32	0.84	0.91

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.3.3 $G(\cdot)$ Student-t with 4 degrees of freedom

Table B15: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.23	0.23	0.42	0.42	0.77	0.77	0.11	0.11	0.31	0.31	0.89	0.89
0	2	0.06	0.05	0.23	0.23	0.94	0.95	0.02	0.01	0.14	0.12	0.98	0.99
1	0	1.00	1.00	0.04	0.04	1.00	1.00	1.00	1.00	0.04	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.07	0.31	0.25	0.46	0.93	0.69	0.07	0.06	0.26	0.24	0.93	0.94
0	2	0.00	0.09	0.04	0.28	1.00	0.91	0.01	0.00	0.08	0.06	0.99	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.99	1.00	0.08	0.04	0.99	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.04	0.28	0.19	0.45	0.96	0.72	0.01	0.13	0.09	0.34	0.99	0.87
0	2	0.00	0.03	0.04	0.18	1.00	0.97	0.00	0.01	0.04	0.08	1.00	0.99
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B16: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	1.99	1.99	0.09	0.09	0.99	0.99	1.98	1.98	0.13	0.13	0.98	0.98
2	1	1.99	1.98	0.13	0.13	0.99	0.98	1.95	1.93	0.23	0.26	0.95	0.94
2	2	1.99	1.99	0.12	0.11	0.99	0.99	1.94	1.95	0.24	0.22	0.94	0.95
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.04	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.06	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00
2	0	1.99	1.99	0.11	0.08	0.99	0.99	1.99	1.99	0.10	0.10	0.99	0.99
2	1	1.95	1.99	0.22	0.10	0.95	0.99	1.99	1.98	0.13	0.14	0.99	0.98
2	2	1.93	1.98	0.26	0.15	0.93	0.98	1.98	1.98	0.13	0.14	0.98	0.98
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.04	0.00	1.00	1.00	2.00	2.00	0.04	0.00	1.00	1.00
2	1	1.99	2.00	0.09	0.04	0.99	1.00	2.00	2.00	0.00	0.00	1.00	1.00
2	2	1.98	1.99	0.15	0.08	0.98	0.99	2.00	2.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B17: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.03	0.02	0.19	0.16	0.97	0.97	0.11	0.13	0.40	0.39	0.91	0.89
0	1	0.10	0.10	0.33	0.33	0.90	0.90	0.21	0.23	0.46	0.49	0.81	0.80
0	2	0.07	0.06	0.29	0.24	0.93	0.95	0.17	0.17	0.46	0.43	0.85	0.86
1	0	1.05	1.04	0.24	0.25	0.94	0.95	1.10	1.11	0.36	0.36	0.90	0.91
1	1	1.07	1.07	0.28	0.30	0.92	0.93	1.18	1.16	0.43	0.44	0.83	0.85
1	2	1.06	1.05	0.24	0.22	0.94	0.95	1.16	1.11	0.46	0.40	0.85	0.91
2	0	1.92	1.95	0.37	0.42	0.86	0.85	2.06	2.06	0.46	0.43	0.83	0.85
2	1	1.89	1.83	0.49	0.51	0.77	0.75	2.06	2.04	0.51	0.56	0.78	0.75
2	2	1.85	1.86	0.43	0.47	0.80	0.77	2.04	2.01	0.57	0.52	0.77	0.77
		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.02	0.02	0.14	0.13	0.98	0.98	0.09	0.14	0.30	0.39	0.91	0.88
0	1	0.04	0.24	0.19	0.44	0.96	0.76	0.16	0.18	0.42	0.44	0.87	0.84
0	2	0.01	0.04	0.12	0.20	0.99	0.97	0.17	0.10	0.46	0.32	0.85	0.90
1	0	1.01	1.02	0.09	0.17	0.99	0.98	1.14	1.13	0.37	0.36	0.87	0.88
1	1	1.04	1.10	0.22	0.30	0.96	0.90	1.14	1.16	0.43	0.40	0.87	0.85
1	2	1.01	1.02	0.12	0.17	0.99	0.97	1.13	1.18	0.40	0.46	0.88	0.85
2	0	2.01	2.02	0.10	0.17	0.99	0.97	2.10	2.09	0.34	0.37	0.90	0.89
2	1	2.03	2.05	0.29	0.32	0.91	0.92	2.14	2.10	0.49	0.43	0.83	0.85
2	2	1.97	1.99	0.30	0.18	0.92	0.97	2.11	2.08	0.44	0.43	0.86	0.84
		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.04	0.02	0.19	0.15	0.96	0.98	0.01	0.02	0.11	0.14	0.99	0.98
0	1	0.06	0.25	0.24	0.47	0.94	0.76	0.04	0.13	0.19	0.33	0.96	0.87
0	2	0.03	0.05	0.16	0.21	0.97	0.95	0.02	0.03	0.14	0.16	0.98	0.97
1	0	1.04	1.03	0.20	0.22	0.96	0.96	1.01	1.02	0.12	0.17	0.99	0.98
1	1	1.05	1.11	0.25	0.31	0.95	0.89	1.03	1.07	0.19	0.25	0.97	0.93
1	2	1.01	1.03	0.12	0.20	0.99	0.97	1.03	1.02	0.19	0.16	0.97	0.98
2	0	2.03	2.02	0.18	0.17	0.97	0.97	2.02	2.02	0.14	0.18	0.98	0.97
2	1	2.08	2.09	0.29	0.30	0.92	0.91	2.05	2.05	0.24	0.21	0.95	0.95
2	2	2.01	2.04	0.16	0.21	0.97	0.96	2.01	2.01	0.11	0.12	0.99	0.99

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.4 Setting $\delta^* = 10^{-1}$ in equation (35)

As discussed in the main paper, δ^* should be very close to zero, and in the main paper we use $\delta^* = 10^{-5}$. The theory predicts that a large δ^* will compress diverging eigenvalues, thus leading (for finite N) to an understatement of r_1 and r_2. As can be seen, this seems particularly evident for r_2.

Table B18: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.11	0.10	0.31	0.30	0.89	0.90	0.05	0.04	0.21	0.21	0.95	0.96
0	2	0.01	0.00	0.11	0.06	0.99	1.00	0.00	0.00	0.04	0.04	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.06	0.00	1.00	1.00	0.99	0.99	0.11	0.10	0.99	0.99
1	2	1.00	1.00	0.04	0.04	1.00	1.00	0.99	0.98	0.08	0.13	0.99	0.98

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.01	0.19	0.08	0.39	0.99	0.81	0.01	0.01	0.12	0.11	0.99	0.99
0	2	0.00	0.01	0.00	0.09	1.00	0.99	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.89	1.00	0.32	0.04	0.89	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.85	1.00	0.36	0.00	0.85	1.00	1.00	1.00	0.04	0.04	1.00	1.00

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.10	0.04	0.29	1.00	0.90	0.00	0.04	0.00	0.20	1.00	0.96
0	2	0.00	0.00	0.00	0.06	1.00	1.00	0.00	0.00	0.00	0.06	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B19: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.06	0.06	1.00	1.00
2	0	1.98	1.98	0.15	0.15	0.98	0.98	1.93	1.94	0.26	0.24	0.93	0.94
2	1	1.90	1.90	0.30	0.31	0.90	0.90	1.75	1.74	0.43	0.44	0.75	0.74
2	2	1.93	1.93	0.27	0.26	0.93	0.93	1.74	1.75	0.44	0.43	0.74	0.75

		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.98	1.00	0.15	0.04	0.98	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.99	1.00	0.10	0.04	0.99	1.00	1.00	1.00	0.04	0.00	1.00	1.00
2	0	1.97	1.98	0.19	0.13	0.97	0.98	1.97	1.98	0.17	0.13	0.98	0.98
2	1	1.78	1.92	0.43	0.29	0.79	0.92	1.94	1.95	0.25	0.22	0.94	0.95
2	2	1.74	1.87	0.45	0.34	0.74	0.87	1.94	1.94	0.23	0.24	0.94	0.94

		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	1.99	1.99	0.09	0.09	0.99	0.99	2.00	2.00	0.04	0.00	1.00	1.00
2	1	1.95	1.99	0.23	0.09	0.95	0.99	1.98	1.99	0.13	0.08	0.98	0.99
2	2	1.93	1.98	0.25	0.13	0.93	0.98	1.99	2.00	0.09	0.04	0.99	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B20: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.08	0.12	0.99	0.99	0.02	0.02	0.16	0.15	0.98	0.98
0	1	0.01	0.01	0.13	0.10	0.98	0.99	0.02	0.03	0.15	0.18	0.98	0.97
0	2	0.01	0.01	0.08	0.10	0.99	0.99	0.01	0.03	0.13	0.16	0.98	0.97
1	0	1.00	0.99	0.09	0.13	0.99	0.99	1.00	1.01	0.15	0.15	0.98	0.98
1	1	0.98	0.98	0.14	0.19	0.98	0.97	1.02	1.01	0.20	0.18	0.96	0.97
1	2	0.99	1.00	0.14	0.17	0.98	0.97	1.00	1.02	0.23	0.18	0.96	0.97
2	0	1.69	1.66	0.50	0.50	0.70	0.66	1.76	1.76	0.45	0.47	0.74	0.74
2	1	1.51	1.45	0.55	0.57	0.51	0.48	1.60	1.55	0.54	0.58	0.59	0.56
2	2	1.48	1.44	0.57	0.57	0.50	0.47	1.58	1.54	0.56	0.56	0.58	0.52

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.04	1.00	1.00	0.02	0.01	0.15	0.10	0.98	0.99
0	1	0.00	0.01	0.04	0.11	1.00	0.99	0.03	0.02	0.18	0.16	0.97	0.97
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.01	0.03	0.13	0.17	0.98	0.97
1	0	1.00	1.00	0.06	0.00	1.00	1.00	1.04	1.03	0.19	0.20	0.96	0.96
1	1	1.00	1.00	0.06	0.10	1.00	0.99	1.03	1.02	0.18	0.15	0.97	0.98
1	2	0.99	1.00	0.13	0.09	0.99	1.00	1.01	1.02	0.16	0.15	0.98	0.98
2	0	1.96	1.99	0.20	0.12	0.96	0.99	1.93	1.93	0.30	0.32	0.90	0.91
2	1	1.77	1.89	0.44	0.32	0.76	0.89	1.86	1.85	0.43	0.42	0.81	0.80
2	2	1.71	1.84	0.48	0.37	0.72	0.85	1.83	1.82	0.39	0.47	0.82	0.77

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.00	0.01	0.04	0.08	1.00	0.99	0.00	0.01	0.04	0.09	1.00	0.99
0	2	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.10	0.09	1.00	1.00	1.00	1.00	0.00	0.06	1.00	1.00
1	1	1.00	1.00	0.10	0.04	1.00	1.00	1.00	1.00	0.00	0.06	1.00	1.00
1	2	1.00	1.00	0.09	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
2	0	2.00	2.00	0.06	0.09	1.00	0.99	2.00	2.00	0.09	0.04	1.00	1.00
2	1	1.96	1.99	0.22	0.13	0.95	0.98	1.99	2.00	0.10	0.06	0.99	1.00
2	2	1.95	1.98	0.22	0.14	0.96	0.98	1.99	1.99	0.15	0.15	0.99	0.99

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.5 Sensitivity analysis to the choice of R_1 and R_2

B.5.1 $R_1 = N$ and $R_2 = N$

Table B21: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.04	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.39	0.39	0.49	0.49	0.61	0.61	0.19	0.18	0.39	0.39	0.81	0.82
0	2	0.32	0.29	0.47	0.46	0.68	0.72	0.05	0.06	0.22	0.23	0.95	0.94
1	0	1.00	1.00	0.06	0.06	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.01	1.01	0.09	0.10	0.99	0.99	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.01	0.06	0.08	1.00	0.99	1.00	1.00	0.00	0.04	1.00	1.00
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.10	0.37	0.31	0.48	0.90	0.63	0.13	0.12	0.34	0.32	0.87	0.88
0	2	0.01	0.10	0.08	0.30	0.99	0.90	0.03	0.03	0.16	0.16	0.97	0.97
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.04	0.32	0.20	0.47	0.96	0.68	0.02	0.17	0.15	0.37	0.98	0.83
0	2	0.00	0.06	0.00	0.23	1.00	0.94	0.00	0.03	0.00	0.16	1.00	0.97
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B22: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.01	0.01	0.10	0.11	0.99	0.99	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.10	0.12	0.30	0.33	0.90	0.88	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.03	0.04	0.16	0.20	0.97	0.96	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.01	1.00	0.09	0.08	0.99	0.99	1.00	1.00	0.04	0.00	1.00	1.00
1	1	1.02	1.02	0.14	0.15	0.98	0.98	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.01	1.01	0.12	0.13	0.99	0.98	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.12	0.00	0.99	1.00	1.99	1.99	0.10	0.11	0.99	0.99
2	1	2.01	2.03	0.11	0.18	0.99	0.97	1.98	1.97	0.17	0.16	0.98	0.97
2	2	2.01	2.00	0.12	0.09	0.99	0.99	1.98	1.98	0.13	0.13	0.98	0.98

		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.06	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.04	0.00	1.00	1.00
2	0	2.00	2.00	0.04	0.00	1.00	1.00	2.00	2.00	0.00	0.09	1.00	1.00
2	1	1.98	2.00	0.13	0.04	0.98	1.00	2.00	2.00	0.00	0.09	1.00	1.00
2	2	1.96	1.99	0.19	0.09	0.96	0.99	2.00	2.00	0.04	0.04	1.00	1.00

		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.04	0.06	1.00	1.00	2.00	2.00	0.04	0.00	1.00	1.00
2	1	1.99	2.00	0.09	0.04	0.99	1.00	2.00	2.00	0.04	0.04	1.00	1.00
2	2	2.00	2.00	0.06	0.00	1.00	1.00	2.00	2.00	0.00	0.09	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B23: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N=50$, $T=100$						$N=100$, $T=100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.01	0.00	0.11	0.08	0.99	0.99	0.03	0.03	0.17	0.17	0.98	0.97
0	1	0.01	0.01	0.12	0.11	0.99	0.99	0.05	0.05	0.21	0.23	0.95	0.95
0	2	0.00	0.01	0.08	0.11	0.99	0.99	0.04	0.03	0.19	0.18	0.96	0.97
1	0	1.00	1.01	0.06	0.08	1.00	0.99	1.02	1.01	0.18	0.11	0.97	0.99
1	1	1.00	1.00	0.11	0.16	0.99	0.98	1.05	1.02	0.25	0.16	0.95	0.97
1	2	1.00	1.01	0.13	0.15	0.98	0.98	1.03	1.02	0.19	0.18	0.97	0.97
2	0	1.82	1.80	0.41	0.41	0.81	0.80	1.90	1.93	0.36	0.36	0.88	0.87
2	1	1.73	1.69	0.48	0.49	0.72	0.69	1.81	1.79	0.45	0.48	0.76	0.74
2	2	1.72	1.71	0.47	0.49	0.72	0.70	1.81	1.81	0.44	0.43	0.78	0.79
		$N=200$, $T=100$						$N=100$, $T=200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.06	1.00	1.00	0.02	0.01	0.14	0.10	0.98	0.99
0	1	0.01	0.07	0.08	0.26	0.99	0.93	0.05	0.03	0.22	0.18	0.96	0.97
0	2	0.00	0.00	0.08	0.00	0.99	1.00	0.02	0.02	0.15	0.14	0.98	0.98
1	0	1.00	1.00	0.04	0.04	1.00	1.00	1.01	1.01	0.09	0.10	0.99	0.99
1	1	1.00	1.00	0.11	0.12	0.99	0.99	1.04	1.04	0.20	0.20	0.96	0.96
1	2	1.00	1.00	0.06	0.10	1.00	1.00	1.02	1.03	0.15	0.18	0.98	0.97
2	0	1.99	1.99	0.14	0.10	0.99	0.99	2.00	1.99	0.19	0.24	0.96	0.96
2	1	1.94	2.00	0.26	0.17	0.93	0.97	1.96	1.94	0.32	0.31	0.91	0.92
2	2	1.92	1.96	0.27	0.20	0.92	0.96	1.98	1.97	0.23	0.30	0.95	0.92
		$N=200$, $T=200$						$N=200$, $T=500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.04	0.04	1.00	1.00
0	1	0.01	0.06	0.10	0.24	0.99	0.94	0.00	0.01	0.00	0.10	1.00	0.99
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.00	0.00	0.06	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.06	1.00	1.00	1.00	1.00	0.04	0.06	1.00	1.00
1	1	1.01	1.01	0.11	0.10	0.99	0.99	1.00	1.00	0.09	0.06	1.00	1.00
1	2	1.01	1.00	0.09	0.04	0.99	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	1.99	0.06	0.13	1.00	0.99	2.00	2.00	0.00	0.06	1.00	1.00
2	1	1.99	2.01	0.17	0.15	0.99	0.98	2.01	2.01	0.13	0.08	0.99	0.99
2	2	1.99	2.00	0.11	0.08	0.99	0.99	2.00	1.99	0.04	0.17	1.00	0.99

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

B.5.2 $R_1 = 3N$ and $R_2 = 3N$

Table B24: Estimated number of factors with linear trend, $\widehat{r}_1$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.13	0.11	0.34	0.32	0.87	0.89	0.09	0.08	0.28	0.28	0.91	0.92
0	2	0.01	0.01	0.10	0.09	0.99	0.99	0.01	0.01	0.08	0.08	0.99	0.99
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.06	0.00	1.00	1.00

		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.04	0.28	0.20	0.45	0.96	0.72	0.04	0.03	0.19	0.18	0.96	0.97
0	2	0.00	0.06	0.04	0.23	1.00	0.94	0.00	0.00	0.04	0.04	1.00	1.00
1	0	1.00	1.00	0.06	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.99	1.00	0.09	0.00	0.99	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	0.99	1.00	0.08	0.00	0.99	1.00	1.00	1.00	0.00	0.00	1.00	1.00

		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_1	r_2	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.01	0.23	0.12	0.42	0.99	0.77	0.00	0.11	0.00	0.31	1.00	0.89
0	2	0.00	0.01	0.04	0.09	1.00	0.99	0.00	0.00	0.00	0.04	1.00	1.00
1	0	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.04	0.00	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_1$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_1 = r_1$.

Table B25: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 0$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	1.99	1.99	0.12	0.12	0.99	0.99	1.98	1.98	0.13	0.14	0.98	0.98
2	1	1.94	1.94	0.24	0.23	0.94	0.94	1.90	1.88	0.31	0.33	0.90	0.89
2	2	1.96	1.94	0.21	0.23	0.96	0.94	1.89	1.89	0.32	0.32	0.89	0.89

		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.04	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	0.99	1.00	0.08	0.00	0.99	1.00	1.00	1.00	0.00	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.04	1.00	1.00
2	0	1.99	2.00	0.09	0.06	0.99	1.00	1.99	1.99	0.10	0.09	0.99	0.99
2	1	1.90	1.98	0.31	0.15	0.90	0.98	1.97	1.97	0.16	0.19	0.97	0.97
2	2	1.91	1.97	0.31	0.17	0.92	0.97	1.98	1.98	0.13	0.15	0.98	0.98

		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	1	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
0	2	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00
1	0	1.00	1.00	0.00	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.04	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.09	0.04	1.00	1.00	2.00	2.00	0.04	0.04	1.00	1.00
2	1	1.98	2.00	0.13	0.04	0.98	1.00	1.99	2.00	0.09	0.00	0.99	1.00
2	2	1.99	2.00	0.09	0.00	0.99	1.00	2.00	2.00	0.04	0.04	1.00	1.00

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.

Table B26: Estimated number of zero-mean $I(1)$ factors, $\widehat{r}_2$, when $r_1 = 1$.

		$N = 50$, $T = 100$						$N = 100$, $T = 100$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.06	0.06	1.00	1.00	0.02	0.03	0.14	0.17	0.98	0.97
0	1	0.02	0.01	0.14	0.12	0.98	0.99	0.04	0.04	0.19	0.19	0.96	0.96
0	2	0.00	0.01	0.08	0.08	0.99	0.99	0.04	0.02	0.20	0.17	0.96	0.98
1	0	1.01	0.99	0.11	0.12	0.99	0.99	1.02	1.03	0.17	0.19	0.98	0.97
1	1	1.01	1.00	0.15	0.08	0.98	0.99	1.04	1.05	0.22	0.22	0.96	0.95
1	2	1.01	0.99	0.12	0.12	0.99	0.99	1.02	1.02	0.17	0.15	0.97	0.98
2	0	1.82	1.82	0.39	0.44	0.82	0.81	1.93	1.91	0.34	0.36	0.89	0.88
2	1	1.73	1.72	0.47	0.48	0.72	0.70	1.83	1.84	0.47	0.43	0.75	0.80
2	2	1.76	1.73	0.47	0.46	0.75	0.71	1.80	1.81	0.45	0.48	0.77	0.77
		$N = 200$, $T = 100$						$N = 100$, $T = 200$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.00	0.00	1.00	1.00	0.03	0.02	0.17	0.15	0.97	0.98
0	1	0.00	0.06	0.00	0.24	1.00	0.94	0.04	0.05	0.21	0.22	0.96	0.95
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.03	0.03	0.18	0.17	0.97	0.97
1	0	1.00	1.00	0.06	0.06	1.00	1.00	1.03	1.03	0.16	0.16	0.97	0.97
1	1	0.99	1.02	0.10	0.15	1.00	0.98	1.03	1.04	0.16	0.20	0.97	0.96
1	2	1.00	1.00	0.00	0.04	1.00	1.00	1.02	1.01	0.21	0.15	0.97	0.98
2	0	2.00	2.00	0.06	0.04	1.00	1.00	2.01	2.00	0.22	0.21	0.96	0.96
2	1	1.96	1.99	0.21	0.18	0.95	0.97	1.95	1.96	0.29	0.29	0.91	0.92
2	2	1.91	1.94	0.29	0.26	0.91	0.95	1.96	1.95	0.29	0.31	0.92	0.91
		$N = 200$, $T = 200$						$N = 200$, $T = 500$					
		average		std. dev.		% correct		average		std. dev.		% correct	
r_2	r_3	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$	$BT1$	$BT2$
0	0	0.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.04	0.00	1.00	1.00
0	1	0.00	0.07	0.06	0.26	1.00	0.93	0.01	0.01	0.10	0.11	0.99	0.99
0	2	0.00	0.00	0.00	0.04	1.00	1.00	0.00	0.00	0.00	0.08	1.00	0.99
1	0	1.00	1.00	0.06	0.04	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
1	1	1.00	1.01	0.06	0.12	1.00	0.99	1.00	1.00	0.04	0.04	1.00	1.00
1	2	1.00	1.00	0.00	0.10	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00
2	0	2.00	2.00	0.04	0.06	1.00	1.00	2.00	2.00	0.00	0.04	1.00	1.00
2	1	2.00	2.01	0.20	0.15	0.98	0.98	1.99	2.00	0.17	0.12	0.99	0.99
2	2	2.00	2.00	0.12	0.04	0.99	1.00	2.00	2.00	0.06	0.15	1.00	0.99

In each cell we report the average and standard deviation of $\widehat{r}_2$ over all Monte Carlo replications, as well as the fraction of times in which $\widehat{r}_2 = r_2$.