Supplement to "Long-Range Dependent Curve Time Series"

Degui Li
Department of Mathematics, University of York

Peter M. Robinson
Department of Economics, London School of Economics

Han Lin Shang
Research School of Finance, Actuarial Studies and Statistics,
Australian National University

Version: March 26, 2019

In this supplement, we first give the detailed proofs of the theoretical results stated in the main document and then present additional numerical results.

A Proofs of the theoretical results

Throughout this appendix, M denotes a generic positive constant. We start with the proof of Proposition 1.

PROOF OF PROPOSITION 1. By the definition of λ_{ni} in (2.3), we have

$$\begin{aligned}\lambda_{ni} &= \int_{\mathcal{C}}\int_{\mathcal{C}} c_n(u,v)\psi_{ni}(v)\psi_{ni}(u)dvdu \\ &= \sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left[\int_{\mathcal{C}}\int_{\mathcal{C}} X_t(u)X_s(v)\psi_{ni}(v)\psi_{ni}(u)dvdu\right] \\ &= \sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left[\int_{\mathcal{C}} X_t(u)\psi_{ni}(u)du\int_{\mathcal{C}} X_s(v)\psi_{ni}(v)dv\right] \\ &= \sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(x_t^i x_s^i\right).\end{aligned}$$

For $i = p_{k-1}+1,\cdots,p_k$ with $k = 1,\cdots,\kappa_0$, using (2.9) in Assumption 2 and standard calculation,

$$\mathsf{E}\left(x_t^i x_s^i\right) \sim \bar{\rho}_i^2 c_{\alpha_k}|t-s|^{1-2\alpha_k} \quad \text{as } |t-s|\to\infty \text{ and } n\to\infty,$$

which implies that

$$\lambda_{ni} = \sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(x_t^i x_s^i\right) \sim \theta_i^2 n^{3-2\alpha_k} \quad \text{as } n\to\infty,$$

completing the proof of (2.12).

For $i \geqslant p_{\kappa_0}+1$, note that

$$\begin{aligned}\sum_{t=1}^{n} x_t^i &= \sum_{t=1}^{n}\sum_{j=0}^{\infty}\langle b_j^i,\eta_t\rangle - \sum_{t=1}^{n}\sum_{j=n-t+1}^{\infty}\langle b_j^i,\eta_t\rangle + \sum_{t=1}^{n}\sum_{j=t}^{\infty}\langle b_j^i,\eta_{t-j}\rangle \\ &=: A_{n1}(i) - A_{n2}(i) + A_{n3}(i).\end{aligned} \tag{A.1}$$

By (2.10) in Assumption 2, we may show that for any t and $i \geqslant p_{\kappa_0}+1$,

$$\sum_{j=\delta(n)}^{\infty}\left(\mathsf{E}\langle b_j^i,\eta_t\rangle^2\right)^{1/2} = O\left(\sum_{j=\delta(n)}^{\infty}\left\|b_j^i\right\|\right) = o(1) \quad \text{as} \quad \delta(n)\to\infty. \tag{A.2}$$

For $A_{n2}(i)$, note that

$$\begin{aligned}A_{n2}(i) &= \sum_{t=1}^{n-\delta_1(n)}\sum_{j=n-t+1}^{\infty}\langle b_j^i,\eta_t\rangle + \sum_{t=n-\delta_1(n)+1}^{n}\sum_{j=n-t+1}^{\infty}\langle b_j^i,\eta_t\rangle \\ &=: A_{n2,1}(i) + A_{n2,2}(i),\end{aligned}$$

where $\delta_1(n) \to \infty$ and $\delta_1(n) = o(n)$ as $n \to \infty$. By the Cauchy-Schwarz inequality and (A.2) with $\delta(n) = \delta_1(n) + 1$, it is easy to see that uniformly for $i \geqslant p_{\kappa_0} + 1$,

$$
\begin{aligned}
\mathsf{E}\left[A_{n2,1}^2(i)\right] &= \sum_{t=1}^{n-\delta_1(n)} \sum_{j_1=n-t+1}^{\infty} \sum_{j_2=n-t+1}^{\infty} \mathsf{E}\left[\langle b_{j_1}^i, \eta_t\rangle \langle b_{j_2}^i, \eta_t\rangle\right] \\
&\leqslant \sum_{t=1}^{n-\delta_1(n)} \sum_{j_1=n-t+1}^{\infty} \sum_{j_2=n-t+1}^{\infty} \left(\mathsf{E}\langle b_{j_1}^i, \eta_t\rangle^2\right)^{1/2} \left(\mathsf{E}\langle b_{j_2}^i, \eta_t\rangle^2\right)^{1/2} \\
&\leqslant \sum_{t=1}^{n-\delta_1(n)} \sum_{j_1=\delta_1(n)+1}^{\infty} \left(\mathsf{E}\langle b_{j_1}^i, \eta_t\rangle^2\right)^{1/2} \sum_{j_2=\delta_1(n)+1}^{\infty} \left(\mathsf{E}\langle b_{j_2}^i, \eta_t\rangle^2\right)^{1/2} \\
&= o\left(n - \delta_1(n)\right) = o(n),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathsf{E}\left[A_{n2,2}^2(i)\right] &= \sum_{t=n-\delta_1(n)+1}^{n} \sum_{j_1=n-t+1}^{\infty} \sum_{j_2=n-t+1}^{\infty} \mathsf{E}\left[\langle b_{j_1}^i, \eta_t\rangle \langle b_{j_2}^i, \eta_t\rangle\right] \\
&\leqslant \sum_{t=n-\delta_1(n)+1}^{n} \sum_{j_1=1}^{\infty} \left(\mathsf{E}\langle b_{j_1}^i, \eta_t\rangle^2\right)^{1/2} \sum_{j_2=1}^{\infty} \left(\mathsf{E}\langle b_{j_2}^i, \eta_t\rangle^2\right)^{1/2} \\
&= O\left(\delta_1(n)\right) = o(n).
\end{aligned}
$$

Hence, we have

$$
\mathsf{E}\left[A_{n2}^2(i)\right] = o(n) \tag{A.3}
$$

uniformly for $i \geqslant p_{\kappa_0} + 1$.

For $A_{n3}(i)$, note that

$$
\begin{aligned}
A_{n3}(i) &= \sum_{s=0}^{\infty} \sum_{j=s+1}^{s+n} \langle b_j^i, \eta_{-s}\rangle \\
&= \sum_{s=0}^{n} \sum_{j=s+1}^{s+n} \langle b_j^i, \eta_{-s}\rangle + \sum_{s=n+1}^{\infty} \sum_{j=s+1}^{s+n} \langle b_j^i, \eta_{-s}\rangle \\
&=: A_{n3,1}(i) + A_{n3,2}(i).
\end{aligned}
$$

Similarly to the proof of (A.3), we may show that uniformly for $i \geqslant p_{\kappa_0} + 1$

$$
\mathsf{E}\left[A_{n3,1}^2(i)\right] = o(n).
$$

On the other hand, by (2.10) in Assumption 2, we have

$$
\begin{aligned}
\mathsf{E}\left[A_{n3,2}^2(i)\right] &= \sum_{s=n+1}^{\infty}\sum_{j_1=s+1}^{s+n}\sum_{j_2=s+1}^{s+n}\mathsf{E}\left[\langle b_{j_1}^i,\eta_{-s}\rangle\langle b_{j_2}^i,\eta_{-s}\rangle\right] \\
&\leqslant \sum_{s=n+1}^{\infty}\sum_{j_1=s+1}^{s+n}\sum_{j_2=s+1}^{s+n}\left(\mathsf{E}\langle b_{j_1}^i,\eta_{-s}\rangle^2\right)^{1/2}\left(\mathsf{E}\langle b_{j_2}^i,\eta_{-s}\rangle^2\right)^{1/2} \\
&= \sum_{s=n+1}^{\infty}\left(\mathsf{E}\langle b_{s+1}^i,\eta_{-s}\rangle^2\right)^{1/2}\sum_{j_1=s+1}^{s+n}\left(\mathsf{E}\langle b_{j_1}^i,\eta_{-s}\rangle^2/\mathsf{E}\langle b_{s+1}^i,\eta_{-s}\rangle^2\right)^{1/2} \\
&\quad \sum_{j_2=s+1}^{s+n}\left(\mathsf{E}\langle b_{j_2}^i,\eta_{-s}\rangle^2\right)^{1/2} \\
&= o(1)\cdot O(n)\cdot o(1) = o(n).
\end{aligned}
$$

Hence, we have

$$\mathsf{E}\left[A_{n3}^2(i)\right] = o(n) \tag{A.4}$$

uniformly for $i \geqslant p_{\kappa_0}+1$. Then, by (2.10), (A.1), (A.3) and (A.4),

$$\sum_{i=p_{\kappa_0}+1}^{\infty}\lambda_{ni} = \sum_{i=p_{\kappa_0}+1}^{\infty}\sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(x_t^i x_s^i\right) = \sum_{i=p_{\kappa_0}+1}^{\infty}\mathsf{E}\left[A_{n1}^2(i)\right] + o\left(n\right) = O(n),$$

which proves (2.13). The proof of Proposition 1 is completed. □

In order to prove Theorem 1, we need to use the following technical lemma which is given in Theorem 18.6.5 of Ibragimov and Linnik (1971). Their proof is slightly in error, but proofs can also be found in a recent book by Giraitis et al. (2012), and in a slightly more general setting in Robinson (1997).

LEMMA A.1. *Suppose that* $\{z_t\}$ *is a stationary linear process defined by*

$$z_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j},$$

where $\{\epsilon_j\}$ *is a sequence of i.i.d. random variables with mean zero and finite variance* σ^2*, and* $\sum_{j=0}^{\infty} a_j^2 < \infty$*. Furthermore,*

$$\mathsf{var}(s_n) \to \infty, \quad s_n = \sum_{t=1}^{n} z_t.$$

Then

$$s_n/\mathrm{var}^{1/2}(s_n) \xrightarrow{d} N(0,1).$$

PROOF OF THEOREM 1. Let

$$\overline{T}_{nk} = n^{-(3/2-\alpha_k)} \cdot T_{nk} = n^{-(3/2-\alpha_k)} \cdot \sum_{t=1}^{n} X_{tk}$$

and $P_{\overline{T}_{nk}}$ be the distribution of $\overline{T}_{nk}$, where X_{tk} and T_{nk} are defined as in Section 2.3. Note that the sub-space $\mathcal{S}_k$ is of dimension $p_k - p_{k-1}$, which is finite. By Proposition 1 and following the proof of Theorem 2.7 in Bosq (2000), we may show that $P_{\overline{T}_{nk}}$ is tight. Therefore, using the orthogonality as shown in (2.3), it suffices to prove that

$$n^{-(3/2-\alpha_k)} \left(\sum_{t=1}^{n} x_t^i,\ i = p_{k-1}+1, \cdots, p_k \right) \xrightarrow{d} \left(\theta_i N_i,\ i = p_{k-1}+1, \cdots, p_k\right), \qquad (A.5)$$

where $k = 1, \cdots, \kappa_0$,

$$\theta_i^2 = \lim_{n\to\infty} \frac{\lambda_{ni}}{n^{3-2\alpha_k}} = \frac{\bar{\rho}_i^2 c_{\alpha_k}}{(1-\alpha_k)(3-2\alpha_k)}, \quad i = p_{k-1}+1, \cdots, p_k, \qquad (A.6)$$

x_t^i is defined in (2.4), and N_i, $i = 1, 2, \cdots, p_{\kappa_0}$, are independent standard normal random variables.

By the Cramér-Wold device (Billingsley 1968), to prove (A.5), we only need to show

$$n^{-(3/2-\alpha_k)} \sum_{i=p_{k-1}+1}^{p_k} a_i \sum_{t=1}^{n} x_t^i \xrightarrow{d} \sum_{i=p_{k-1}+1}^{p_k} a_i \theta_i N_i, \qquad (A.7)$$

for any $\left(a_{p_{k-1}+1}, \cdots, a_{p_k}\right)^\tau \in \mathcal{R}^{p_k - p_{k-1}}$. By orthogonality as shown in (2.3), we may show the independence of N_i and

$$\mathrm{var}\left(\sum_{i=p_{k-1}+1}^{p_k} a_i \sum_{t=1}^{n} x_t^i \right) = \sum_{i=p_{k-1}+1}^{p_k} a_i^2 \mathrm{var}\left(\sum_{t=1}^{n} x_t^i \right). \qquad (A.8)$$

Furthermore, as in the proof of Proposition 1, for each $i = p_{k-1}+1, \cdots, p_k$,

$$E\left(x_1^i x_{r+1}^i\right) \sim \bar{\rho}_i^2 c_{\alpha_k} r^{1-2\alpha_k} \quad \text{as } r \to \infty \text{ and } n \to \infty,$$

indicating that

$$
\begin{aligned}
\frac{1}{n^{3-2\alpha_k}} \cdot \mathsf{var}\left(\sum_{t=1}^{n} x_t^i\right) &= \frac{2}{n^{3-2\alpha_k}} \sum_{r=1}^{n-1} (n-r)\mathsf{E}\left(x_1^i x_{r+1}^i\right) + O(n^{2\alpha_k - 2}) \\
&\sim \frac{2\bar{\rho}_i^2 c_{\alpha_k}}{n^{3-2\alpha_k}} \left(n \sum_{r=1}^{n-1} r^{1-2\alpha_k} - \sum_{r=1}^{n-1} r^{2-2\alpha_k}\right) \\
&\sim \frac{\bar{\rho}_i^2 c_{\alpha_k}}{(1-\alpha_k)(3-2\alpha_k)}. \qquad \text{(A.9)}
\end{aligned}
$$

Then, in view of (A.8), (A.9) and Lemma A.1, we readily prove (A.7), completing the proof of Theorem 1. $\square$

We next give the detailed proof of Proposition 2.

PROOF OF PROPOSITION 2. Define

$$
\bar{r}_{k1}(u,v) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-k} X_t(u) X_{t+k}(v), & k \geqslant 0, \\ \frac{1}{n} \sum_{t=1}^{n-|k|} X_{t+|k|}(u) X_t(v), & k < 0, \end{cases}
$$

and

$$
\begin{aligned}
\bar{r}_{k2}(u,v) &= \bar{r}_k(u,v) - \bar{r}_{k1}(u,v) \\
&= \begin{cases} \frac{n-k}{n}\overline{X}_n(u)\overline{X}_n(v) - \left[\frac{1}{n}\sum_{t=1}^{n-k} X_t(u)\right]\overline{X}_n(v) - \overline{X}_n(u)\left[\frac{1}{n}\sum_{t=k+1}^{n} X_t(v)\right], & k \geqslant 0, \\ \frac{n+k}{n}\overline{X}_n(u)\overline{X}_n(v) - \left[\frac{1}{n}\sum_{t=1-k}^{n} X_t(u)\right]\overline{X}_n(v) - \overline{X}_n(u)\left[\frac{1}{n}\sum_{t=1}^{n+k} X_t(v)\right], & k < 0, \end{cases}
\end{aligned}
$$

where $|k| \leqslant m$. Let $H_1 = 3/2 - \alpha_1$ throughout this proof. Note that

$$
\begin{aligned}
\bar{c}_m(u,v) &= \frac{1}{m^{2H_1}} \sum_{|k| \leqslant m} (m - |k|)\, \bar{r}_k(u,v) \\
&= \frac{1}{m^{2H_1}} \sum_{|k| \leqslant m} (m - |k|)\, \bar{r}_{k1}(u,v) + \frac{1}{m^{2H_1}} \sum_{|k| \leqslant m} (m - |k|)\, \bar{r}_{k2}(u,v) \\
&=: \bar{c}_{m1}(u,v) + \bar{c}_{m2}(u,v). \qquad \text{(A.10)}
\end{aligned}
$$

Let $\overline{C}_{mi}$, $i = 1, 2$, be operators defined by

$$
\overline{C}_{mi}(x)(u) = \int_{\mathcal{C}} \bar{c}_{mi}(u,v) x(v) dv, \quad x \in \mathcal{H}.
$$

Noting that $\mathsf{E}[X_t] = 0$ by Assumption 1 and using Corollary 1, it is easy to show that

$$\int_{\mathcal{C}}\int_{\mathcal{C}} \overline{r}_{k2}^2(u,v)\,du\,dv = O_P\left(n^{2(2H_1-2)}\right),$$

which leads to

$$\begin{aligned}\left\|\overline{C}_{m2}\right\|_{\mathcal{S}} &= O_P\left(m^{2-2H_1}n^{2H_1-2}\right) = O_P\left((m/n)^{2-2H_1}\right)\\ &= O_P\left((m/n)^{2\alpha_1-1}\right) = o_P(1)\end{aligned} \tag{A.11}$$

as $m = o(n)$ and $\alpha_1 > 1/2$. By (A.10) and (A.11), in order to complete the proof of Proposition 2, we only need to show that

$$\left\|\overline{C}_{m1} - C\right\|_{\mathcal{S}} = o_P(1). \tag{A.12}$$

Observe that

$$\begin{aligned}\overline{c}_{m1}(u,v) - c(u,v) &= \left(\frac{1}{m^{2H_1}}\sum_{|k|\leqslant m}(m-|k|)\left\{\overline{r}_{k1}(u,v) - \mathsf{E}[\overline{r}_{k1}(u,v)]\right\}\right) + \\ &\quad \left(\frac{1}{m^{2H_1}}\sum_{|k|\leqslant m}(m-|k|)\mathsf{E}[\overline{r}_{k1}(u,v)] - c(u,v)\right)\\ &=: \overline{c}_{m3}(u,v) + \overline{c}_{m4}(u,v).\end{aligned} \tag{A.13}$$

Let $\overline{C}_{m3}$ and $\overline{C}_{m4}$ be the operators defined similarly to $\overline{C}_{m1}$ above but with $\overline{c}_{m1}(u,v)$ replaced by $\overline{c}_{m3}(u,v)$ and $\overline{c}_{m4}(u,v)$, respectively.

We first consider $\overline{C}_{m4}$. By Assumption 2, we may show that for $k \geqslant 0$

$$\begin{aligned}\mathsf{E}[\overline{r}_{k1}(u,v)] &= \frac{1}{n}\sum_{t=1}^{n-|k|}\mathsf{E}[X_t(u)X_{t+k}(v)]\\ &= \mathsf{E}[X_1(u)X_{1+k}(v)] - \frac{k}{n}\cdot\mathsf{E}[X_1(u)X_{1+k}(v)]\\ &= \mathsf{E}[X_1(u)X_{1+k}(v)] + O\left(k^{2-2\alpha_1}/n\right)\end{aligned}$$

for k sufficiently large; and for $k < 0$, $\mathsf{E}[\overline{r}_{k1}(u,v)] = \mathsf{E}\left[X_{1+|k|}(u)X_1(v)\right] + O\left(|k|^{2-2\alpha_1}/n\right)$ for $|k|$ sufficiently large. By definition of $c(u,v)$ in (3.1), we then have

$$\left\|\overline{C}_{m4}\right\|_{\mathcal{S}} = O(m/n) + o(1) = o(1). \tag{A.14}$$

We finally consider $\overline{C}_{m3}$. Define

$$\begin{aligned}\tilde{r}_{k1}(u,v) &= \begin{cases} \frac{1}{n}\sum_{t=1}^{n-k} X_{t,1}(u)X_{t+k,1}(v), & k \geqslant 0, \\ \frac{1}{n}\sum_{t=1}^{n-|k|} X_{t+|k|,1}(u)X_{t,1}(v), & k < 0, \end{cases} \\ \widehat{r}_{k1}(u,v) &= \overline{r}_{k1}(u,v) - \tilde{r}_{k1}(u,v),\end{aligned}$$

where $X_{t,1}(\cdot) = X_{t1}(\cdot)$ is the projection of $X_t(\cdot)$ onto $\mathcal{S}_1$ which is defined as in (2.11) of the main document. Let

$$\overline{c}_{m5}(u,v) = \frac{1}{m^{2H_1}} \sum_{|k|\leqslant m} (m-|k|)\left\{\tilde{r}_{k1}(u,v) - \mathsf{E}[\tilde{r}_{k1}(u,v)]\right\}.$$

and $\overline{C}_{m5}$ be the corresponding operator. If $\kappa_0 \geqslant 2$, by (2.11) and Assumption 2, we can prove that

$$\int_{\mathcal{C}}\int_{\mathcal{C}} \widehat{r}_{k1}^2(u,v)dudv = O\left(|k|^{2(1-\alpha_1-\alpha_2)}\right),$$

which indicates that

$$\left\|\overline{C}_{m3} - \overline{C}_{m5}\right\|_{\mathcal{S}} = O_P\left(m^{\alpha_1-\alpha_2}\right) = o_P(1) \tag{A.15}$$

as $\alpha_1 < \alpha_2$ in Assumption 2. The proof of (A.15) is analogous when $\kappa_0 = 1$ by replacing α_2 by 1.

We next derive the asymptotic order for the Hilbert-Schmidt norm of $\overline{C}_{m5}$. For $i = 1, \cdots, p_1$ and $|k| \leqslant m$, we let

$$\tilde{r}_k^i = \begin{cases} \frac{1}{n}\sum_{t=1}^{n-k} x_t^i x_{t+k}^i & k \geqslant 0, \\ \frac{1}{n}\sum_{t=1}^{n-|k|} x_t^i x_{t-k}^i & k < 0, \end{cases}$$

and define

$$\overline{c}_m^i = \frac{1}{m^{2H_1}} \sum_{|k|\leqslant m} (m-|k|)\left[\tilde{r}_k^i - \mathsf{E}(\tilde{r}_k^i)\right].$$

Then, by (2.11), we have

$$\int_{\mathcal{C}}\int_{\mathcal{C}} \overline{c}_{m5}^2(u,v)dudv = O_P\left(\sum_{i=1}^{p_1} \mathsf{Var}\left(\overline{c}_m^i\right)\right). \tag{A.16}$$

As p_1 is fixed, it remains to derive the order of $\mathsf{Var}(\bar{c}_m^i)$, $i = 1, \cdots, p_1$. Note that

$$\mathsf{Var}(\bar{c}_m^i) \leqslant m^{3-4H_1} \sum_{|k| \leqslant m} \mathsf{Var}\left(\tilde{r}_k^i\right), \tag{A.17}$$

and

$$\mathsf{E}\left[x_t^i x_{t+j}^i\right] = \sum_{s=0}^{\infty} \mathsf{E}\left[\langle b_s^i, \eta_{t-s}\rangle \langle b_{s+j}^i, \eta_{t-s}\rangle\right].$$

We next consider only $k \geqslant 0$ as $k < 0$ can be handled in the same way. Using Isserlis' formula, we have

$$\mathsf{Var}(\tilde{r}_k^i) \leqslant \frac{4}{n} \cdot \sum_{j=1}^{n} \left\{ \sum_{s=0}^{\infty} \mathsf{E}\left[\langle b_s^i, \eta_{t-s}\rangle \langle b_{s+j}^i, \eta_{t-s}\rangle\right] \right\}^2 + \frac{1}{n^2} \cdot \sum_{s,t=1}^{n-|k|} \mathsf{cum}\left(x_s^i, x_{s+k}^i, x_t^i, x_{t+k}^i\right), \tag{A.18}$$

where "cum" denotes the cumulant. The second term on the right hand side of (A.18) is of order $O\left(\sum_{s=0}^{\infty} s^{-4\alpha_1}/n\right) = O(1/n)$, so its contribution to $\mathsf{Var}(\bar{c}_m^i)$ in (A.17) is $O\left(m^{4-4H_1}/n\right) = O\left(m^{4\alpha_1-2}/n\right) = o(1)$ noting that $m \propto n^{\gamma}$ with $0 < \gamma < 1/(4\alpha_1 - 2)$ assumed in Proposition 2. By (2.9) in Assumption 2,

$$\begin{aligned}
& \sum_{s=0}^{\infty} \mathsf{E}\left[\langle b_s^i, \eta_{t-s}\rangle \langle b_{s+j}^i, \eta_{t-s}\rangle\right] \\
\leqslant\ & M\left(j^{-\alpha_1} + \sum_{s=1}^{j} i^{-\alpha_1}(i+j)^{-\alpha_1} + \sum_{s=j+1}^{\infty} i^{-\alpha_1}(i+j)^{-\alpha_1}\right) \\
\leqslant\ & M\left(j^{-\alpha_1} + j^{-\alpha_1}\sum_{s=1}^{j} i^{-\alpha_1} + \sum_{s=j+1}^{\infty} i^{-2\alpha_1}\right) \\
=\ & O\left(j^{1-2\alpha_1}\right)
\end{aligned}$$

when j is sufficiently large. For $1/2 < \alpha_1 < 3/4$, the first term on the right hand side of (A.18) is $O(n^{2-4\alpha_1})$ so its contribution to $\mathsf{Var}(\bar{c}_m^i)$ in (A.17) is $O\left(m^{4-4H_1} n^{2-4\alpha_1}\right) = O\left((m/n)^{4\alpha_1-2}\right) = o(1)$; for $\alpha_1 = 3/4$, it is $O\left(\log n/n\right)$, giving contribution $O\left(m \log n/n\right) = o(1)$; and for $3/4 < \alpha_1 < 1$, it is $O\left(1/n\right)$, giving contribution $O\left(m^{4-4H_1}/n\right) = O\left(m^{4\alpha_1-2}/n\right) = o(1)$. Therefore, we have proved that $\mathsf{Var}(\bar{c}_m^i) = o(1)$, $i = 1, \cdots, p_1$, which together with (A.16) implies that

$$\|\overline{C}_{m5}\|_{\mathcal{S}} = o_P(1). \tag{A.19}$$

Then we can complete the proof of (A.12) by (A.13)–(A.15) and (A.19), and thus complete the proof of Proposition 2. $\square$

PROOF OF THEOREM 2. Note that $\overline{c}_m(u, v)$ defined in (3.4) is proportional to $\widehat{c}_m(u, v)$ defined in (3.5). So we may show that for $i = 1, \cdots, p_1$,

$$\widehat{\lambda}_{m,i}\widehat{\psi}_i(u) = \int_{\mathcal{C}} \widehat{c}_m(u, v)\widehat{\psi}_i(v)dv$$

and

$$\widehat{\lambda}_i\widehat{\psi}_i(u) = \int_{\mathcal{C}} \overline{c}_m(u, v)\widehat{\psi}_i(v)dv,$$

where $\widehat{\lambda}_{m,i}$ is the i^{th} largest eigenvalue of $\widehat{c}_m(u, v)$ and $\widehat{\lambda}_i = \widehat{\lambda}_{mi}/m^{3-2\alpha_1}$. This, together with Lemma 4.3 in Bosq (2000), indicates that

$$\left\|\widehat{\psi}_i - \tau_i\psi_i\right\| \leqslant M \cdot \left\|\overline{C}_m - C\right\|_{\mathcal{S}} \Big/ \xi_i, \tag{A.20}$$

uniformly for $i = 1, \cdots, p_1$, where $\xi_1 = \lambda_1 - \lambda_2$, $\xi_i = \min\{\lambda_{i-1} - \lambda_i, \lambda_i - \lambda_{i+1}\}$ for $i = 2, \cdots, p_1$. Then, by Proposition 2 and (A.20), we can prove (3.9) in Theorem 2. The proof of (3.10) is similar, so details are omitted here. $\square$

Before proving Proposition 3, we give a lemma on strong approximation of scalar long-range dependent linear processes, which is a corollary in Wang et al. (2003).

LEMMA A.2. *Suppose that* $\{z_t\}$ *is a scalar long-range dependent linear process defined by*

$$z_t = \sum_{j=0}^{\infty} b_j\epsilon_{t-j}, \qquad b_0 = 1, \quad b_j \sim j^{-\alpha} \text{ as } j \to \infty,$$

where $\{\epsilon_j\}$ *is a sequence of i.i.d. random variables with* $\mathsf{E}[\epsilon_1] = 0$*,* $\mathsf{E}[\epsilon_1^2] = \sigma^2$ *and* $\mathsf{E}[|\epsilon_1|^{2+\delta}] < \infty$*,* $\delta > 0$*. Then, on an appropriate probability space of* $\{\epsilon_j\}$*, we can construct a fractional Brownian motion* $\mathsf{B}_H(\cdot)$ *with* $H = 3/2 - \alpha$ *such that*

$$\sup_{0\leqslant r\leqslant 1}\left|\frac{1}{n^H}\sum_{t=1}^{\lfloor nr\rfloor} z_t - \theta_\alpha \mathsf{B}_H(r)\right| = o\left(n^{(3+\delta)/(2+\delta)-3/2}\right) \quad a.s.,$$

where $\theta_\alpha^2 = \frac{c_\alpha\sigma^2}{(1-\alpha)(3-2\alpha)}$.

PROOF OF PROPOSITION 3. Let $H_1 = 3/2 - \alpha_1$. By Lemma A.2 above, we have

$$\sup_{0\leqslant r\leqslant 1}\left|\frac{1}{n^{H_1}}\sum_{t=1}^{\lfloor nr\rfloor} x_t^1 - \theta_1 \mathsf{B}_{H_1}(r)\right| = o_P(1),$$

which indicates that

$$\frac{1}{n^{H_1}}\max_{1\leqslant k\leqslant n}\sum_{t=1}^{k}(x_t^1-\overline{x}^1) \xrightarrow{P} \theta_1 \sup_{0\leqslant r\leqslant 1}\left[\mathsf{B}_{H_1}(r) - r\mathsf{B}_{H_1}(1)\right]$$

and

$$\frac{1}{n^{H_1}}\min_{1\leqslant k\leqslant n}\sum_{t=1}^{k}(x_t^1-\overline{x}^1) \xrightarrow{P} \theta_1 \inf_{0\leqslant r\leqslant 1}\left[\mathsf{B}_{H_1}(r) - r\mathsf{B}_{H_1}(1)\right].$$

Thus, we have

$$\frac{1}{n^{H_1}}R_n \xrightarrow{P} \theta_1 \left\{\sup_{0\leqslant r\leqslant 1}\left[\mathsf{B}_{H_1}(r) - r\mathsf{B}_{H_1}(1)\right] - \inf_{0\leqslant r\leqslant 1}\left[\mathsf{B}_{H_1}(r) - r\mathsf{B}_{H_1}(1)\right]\right\}. \tag{A.21}$$

On the other hand, by the ergodic theorem, we may show that

$$(S_n^*)^2 = \frac{1}{n}\sum_{t=1}^{n}\left(x_t^1\right)^2 - \left(\overline{x}^1\right)^2 \xrightarrow{P} \mathsf{E}\left[(x_t^1)^2\right] > 0. \tag{A.22}$$

By (A.21) and (A.22), we have

$$\frac{1}{n^{H_1}}\frac{R_n}{S_n^*} \xrightarrow{P} \mathsf{V}, \tag{A.23}$$

where

$$\mathsf{V} = \left\{\mathsf{E}\left[(x_t^1)^2\right]\right\}^{-1/2}\theta_1 \left\{\sup_{0\leqslant r\leqslant 1}\left[\mathsf{B}_{H_1}(r) - r\mathsf{B}_{H_1}(1)\right] - \inf_{0\leqslant r\leqslant 1}\left[\mathsf{B}_{H_1}(r) - r\mathsf{B}_{H_1}(1)\right]\right\}.$$

Using Theorem 2, we may show that

$$\frac{1}{n^{H_1}}\frac{\widehat{R}_n}{\widehat{S}_n^*} - \frac{1}{n^{H_1}}\frac{R_n}{S_n^*} = o_P(1),$$

which together with (A.23) implies that

$$\frac{1}{n^{H_1}}\frac{\widehat{R}_n}{\widehat{S}_n^*}\xrightarrow{P} V, \tag{A.24}$$

completing the proof of Proposition 3. $\square$

PROOF OF PROPOSITION 4. Using Lemma 4.2 in Bosq (2000) and Proposition 2 in the main document and noting that

$$\lambda_i = \lim_{m\to\infty}\frac{\lambda_{m,i}}{m^{3-2\alpha_1}} = 0, \ \ i = p_1+1, \cdots,$$

where $\lambda_{m,i} = \lambda_{mi}$, we may show that

$$\frac{\widehat{\lambda}_{m,i}-\lambda_{m,i}}{m^{3-2\alpha_1}} = o_P(1) \ \text{ for } i = 1, \cdots, p_1, \tag{A.25}$$

and

$$\widehat{\lambda}_{m,i} = o_P\left(m^{3-2\alpha_1}\right) \ \text{ for } i = p_1+1, \cdots. \tag{A.26}$$

By (A.25), we readily have that for $i = 1, \cdots, p_1-1$,

$$\frac{\widehat{\lambda}_{m,i+1}}{\widehat{\lambda}_{m,i}} = \frac{\widehat{\lambda}_{m,i+1}/m^{3-2\alpha_1}}{\widehat{\lambda}_{m,i}/m^{3-2\alpha_1}} \xrightarrow{P} \frac{\lambda_{i+1}}{\lambda_i}, \tag{A.27}$$

which, together with (3.12) in Proposition 1, indicates that for $i = 1, \cdots, p_1-1$, $\widehat{\lambda}_{m,i+1}/\widehat{\lambda}_{m,i}$ is strictly larger than a positive constant with probability approaching one. Similarly, when $i = p_1$, by (A.25) and (A.26), we can prove that

$$\frac{\widehat{\lambda}_{m,p+1}}{\widehat{\lambda}_{m,p}} = \frac{\widehat{\lambda}_{m,p+1}/m^{3-2\alpha_1}}{\widehat{\lambda}_{m,p}/m^{3-2\alpha_1}} = o_P(1)/\lambda_p = o_P(1). \tag{A.28}$$

Finally, for $i = p_1+1, \cdots, \bar{P}$, by (A.26), we readily have that

$$\mathsf{P}\left(\left|\widehat{\lambda}_{m,i+1}/\widehat{\lambda}_{m,1}\right| < \epsilon_\star, \ \left|\widehat{\lambda}_{m,i}/\widehat{\lambda}_{m,1}\right| < \epsilon_\star\right) \to 1,$$

which, together with (3.15), indicates that

$$\mathsf{P}\left(\widehat{\lambda}_{m,i+1}/\widehat{\lambda}_{m,i} = 0/0 = 1\right) \to 1. \tag{A.29}$$

With (A.27)–(A.29), we complete the proof of Proposition 4. $\square$

PROOF OF PROPOSITION 5. (i) Observe that

$$\begin{aligned}\overline{X}_t &= \sum_{i=0}^{\infty} \beta_{i,-d} \cdot \pi\left[\mathbf{A}_\infty(\overline{\boldsymbol{\eta}}_{t-i})\right] \\ &= \sum_{i=0}^{\infty} \beta_i^* \cdot \pi\left[\mathbf{A}_\infty(\overline{\boldsymbol{\eta}}_{t-i})\right] + \sum_{i=0}^{\infty} \beta_i^\diamond \cdot \pi\left[\mathbf{A}_\infty(\overline{\boldsymbol{\eta}}_{t-i})\right] \\ &=: \overline{X}_t^* + \overline{X}_t^\diamond,\end{aligned}$$

where β_i^* and $\beta_i^\diamond$ are defined in (4.7). Define

$$\overline{x}_t^{j,*} = \langle \overline{X}_t^*, \psi_j^\diamond \rangle = \sum_{i=0}^{\infty} \beta_i^* \cdot \left\langle \pi\left[\mathbf{A}_\infty(\overline{\boldsymbol{\eta}}_{t-i})\right], \psi_j^\diamond \right\rangle =: \sum_{i=0}^{\infty} \beta_i^* \eta_{t-i}^j$$

and

$$\overline{x}_t^{j,\diamond} = \langle \overline{X}_t^\diamond, \psi_j^\diamond \rangle = \sum_{i=0}^{\infty} \beta_i^\diamond \cdot \left\langle \pi\left[\mathbf{A}_\infty(\overline{\boldsymbol{\eta}}_{t-i})\right], \psi_j^\diamond \right\rangle =: \sum_{i=0}^{\infty} \beta_i^\diamond \eta_{t-i}^j,$$

where $\psi_j^\diamond(\cdot)$, $j = 1, \cdots, p_\diamond$, are defined in (4.10). By the definitions of β_i^* and $\beta_i^\diamond$, $\{\overline{x}_t^{j,*} : t \in \mathcal{Z}\}$ is a sequence of stationary long-range dependent random variables, whereas $\{\overline{x}_t^{j,\diamond} : t \in \mathcal{Z}\}$ is a sequence of stationary short-range dependent random variables.

Note that

$$c_n^\diamond(u,v) = \mathsf{E}\left[\sum_{t=1}^{n}\sum_{s=1}^{n} \overline{X}_t(u)\overline{X}_s(v)\right], \quad u, v \in \mathcal{C}.$$

By Assumption 4, the operator $C_n^\diamond$, which is defined by

$$C_n^\diamond(x)(u) = \int_{\mathcal{C}} c_n^\diamond(u,v)x(v)dv, \ \ x \in \mathcal{H},$$

is $p_\diamond$-dimensional with $(\lambda_{nj}^\diamond, \psi_j^\diamond)$ as pairs of eigenvalue and eigenfunction, where for $j = 1, \cdots, p_\diamond$,

$$\begin{aligned}\lambda_{nj}^\diamond &= \int_{\mathcal{C}}\int_{\mathcal{C}} c_n^\diamond(u,v)\psi_j^\diamond(v)\psi_j^\diamond(u)dvdu \\ &= \sum_{t=1}^{n}\sum_{s=1}^{n} \mathsf{E}\left[\left(\overline{x}_t^{j,*} + \overline{x}_t^{j,\diamond}\right)\left(\overline{x}_s^{j,*} + \overline{x}_s^{j,\diamond}\right)\right].\end{aligned}$$

By some elementary calculations, we may show that

$$\sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(\overline{x}_t^{j,*}\overline{x}_s^{j,*}\right)\sim\frac{c_{1-d}\lambda_j^{\diamond}}{d(1+2d)\Gamma^2(d)}n^{2d+1},$$
$$\sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(\overline{x}_t^{j,\diamond}\overline{x}_s^{j,*}\right)+\sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(\overline{x}_t^{j,*}\overline{x}_s^{j,\diamond}\right)+\sum_{t=1}^{n}\sum_{s=1}^{n}\mathsf{E}\left(\overline{x}_t^{j,\diamond}\overline{x}_s^{j,\diamond}\right)=O(n),$$

which leads to (4.12). Therefore, Proposition 1 holds with $p_1=p_\diamond$ and $\kappa_0=1$.

(ii) By the definition of $\widetilde{X}_t$, we have

$$\begin{aligned}
\widetilde{X}_t &= \sum_{i=0}^{\infty}\beta_{i,-d}\sum_{j=0}^{\infty}\pi\left[\mathbf{A}_j(\overline{\boldsymbol{\eta}}_{t-i-j})\right]-\sum_{i=0}^{\infty}\beta_{i,-d}\sum_{j=0}^{\infty}\pi\left[\mathbf{A}_j(\overline{\boldsymbol{\eta}}_{t-i})\right]\\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{i}\beta_{j,-d}\pi\left[\mathbf{A}_{i-j}(\overline{\boldsymbol{\eta}}_{t-i})\right]-\sum_{i=0}^{\infty}\beta_{i,-d}\sum_{j=0}^{i}\pi\left[\mathbf{A}_j(\overline{\boldsymbol{\eta}}_{t-i})\right]-\sum_{i=0}^{\infty}\beta_{i,-d}\sum_{j=i+1}^{\infty}\pi\left[\mathbf{A}_j(\overline{\boldsymbol{\eta}}_{t-i})\right]\\
&= -\sum_{i=0}^{\infty}\beta_{i,-d}\sum_{j=i+1}^{\infty}\pi\left[\mathbf{A}_j(\overline{\boldsymbol{\eta}}_{t-i})\right]+\sum_{i=0}^{\infty}\sum_{j=0}^{i}\beta_{j,-d}\pi\left[\mathbf{A}_{i-j}(\overline{\boldsymbol{\eta}}_{t-i})\right]-\sum_{i=0}^{\infty}\sum_{j=0}^{i}\beta_{i,-d}\pi\left[\mathbf{A}_{i-j}(\overline{\boldsymbol{\eta}}_{t-i})\right]\\
&= -\sum_{i=0}^{\infty}\beta_{i,-d}\cdot\pi\left[\tilde{\mathbf{A}}_i(\overline{\boldsymbol{\eta}}_{t-i})\right]+\sum_{i=0}^{\infty}\pi\left[\mathbf{A}_i^{\beta}(\overline{\boldsymbol{\eta}}_{t-i})\right], \qquad \text{(A.30)}
\end{aligned}$$

where

$$\widetilde{\mathbf{A}}_i=\sum_{j=i+1}^{\infty}\mathbf{A}_j,\quad \mathbf{A}_i^{\beta}=\sum_{j=0}^{i}(\beta_{j,-d}-\beta_{i,-d})\mathbf{A}_{i-j}.$$

By Assumption 3 and Minkowski's inequality, we have

$$\|\widetilde{\mathbf{A}}_i\|_{\mathcal{S}}\leqslant\sum_{j=i+1}^{\infty}\|\mathbf{A}_j\|_{\mathcal{S}}\leqslant M\cdot\rho_\star^i,\quad 0<\rho_\star<1, \qquad \text{(A.31)}$$

when i is sufficiently large and M is a positive constant. On the other hand,

$$\begin{aligned}
\mathbf{A}_i^{\beta} &= \sum_{j=0}^{i}(\beta_{j,-d}-\beta_{i,-d})\mathbf{A}_{i-j}\\
&= \sum_{j=0}^{\lfloor i/2\rfloor}(\beta_{j,-d}-\beta_{i,-d})\mathbf{A}_{i-j}+\sum_{j=\lfloor i/2\rfloor+1}^{i}(\beta_{j,-d}-\beta_{i,-d})\mathbf{A}_{i-j}. \qquad \text{(A.32)}
\end{aligned}$$

By (4.7) and the mean-value theorem, for $j = \lfloor i/2 \rfloor + 1, \cdots, i$, we have

$$|\beta_{j,-d} - \beta_{i,-d}| \leqslant M \cdot i^{-2+d}(i - j). \tag{A.33}$$

Then, by (A.31)–(A.33) and Assumption 3, we readily have

$$\|\mathbf{A}_i^{\beta}\|_{\mathcal{S}} \leqslant M \left(i\rho_\star^{i/2} + i^{-2+d} \sum_{j=0}^{\infty} j\rho_\star^j \right) \leqslant M \cdot i^{-2+d} \tag{A.34}$$

when i is sufficiently large. By (A.30), (A.31) and (A.33), we can prove that $\{\widetilde{X}_t : t \in \mathcal{Z}\}$ is a sequence of short-range dependent curve time series and the Hilbert-Schmidt norm of $C_n^\star$ has the order of $O(n)$, where

$$C_n^\star(x)(u) = \int_{\mathcal{C}} c_n^\star(u,v)x(v)dv, \;\; c_n^\star(u,v) = \mathsf{E}\left[\sum_{t=1}^{n}\sum_{s=1}^{n} \widetilde{X}_t(u)\widetilde{X}_s(v)\right].$$

The proof of Proposition 5 has been completed. □

B Additional numerical results

In this appendix, we report additional numerical result obtained from the simulation studies and empirical data analysis.

B.1 R/S estimation result using normalised scores

From Tables 1 and 4 in the main document (Li et al. 2018), $\widehat{p}_1$, the estimated number of orthonormal functions which span $\mathcal{S}_1$, could be larger than 1. Letting $\widehat{x}_t^i$, $i = 1, \cdots, \widehat{p}_1$, be the approximate score processes defined as $\widehat{x}_t^1$ in Section 3.2, it is a natural idea to normalise components in the multivariate score process $\left(\widehat{x}_t^i,\ i = 1, \cdots, \widehat{p}_1\right)$ and then apply the R/S method to the normalised scores. Specifically, define

$$\widehat{x}_t^N = \sqrt{\sum_{i=1}^{\widehat{p}_1} (\widehat{x}_t^i)^2}, \qquad t = 1, \cdots, n,$$

and

$$\widehat{\alpha}_N = 3/2 - \widehat{H}_N, \;\; \widehat{H}_N = \log\left(\widehat{R}_n^N/\widehat{S}_n^N\right)/\log n,$$

where $\widehat{R}_n^N$ and $\widehat{S}_n^N$ are defined similarly to $\widehat{R}_n$ and $\widehat{S}_n^*$ with $\widehat{x}_t^1$ replaced by $\widehat{x}_t^N$. Table 1 reports the mean and median of the estimates of $d = 1 - \alpha_1$ (with 0.2 as the true value) in the simulation studies using R/S for the normalised scores, where both the CPV and ratio methods are used to select $\widehat{p}_1$. For comparison, we also present R/S estimation result in Table 2 using only the first set of scores. Comparing the results between Tables 1 and 2, the finite-sample performance of R/S using normalised scores is similar to that using only the first score.

Table 1: Mean and median of the R/S estimates of d using normalised scores.

Method	Model		$n = 500$	$n = 1000$	$n = 2000$
CPV	FARIMA(1,0.2, 0)	Mean	0.1572	0.1552	0.1526
		Median	0.1611	0.1617	0.1629
	FARIMA(1, 0.2,1)	Mean	0.1774	0.1778	0.1775
		Median	0.1773	0.1777	0.1780
Ratio	FARIMA(1, 0.2, 0)	Mean	0.1620	0.1616	0.1593
		Median	0.1620	0.1621	0.1625
	FARIMA(1, 0.2, 1)	Mean	0.1774	0.1777	0.1776
		Median	0.1773	0.1777	0.1780

Table 2: Mean and median of the R/S estimates of d using the first set of scores.

Model		$n = 500$	$n = 1000$	$n = 2000$
FARIMA(1,0.2,0)	Mean	0.1587	0.1630	0.1648
	Median	0.1589	0.1627	0.1663
FARIMA(1,0.2,1)	Mean	0.1738	0.1777	0.1785
	Median	0.1740	0.1777	0.1794

B.2 Estimation of p_1 in incremental fertility rate data analysis

In Table 3, we present the retained number of orthonormal functions spanning $\mathcal{S}_1$ in the incremental fertility rate data analysis (see Section 5.3 in the main document) based on the CPV and ratio methods.

Table 3: Estimates of p_1 for the incremental fertility rate data in 14 developed countries.

Country	CPV	Ratio	Country	CPV	Ratio
Australia	2	2	Canada	2	2
Denmark	2	2	Finland	2	2
France	2	2	Germany	1	1
Iceland	3	1	Italy	2	2
Netherland	1	1	Spain	2	2
Sweden	2	3	Switzerland	2	2
UK	2	2	USA	2	2

B.3 Confidence interval via maximum entropy bootstrap

In Tables 3 and 6 of the main document, we present the pointwise estimated values of α_1 for the two empirical data sets by applying the R/S method to the first set of estimated scores. In order to quantify estimation uncertainty, one typically constructs the confidence intervals. Based on the first set of the estimated scores, we next introduce a maximum entropy bootstrap method (e.g., Vinod 2004, Shang 2018). The advantages of the maximum entropy bootstrap are: (1) the stationarity condition is not required; (2) the bootstrap technique computes ranks of a time series, robust against outliers of the univariate time series; (3) the bootstrap samples satisfy the ergodic theorem, central limit theorem and mean preserving constraint; (4) the bootstrap samples are adjusted so that the population variance of the maximum entropy density equals that of the original data. The detailed steps of the bootstrap algorithm are described as follows (Vinod and de Lacalle 2009).

1. Sort the first set of estimated principal component scores in increasing order to create order statistics $\widehat{x}^1_{(t)}$, $t = 1, \cdots, n$, and store the ordering index vector.

2. Compute immediate points $z_t = \left(\widehat{x}^1_{(t)} + \widehat{x}^1_{(t+1)}\right)/2$, $t = 1, \cdots, n-1$, from the order statistics constructed in Step 1.

3. Compute the trimmed mean of deviations $\widehat{x}_t^1 - \widehat{x}_{t-1}^1$ among our consecutive observations, i.e.,
$$m_{\text{trim}} = \frac{1}{n-1} \sum_{t=2}^{n-1} \left(\widehat{x}_t^1 - \widehat{x}_{t-1}^1 \right).$$
Compute the lower limit for the left tail as $z_0 = \widehat{x}_{(1)}^1 - m_{\text{trim}}$ and the upper limit for the right tail as $z_n = \widehat{x}_{(n)}^1 + m_{\text{trim}}$. These limits become the limiting intermediate points.

4. Compute the mean of the maximum entropy density within each interval such that the "mean-preserving constraint" is satisfied (Vinod 2004). The interval means $m_1, \ldots, m_n$ are given by
$$\begin{aligned} m_1 &= \tfrac{3}{4}\widehat{x}_{(1)}^1 + \tfrac{1}{4}\widehat{x}_{(2)}^1, \\ m_t &= \tfrac{1}{4}\widehat{x}_{(t-1)}^1 + \tfrac{1}{2}\widehat{x}_{(t)}^1 + \tfrac{1}{4}\widehat{x}_{(t+1)}^1, \quad t = 2, \cdots, n-1, \\ m_n &= \tfrac{1}{4}\widehat{x}_{(n-1)}^1 + \tfrac{3}{4}\widehat{x}_{(n)}^1. \end{aligned}$$

5. Generate random numbers from Uniform $[0, 1]$, compute sample quantiles of the maximum entropy density at those points and sort them.

6. Re-order the sorted sample quantiles by using the ordering index of Step 1. This step recovers the temporal dependence of the initial data.

7. Adjust the variance of bootstrap samples so that the population variance of the maximum entropy density equals that of original data (e.g., Vinod 2004).

8. Repeat Steps 2–7 for 1000 times and apply the R/S method to estimate the long memory parameter for each time. Finally, with the 1000 estimated values, we construct the 95% pointwise confidence interval via sample quantiles.

In Tables 4 and 5, we present the 95% pointwise confidence intervals of the memory parameters for the stock and fertility data sets, respectively. The results in Tables 4 and 5 provide further support to that the US stock prices and age-specific fertility rates in our empirical analysis are long-range dependent over time.

Table 4: 95% pointwise confidence intervals of α_1 for the log returns of the six stocks.

Stock	Confidence interval	Stock	Confidence interval
S&P 500	(0.8539, 0.9027)	AAPL	(0.7948, 0.8052)
EC	(0.7724, 0.7980)	WFC	(0.7932, 0.8089)
CL	(0.7450, 0.7616)	XOM	(0.8054, 0.8201)

Table 5: 95% pointwise confidence intervals of α_1 for the incremental fertility rate.

Country	95% confidence interval	Country	95% confidence interval
Australia	(0.7576, 0.7825)	Canada	(0.7567, 0.7762)
Denmark	(0.7255, 0.7414)	Finland	(0.7666, 0.8229)
France	(0.7113, 0.7577)	Germany	(0.7152, 0.7458)
Iceland	(0.8907, 0.9566)	Italy	(0.7355, 0.7648)
Netherland	(0.7129, 0.7482)	Spain	(0.8261, 0.8567)
Sweden	(0.7125, 0.7321)	Switzerland	(0.7340, 0.7499)
UK	(0.7661, 0.8253)	USA	(0.7595, 0.7770)

B.4 Local Whittle estimation

The local Whittle estimation is a semiparametric method to estimate the Hurst parameter based on the periodogram. It is introduced by Künsch (1987) and later developed by Robinson (1995), Velasco (1999) and subsequent authors. Note that the spectral density $f(\lambda)$ of stationary time series is usually assumed to satisfy that

$$f(\lambda) \sim G\lambda^{-2d} \text{ as } \lambda \to 0+, \tag{B.1}$$

where $0 < G < \infty$ and $-1/2 < d < 1/2$. Like in Section 4 of the main document, we only consider $0 < d < 1/2$ as the main interest lies in long-range dependence. Recall that the first score process $\{x_t^1 : t \in \mathbb{Z}\}$ defined in (2.4) of the main document is a univariate long-range dependent linear process. So it is sensible to assume that its spectral density satisfies (B.1) with $d = 1 - \alpha_1$. We next introduce the local Whittle method to estimate the unknown parameters G and d (or equivalently α_1) in (B.1).

Define the periodogram of the first score process $\{x_t^1 : t \in \mathbb{Z}\}$ as

$$w(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} x_t^1 e^{it\lambda}, \quad I(\lambda) = |\omega(\lambda)|^2.$$

Following the notation of Velasco (1999), we consider the objective function:

$$Q(G,d) = \frac{1}{m_\diamond}\sum_{j=1}^{m_\diamond}\left\{\log(G\lambda_j^{-2d}) + \frac{I(\lambda_j)}{G\lambda_j^{-2d}}\right\}, \tag{B.2}$$

where $\lambda_j = (2\pi j)/n$, $j = 1, \cdots, m_\diamond$, and $m_\diamond$ is a positive integer satisfying $m_\diamond \to \infty$ and $m_\diamond = o(n)$ (e.g., Robinson 1995). As in Robinson (1995), we define the estimates

$$\left(\widehat{G}, \widehat{d}\right) = \underset{0<G<\infty, d\in\Theta}{\arg\min}\; Q(G,d), \tag{B.3}$$

where $\Theta = [\Delta_1, \Delta_2]$, Δ_1 and Δ_2 are numbers picked such that $-1/2 < \Delta_1 < \Delta_2 < 1/2$. Alternatively, we may obtain

$$\widehat{d} = \underset{d\in\Theta}{\arg\min}\, R(d) \tag{B.4}$$

where

$$R(d) = \log\widehat{G}(d) - \frac{2d}{m_\diamond}\sum_{j=1}^{m_\diamond}\log\lambda_j, \quad \widehat{G}(d) = \frac{1}{m_\diamond}\sum_{j=1}^{m_\diamond}\lambda_j^{2d} I(\lambda_j).$$

Then, the local Whittle estimate of α_1 can be constructed as $1-\widehat{d}$. In practice, as x_t^1, $t = 1, \cdots, n$, are unobservable, we approximate them by $\widehat{x}_t^1$ in (B.3) and (B.4).

Table 6 below presents the local Whittle estimated values of α_1 for the log returns of the six US stocks, where $m_\diamond = \lfloor 1 + n^{0.65}\rfloor$. These values are generally smaller than those obtained by the R/S method reported in Table 3 of the main document.

Table 6: Local Whittle estimates for α_1 of the log returns of the six stocks.

Stock	Local Whittle estimates
S&P 500	0.7580
EC	0.6979
CL	0.5441
AAPL	0.6946
WFC	0.6761
XOM	0.6989

References

Billingsley, P. (1968), *Convergence of Probability Measure*, Wiley, New York.

Bosq, D. (2000), *Linear Processes in Function Spaces*, Springer, New York.

Giraitis, L., Koul, H. & Surgailis, D. (2012), *Large Sample Inference for Long Memory Processes*, Imperial College Press, London.

Ibragimov, I. A. & Linnik, Y. V. (1971), *Independence and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen.

Künsch, H. R. (1987), Statistical aspects of self-similar processes, *in* Y. A. Prohorov & V. V. Sazonov, eds, 'Proceedings 1st World Congress of the Bernoulli Society', VNU Science Press, Utrecht, pp. 67–74.

Li, D., Robinson, P. M. & Shang, H. L. (2018), Long-range dependent curve time series, Technical report, University of York.
URL: https://www.researchgate.net/publication/329209891_Long-Range_Dependent_Curve_Time_Series

Robinson, P. M. (1995), 'Gaussian semiparametric estimation of long range dependence', *The Annals of Statistics* **23**(5), 1630–1661.

Robinson, P. M. (1997), 'Large-sample inference for nonparametric regression with dependent errors', *The Annals of Statistics* **25**(5), 2054–2083.

Shang, H. L. (2018), 'Bootstrap methods for stationary functional time series', *Statistics and Computing* **28**(1), 1–10.

Velasco, C. (1999), 'Gaussian semiparametric estimation of non-stationary time series', *Journal of Time Series Analysis* **20**(1), 87–127.

Vinod, H. D. (2004), 'Ranking mutual funds using unconventional utility theory and stochastic dominance', *Journal of Empirical Finance* **11**(3), 353–377.

Vinod, H. D. & de Lacalle, J. L. (2009), 'Maximum entropy bootstrap for time series: the meboot R package', *Journal of Statistical Software* **29**(5).

Wang, Q., Lin, Y. & Gulati, C. (2003), 'Strong approximation for long memory processes with applications', *Journal of Theoretical Probability* **16**(2), 377–389.