Supplementary appendix to:
Learning can generate Long Memory*

March 19, 2015

Contents

*Emails for correspondance: chevillon@essec.edu and sophocles.mavroeidis@economics.ox.ac.uk

1 Empirical results with hyperbolic gain

This section considers learning algorithm with hyperbolic weights such that $y_{t+1}^{e} = \kappa(L) y_t$ with $\kappa(L) = 1 - (1-L)^g$ for $g \in (0,1)$; hence $\delta_\kappa = 1 - g$. Table S.1 reports the minimum parameter g for which the null hypothesis is not rejected, *i.e.*, the minimum value of g that is consistent with the memory of y_t under hyperbolically discounted least squares learning when x_t has short memory. The minimum g that satisfy this property is smaller, by a range of 0.10-0.20, than the degrees of long memory reported in Table 5. The UK is an exception for which the minimum g is much smaller, between 0.07 and 0.15 depending on the test statistic considered.

2 Proof of Lemma 6 in the paper

For $\delta_\kappa \in (0,1)$, $\delta_\kappa - 2 \in (-2,-1)$. Yong (1974), Theorems III-24 and -27, show that for $\delta_\kappa \in (0,1)$, if there exists a function S slowly varying at infinity such that $a_j = j^{\delta_\kappa - 2} S(j)$, then

$$\sum_{j=1}^{\infty} a_j \cos(j\omega) - \sum_{j=1}^{\infty} a_j \underset{\omega\to0^+}{\sim} \frac{\pi}{2\Gamma(2-\delta_k)\cos\frac{(2-\delta_k)\pi}{2}} \omega^{1-\delta_\kappa} S\left(\frac{1}{\omega}\right)$$
$$\sum_{j=1}^{\infty} a_j \sin(j\omega) \underset{\omega\to0^+}{\sim} \frac{\pi}{2\Gamma(2-\delta_k)\sin\frac{(2-\delta_k)\pi}{2}} \omega^{1-\delta_\kappa} S\left(\frac{1}{\omega}\right).$$

Define, for $x \geq 1$, $S(x) = \kappa_{\lfloor x\rfloor} / \lfloor x \rfloor^{\delta_\kappa - 2}$, where $\lfloor x \rfloor$ is the integer part of x. Then as $x \to \infty$ and for $\lambda \geq 1/x$,

$$S(\lambda x)/S(x) = \frac{\kappa_{\lfloor \lambda x\rfloor}}{\kappa\lfloor x\rfloor} \frac{\lfloor x\rfloor^{\delta_\kappa-2}}{\lfloor \lambda x\rfloor^{\delta_\kappa-2}} \to 1,$$

so S is slowly varying with $S\left(\frac{1}{\omega}\right) \to c_\kappa$ as $\omega \to 0$. This implies, using $\kappa_j = j^{\delta_\kappa-2}S(j)$ for $j \geq 1$ and $\sum_{j=0}^{\infty} \kappa_j = \kappa(1) = 1$, that

$$\kappa\left(e^{-i\omega}\right) - 1 \underset{\omega\to0^+}{\sim} \frac{\pi c_\kappa}{2\Gamma(2-\delta_\kappa)} \left[-\frac{1}{\cos\frac{\pi\delta_\kappa}{2}} + i\frac{1}{\sin\frac{\pi\delta_\kappa}{2}}\right] \omega^{1-\delta_\kappa},$$

i.e. the result holds for $\mathrm{Re}\left(\kappa\left(e^{i\omega}\right) - 1\right)$ setting $c_\kappa^* = \frac{\pi c_\kappa}{2\Gamma(2-\delta_\kappa)} \frac{1}{\cos\frac{\pi\delta_\kappa}{2}}$ such that $c_\kappa c_\kappa^* > 0$. Also, using

$$\left|\frac{-1}{\cos\frac{\pi\delta_\kappa}{2}} + i\frac{1}{\sin\frac{\pi\delta_\kappa}{2}}\right|^2 = \left(\cos\frac{\pi\delta_\kappa}{2}\sin\frac{\pi\delta_\kappa}{2}\right)^{-2} = \left(\frac{\sin\pi\delta_\kappa}{2}\right)^{-2},$$

and $\Gamma(1+z) = z\Gamma(z)$ together with $\Gamma(1-z)\Gamma(z) = \frac{\pi}{\sin \pi z}$, we obtain

$$\left|1-\kappa\left(e^{i\omega}\right)\right|^2 \sim \frac{c_\kappa^2 \Gamma(\delta_\kappa)^2}{(1-\delta_\kappa)^2}\omega^{2(1-\delta_\kappa)}, \tag{1}$$

and $c_\kappa^{**} = \frac{c_\kappa^2 \Gamma(\delta_\kappa)^2}{(1-\delta_\kappa)^2} > 0$.

3 Proof of Lemma 7 in the paper

Consider

$$\begin{aligned} f_y(\omega) - f_y(0) &= \frac{(f_x(\omega) - f_x(0))(1-\beta)^2}{(1-\beta)^2 \left|1-\beta+\beta\left(1-\kappa\left(e^{-i\omega}\right)\right)\right|^2} \\ &\quad - f_x(0)\frac{2\beta(1-\beta)\,\mathrm{Re}\left[1-\kappa\left(e^{-i\omega}\right)\right] + \beta^2\left|1-\kappa\left(e^{-i\omega}\right)\right|^2}{(1-\beta)^2\left|1-\beta+\beta\left(1-\kappa\left(e^{-i\omega}\right)\right)\right|^2}, \end{aligned}$$

since

$$\left|1-\beta+\beta\left(1-\kappa\left(e^{-i\omega}\right)\right)\right|^2 = (1-\beta)^2 - 2\beta(1-\beta)\,\mathrm{Re}\left[\kappa\left(e^{-i\omega}\right)-1\right] - \beta^2\left|\kappa\left(e^{-i\omega}\right)-1\right|^2.$$

Under Assumption B, $|f_x'(0)| < \infty$ (see Stock, 1994), so $f_x(\omega) - f_x(0) = O(\omega)$. Now if $\delta_\kappa > 0$, under constant learning, $\kappa_j \sim c_\kappa j^{\delta_\kappa - 2}$ for some $c_\kappa \neq 0$. Lemma 6 implies that there exist $c_\kappa^* \neq 0$, with $c_\kappa c_\kappa^* > 0$, and $c_\kappa^{**} > 0$ such that

$$\begin{aligned} f_y(\omega) - f_y(0) &\underset{\omega \to 0^+}{\sim} \frac{-2\beta(1-\beta) f_x(0) c_\kappa^* \omega^{1-\delta_\kappa} + \beta^2 f_x(0) c_\kappa^{**}\omega^{2(1-\delta_\kappa)}}{(1-\beta)^4} \\ &\underset{\omega \to 0^+}{\sim} -\frac{2\beta f_x(0) c_\kappa^*}{(1-\beta)^3}\omega^{1-\delta_\kappa}. \end{aligned} \tag{2}$$

We first note that by definition of the population spectrum

$$f_y(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k e^{-i\omega k} = \frac{1}{2\pi}\left\{\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \cos \omega k\right\},$$

where $\gamma_k = \mathsf{Cov}(y_t, y_{t-k})$, is symmetric since y_t is stationary. We assume for now that γ_k is of bounded variation, then, for $\omega \neq 0$, the series $\sum_{k=1}^n \gamma_k \cos \omega k$ converges uniformly as $n \to \infty$ (see Zygmund, 1935, Section. 1.23). It follows that the derivative of f_y satisfies:

$$f_y'(\omega) = -\frac{1}{\pi}\sum_{k=1}^{\infty} k\gamma_k \sin k\omega. \tag{3}$$

We now use Theorem III-11 of Yong (1974) who works under the assumption that $\{a_k\}_{k\in\mathbb{N}}$ is a sequence of positive numbers that is quasi-monotonically convergent to zero (i.e. $a_k \to 0$

and there exist $M \geq 0$, such that $a_{k+1} \leq a_k \left(1 + \frac{M}{k}\right)$ for all $k \geq k_0(M)$) and that $\{a_k\}$ is also of bounded variation, i.e. $\sum_{k=1}^{\infty} |\Delta a_k| < \infty$. The theorem states that for $a \in (0, 1)$, $a_k \sim k^{-a} S^*(k)$ as $k \to \infty$, with S^* slowly varying, if and only if

$$\sum_{k=1}^{\infty} a_k \sin k\omega \sim \frac{\pi}{2\Gamma(a) \sin \frac{\pi a}{2}} \omega^{a-1} S^* \left(\frac{1}{\omega}\right) \text{ as } \omega \to 0^+.$$

We apply this theorem to (3), using expression (2) that $f_y'(\omega) \underset{\omega \to 0^+}{\sim} -\frac{2\beta f_x(0) c_\kappa^*}{(1-\beta)^3} \omega^{-\delta_\kappa}$

$$-\frac{1}{\pi} \sum_{k=1}^{\infty} k\gamma_k \sin k\omega \underset{\omega \to 0^+}{\sim} -\frac{2\beta f_x(0) c_\kappa^*}{(1-\beta)^3} \omega^{-\delta_\kappa}.$$

We let $a = 1 - \delta_\kappa$ in the theorem of Yong above, defining

$$a_k = \frac{k\gamma_k}{\pi} \frac{(1-\beta)^3}{2\beta f_x(0) c_\kappa^*}.$$

This implies that $a_k \sim \frac{\pi}{2\Gamma(a) \sin \frac{\pi a}{2}} k^{-(1-\delta_\kappa)}$, with $\Gamma(1-\delta_\kappa) \sin \frac{\pi(1-\delta_\kappa)}{2} = \frac{\pi}{2\Gamma(\delta_\kappa) \sin \frac{\pi \delta_\kappa}{2}}$, i.e.

$$\gamma_k \sim \frac{2\pi\beta f_x(0) c_\kappa^* \Gamma(\delta_\kappa) \sin \frac{\pi\delta_\kappa}{2}}{(1-\beta)^3} k^{-(2-\delta_\kappa)}. \tag{4}$$

To apply Theorem III-11 of Yong (1974), we check that a_k thus defined is a quasi-monotonic sequence with bounded variation. The first holds since $a_{k+1}/a_k \sim (1 + 1/k)^{-(1-\delta_\kappa)} < 1$, so choose M such that $a_{k+1}/a_k < 1$ for $k > M$. Also $k\gamma_k$ is clearly of bounded variation since it is asymptotically positive and

$$\Delta(k\gamma_k) = O\left(k^{-(2-\delta_\kappa)}\right)$$

is summable. Finally, we check the uniform convergence condition in Zygmund (1935): $|\Delta\gamma_k| = O\left(k^{-(3-\delta_\kappa)}\right)$ so γ_k is of bounded variation.

4 Proof of Theorem 8 in the paper

We present in turn the proofs for the spectral density and the autocorrelation. In the proofs we omit the dependence of β on T for ease of notation.

4.1 Spectral density

We consider the behavior of the spectral density of y_t about the origin under the assumption that $\kappa_j \sim c_\kappa j^{\delta_\kappa - 2}$ so define $(c^*_\kappa, c^{**}_\kappa)$ as in Lemma 6. Let $\beta = 1 - c_\beta T^{-\nu}$, $\nu \in [0,1]$. As $\omega \to 0^+$, the spectral density of f_y is, for $\delta_\kappa \in (1/2, 1)$:

$$f_y(\omega) = \frac{f_x(\omega)}{\left|1 - \beta\kappa\left(e^{-i\omega}\right)\right|^2} = \frac{f_x(\omega)}{\left|1 - \beta + \beta\left(1 - \kappa\left(e^{-i\omega}\right)\right)\right|^2}, \tag{5}$$

which implies

$$f_y(\omega) \tag{6}$$
$$= \frac{f_x(\omega)}{(1-\beta)^2 - 2\beta c^*_\kappa (1-\beta)\,\omega^{1-\delta_\kappa} + \beta^2 c^{**}_\kappa \omega^{2(1-\delta_\kappa)} + o\left((1-\beta)\,\omega^{1-\delta_\kappa}\right) + o\left(\omega^{2(1-\delta_\kappa)}\right)}.$$

Hence when $\delta_\kappa \in (0, 1/2)$:

$$f_{\Delta y}(\omega) = \frac{f_x(\omega)\left(\omega^2 + o\left(\omega^2\right)\right)}{(1-\beta)^2 - 2\beta c^*_\kappa (1-\beta)\,\omega^{1-\delta_\kappa} + \beta^2 c^{**}_\kappa \omega^{2(1-\delta_\kappa)} + o\left((1-\beta)\,\omega^{1-\delta_\kappa}\right) + o\left(\omega^{2(1-\delta_\kappa)}\right)}.$$

Consider the Fourier frequencies $\omega_j = 2\pi j/T$ for $j = 1, ..., n$ with $n = o(T)$. If $\nu > 1 - \delta_\kappa$, then for $j = 1, ..., n$, $(1-\beta) = o\left(\omega_j^{1-\delta_\kappa}\right)$ and

$$f_y(\omega_j) \underset{\omega_j \to 0^+}{\sim} \frac{1}{c^{**}_\kappa}\omega_j^{-2(1-\delta_\kappa)},$$

which also implies that $f_{\Delta y}(\omega_j) \underset{\omega_j \to 0^+}{\sim} \frac{1}{c^{**}_\kappa}\omega_j^{-2\delta_\kappa}$. Notice that setting $\beta = 1$ yields a similar result.

4.2 Autocorrelations

To derive our results we first need the following lemma whose proof is in the next section

Lemma A *Let f a spectral density with f, f' and f'' bounded, $f > 0$ in a neighborhood of the origin and $f'(0) = 0$. Let $|\lambda| \in (0,1)$ and $\omega_k = 2\pi k/T$, $k = o(T)$. Then,*

$$T^{\lambda-1}\sum_{j=1}^{T} j^{-\lambda} f(\omega_j)\cos(j\omega_k) = \mathcal{O}\left(k^{\lambda-1}\right). \tag{7}$$

The autocovariance function of y_t satisfies

$$\gamma_k = \frac{1}{2\pi}\int_0^{2\pi} f_y(\omega)\, e^{ik\omega} d\omega \tag{8}$$
$$= \frac{1}{2\pi}\int_0^{2\pi} f_y(\omega)\cos(k\omega)\, d\omega, \tag{9}$$

to which the following finite sum converges (when it does converge)

$$\frac{1}{2\pi T}\sum_{j=1}^{T} f_y\left(\frac{2\pi j}{T}\right)\cos\frac{2\pi jk}{T} \underset{T\to\infty}{\to} \gamma_y\left(k\right). \tag{10}$$

In the proof, we use the theorems of Yong (1974) regarding sums of the type $\sum_{j=1}^{\infty} a_j \cos j\omega_k$ with $a_j \sim j^{-\alpha}$ for some $\alpha > 0$ as $\omega_k \to 0^+$. In the summation, we let the number of summation terms T and the evaluation value $\omega_k = 2\pi k T^{-1}$ tend to their limits, including $k \to \infty$. Zygmund (1935), Section.1.23, shows that the sum (10) converges uniformly if $f_y(\omega_j)$ is of bounded variation and the latter follows from the results of Section. 4.1. Our proof is therefore related to the method of Erdélyi (1956) considered by Lieberman and Phillips (2008) – the difference is that we consider the limit of a finite sum (10) evaluated at Fourier frequencies where Lieberman and Phillips work with the integral representation (8).

We apply Lemma A to expression (10) together with (6). The spectral density f_x is assumed to admit a bounded second order derivative.

When $\nu > 1-\delta_\kappa$, then $1-\beta = o(\omega_j)$ for all Fourier frequencies ω_j, $j = 1, ..., T$. Expression (6) hence implies that

$$\begin{aligned} \delta_\kappa \in (1/2, 1) &: f_y(\omega_j) \sim \frac{f_x(\omega_j)}{(2\pi)^{2(1-\delta_\kappa)} c_\kappa^{**} (j/T)^{2(1-\delta_\kappa)}}; \\ \delta_\kappa \in (0, 1/2) &: f_{\Delta y}(\omega_j) \sim \frac{f_x(\omega_j)}{(2\pi)^{-2\delta_\kappa} c_\kappa^{**} (j/T)^{-2\delta_\kappa}}. \end{aligned} \tag{11}$$

We refer to Lemma A where we let $\lambda = 2(1-\delta_\kappa)$ if $\delta_\kappa \in (1/2, 1)$ and $\lambda = -2\delta_\kappa$ if $\delta_\kappa \in (0, 1/2)$. Then for $k = o(T)$:

$$\delta_\kappa \in (1/2, 1) : \gamma_y(k) = \begin{cases} \mathcal{O}\left(k^{1-2\delta_\kappa}\right), & k \neq 0; \\ O(1), & k = 0. \end{cases}$$

$$\delta_\kappa \in (0, 1/2) : \gamma_{\Delta y}(k) = \begin{cases} \mathcal{O}\left(k^{-1-2\delta_\kappa}\right), & k \neq 0; \\ O(1), & k = 0. \end{cases}$$

5 Proof of Lemma A

5.1 Preliminary results

To derive our results we first need the following lemmas, S.1 and S.2.

Lemma S.1. *Let f a spectral density with f, f' and f'' bounded, $f>0$ in a neighborhood of the origin and $f'(0)=0$. Let $\lambda \in (0,1)$ and define the sequence*

$$a_j = \begin{cases} f\left(\frac{2\pi}{T}j\right) j^{-\lambda}, & j \leq T; \\ f(0)\, j^{-\lambda}, & j > T. \end{cases} \tag{S.7}$$

Then $\{a_j\}$ is of pure bounded variation, defined as $\sum_{j=n}^{\infty} |\Delta a_j| = O(a_n)$ as $n \to \infty$ (Yong, 1974, Definition I-4).

Lemma S.2. *For $\lambda \in (0,1)$, as $\left(T\omega, \omega^{-1}\right) \to (\infty, \infty)$*

$$\sum_{j=1}^{T} a_j \cos j\omega \sim \frac{\pi f(0)}{2\Gamma(\lambda)\cos\frac{\lambda\pi}{2}} \omega^{\lambda-1};$$
$$\sum_{j=1}^{T} a_j \sin j\omega \sim \frac{\pi f(0)}{2\Gamma(\lambda)\sin\frac{\lambda\pi}{2}} \omega^{\lambda-1},$$

where $\{a_j\}$ is defined in Lemma S.1.

5.1.1 Proof of Lemma S.1

When $j \leq T$, the difference Δa_j can be decomposed into

$$\Delta a_j = \frac{f(\omega_j)\left[(j-1)^{\lambda} - j^{\lambda}\right]}{j^{\lambda}(j-1)^{\lambda}} + \frac{f(\omega_j) - f(\omega_{j-1})}{(j-1)^{\lambda}},$$

where a Taylor expansion yields

$$f(\omega_j) - f(\omega_{j-1}) = \frac{1}{2} f''(\omega_{j-1})(\Delta\omega_j)^2 + o\left((\Delta\omega_j)^2\right),$$

and

$$\frac{f(\omega_j)\left[(j-1)^{\lambda} - j^{\lambda}\right]}{j^{\lambda}(j-1)^{\lambda}} = f(\omega_j)\frac{\left[(1-1/j)^{\lambda} - 1\right]}{(j-1)^{\lambda}} = f(\omega_j)\frac{-\lambda}{j^{1+\lambda}} + o\left(\frac{1}{j^{1+\lambda}}\right).$$

Since f and f'' are bounded, there exist $m_0, m_2 >$ such that

$$|\Delta a_j| \leq \frac{m_0\lambda}{j^{1+\lambda}} + \frac{m_2}{2j^{\lambda}}\left(\frac{2\pi}{T}\right)^2.$$

When $j > T$, the latter expression also holds with $m_2 = 0$ so for all j :

$$|\Delta a_j| \leq \frac{m_0\lambda}{j^{1+\lambda}} + \frac{m_2}{2j^{\lambda}}\left(\frac{2\pi}{T}\right)^2 1_{\{j\leq T\}},$$

where $1_{\{\cdot\}}$ is the indicator function.

Now consider $N \geq T \geq n$, then

$$\sum_{j=n}^{N} |\Delta a_j| = \sum_{j=n}^{T} |\Delta a_j| + \sum_{j=T+1}^{N} |\Delta a_j|, \tag{S.8}$$

since for $j \leq T$, $T^{-2} \leq j^{-2}$, the previous expression rewrites as

$$\begin{aligned}\sum_{j=n}^{N} |\Delta a_j| &\leq \sum_{j=n}^{T} \left(\frac{m_0\lambda}{j^{1+\lambda}} + \frac{2\pi^2 m_2}{j^{2+\lambda}}\right) + \sum_{j=T+1}^{N} |\Delta a_j| \\ &\leq m_0 \left[n^{-\lambda} - (N+1)^{-\lambda}\right] + \frac{2\pi^2 m_2}{1+\lambda}\left[n^{-1-\lambda} - (T+1)^{-1-\lambda}\right].\end{aligned}$$

Letting $N \to \infty$, there exists M such that

$$\sum_{j=n}^{\infty} |\Delta a_j| \leq n^{-\lambda}\left[m_0 + \frac{2\pi^2 m_2}{1+\lambda}\frac{1}{n}\right] + \frac{M}{T^{\lambda+1}},$$

and then if $a_n \neq 0$,

$$\frac{\sum_{j=n}^{\infty} |\Delta a_j|}{a_n} \leq \frac{\left[m_0 + \frac{2\pi^2 m_2}{1+\lambda}\right] + \frac{M}{n^{-\lambda}T^{\lambda+1}}}{f\left(\frac{2\pi}{T}\min\{n,T\}\right)}.$$

We now use the fact that f is bounded above zero in a neighborhood of the origin so for T large enough, $f\left(\frac{2\pi}{T}n\right) > 0$. Also for $n \leq T$, $n^{-\lambda}T^{\lambda+1} \to \infty$ as $n \to \infty$. Therefore, there exist (n_0, M_0) such that for all $n \geq n_0$ with $a_n \neq 0$ and for $T \geq n$, $\frac{\sum_{j=n}^{\infty}|\Delta a_{T,j}|}{a_n} \leq M_0$, i.e. $\sum_{j=n}^{\infty} |\Delta a_j| = O(a_n)$.

Now assume that $T \leq n$, the previous expressions yield $\sum_{j=n}^{\infty} |\Delta a_j| \leq f(0)\, n^{-\lambda} = a_n$ so the results also hold.

5.1.2 Proof of Lemma S.2

Introduce the Dirichlet kernel:

$$D_T(\omega) = \frac{1}{2} + \sum_{j=1}^{T} \cos j\omega = \frac{\sin(T+1/2)\,\omega}{2\sin\frac{\omega}{2}},$$

where $|D_T(\omega) - 1/2| \leq \pi\omega^{-1}$ for $0 < \omega \leq \pi$ (Yong, 1974, p.39). Then, for all $\omega > 0$ and $T < N$,

$$\sum_{j=T+1}^{N} a_j \cos j\omega = f(0) \sum_{j=T+1}^{N} j^{-\lambda} \cos j\omega.$$

Then

$$\left|\sum_{j=T+1}^{N} j^{-\lambda}\cos j\omega\right|$$
$$= \left|\sum_{j=T}^{N-1}\left(j^{-\lambda}-(j+1)^{-\lambda}\right)\left(D_j(\omega)-\frac{1}{2}\right)+N^{-\lambda}\left(D_N(\omega)-\frac{1}{2}\right)-T^{-\lambda}\left(D_T(\omega)-\frac{1}{2}\right)\right|$$
$$\leq \sum_{j=T}^{N-1}\left|j^{-\lambda}-(j+1)^{-\lambda}\right|\left|D_j(\omega)-\frac{1}{2}\right|+N^{-\lambda}\left|D_N(\omega)-\frac{1}{2}\right|+T^{-\lambda}\left|D_T(\omega)-\frac{1}{2}\right|. \quad \text{(S.9)}$$

We notice that $j^{-\lambda}-(j+1)^{-\lambda}=j^{-\lambda}\left(1-\left(1+j^{-1}\right)^{-\lambda}\right)\in\left[j^{-\lambda}\left(\lambda j^{-1}-\frac{\lambda(\lambda+1)}{2j^2}\right),j^{-\lambda}\lambda j^{-1}\right]$ so $\left|j^{-\lambda}-(j+1)^{-\lambda}\right|\leq\lambda j^{-\lambda-1}$. It follows that

$$\left|\sum_{j=T+1}^{N} j^{-\lambda}\cos j\omega\right| \leq \lambda\pi\omega^{-1}\left(T^{-\lambda}-N^{-\lambda}\right)+N^{-\lambda}\pi\omega^{-1}+T^{-\lambda}\pi\omega^{-1}$$
$$= (1+\lambda)\pi\omega^{-1}T^{-\lambda}+(1-\lambda)\pi\omega^{-1}N^{-\lambda}.$$

Hence letting $N\to\infty$, for all $T,\omega>0$,

$$\left|\sum_{j=1}^{T} j^{-\lambda}\cos j\omega-\sum_{j=1}^{\infty} j^{-\lambda}\cos j\omega\right| \leq (1+\lambda)\pi\omega^{-1}T^{-\lambda}. \quad (12)$$

Now, Yong (1974), Theorem III-17, states that for $\lambda\in(0,1)$ since $\{a_j\}$ is of pure bounded variation

$$\sum_{j=1}^{\infty} a_j\cos j\omega \underset{\omega\to0^+}{\sim} \frac{\pi f(0)}{2\Gamma(\lambda)\cos\frac{\lambda\pi}{2}}\omega^{\lambda-1}.$$

Hence, in expression (12), as $\left(T,\omega^{-1}\right)\to(\infty,\infty)$: $\omega^{-1}T^{-\lambda}=o\left(\omega^{\lambda-1}\right)$ if $\omega^{-1}=o(T)$, in which case

$$\sum_{j=1}^{T} a_j\cos j\omega \underset{(T\omega,\omega)\to(\infty,0)}{\sim} \frac{\pi f(0)}{2\Gamma(\lambda)\cos\frac{\lambda\pi}{2}}\omega^{\lambda-1}.$$

When the summation involves a sine function, the proof is similar: $\left|\sum_{j=T+1}^{\infty} j^{-\lambda}\sin j\omega\right|<(1+\lambda)\pi\omega^{-1}T^{-\lambda}$ using the conjugate Dirichlet kernel:

$$\overline{D}_T(\omega)=\sum_{j=1}^{T}\sin j\omega=\frac{\cos\frac{\omega}{2}-\cos\left(T+\frac{1}{2}\right)\omega}{2\sin\frac{\omega}{2}},$$

with $\left|\overline{D}_T(\omega)\right|\leq\pi\omega^{-1}$ (also Yong, 1974, p. 39). Then an equivalent of expression (12) holds for the summation involving sine functions and we also refer to Theorem III-17 of Yong.

5.2 Proof of Lemma A

First, if $k=0$, $\sum_{j=1}^{T} j^{-\lambda} f(\omega_j) = O\left(T^{1-\lambda}\right)$ since $f(\omega_T) = f(0) \neq 0$ so for $\lambda \neq 1$ the results follows.

We now assume $k \neq 0$. Lemma S.2 implies as $(k, T/k) \to (\infty, \infty)$ that for $\lambda \in (0,1)$,

$$T^{\lambda} \sum_{j=1}^{T} \frac{f(\omega_j)\cos(j\omega_k)}{j^{\lambda}} \sim \frac{\pi f(0)}{2\Gamma(\lambda)\cos\frac{\lambda\pi}{2}} \left(2\pi \frac{k}{T}\right)^{\lambda-1} T^{\lambda}. \tag{13}$$

Now, let $\lambda \in (-1,0)$. We use a procedure similar to integration by part to express (13) in terms of $\sum_{j=1}^{T} j^{-(\lambda+1)} f_x(\omega_j)\sin(j\omega_k)$ where $\lambda+1 \in (0,1)$ which allows to work alongside the previous results. Start with

$$\begin{aligned}
&(j+1)^{-\lambda}\sin((j+1)\omega_k) - j^{-\lambda}\sin(j\omega_k)\\
&= \left((j+1)^{-\lambda} - j^{-\lambda}\right)\sin((j+1)\omega_k) + j^{-\lambda}\left\{\sin((j+1)\omega_k) - \sin(j\omega_k)\right\},
\end{aligned}$$

where $(j+1)^{-\lambda} - j^{-\lambda} = -\lambda(j+1)^{-\lambda-1} + o\left(j^{-\lambda-1}\right)$ and

$$\sin((j+1)\omega_k) - \sin(j\omega_k) = \sin(j\omega_k)(\cos(\omega_k)-1) + \cos(j\omega_k)\sin(\omega_k).$$

Hence

$$\begin{aligned}
&\sum_{j=1}^{T}\left[(j+1)^{-\lambda}\sin((j+1)\omega_k) - j^{-\lambda}\sin(j\omega_k)\right]\\
&= \sum_{j=1}^{T}\left\{-\lambda(j+1)^{-\lambda-1} + o\left((j+1)^{-\lambda-1}\right)\right\}\sin((j+1)\omega_k)\\
&+ \sin(\omega_k)\sum_{j=1}^{T} j^{-\lambda}\cos(j\omega_k) + (\cos(\omega_k)-1)\sum_{j=1}^{T} j^{-\lambda}\sin(j\omega_k),
\end{aligned}$$

i.e. since $k \neq 0$:

$$\begin{aligned}
\sum_{j=1}^{T} j^{-\lambda}\cos(j\omega_k) &= (T+1)^{-\lambda}\frac{\sin(T+1)\omega_k}{\sin(\omega_k)} - 1 - \lambda\\
&\quad - \frac{\cos(\omega_k)-1}{\sin(\omega_k)}\sum_{j=1}^{T} j^{-\lambda}\sin(j\omega_k) + \frac{\lambda}{\sin(\omega_k)}\sum_{j=1}^{T+1} j^{-(1+\lambda)}\sin(j\omega_k)\\
&\quad + o\left(\frac{\sum_{j=2}^{T+1} j^{-(1+\lambda)}\sin(j\omega_k)}{\sin(\omega_k)}\right).
\end{aligned}$$

Similarly, since $\cos((j+1)\omega_k) - \cos(j\omega_k) = \cos(j\omega_k))(\cos(\omega_k) - 1) - \sin(j\omega_k)\sin\omega_k$, we can express:

$$\sum_{j=1}^{T} j^{-\lambda} \sin(j\omega_k) = -(T+1)^{-\lambda} \frac{\cos(T+1)\omega_k}{\sin(\omega_k)} + (1+\lambda) \frac{\cos\omega_k}{\sin\omega_k} + \frac{\cos(\omega_k) - 1}{\sin(\omega_k)} \sum_{j=1}^{T} j^{-\lambda} \cos(j\omega_k) - \frac{\lambda}{\sin(\omega_k)} \sum_{j=1}^{T+1} j^{-(1+\lambda)} \cos(j\omega_k) + o\left(\frac{\sum_{j=2}^{T+1} j^{-(1+\lambda)} \cos(j\omega_k)}{\sin(\omega_k)}\right).$$

Plugging $\sum_{j=1}^{T} j^{-\lambda} \sin(j\omega_k)$ in the expression for $\sum_{j=1}^{T} j^{-\lambda} \cos(j\omega_k)$ yields:

$$\left(1 + \frac{(\cos(\omega_k) - 1)^2}{\sin^2(\omega_k)}\right) \sum_{j=1}^{T} j^{-\lambda} \cos(j\omega_k)$$
$$= (T+1)^{-\lambda} \frac{\sin(\omega_k)\sin[(T+1)\omega_k] + (\cos(\omega_k) - 1)\cos[(T+1)\omega_k]}{\sin^2(\omega_k)}$$
$$- (1+\lambda)\left(1 + \frac{\cos(\omega_k) - 1}{\sin^2(\omega_k)} \cos\omega_k\right)$$
$$+ \frac{\lambda}{\sin(\omega_k)} \sum_{j=1}^{T+1} j^{-(1+\lambda)} \sin(j\omega_k) + \lambda \frac{\cos(\omega_k) - 1}{\sin^2(\omega_k)} \sum_{j=1}^{T+1} j^{-(1+\lambda)} \cos(j\omega_k)$$
$$+ o\left(\frac{\sum_{j=2}^{T+1} j^{-(1+\lambda)} \sin(j\omega_k)}{\sin(\omega_k)}\right) + o\left(\frac{\cos(\omega_k) - 1}{\sin^2(\omega_k)} \sum_{j=2}^{T+1} j^{-(1+\lambda)} \cos(j\omega_k)\right).$$

Using Lemma S.2, this leads, as $\left(T\omega_k, \omega_k^{-1}\right) \to (\infty, \infty)$, and since $1 + \lambda \in (0,1)$ to:

$$\sum_{j=1}^{T} j^{-\lambda} \cos(j\omega_k) = \frac{1}{T^{\lambda}\omega_k} \sin[(T+1)\omega_k] - (1+\lambda)\left(1 + \frac{1}{2}\right) + \frac{\lambda\pi f(0)}{2\Gamma(\lambda+1)\cos\frac{(\lambda+1)\pi}{2}} \omega_k^{\lambda-1} + o\left(\omega_k^{\lambda-1}\right),$$

where $T^{-\lambda}\omega_k^{-1} = O\left(\omega_k^{\lambda-1} k^{-\lambda}\right)$. Hence

$$\sum_{j=1}^{T} j^{-\lambda} \cos(j\omega_k) = \frac{\lambda\pi f(0)}{2\Gamma(\lambda+1)\cos\frac{(\lambda+1)\pi}{2}} \omega_k^{\lambda-1} + o\left(\omega_k^{\lambda-1}\right), \tag{S.11}$$

which simplifies to $T^{\lambda-1} \sum_{j=1}^{T} j^{-\lambda} \cos(j\omega_k) = \mathcal{O}\left(k^{\lambda-1}\right)$.

6 Additional Simulations

6.1 Standard deviation of the sample mean

We present complementary figures that report the log of $\mathsf{sd}\left(T^{-1/2}\sum_{t=1}^{T} y_t\right)$ and the growth rate of $\mathsf{sd}\left(T^{-1/2}\sum_{t=1}^{T} y_t\right)/\log T$ under RLS and CGLS learning. These figures exemplify the results of Theorem 2 in the paper (figures S.1 and S.3) and Theorem 3 (figures S.2 and S.4), i.e. with local-asymptotic parameters in the case of CGLS ($c_\beta = c_g = 1/2$). In all figures, the horizontal axes report the log of the sample size. The sample sizes range from 200 to 50,000 for which we produce 10,000 Monte Carlo replications. The data generating process is the same as in the paper, Section 4, with $\sigma_0^2 = \infty$.

The figures show that, as the sample size increases $\mathsf{sd}\left(T^{-1/2}\sum_{t=1}^{T} y_t\right)$ increase linearly and the rate of growth are as the theorems imply.

6.2 Distribution of the log periodogram estimators

Figures S.5 and S.6 report the densities of the GPH and local Whittle likelihood estimators $\widehat{d}$ of the degree of fractional integration of y_t. The local Whittle estimator is obtained by constrained maximization over the range $d \in (-1, 2)$. The model is the 'mean plus noise' model of the paper and the simulation settings are the same as in Figure 1 and Tables 2 and 3 of the paper.

	$\log(D_t/P_t)$	Canada	France	Germany	Italy	Japan	UK
2ELW	0.61	0.36	0.22	0.62	0.63	0.55	0.15
FELW	0.64	0.38	0.20	0.63	0.61	0.51	0.07

Table S.1: The table reports the minimum value of the parameter g such that a t-test of $\mathsf{H}_0 : d = 0$ versus $\mathsf{H}_1 : d > 0$ is not rejected for $x_t = y_t - \beta y^e_{t+1}$, where $y^e_{t+1} = (1 - (1 - L)^g)\, y_t$, at a 5% asymptotic nominal level of significance. For details of estimators and data, see Table 4.

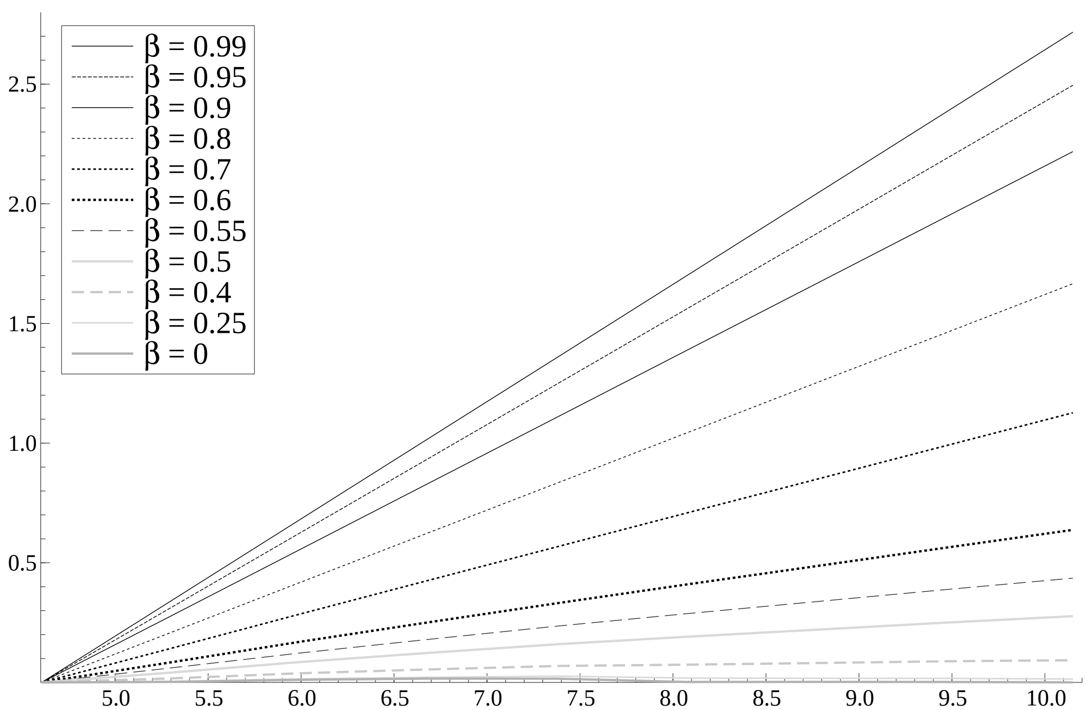

Figure S.1: Magnitude of the log of the Monte Carlo standard deviation of the sample mean, $\log \widehat{\mathsf{sd}}\left(T^{-1/2} \sum y_t\right)$, as a function of the log sample size when agents learn using RLS.

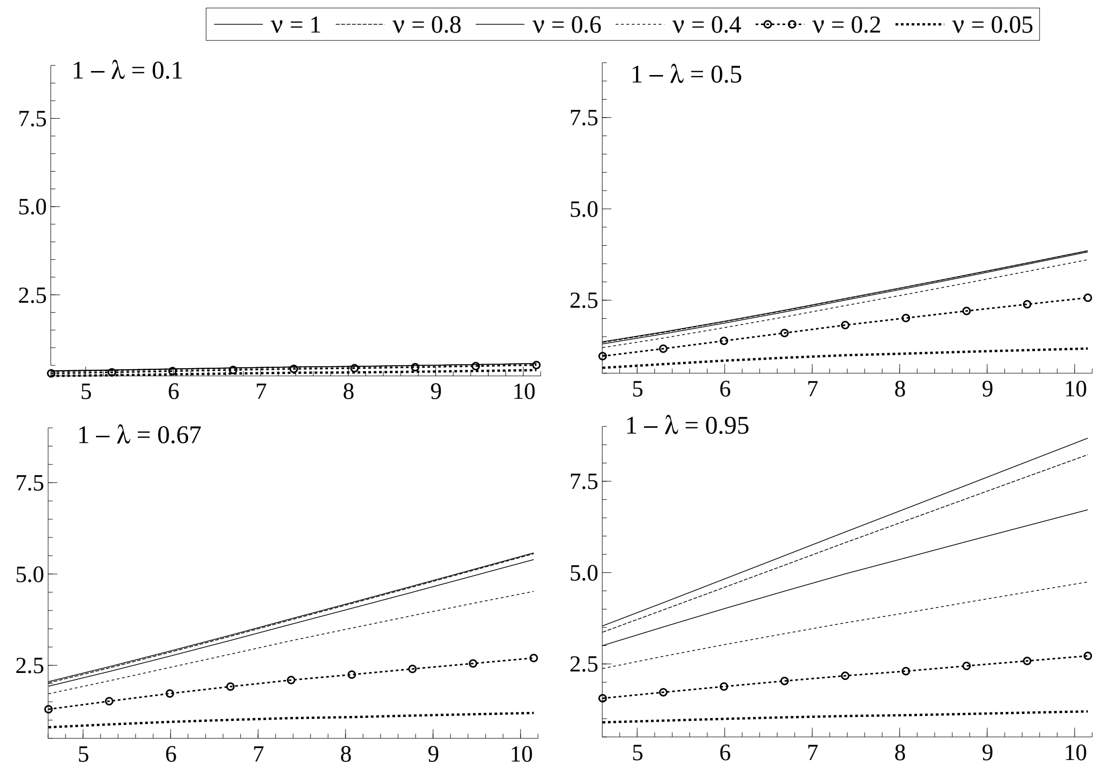

Figure S.2: Magnitude of the log of the Monte Carlo standard deviation of the sample mean, $\log \widehat{\mathsf{sd}}\left(T^{-1/2}\sum y_t\right)$, as a function of the log sample size when agents learn using CGLS. The parameters are local asymptotic in the sense that there exist $c_\beta, c_g > 0$ such that $\beta = 1 - c_\beta T^{-\nu}$ and $\bar{g} = c_g T^{-\lambda}$, with $(\nu, \lambda) \in (0, 1]^2$.

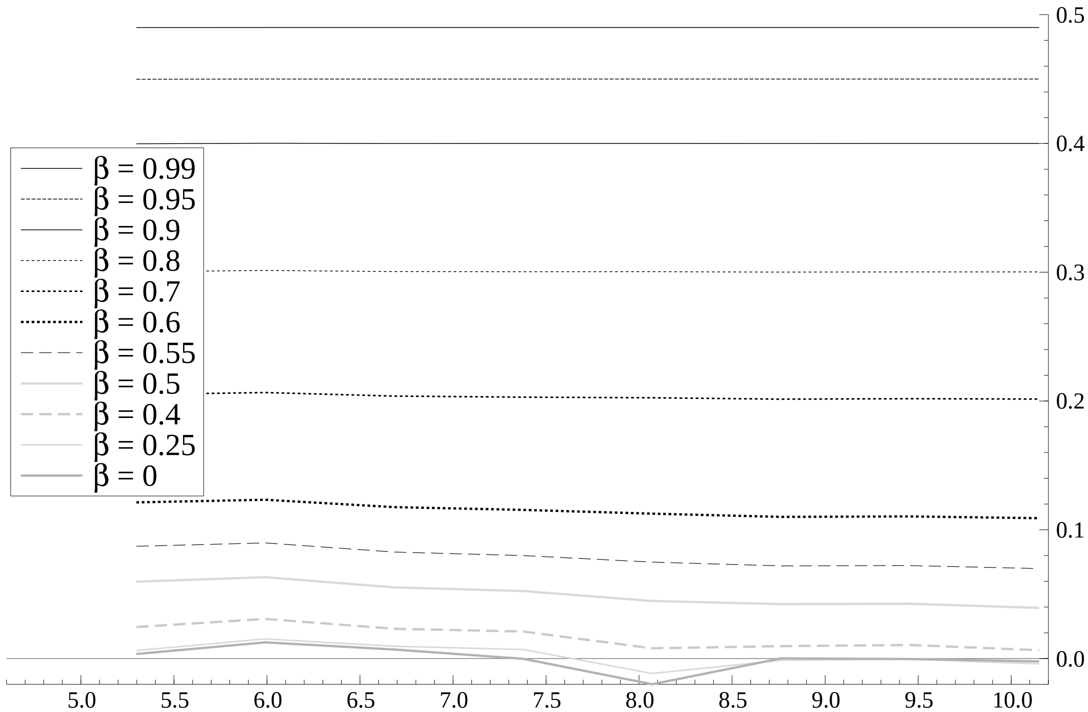

Figure S.3: Growth of the ratio $\log \widehat{\mathsf{sd}}\left(T^{-1/2}\sum y_t\right) / \log T$ as a function of the log sample size T when agents learn using RLS.

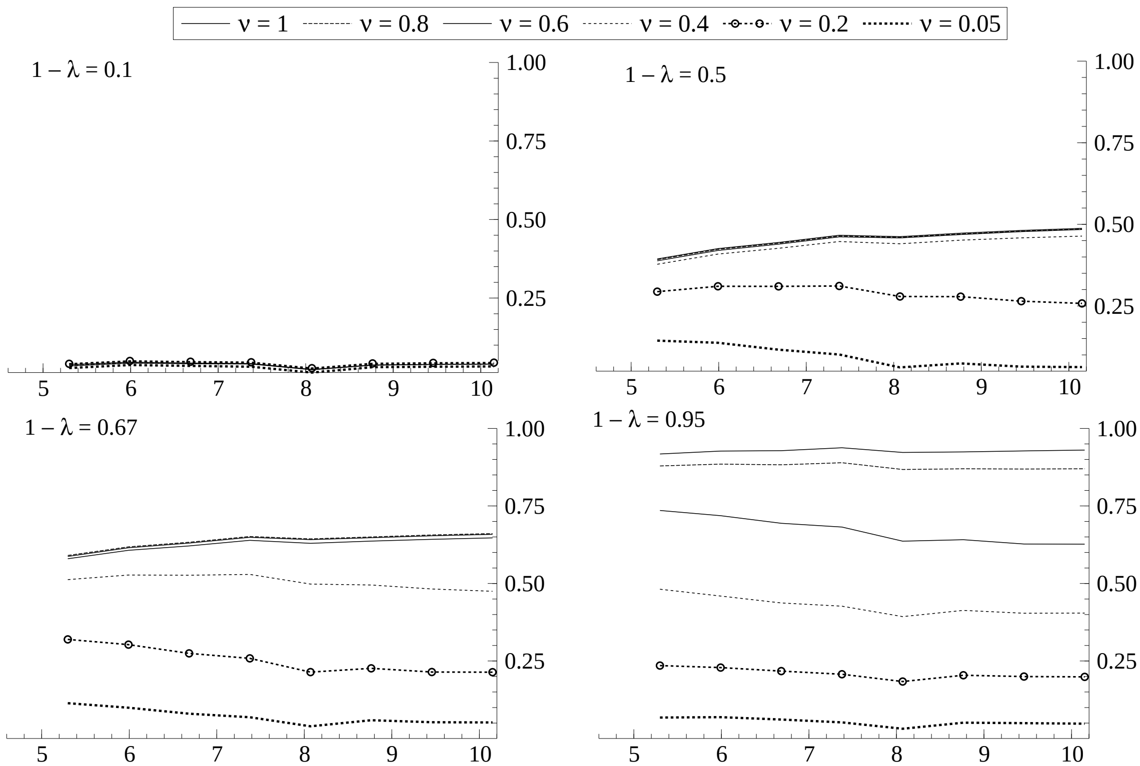

Figure S.4: Growth of the ratio $\log \widehat{\mathsf{sd}}\left(T^{-1/2}\sum y_t\right)/\log T$ as a function of the log sample size T when agents learn using CGLS. The parameters are local asymptotic in the sense that there exist $c_\beta, c_g > 0$ such that $\beta = 1 - c_\beta T^{-\nu}$ and $\bar{g} = c_g T^{-\lambda}$, with $(\nu, \lambda) \in (0, 1]^2$.

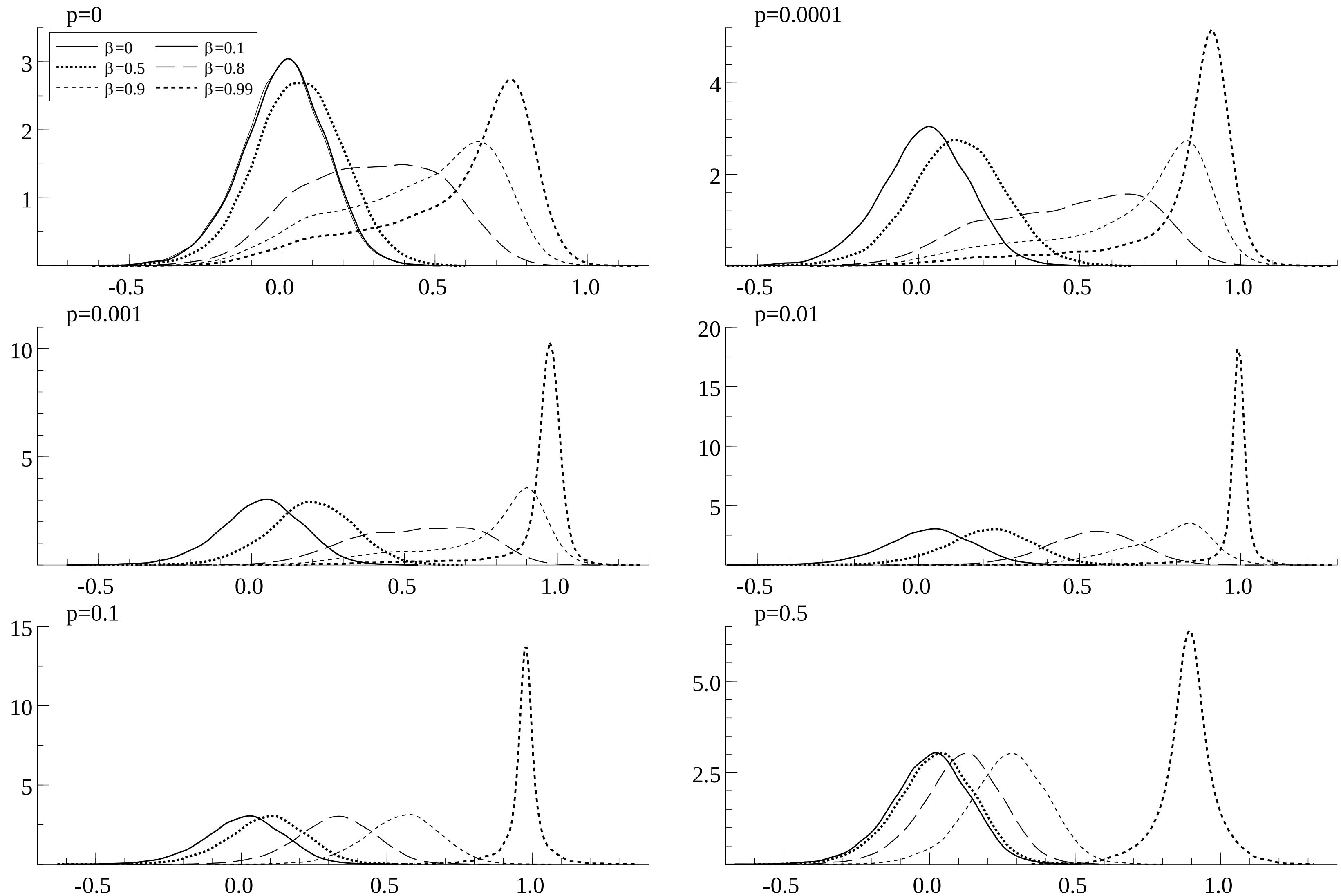

Figure S.5: Density of the GPH log periodogram estimator of the fractional integration parameter d using $n = \sqrt{T}$ Fourier frequencies over samples of $T = 1000$ observations. The model is the 'mean plus noise' perceived law of motion presented in section 2 of the paper. g is 0 under RLS learning and $\bar{g}$ otherwise. The number of Monte Carlo replications is 10,000.

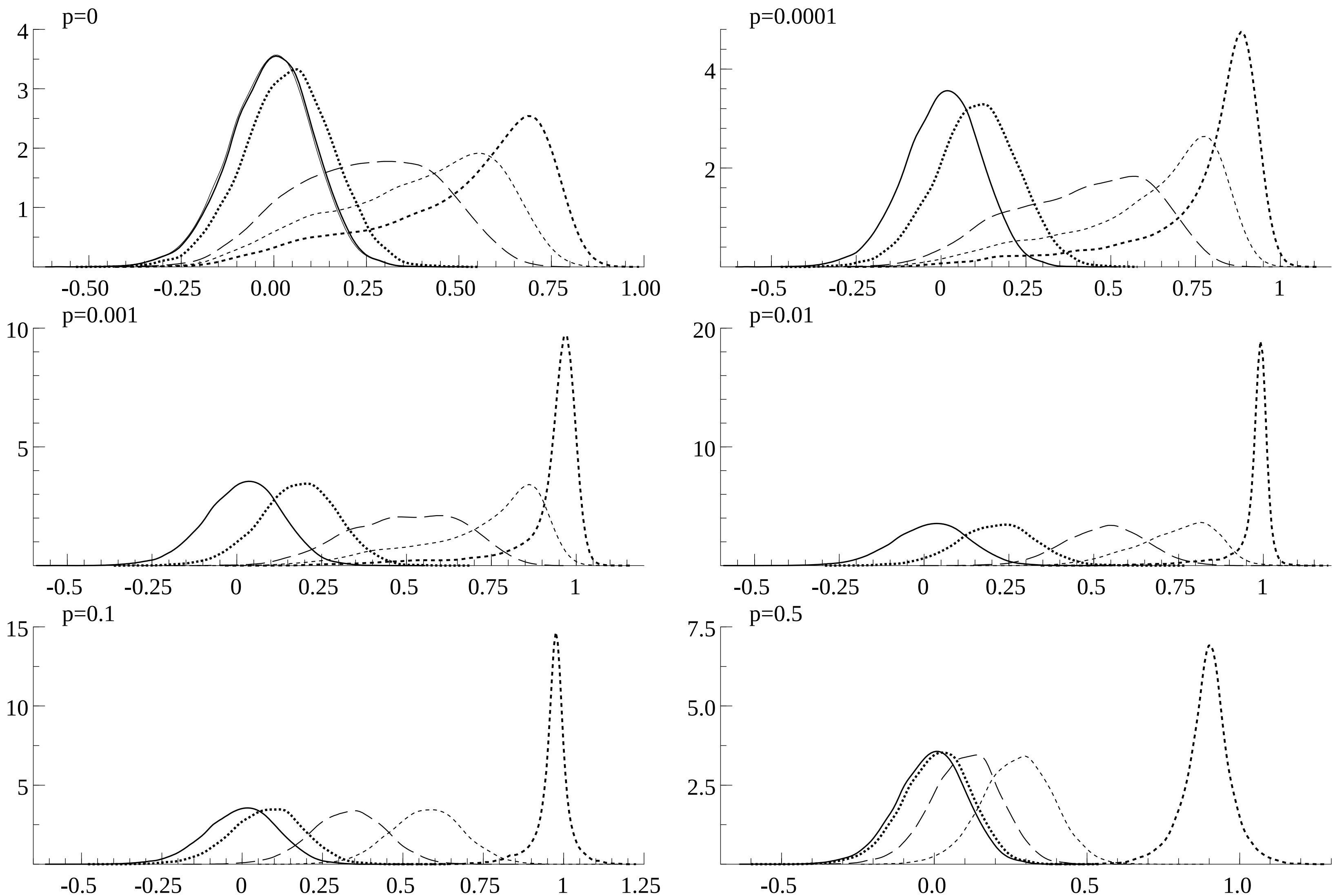

Figure S.6: Density of the Whittle likelihood estimator of the fractional integration parameter d using $n = \sqrt{T}$ Fourier frequencies over samples of $T = 1000$ observations. The model is the 'mean plus noise' perceived law of motion presented in section 2 of the paper. g is 0 under RLS learning and $\bar{g}$ otherwise. The number of Monte Carlo replications is 10,000.