Return predictability, dividend growth, and the persistence of the price–dividend ratio

Golinski, Adam; Madeira, Joao; Rambaccussing, Dooruj

Published in:
International Journal of Forecasting

DOI:
10.1016/j.ijforecast.2024.03.005

Publication date:
2024

Document Version
Other version

Appendix A Possible origins of long memory in the price–dividend ratio

In this section we consider several explanations that could account for the presence of long memory in the price–dividend ratio. Robinson (1978) and Granger (1980) showed that fractionally integrated series can occur from realistic aggregation situations. For instance, independent series generated by a first-order autoregressive process can result in a fractionally integrated series when aggregated. Considering that the aggregate stock market index consists of hundreds of individual stock series, the aggregation result seems a very plausible explanation of the long-range persistence in the price–dividend ratio.

Alternatively, long memory could be induced within a learning framework in which agents update their beliefs using linear algorithms. Chevillon and Mavroeidis (2017) showed that a realistic degree of long memory can arise in the price–dividend ratio without any persistence in the exogenous shocks, if agents learn according to the constant gain least squares algorithm. Chevillon and Mavroeidis (2017), with annual S&P 500 data for the period from 1871–2011 and using two semiparametric estimators, found the long-memory parameter at about 0.8, which is close to our estimates in Table 1, Panel C.

Another explanation is that long memory in the price–dividend ratio could be a result of structural breaks. As shown by Diebold and Inoue (2001), in some circumstances rare structural breaks and long memory are really two sides of the same coin, and they cannot be distinguished from each other in finite samples. Lettau and Van Nieuwerburgh (2008) claimed that strong persistence in the price–dividend ratio could be generated by structural breaks (or shifts) in the steady-state mean of the economy. They showed that if the shifts are accounted for, then the return forecasting ability of the price–dividend ratio is stable over time. However, in an earlier study, Granger and Hyung (2004) established that if the true series is a long-memory process, it is very likely that spurious breaks will be detected. Conversely, even if the true process was generated by occasional breaks, the long-memory process can successfully reproduce many features of the true series and (under some condi-

tions) can yield better forecasts. Indeed, Lettau and Van Nieuwerburgh (2008) report that difficulties with detecting the breaks in real time make it hard to forecast stock returns. These findings reinforce the long-memory argument in the price–dividend ratio.

As a result of the price–dividend being persistent and exhibiting long memory, expected returns and the expected dividend growth rate are likely to be mean-reverting and anit-persistent. This results from the first difference of the price–dividend ratio. Since both expected returns and dividend growth are latent series, the antipersistence can be observed by applying the first difference on the price–dividend ratio.

Appendix B Proofs of Theorems 1 and 2

Proof of Theorem 1 i. Re-write the spectral density (10) as:

$$f(\lambda) = \frac{1}{2\pi} \left| \sum_{j=0}^{\infty} \psi_j e^{-i\lambda j} \right|^2, \tag{A-1}$$

which shows explicitly that at $\lambda = 0$, the spectral density at zero is determined by the sum of moving-average coefficients. Taking the first derivative of (A-1) with respect to λ yields:

$$2\pi f^{'}(\lambda) = i \left(\sum_{j=0}^{\infty} j\psi_j e^{i\lambda j} \right) \left(\sum_{j=0}^{\infty} \psi_j e^{-i\lambda j} \right) - i \left(\sum_{j=0}^{\infty} \psi_j e^{i\lambda j} \right) \left(\sum_{j=0}^{\infty} j\psi_j e^{-i\lambda j} \right). \tag{A-2}$$

Taking the limit at $\lambda \to 0^+$ and applying the algebraic limit theorem gives:

$$\begin{aligned}
\lim_{\lambda \to 0^+} 2\pi f^{'}(\lambda) &= \lim_{\lambda \to 0^+} \left[i \left(\sum_{j=0}^{\infty} j\psi_j e^{i\lambda j} \right) \left(\sum_{j=0}^{\infty} \psi_j e^{-i\lambda j} \right) - i \left(\sum_{j=0}^{\infty} \psi_j e^{i\lambda j} \right) \left(\sum_{j=0}^{\infty} j\psi_j e^{-i\lambda j} \right) \right] \\
&= i \lim_{\lambda \to 0^+} \left(\sum_{j=0}^{\infty} j\psi_j e^{i\lambda j} \right) \lim_{\lambda \to 0^+} \left(\sum_{j=0}^{\infty} \psi_j e^{-i\lambda j} \right) - i \lim_{\lambda \to 0^+} \left(\sum_{j=0}^{\infty} \psi_j e^{i\lambda j} \right) \lim_{\lambda \to 0^+} \left(\sum_{j=0}^{\infty} j\psi_j e^{-i\lambda j} \right) \\
&= i \left(\sum_{j=0}^{\infty} \psi_j \right) \lim_{\lambda \to 0^+} \left[\left(\sum_{j=0}^{\infty} j\psi_j e^{i\lambda j} \right) - \left(\sum_{j=0}^{\infty} j\psi_j e^{-i\lambda j} \right) \right]
\end{aligned}$$

$$
\begin{aligned}
&= i\left(\sum_{j=0}^{\infty}\psi_j\right)\lim_{\lambda\to 0^+}\left[\left(\sum_{j=0}^{\infty}j\psi_j\left(\cos\lambda j+i\sin\lambda j\right)\right)-\left(\sum_{j=0}^{\infty}j\psi_j\left(\cos\lambda j-i\sin\lambda j\right)\right)\right]\\
&= i\left(\sum_{j=0}^{\infty}\psi_j\right)\lim_{\lambda\to 0^+}\left[2i\left(\sum_{j=0}^{\infty}j\psi_j\sin\lambda j\right)\right]\\
&= -2\left(\sum_{j=0}^{\infty}\psi_j\right)\lim_{\lambda\to 0^+}\left(\sum_{j=0}^{\infty}j\psi_j\sin\lambda j\right) \qquad \text{(A-3)}
\end{aligned}
$$

Since $\varphi_j \sim c_0 j^{\delta-1}$ and

$$\lim_{\lambda\to 0^+}\sum_{j=0}^{\infty}j^{\delta-1}\sin\lambda j\to\infty \text{ for } \delta>0,$$

we find that

$$\lim_{\lambda\to 0^+}\sum_{j=0}^{\infty}j\varphi_j\sin\lambda j\to\begin{cases}+\infty \text{ for } \delta>0\\ -\infty \text{ for } \delta\in(-1,0)\end{cases} \qquad \text{(A-4)}$$

and therefore, for $\delta\in(-1,0)$,

$$
\begin{aligned}
\lim_{\lambda\to 0^+}\sum_{j=0}^{\infty}j\psi_j\sin\lambda j &= \lim_{\lambda\to 0^+}\sum_{j=0}^{\infty}j\left(\sum_{k=0}^{\infty}\rho^k\varphi_{k+j}\right)\sin\lambda j\\
&= \lim_{\lambda\to 0^+}\sum_{k=0}^{\infty}\rho^k\left(\sum_{j=0}^{\infty}j\varphi_{k+j}\sin\lambda j\right) \qquad \text{(A-5)}
\end{aligned}
$$

diverges to $-\infty$. Since $\sum_{j=0}^{\infty}\psi_j<0$, we conclude that

$$\lim_{\lambda\to 0^+}f'(\lambda)\to-\infty. \qquad \text{(A-6)}$$

Q.E.D.

Proof of Theorem 1 ii. Under Assumptions 1 and 2, the sum of the moving-average coefficients is:

$$\sum_{j=0}^{\infty}\psi_j = \psi_0+\sum_{j=1}^{\infty}\sum_{i=0}^{\infty}\rho^i\varphi_{i+j}$$

$$
\begin{aligned}
&\sim \sum_{j=1}^{\infty}\sum_{i=0}^{\infty}\rho^i c(i+j)^{\delta-1} \\
&\sim c\sum_{i=0}^{\infty}\rho^j\sum_{j=1}^{\infty}(i+j)^{\delta-1} \\
&\sim -\frac{c}{\delta}\sum_{j=1}^{\infty}\rho^j j^{\delta}, \qquad \text{(A-7)}
\end{aligned}
$$

where the last line follows from approximating the sum by an integral (see Abadir, Heijmans, and Magnus, 2018, Section A.4.1). The sum in (A-7) is finite by the usual convergence criteria. *Q.E.D.*

Proof of Theorem 1 iii. Under Assumptions 1 and 2′, the sum of the moving-average coefficients is:

$$
\begin{aligned}
\sum_{j=0}^{\infty}\psi_j &= \sum_{j=0}^{\infty}\rho^j\varphi_j + \sum_{j=0}^{\infty}\rho^j\varphi_{j+1} + \sum_{j=0}^{\infty}\rho^j\varphi_{j+2} + \dots \\
&= \rho^0\sum_{j=0}^{\infty}\varphi_j + \rho^1\sum_{j=1}^{\infty}\varphi_j + \rho^2\sum_{j=2}^{\infty}\varphi_j + \dots \\
&= 0 - \rho^1\sum_{j=0}^{0}\varphi_j - \rho^2\sum_{j=0}^{1}\varphi_j - \dots \\
&= -\rho\left[\varphi_0 + \rho\sum_{j=0}^{1}\varphi_j + \rho^2\sum_{j=0}^{2}\varphi_j + \dots\right] \\
&= -\rho\left[\varphi_0\sum_{j=0}^{\infty}\rho^j + \varphi_1\sum_{j=1}^{\infty}\rho^j + \varphi_2\sum_{j=2}^{\infty}\rho^j + \dots\right] \\
&= -\rho\left[\frac{\varphi_0}{1-\rho} + \frac{\varphi_1\rho}{1-\rho} + \frac{\varphi_2\rho^2}{1-\rho} + \dots\right] \\
&= \frac{-\rho}{1-\rho}\sum_{j=0}^{\infty}\rho^j\varphi_j \\
&= \frac{-\rho}{1-\rho}\psi_0. \qquad \text{(A-8)}
\end{aligned}
$$

The third equality follows from the observation that for an ARFIMA(p,δ,q) process with

$\delta < 0$, the moving-average coefficients sum up to zero, i.e. $\sum_{j=0}^{\infty} \varphi_j = 0$, and the sixth line follows from applying the sum of a geometric series. The other equalities follow from straightforward algebra. Substituting this result in the spectral density equation at zero frequency, we obtain:

$$f(0) = \frac{1}{2\pi} \frac{\rho^2 \psi_0^2}{(1-\rho)^2}. \tag{A-9}$$

Q.E.D.

Proof of Theorem 2. Under Assumptions 1 and 2, the moving-average coefficients φ_k in (5) decay hyperbolically at the following rate:

$$\ln\left(\frac{\varphi_{k+1}}{\varphi_k}\right) \sim \frac{\delta - 1}{k} \quad \text{as} \quad k \to \infty. \tag{A-10}$$

Since ψ_k is asymptotically monotone in k, it can be approximated with the continuous-time limit:

$$\psi_k = \sum_{j=0}^{\infty} \rho^j (j+k)^{\delta-1} \sim \int_0^{\infty} \rho^j (j+k)^{\delta-1} \quad \text{as} \quad k \to \infty. \tag{A-11}$$

Its rate of decay is then:

$$\begin{aligned} \frac{\partial \psi_k}{\psi_k \partial k} &= \frac{\delta - 1}{\psi_k} c_0 \int_0^{\infty} \rho^j (j+k)^{\delta-2} \\ &\sim \frac{\delta - 1}{k} \frac{1}{\psi_k} c_0 \int_0^{\infty} \rho^j (j+k)^{\delta-1} \\ &\sim \frac{\delta - 1}{k} \quad \text{as} \quad k \to \infty, \end{aligned} \tag{A-12}$$

which is the same as for φ_k. *Q.E.D.*

Appendix C Identification of $\boldsymbol{\Sigma}$

In this appendix we discuss the identification of the model parameters, focusing in particular on the elements of $\boldsymbol{\Sigma}$. Consider a generalized version of the model presented in Section 3.2, with the expected returns and expected dividend growth following stationary and invertible processes ARFIMA(p_m, δ_m, q_m) and ARFIMA(p_g, δ_g, q_g), respectively, such that they have a state-space moving-average representation as in Eqs. (14a)–(14b) and (15a)–(15b).

The unconditional mean parameters μ_m and μ_g are identified trivially from $\mu_m = E[r_t]$ and $\mu_g = E[\Delta d_t]$, respectively. The identifiability of other conditional mean parameters follows from applying the Yule–Walker argument to the observable moments $\text{cov}(r_t, r_{t-j})$ and $\text{cov}(\Delta d_t, \Delta d_{t-j})$ for $j \geq 1$.

To determine the identifiability of the $\boldsymbol{\Sigma}$ parameters, first we need to determine the relation between the shocks. Substituting the price–dividend Eq. (4) back to Eq. (3), we obtain:

$$\begin{aligned} r_{t+1} - \Delta d_{t+1} &= \kappa + \frac{\rho\kappa}{1-\rho} - pd_t + \rho E_{t+1} \sum_{j=0}^{\infty} \rho^j \left(g_{t+1+j} - m_{t+1+j}\right) \qquad \text{(A-13)} \\ &= \kappa + \frac{\rho\kappa}{1-\rho} - pd_t + \rho \mathbf{w}' E_{t+1} \sum_{j=0}^{\infty} \rho^j \left(\mathbf{C}_{g,t+1+j} - \mathbf{C}_{m,t+1+j}\right) \\ &= \kappa + \frac{\rho\kappa}{1-\rho} - pd_t + \rho \mathbf{w}' \sum_{j=0}^{\infty} \rho^j \mathbf{F}^j \left(\mathbf{C}_{g,t+1} - \mathbf{C}_{m,t+1}\right) \\ &= \kappa + \frac{\rho\kappa}{1-\rho} - pd_t + \rho \mathbf{w}' \sum_{j=0}^{\infty} \rho^j \left[\mathbf{F}^{j+1} \left(\mathbf{C}_{g,t} - \mathbf{C}_{m,t}\right) + \mathbf{F}^j \left(\mathbf{h}_g \varepsilon_{g,t+1} - \mathbf{h}_m \varepsilon_{m,t+1}\right)\right], \end{aligned}$$

where we used the state-space representation and evaluated the expectation at time $t+1$. Subtracting from Eq. (A-13) its conditional expectation at time t yields:

$$\begin{aligned} \varepsilon_{r,t+1} - \varepsilon_{d_t+1} &= \rho \mathbf{w}' \sum_{j=0}^{\infty} \rho^j \mathbf{F}^j \left(\mathbf{h}_g \varepsilon_{g,t+1} - \mathbf{h}_m \varepsilon_{m,t+1}\right) \qquad \text{(A-14)} \\ &= \rho \mathbf{b}' \mathbf{h}_g \varepsilon_{g,t+1} - \rho \mathbf{b}' \mathbf{h}_m \varepsilon_{m,t+1}, \end{aligned}$$

or

$$\varepsilon_{r,t} = q_1 \varepsilon_{m,t} + q_2 \varepsilon_{g,t} + \varepsilon_{d,t}, \tag{A-15}$$

where $q_1 = -\rho \mathbf{b}' \mathbf{h}_m = -\sum_{k=0}^{\infty} \rho^{k+1} \varphi_{m,k}$ and $q_2 = \rho \mathbf{b}' \mathbf{h}_g = \sum_{k=0}^{\infty} \rho^{k+1} \varphi_{g,k}$.

As in Rytchkov (2012), we define the transformed variables:

$$\begin{aligned} \widetilde{\mathbf{y}}_t &= \begin{bmatrix} \widetilde{r}_t \\ \widetilde{\Delta d}_t \end{bmatrix} \\ &= \begin{bmatrix} \left(\sum_{j=0}^{\infty} \varphi_{m,j} L^j\right)^{-1} (r_t - \mu_m) \\ \left(\sum_{j=0}^{\infty} \varphi_{g,j} L^j\right)^{-1} (\Delta d_t - \mu_g) \end{bmatrix} \end{aligned} \tag{A-16}$$

which follows the process:

$$\begin{aligned} \widetilde{\mathbf{y}}_t &= \begin{bmatrix} \left(\sum_{j=0}^{\infty} \varphi_{m,j} L^j\right)^{-1} (m_t - \mu_m + \varepsilon_r) \\ \left(\sum_{j=0}^{\infty} \varphi_{g,j} L^j\right)^{-1} (g_t - \mu_g + \varepsilon_d) \end{bmatrix} \\ &= \begin{bmatrix} \left(\sum_{j=0}^{\infty} \varphi_{m,j} L^j\right)^{-1} \varepsilon_{r,t} + \varepsilon_{m,t-1} \\ \left(\sum_{j=0}^{\infty} \varphi_{g,j} L^j\right)^{-1} \varepsilon_{d,t} + \varepsilon_{g,t-1} \end{bmatrix} \\ &= \begin{bmatrix} (1 - \sum_{j=1}^{\infty} \xi_{m,j} L^j) \varepsilon_{r,t} + \varepsilon_{m,t-1} \\ (1 - \sum_{j=1}^{\infty} \xi_{g,j} L^j) \varepsilon_{d,t} + \varepsilon_{g,t-1} \end{bmatrix}, \end{aligned} \tag{A-17}$$

where $\xi_{m,j}$ and $\xi_{g,t}$ are the coefficients of the autoregressive representation of m_t and g_t, respectively. Using Eq. (A-14) gives:

$$\begin{aligned} \widetilde{\mathbf{y}}_t &= \begin{bmatrix} \left(1 - \sum_{j=1}^{\infty} \xi_{m,j} L^j\right) (q_1 \varepsilon_{m,t} + q_2 \varepsilon_{g,t} + \varepsilon_{d,t}) + \varepsilon_{m,t-1} \\ \left(1 - \sum_{j=1}^{\infty} \xi_{g,j} L^j\right) \varepsilon_{d,t} + \varepsilon_{g,t-1} \end{bmatrix} \\ &= \sum_{j=0}^{\infty} \mathbf{A}_j L^j \begin{bmatrix} \varepsilon_{m,t} \\ \varepsilon_{g,t} \\ \varepsilon_{d,t} \end{bmatrix}, \end{aligned} \tag{A-18}$$

where

$$\mathbf{A}_0 = \begin{bmatrix} q_1 & q_2 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}_1 = \begin{bmatrix} (1-\xi_{m,1}q_1) & -\xi_{m,1}q_2 & -\xi_{m,1} \\ 0 & 1 & -\xi_{g,1} \end{bmatrix}$$

and

$$\mathbf{A}_j = \begin{bmatrix} -\xi_{j,1}q_1 & -\xi_{m,j}q_2 & -\xi_{m,j} \\ 0 & 0 & -\xi_{g,j} \end{bmatrix} \quad \text{for} \quad j \geq 2.$$

Thus, given ρ and the conditional mean parameters ξ, from the moments of $\widetilde{\mathbf{y}}_t$:

$$\operatorname{var}(\widetilde{\mathbf{y}}_t) = \sum_{k=0}^{\infty} \mathbf{A}_k \boldsymbol{\Sigma} \mathbf{A}_k', \tag{A-19}$$

$$\operatorname{cov}(\widetilde{\mathbf{y}}_t, \widetilde{\mathbf{y}}_{t-j}) = \sum_{k=0}^{\infty} \mathbf{A}_k \boldsymbol{\Sigma} \mathbf{A}_{k+j}', \tag{A-20}$$

we can find a system of equations that allows us to identify the elements of $\boldsymbol{\Sigma}$:

$$\operatorname{vech}(\operatorname{var}(\widetilde{\mathbf{y}}_t)) = \mathbf{D}_2^{+} \left(\sum_{k=0}^{\infty} \mathbf{A}_k \otimes \mathbf{A}_k \right) \mathbf{D}_3 \operatorname{vech}(\boldsymbol{\Sigma}) = \overline{\mathbf{A}}_0 \operatorname{vech}(\boldsymbol{\Sigma}), \tag{A-21}$$

$$\operatorname{vec}(\operatorname{cov}(\widetilde{\mathbf{y}}_t, \widetilde{\mathbf{y}}_{t-j})) = \left(\sum_{k=0}^{\infty} \mathbf{A}_k \otimes \mathbf{A}_{k+j} \right) \mathbf{D}_3 \operatorname{vech}(\boldsymbol{\Sigma}) = \overline{\mathbf{A}}_j \operatorname{vech}(\boldsymbol{\Sigma}), \tag{A-22}$$

where $\otimes$ denotes the Kronecker product, vec and vech are the vec and half-vec operators, respectively, $\mathbf{D}_n$ denotes the duplication matrix such that, for an $n \times n$ symmetric $\mathbf{X}$, $\mathbf{D}_n \operatorname{vech}(\mathbf{X}) = \operatorname{vec}(\mathbf{X})$, and $\mathbf{D}^{+}$ is its Moore–Penrose inverse.[22] The 3×7 matrix $\overline{\mathbf{A}}_0$ and 4×7 matrices $\overline{\mathbf{A}}_j$ (for $j \geq 1$) are defined in Eqs. (A-21) and (A-22). The system can be stacked into a system of $3 + 4p$ equations:

$$\mathbf{M} \operatorname{vech}(\boldsymbol{\Sigma}) = \left[\overline{\mathbf{A}}_0', \overline{\mathbf{A}}_1', \ldots, \overline{\mathbf{A}}_p' \right]' \operatorname{vech}(\boldsymbol{\Sigma}). \tag{A-23}$$

The explicit expression of $\mathbf{M}$ for the AR(1) model is given in Rytchkov (2012). In this specific case, the column rank of $\mathbf{M}$ is five, which indicates that only five parameters of $\boldsymbol{\Sigma}$

[22] See Abadir and Magnus (2005), ch. 11.

can be identified. Rytchkov (2012) shows that in this case, there exists a matrix $\mathbf{\Omega}$ that satisfies the restrictions: $\sum_{k=0}^{1} \mathbf{A}_k \mathbf{\Sigma} \mathbf{A}_k' = \mathbf{0}$ and $\mathbf{A}_0 \mathbf{\Sigma} \mathbf{A}_1' = \mathbf{0}$, such that replacing $\mathbf{\Sigma}$ with $\widetilde{\mathbf{\Sigma}} = \mathbf{\Sigma} + \mathbf{\Omega}$ yields observationally equivalent statistics.

For systems with richer autoregressive representation, such as AR(p) with $p \geq 2$, the explicit algebraic solution is not readily available, and we turn to brute-force numerical methods to determine the rank of $\mathbf{M}$.[23] In this case, we verify that all $\mathbf{\Sigma}$ coefficients are indeed identified. That is, the rank of $\mathbf{M}$ is six if we consider a sufficient number of moments. In the special case when $\phi_{m,j} = \phi_{g,j}$ for all j, to achieve identification we need the variance and any two autoregressive moments: $\text{rank}(\left[\overline{\mathbf{A}}_0', \overline{\mathbf{A}}_j', \overline{\mathbf{A}}_k'\right]') = 6 \ \forall \ 1 \leq j, k \leq p, \ j \neq k$. In all other cases, as long as the dynamics of m_t are different from g_t, i.e. $\phi_{m,j} \neq \phi_{g,j}$ for any j, all parameters of $\mathbf{\Sigma}$ can be identified from the variance of $\widetilde{\mathbf{y}}_t$ and any other autoregressive moment, i.e. $\text{rank}(\left[\overline{\mathbf{A}}_0', \overline{\mathbf{A}}_j'\right]') = 6 \ \forall 1 \leq j \leq p$. The same result holds if m_t and/or g_t follow a fractionally integrated process, even when there are no short-memory dynamics (i.e. fractional noise) and $\delta_m = \delta_g$.

Appendix D Kalman equations

In this section we discuss the Kalman filtering procedure and then present the log-likelihood function which will subsequently be maximized.

In order to obtain the Kalman equations, it is convenient to write the measurement equations in the form where the shocks are lagged relatively to the state vector. Therefore we define the new state variables $\mathbf{x}_{m,t+1} = \mathbf{C}_{m,t}$ and $\mathbf{x}_{g,t+1} = \mathbf{C}_{g,t}$, so the transition equations are now

$$\mathbf{x}_{m,t+1} = \mathbf{F}\mathbf{x}_{m,t} + \mathbf{h}_m \varepsilon_{m,t}, \tag{A-24}$$

$$\mathbf{x}_{g,t+1} = \mathbf{F}\mathbf{x}_{g,t} + \mathbf{h}_g \varepsilon_{g,t}, \tag{A-25}$$

[23] We use Matlab's Symbolic Math Toolbox.

and the measurement equations are:

$$\Delta d_t = \mu_g + \mathbf{w}'\mathbf{x}_{g,t} + \varepsilon_{d,t}, \tag{A-26}$$

$$pd_t = A + \mathbf{b}'\mathbf{F}\mathbf{x}_{g,t} - \mathbf{b}'\mathbf{F}\mathbf{x}_{m,t} + \mathbf{b}'\mathbf{h}_g\varepsilon_{g,t} - \mathbf{b}'\mathbf{h}_m\varepsilon_{m,t}. \tag{A-27}$$

In general notation, the transition and measurement equations are

$$\mathbf{x}_{t+1} = \overline{\mathbf{F}}\mathbf{x}_t + \mathbf{v}_t, \tag{A-28}$$

$$\mathbf{y}_t = \mathbf{e} + \overline{\mathbf{W}}\mathbf{x}_t + \mathbf{z}_t, \tag{A-29}$$

with

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{x}_{m,t} \\ \mathbf{x}_{g,t} \end{bmatrix}, \overline{\mathbf{F}} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}, \mathbf{v}_t = \begin{bmatrix} \mathbf{h}_m\varepsilon_{m,t} \\ \mathbf{h}_g\varepsilon_{g,t} \end{bmatrix},$$

$$\mathbf{y}_t = \begin{bmatrix} \Delta d_t \\ pd_t \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} \mu_g \\ A \end{bmatrix}, \quad \overline{\mathbf{W}} = \begin{bmatrix} \mathbf{w}' & \mathbf{0} \\ \mathbf{b}'\mathbf{F} & -\mathbf{b}'\mathbf{F} \end{bmatrix}, \quad \mathbf{z}_t = \begin{bmatrix} \varepsilon_{d,t} \\ \mathbf{b}'\mathbf{h}_g\varepsilon_{g,t} - \mathbf{b}'\mathbf{h}_m\varepsilon_{m,t} \end{bmatrix},$$

where $\mathbf{0}$ is an infinite-dimensional matrix of zeros.

The Kalman recursive equations of the model are:

$$\left\{ \begin{array}{c} \boldsymbol{\Delta}_t = \overline{\mathbf{W}}\boldsymbol{\Omega}_t\overline{\mathbf{W}}' + \mathbf{R} \\ \boldsymbol{\Theta}_t = \overline{\mathbf{F}}\boldsymbol{\Omega}_t\overline{\mathbf{W}}' + \mathbf{S} \\ \boldsymbol{\Omega}_{t+1} = \overline{\mathbf{F}}\boldsymbol{\Omega}_t\overline{\mathbf{F}}' + \mathbf{Q} - \boldsymbol{\Theta}_t\boldsymbol{\Delta}_t^{-1}\boldsymbol{\Theta}_t' \\ \widehat{\mathbf{x}}_{t+1} = \overline{\mathbf{F}}\widehat{\mathbf{x}}_t + \boldsymbol{\Theta}_t\boldsymbol{\Delta}_t^{-1}\left(\mathbf{y}_t - \mathbf{e} - \overline{\mathbf{W}}\widehat{\mathbf{x}}_t\right) \end{array} \right\} \tag{A-30}$$

where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{h}_g\mathbf{h}_g'\sigma_g^2 & \mathbf{h}_g\mathbf{h}_m'\rho_{mg}\sigma_m\sigma_g \\ \mathbf{h}_m\mathbf{h}_g'\rho_{mg}\sigma_m\sigma_g & \mathbf{h}_m\mathbf{h}_m'\sigma_m^2 \end{bmatrix},$$

$$\mathbf{R} = \begin{bmatrix} \sigma_d^2 & \mathbf{b}'\mathbf{h}_g\rho_{gd}\sigma_g\sigma_d - \mathbf{b}'\mathbf{h}_m\rho_{md}\sigma_m\sigma_d \\ \mathbf{b}'\mathbf{h}_g\rho_{gd}\sigma_g\sigma_d - \mathbf{b}'\mathbf{h}_m\rho_{md}\sigma_m\sigma_d & (\mathbf{b}'\mathbf{h}_g)^2\sigma_g^2 + (\mathbf{b}'\mathbf{h}_m)^2\sigma_m^2 - 2\mathbf{b}'\mathbf{h}_g\mathbf{b}'\mathbf{h}_m\rho_{mg}\sigma_m\sigma_g \end{bmatrix},$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{h}_g\rho_{gd}\sigma_g\sigma_d & \mathbf{h}_g\mathbf{b}'\mathbf{h}_g\sigma_g^2 - \mathbf{h}_g\mathbf{b}'\mathbf{h}_m\rho_{mg}\sigma_m\sigma_g \\ \mathbf{h}_m\rho_{md}\sigma_m\sigma_d & \mathbf{h}_m\mathbf{b}'\mathbf{h}_g\rho_{mg}\sigma_m\sigma_g - \mathbf{h}_m\mathbf{b}'\mathbf{h}_m\sigma_m^2 \end{bmatrix},$$

using as initial condition $\mathbf{x}_1 = \mathbf{0}$.

The log-likelihood function is then given by:

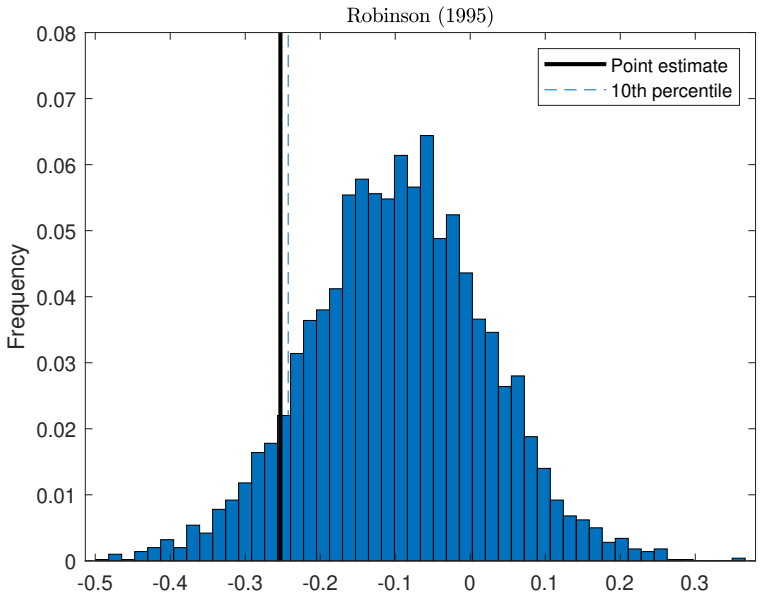
$$\ell = (2\pi)^{-kT/2}\left(\prod_{t=1}^{T}\det\boldsymbol{\Delta}_t\right)^{-1/2}\exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\mathbf{y}_t - \widehat{\mathbf{y}}_t\right)'\boldsymbol{\Delta}_j^{-1}\left(\mathbf{y}_t - \widehat{\mathbf{y}}_t\right)\right) \tag{A-31}$$

with $\widehat{\mathbf{y}}_t = \mathbf{e} + \overline{\mathbf{W}}\widehat{\mathbf{x}}_t$, where T is the sample size, and k is the size of the system.

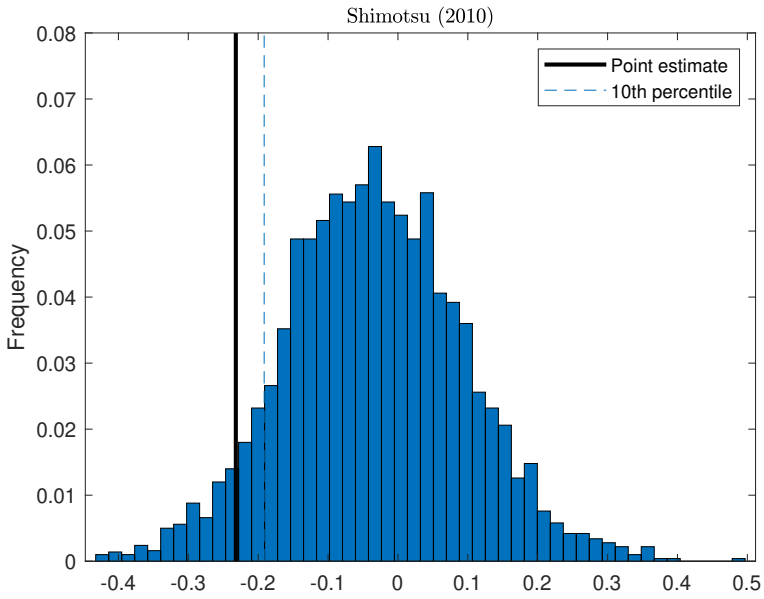
Appendix E Simulation results

Although fractional integration is a phenomenon that, by definition, is manifested through long-range dependence, the estimation techniques allow us to estimate the parameters of the process reliably from relatively short samples.

To that end, we simulated 5000 samples of the variables using the estimates of the PV-AR model and the PV-ARFIMA model set out in Section 3, using the truncation lag $l = 1000$. For each simulated dividend growth series, we estimated the fractional integration parameter. The graphical results for the Robinson (1995) estimator and the Shimotsu (2010) estimator are presented in Figures A-6(a) and (b), respectively. (The results for the Geweke and Porter-Hudak (1983) estimator look very similar and are available upon request.) The results show that the hypothesis of no fractional integration in expected dividend growth is rejected by both estimators at the 5% significance level (p-values equal to 8.78% and 5.52% for the Robinson and the Shimotsu estimators, respectively). Thus, the results for the dividend growth process provide strong evidence in favor of including a fractional integration component in the present-value model.



(a)



(b)

Figure A-6: Histogram for the Robinson (1995) (a) and Shimotsu (2010) (b) estimates for the return series generated by 5000 bootstrap simulations of the PV-AR model. The black line indicates the empirical estimate and the dashed line represents the 10th percentile of the distribution.

In addition, we calculated the bootstrap p-value for the likelihood ratio test reported in Table 4. In Figure A-7 we plot the histogram of the bootstrap LR test, with vertical lines indicating the value of the LR test and the 10% bootstrap cut-off. The bootstrap p-value amounts to 6.38%.

Taken together, the results indicate that it is unlikely that we could observe the results for the long-memory model as reported in the paper if they were a random feature of the short-memory data-generating process.

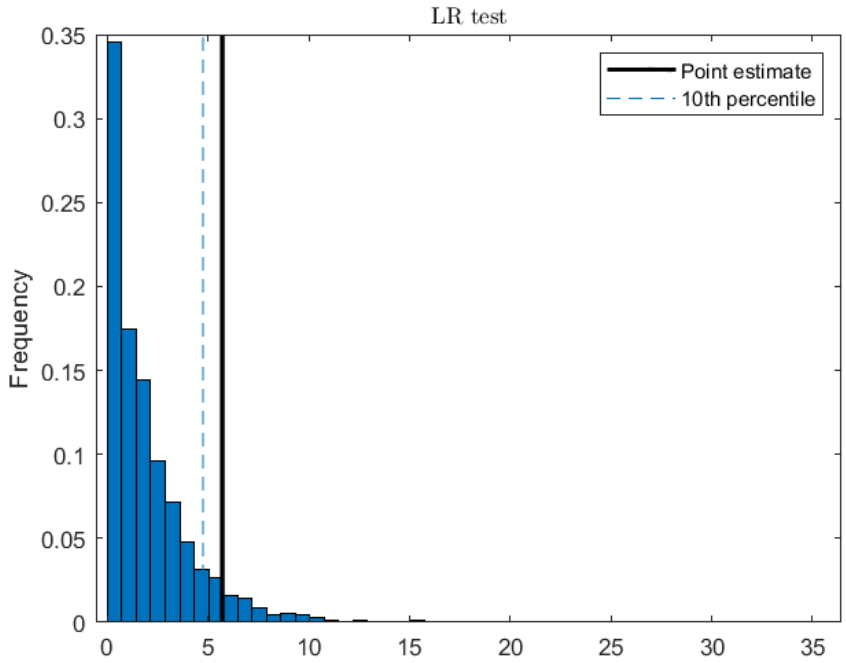


Figure A-7: Histogram of the likelihood ratio test generated by 10,000 bootstrap simulations of the PV-AR model. The black line represents the empirical estimate and the dashed line shows the 10th percentile of the distribution.