Article:
Golinski, Adam, Rodrigues Madeira, Joao Antonio and Rambaccussing, Dooruj (2025) Return Predictability, Dividend Growth and the Persistence of the Price-Dividend Ratio. International journal of forecasting. pp. 92-110. ISSN: 0169-2070

https://doi.org/10.1016/j.ijforecast.2024.03.005

International Journal of Forecasting 41 (2025) 92–110

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

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

Adam Goliński[a,b,*], João Madeira[c], Dooruj Rambaccussing[d]

[a] Banque de France, Financial Economics Research Division, France
[b] University of York, Department of Economics and Related Studies, United Kingdom
[c] Department of Economics and BRU-Business Research Unit, ISCTE-University Institute of Lisbon, Portugal
[d] Economic Studies, School of Business, University of Dundee, United Kingdom

ARTICLE INFO

Keywords:
Price–dividend ratio
Persistence
Fractional integration
Return predictability
Present-value model

ABSTRACT

Empirical evidence shows that the order of integration of returns and dividend growth is approximately equal to the order of integration of the first-differenced price–dividend ratio, which is about 0.7. Yet the present-value identity implies that the three series should be integrated of the same order. We reconcile this puzzle by showing that the aggregation of antipersistent expected returns and expected dividends gives rise to a price–dividend ratio with properties that mimic long memory in finite samples. In an empirical implementation, we extend and estimate the state-space present-value model by allowing for fractional integration in expected returns and expected dividend growth. This extension improves the model's forecasting power in-sample and out-of-sample. In addition, expected returns and expected dividend growth modeled as ARFIMA processes are more closely related to future macroeconomic variables, which makes them suitable as leading business cycle indicators.

1. Introduction

The prediction properties of the price–dividend ratio have strong theoretical foundations grounded in the present-value (PV) identity, popularized in the log-linear form by Campbell and Shiller (1988). As argued by Cochrane (2008a), the fact that the price–dividend ratio is not constant means that either expected returns or expected dividend growth is predictable, or that there is a bubble, so that the price–dividend ratio is non-stationary and not mean-reverting. In fact, the price–dividend ratio has been shown to have strong forecasting power for returns, especially at long horizons (see Fama & French, 1988 and Cochrane, 1999). However, many studies have pointed out that return predictability has been overstated due to the high persistence of the price–dividend ratio (e.g. Mankiw & Shapiro, 1986, Stambaugh, 1999, Goyal & Welch, 2003).

Indeed, the price–dividend ratio is highly persistent. Although traditional unit root tests reject the null hypothesis that the series is integrated of order one, $I(1)$ (see Table 1, Panel A; see also Campbell & Shiller, 1988), stationarity tests also reject the $I(0)$ hypothesis (Table 1, Panel B). Moreover, the semiparametric estimates of the fractional integration parameter δ (Table 1, Panel C) suggest that the price–dividend ratio is integrated of order 0.7, approximately, which is a value close to that found

☆ The views expressed in this article are those of the authors and do not necessarily reflect the views of the Banque de France.

* Correspondence to: Banque de France, Division of Monetary and Financial Studies, 31 rue Croix-des-petits-champs, 75049, Paris, France.
E-mail addresses: adam.golinski@banque-france.fr (A. Goliński), joao.antonio.madeira@iscte-iul.pt (J. Madeira), d.rambaccussing@dundee.ac.uk (D. Rambaccussing).

https://doi.org/10.1016/j.ijforecast.2024.03.005

Table 1
Unit root, $I(0)$, and fractional difference estimates for returns, dividend growth, and price–dividend ratio.
In Panel A we present the results of the unit-root tests: ADF, Phillips–Perron, and fractional ADF tests. In Panel B we report the results of the $I(0)$ tests: KPSS and Lobato–Robinson tests. In Panel C we present the estimates of the fractional integration parameter obtained by the semiparametric estimators: GPH, (Robinson, 1995), and Shimotsu (2010) with the bandwidth equal to 22 observations. The standard errors are reported in small font. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is 1934–2022.

	r	Δd	pd
	Panel A: $I(1)$ tests		
ADF	−8.7643***	−6.6391***	−3.6063**
Phillips–Perron	−8.7643***	−8.1717***	−3.6063**
Fractional ADF			−9.3648***
	Panel B: $I(0)$ test		
KPSS	0.0526	0.0880	0.5761**
Lobato–Robinson	0.4549	0.5484	5.1713**
	Panel C: Fractional difference estimates		
GPH	−0.3370***	−0.2213*	0.6952***
	0.1212	0.1212	0.1212
Robinson	−0.2531***	−0.1387	0.6805***
	0.0928	0.0928	0.0928
Shimotsu	−0.2320**	−0.0226	0.7047***
	0.0918	0.0918	0.0918

in Chevillon and Mavroeidis (2017), about 0.8. These results could suggest that the price–dividend ratio is a long memory process, that is, fractionally integrated of order $0 < \delta < 1$ (with $\delta \geq 1/2$ implying non-stationarity), with slow (typically hyperbolic) decay of the autocorrelation function at long lags and an infinite spike at frequency zero. Such a finding on the price–dividend ratio may be difficult to understand from the traditional asset pricing perspective, since it would imply bubble-like behavior.

The finding of long memory in the price–dividend ratio is nonetheless puzzling. The present-value identity, popularized in the log-linear form by Campbell and Shiller (1988), implies that the price–dividend ratio is a linear function of a discounted stream of expected future dividend growth and stock returns. Since dividend growth and returns are close to being serially uncorrelated, it is hard to think that the unobserved expectations are very persistent. Indeed, we estimate a fractional integration coefficient for both returns and dividend growth of about −0.25, implying the series are antipersistent (that is $\delta < 0$). As noted by Maynard and Phillips (2001), a different order of integration of returns and the price–dividend ratio invalidates statistical inference in predictive regressions. As such, the finding of long memory in the price–dividend ratio poses a question that we address in this paper.

We re-examine the hypothesis of predictability in expected returns from the perspective of the log price–dividend ratio as aggregated expectations of future returns and dividend growth. We show that the apparent long memory in the price–dividend ratio can be generated from antipersistent expected returns and expected dividend growth. If we allow the expected dividend growth and expected returns to be integrated of order $I(\delta) < 0$, then discounting future expectations of these series with the discount factor given by the log-linearization constant gives rise to a price–dividend series that exhibits a spike at frequency zero characteristic for a long memory series process, but finite. In the limiting case, if the discount factor were one, the price–dividend ratio would become a true long memory process $I(\delta + 1)$. Since future expectations are discounted at a rate smaller than one, it follows that the rate of decay of the moving-average coefficients of the price–dividend ratio is asymptotically the same as that of expected returns and expected dividend growth (assuming no cointegration between these two series) for any δ. However, with antipersistent expected returns and expected dividend growth, the price–dividend ratio can appear to be long memory in finite samples. In particular, we show that the process can exhibit a slow decay of autocorrelations and have a convex shape of the spectral density close to zero frequency. In a sense, our explanation of the mechanism generating spurious long memory is similar in spirit to the rare break mechanism proposed by Diebold and Inoue (2001).

In an empirical implementation, we specify expected returns and expected dividend growth as autoregressive fractionally integrated moving-average (ARFIMA) processes. This allows us to reconcile the antipersistent expectation series on one hand with a quasi-persistent price–dividend ratio on the other hand. Within the state-space system, we specify expected returns and expected dividend growth as latent variables and estimate them using the Kalman filter with the maximum likelihood estimator. As such, our empirical approach is similar to Van Binsbergen and Koijen (2010) and Rytchkov (2012), who, using an AR(1) specification, suggest that adopting the state-space present-value framework increases the predictability of returns beyond that from the price–dividend ratio predictive regressions.

Within the present-value framework, we find a negative and statistically significant fractional integration estimate (−0.3) for expected dividend growth and expected returns that roughly matches the semiparametric estimates (about −0.25). The forecasts from the filtered expectations series of the fractionally integrated model are preferable over those obtained with the standard AR(1) model of Van Binsbergen and Koijen (2010). The present-value model with the fractionally integrated component also outperforms the classical forecasting regressions with the price–dividend ratio. Several forecasting exercises on the last 20 years of data confirm the relevance of using a model which allows for fractional integration in improving the forecasting ability of the present-value model, both in-sample and out-of-sample. From a macroeconomic perspective, our filtered series of expected returns is clearly countercyclical, which is in line with many other studies (for a survey, see Campbell & Diebold, 2009). Moreover, we find that the series of expected returns and expected dividend growth filtered from the ARFIMA model predict consumption growth and industrial production growth better than if modeled as AR(1) processes.

The predominance of the spectral density shape that signifies high persistence in economic data was first noted by Granger (1966), who referred to it as 'the typical spectral shape of an economic variable'. The origin of long memory can be plausibly explained by aggregation (Granger, 1980), learning (Chevillon & Mavroeidis, 2017), or structural breaks (Diebold & Inoue, 2001). In Appendix A, we provide an overview of possible origins of long memory in time series. Our paper contributes to this literature by showing that the long-memory-like behavior of the price–dividend ratio can be explained by the aggregation of antipersistent expectations of future returns and dividend growth. Antipersistence, in turn, occurs as a result of overdifferencing a long memory series. Although the formal explanation of the origin of antipersistence in returns and dividend growth is beyond the scope of this paper, we note that it is consistent with the evidence that investors overreact to news and make systematic errors in their expectations (see Poterba & Summers, 1988, Lakonishok, Shleifer, & Vishny, 1994, and La Porta, 1996). Consequently, a slow resolution of uncertainty (possibly combined with another aggregation mechanism) can lead to negative and slowly decaying autocorrelation lags.

The remainder of the paper is organized as follows. In Section 2, we present evidence on long memory in the price–dividend ratio and discuss its implications for related issues, such as log linearization and price bubbles. In Section 3, we propose a theoretical model of the price–dividend ratio with antipersistent expected returns and expected dividend growth, and cast it in the state-space system. The estimation methodology and results are presented in Section 4. In Section 5, we compare the in-sample and out-of-sample performance of the PV-ARFIMA model with that of a nested model and of classical forecasting regressions. In this section, we also show the analysis of the business cycle properties of expected returns and dividend growth obtained from the PV-ARFIMA model. Section 6 concludes.

2. Persistence of the price–dividend ratio

2.1. Data

In our empirical investigation we use value-weighted NYSE/Amex/Nasdaq/Arca index data, including all distributions,[1] available from the Center for Research in Securities Prices (CRSP). Similarly to Van Binsbergen and Koijen (2010) we adopt the assumption of reinvesting the dividends at the risk-free rate.

We construct the annual dividends and price–dividend ratio from the monthly series using the risk-free reinvestment strategy. As the risk-free rate we use the three-month T-bill rate, which is available from January 1934.[2] Our sample ends in 2022, which gives us 89 annual observations. Although monthly or quarterly data would be preferable, we found a strong seasonal pattern in the correlogram of dividend growth series at monthly and quarterly frequencies, which, if not accounted for, invalidates the analysis of the time series dynamics.[3] We then obtain real returns and real dividend growth series by using the consumer price index (CPI) from the U.S. Bureau of Labor Statistics.

2.2. Order of integration of the price–dividend ratio

We begin the analysis with an examination of the time series properties of the stock market data. The total stock market log return (r_{t+1}) and log dividend growth rate (Δd_{t+1}) are defined as:

$$r_{t+1} \equiv \log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right), \tag{1}$$

$$\Delta d_{t+1} \equiv \log\left(\frac{D_{t+1}}{D_t}\right) \tag{2}$$

and the price–dividend ratio (PD_t) is:

$$PD_t \equiv \frac{P_t}{D_t}.$$

Using $pd_t \equiv \log(PD_t)$ and (2), one can re-write the log-linearized return (1) as:

$$r_{t+1} - \Delta d_{t+1} \simeq \kappa + \rho pd_{t+1} - pd_t, \tag{3}$$

with $\overline{pd} = E(pd_t)$, $\kappa = \log(1+\exp(\overline{pd})) - \rho pd$, and $\rho = \frac{\exp(\overline{pd})}{1+\exp(\overline{pd})}$ (see Campbell & Shiller, 1988). In our sample $\overline{pd} = 3.448$ and thus $\rho = 0.968$.

Fig. 1 shows the time series of the logarithm of the price–dividend ratio. From graphical inspection it is clear that the ratio is highly persistent. We therefore explore the hypothesis that the price–dividend ratio may be non-stationary.

In Table 1, Panel A, we report three unit-root tests (the null hypothesis being that the process is $I(1)$): the augmented Dickey and Fuller (1979) (hence ADF) test, the Phillips and Perron (1988) test, and the fractional augmented Dickey–Fuller test by Dolado, Gonzalo, and Mayoral (2002). In Panel B we report two stationarity tests (the null hypothesis being that the process is $I(0)$): Kwiatkowski, Phillips, Schmidt, and Shin (1992) (hence KPSS) and Lobato and Robinson (1998). Although the ADF, Phillips–Perron, and KPSS tests are consistent under the long-memory alternative (Sowell, 1990, Kramer, 1998), it is well known that they have small power properties (see e.g. Diebold & Rudebusch, 1991). The fractional ADF and Lobato–Robinson tests are designed explicitly against the fractional alternative and therefore are expected to exhibit superior in-sample behavior. The number of lags included in the ADF and fractional ADF tests is selected based on the minimum Bayesian information criterion (BIC), while in the other tests we use an automatic lag selection procedure based on the Bartlett kernel, as in Stock (1986). The ADF and Phillips–Perron unit-root tests of the price–dividend ratio include an intercept and time trend, but we exclude them from the tests of returns and dividend growth. The fractional integration parameter in the fractional ADF test was obtained by the (Shimotsu, 2010) estimator reported in Panel C of Table 1.

From Panel A of Table 1 we can see that the unit-root hypothesis is strongly rejected for returns and dividend growth. It is also rejected for the price–dividend ratio at the 5% significance level by the ADF and Phillips–Perron tests. The fractional ADF test, however, which is designed to deal with the fractional alternative, rejects the null of the unit root at the 1% significance level. At the same

1 These include both ordinary and special dividends, but not share repurchases.

2 The data are available at the website of the Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org.

3 See also Ang and Bekaert (2007) and Cochrane (2011), Appendix A.1.

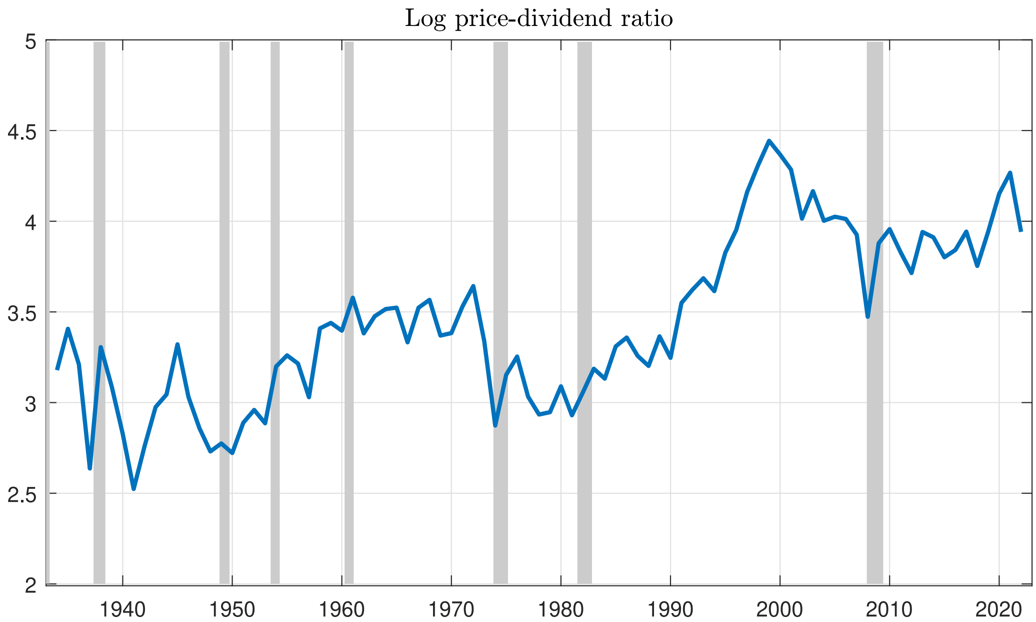

Fig. 1. Time series of log price–dividend ratio.
The figure plots the logarithm of the price–dividend ratio (y-axis) against the year (x-axis). The grey areas refer to the NBER recessionary periods (only those longer than nine months).

time, as indicated in Panel B, the $I(0)$ tests reject the null for the price–dividend ratio but not for stock returns or dividend growth. In summary, the results indicate that stock returns and dividend growth are consistent with the $I(0)$ assumption, while the price–dividend ratio is neither $I(0)$ nor $I(1)$.

We now examine the hypothesis that the price–dividend ratio is integrated of order higher than zero but smaller than one. Table 1, Panel C reports the estimates of the order of fractional integration (δ) of the price–dividend ratio, returns, and dividend growth series obtained using three different semiparametric estimators: the periodogram regression proposed by Geweke and Porter-Hudak (1983) (hence GPH), the Gaussian semiparametric estimator introduced by Robinson (1995), and the two-step exact local Whittle estimator proposed by Shimotsu (2010). The GPH and Robinson estimators were designed for stationary time series ($\delta < 1/2$). Therefore, when using these estimators we first-difference the price–dividend series to estimate $\delta - 1$. The Shimotsu estimator is valid for both stationary and non-stationary time series ($\delta \geq 1/2$) as well as in the presence of structural instability. If the fractional integration parameter is larger than zero, the series is said to exhibit long memory, which means that it displays slowly decaying positive autocorrelation at long lags: $\psi_j \sim (\Gamma(\delta))^{-1} j^{\delta-1}$ for $j \to \infty$, where Γ denotes the gamma function. On the other hand, if $\delta < 0$, we say that the series is antipersistent; in this case the series has asymptotically negative autocorrelations decaying at a hyperbolic rate. For $\delta = 0$ the series is a short-memory process. Moreover, the series is stationary if $\delta < 1/2$ and invertible if $\delta > -1/2$.[4]

The semiparametric estimators do not make any assumptions regarding the dynamics away from very low frequencies. Specifically, we use bandwidth equal to $T^{0.75}$ for the GPH estimator and one-third of the sample for the Robinson and Shimotsu (2010) estimators, which gives bandwidth of 29 and 30 observations, respectively. We checked that the choice of bandwidth has little effect on the estimation results.[5] The semiparametric estimates all conclude that the estimates of the price–dividend ratio fractional parameter (δ) are about 0.7. Table 1, Panel C also shows that the time series of returns and dividend growth seem to be integrated of order smaller than zero.

Using log linearization for a non-stationary variable could be questionable, since the approximation error can become big when the function is far from the point of linearization. Given the evidence on the fractional integration coefficient of the price–dividend ratio being larger than 0.5, the concern about the stationarity of the series seems valid. However, in the next section we argue that the non-stationarity of the price–dividend ratio is a small-sample phenomenon that disappears asymptotically. This means that the technique of log linearization around the mean remains valid. Also, it should be noted that, even for non-stationary series, in finite samples log linearization could be used for approximation around any point, such as the sample mean, although it can invalidate asymptotic inference. Moreover, in the rational bubbles framework with locally explosive expectations, Engsted, Pedersen, and Tanggaard (2012) document that

[4] See Granger and Joyeux (1980) and Hosking (1981).

[5] The exceptions here are the Shimotsu (2010) estimates for the dividend growth process; using one-quarter of the sample (22 observations) yields an estimate of the fractional integration parameter equal to −0.1504.

the approximation is “very accurate even in the presence of large explosive bubbles”. Some other examples of using Campbell and Shiller (1988) log linearization with locally non-stationary variables are Balke and Wohar (2009), Wu (1997), and Phillips, Wu, and Yu (2011).

2.3. Non-stationarity of the price–dividend ratio and price bubbles

Another aspect of the apparent non-stationarity of the price–dividend ratio is that it could point towards a price bubble. If rational bubbles exist, then the price and dividend levels should not be cointegrated. Since we should expect prices and dividends to be unit-root processes, if these series are not cointegrated, then the price–dividend ratio will also be non-stationary. In the next section we argue that, due to the aggregation of antipersistent expectations, the non-stationarity of the price–dividend ratio is spurious and, in fact, the price–dividend ratio is asymptotically a stationary process.

Nonetheless, we formally examine the fractional cointegration hypothesis between prices and dividends by applying the nonparametric variance ratio test developed by Nielsen (2010). We say that two series integrated of order δ are fractionally cointegrated if there exists a linear combination of them that is integrated of order $I(\delta - b)$ for some $b > 0$. As such, it could be expected that the estimated fractional difference parameter for the pd_t series should be smaller than that for prices and dividends.[6] In fact, since the order of integration of the price–dividend ratio reported in Table 1 is smaller than and statistically different from 1, it indicates that prices and dividends are actually cointegrated.

The results of the Nielsen test are reported in Table 2. Although the order of integration of the series is not needed to find the value of the test, it does affect the distribution of the statistic. Therefore we simulate the critical values for the assumed order of integration for prices and dividends equal to 0.8, 0.9, and 1. Despite a relatively small number of observations, as for the nonparametric test, we reject the null hypothesis of no cointegration when the order of integration of prices and dividends is 0.9 or higher, arguably the most economically relevant cases.[7] As such, we find that there is little evidence of a price bubble in our data sample.

Table 2
Nielsen (2010) ($\Lambda_{2,0}(0.1)$) variance ratio test for cointegration.
The test for price and dividends levels includes a non-zero mean and linear trend. The top panel shows the results of the test for price levels and dividends and the bottom panel for returns and dividend growth. The p-values for the Nielsen test were obtained from the simulated distribution for different values of delta. The sample period is 1934–2022.

	$p - d$		
Nielsen	3.87		
δ_0	0.8	0.9	1.0
p-value	0.25	0.09	0.03
	$r - \Delta d$		
Nielsen	5.03		
δ_0	−0.2	−0.1	0.0
p-value	0.92	0.48	0.07

3. Present-value model with fractional integration

3.1. Aggregation of expectations in the log price–dividend ratio

At this point, we should consider the order of integration of the series in the log-linearized return Eq. (3). Since returns and dividend growth are stationary with approximately the same order of integration (see Table 1), the apparent non-stationarity of the price–dividend ratio poses a puzzle, which we now address.

After rearranging (3) for the price–dividend ratio, iterating forward and taking conditional expectations, we obtain the PV identity:

$$\begin{aligned} pd_t &= \frac{\kappa}{1-\rho} + E_t\left[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}\right] - E_t\left[\sum_{j=0}^{\infty} \rho^j r_{t+1+j}\right] \quad (4) \\ &= \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^j E_t\left(g_{t+j} - m_{t+j}\right), \end{aligned}$$

where $m_t \equiv E_t[r_{t+1}]$ and $g_t \equiv E_t[\Delta d_{t+1}]$. The PV identity above reveals that the log price–dividend ratio is determined by expected future dividend growth and returns discounted at rate ρ.

Consider the implications of the aggregation of expectations in (4). Under rational expectations, if the expected returns and expected dividend growth are constant and there are no rational bubbles, then the price–dividend ratio should be constant, a result emphasized by Cochrane (2008a). On the other hand, the fact that the price–dividend ratio is time-varying can be interpreted as evidence that either expected dividend growth, expected returns, or both, are time-varying. Generally, we can re-write their difference using an infinite-order moving-average representation.

Assumption 1. The expected dividend growth and expected returns allow a linear representation

$$g_t - m_t = \mu + \varphi_0 \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots, \quad (5)$$

where $\varepsilon_t \sim i.i.d.N(0, 1)$ and $0 < \varphi_0 < \infty$.

Skipping for simplicity the constant term, the price–dividend ratio is then:

$$\begin{aligned} pd_t &= \varepsilon_t \sum_{j=0}^{\infty} \rho^j \varphi_j + \varepsilon_{t-1} \sum_{j=0}^{\infty} \rho^j \varphi_{j+1} + \varepsilon_{t-2} \sum_{j=0}^{\infty} \rho^j \varphi_{j+2} + \cdots \\ &= \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots, \quad (6) \end{aligned}$$

where $\psi_i = \sum_{j=0}^{\infty} \rho^j \varphi_{j+i}$.

Consider for a moment the counterfactual case when the discount factor ρ is equal to unity. When expected returns and expected dividend growth are ARMA processes, such that the decay of the moving-average terms φ_j is

6 See also Cunado, Gil-Alana, and de Gracia (2005) and Koustas and Serletis (2005).

7 See Nielsen (2010), n. 1.

geometric, i.e. $\varphi_j \sim ca^j$ as $j \to \infty$, where $|a| < 1$ and c is a constant, then the partial sums ψ_j will also decay geometrically as j becomes large. For example, assuming that $g_t - m_t$ is an AR(1) process, so that $\varphi_j = \varphi^j$, then $\psi_j = \varphi^j/(1-\varphi)$. On the other hand, when the moving-average coefficients in (5) decay hyperbolically, i.e. $\varphi_j \sim cj^{\delta-1}$ as $j \to \infty$ with $\delta < 0$, where c is again a generic constant, then, by applying Lemma D.2 from Goliński and Zaffaroni (2016), $\psi_j \sim cj^{\delta}$. Therefore, the order of integration of the price–dividend ratio depends on whether the rate of decay of φ's is geometric or hyperbolic.

In the present-value context of (4), $\rho < 1$ by construction. Thus, in the limit, the rate of decay of ψ_j is the same as φ_j. However, interesting behavior of the price–dividend ratio emerges when the order of integration of the underlying series $g_t - m_t$ is negative, i.e. when $\delta < 0$. We maintain this assumption throughout this section.

Assumption 2. The moving-average terms in Eq. (5) decay hyperbolically, i.e. $\varphi_j \sim c_0 j^{\delta-1}$ with $\delta \in (-1, 0)$ and c_0 a constant.

In the empirical part, we make a stronger assumption about the data-generating process by assuming that expected returns and expected dividend growth follow an ARFIMA(p, δ, q) process, which implies a similar representation for $g_t - m_t$.

Assumption 2'. The joint process of expected returns and expected dividend growth in Eq. (5) follows an autoregressive fractionally integrated moving-average process ARFIMA(p, δ, q) with $\delta \in (-1, 0)$.

Remark 1. There is no assumption of cointegration between expected dividend growth and expected returns, or lack thereof. What is required, however, is the antipersistence of the joint process $g_t - m_t$. For example, expected returns and expected dividend growth can both be fractionally integrated processes with $\delta_g = \delta_d$, but due to cointegration, the fractional parameter for the series $g_t - m_t$ is smaller than zero. Alternatively, the two series can have different orders of integration, $\delta_g \neq \delta_d$, that satisfy $\delta_g < 0$ and $\delta_d < 0$; in this case the order of integration of the joint process $g_t - m_t$ will be $\max[\delta_g, \delta_d]$.

If the process $g_t - m_t$ exhibits antipersistence as assumed above, then in general the sum of autocovariances of the price–dividend ratio, $\sum_{j=-\infty}^{\infty}\gamma(j)$, will be finite and different from zero. Consequently, the spectral density function will exhibit a finite spike at zero frequency. The behavior of the spectral density near zero frequency is summarized in the following theorem.

Theorem 1. *Let Assumptions 1 and 2 hold. Then, the spectral density of the log-linearized price–dividend ratio has the following properties:*

(i)

$$\lim_{\lambda \to 0^+} \frac{\partial f(\lambda)}{\partial \lambda} = -\infty. \tag{7}$$

(ii)

$$0 < f(0) < \infty; \tag{8}$$

(iii) Under assumptions 1 and 2', the spectral density at frequency zero is:

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

Proof. See Appendix B.[8] □

Remark 2. Theorem 1 describes the behavior of the spectral density near frequency zero, which is different from that of the typical supposition about the behavior of the price–dividend ratio. In particular, if the price–dividend ratio follows a short-memory autoregressive process, then the spectral density

$$f(\lambda) = \frac{1}{2\pi}\left(\sum_{j=0}^{\infty}\psi_j e^{i\lambda j}\right)\left(\sum_{j=0}^{\infty}\psi_j e^{-i\lambda j}\right) \tag{10}$$

is finite and bounded away from zero with $\partial f(\lambda)/\partial\lambda = 0$ at $\lambda = 0$. If the price–dividend ratio follows a long-memory process, such as ARFIMA(p, δ_{pd}, q) with $\delta_{pd} > 0$, then the sum of the moving-average coefficients and the spectral density at frequency zero are unbounded. If the price–dividend ratio follows an ARFIMA process with negative memory, i.e. $\delta_{pd} < 0$, then $\sum_{j=0}^{\infty}\psi_j = 0$ and the spectral density at frequency zero is zero. Thus, the behavior described in Theorem 1 is distinctly different from any of these cases.

For the sake of exposition, assume that $g_t - m_t$ in (5) follows an ARFIMA(0, δ, 0) process with $\delta = -0.2$, which is approximately the value of the fractional integration parameter estimated for the dividend growth and return series in Table 1. In Fig. 2 we plot the normalized rate of decay of ψ_k calculated as $1 + \log(\psi_{k+1}/\psi_k) \times k$ for different values of ρ. Recall that for large k, $\varphi_k \simeq ck^{\delta-1}$. When $\rho = 1$ we can see that the pd_t process becomes $I(0.8)$. On the other hand, for values of $\rho < 1$ the asymptote of the normalized rate of decay is the same as for the underlying process, -0.2. The convergence to this asymptote, however, is very slow.

In Fig. 3 we plot the spectral density of the price–dividend ratio for different values of ρ. For $\rho = 1$ the shape of the spectral density becomes unbounded and corresponds to the long-memory process with δ equal to 0.8. For $\rho < 1$ near frequency zero, the series shows a sharp increase reminiscent of the spike displayed by a genuine long-memory process. As such, in a finite sample, the price–dividend ratio is likely to appear as a long-memory series. Using the property of an ARFIMA model that $\sum_{j=0}^{\infty}\varphi_j = 0$ when $\delta < 0$, we can show that the value

[8] We are grateful to Karim Abadir for presenting us with the proof of Theorem 1(ii) under the more general Assumption 2. See also Abadir, Heijmans, and Magnus, Section A.4.1

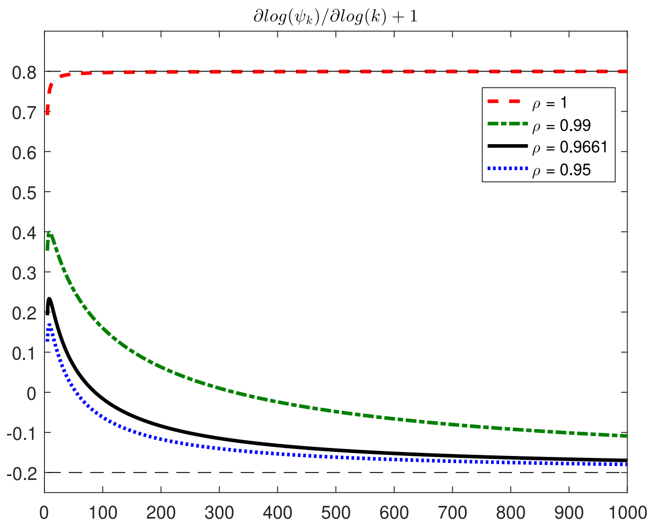

Fig. 2. Partial derivative plot against sample size.
The differential of $\log(\psi_k)$ in Eq. (5) with respect to $\log(k) + 1$ (y-axis) is plotted against the sample size (x-axis). Different values of ρ are represented by different colors and patterns. The asymptotes denote the order of integration of the series.

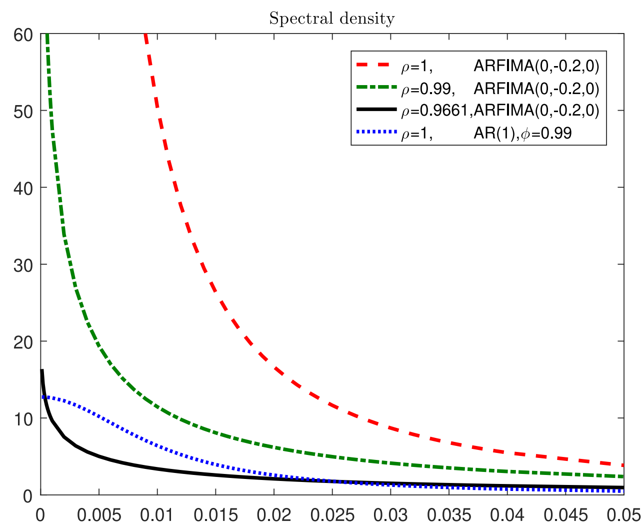

Fig. 3. Spectral density of the price–dividend ratio.
The figure shows the implied spectral density for the underlying processes in Eq. (5) following the ARFIMA(0, −0.2, 0) or AR(1) process for different values of ρ. The spectral densities for different values of ρ are represented by different colors and patterns.

of the spectral density at the origin is given by[9]:

$$f(0) = \frac{1}{2\pi}\left|\sum_{j=0}^{\infty}\psi_j\right|^2 = \frac{1}{2\pi}\left|\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\rho^j\varphi_{i+j}\right|^2 = \frac{1}{2\pi}\frac{\rho^2\psi_0^2}{(1-\rho)^2}, \tag{11}$$

and, as should be expected, it is higher for larger values of ρ. For comparison we also plot the spectral density generated by an AR(1) process with the autocorrelation coefficient equal to 0.99. The spectral density of the AR(1) series near the origin becomes flat, which is plainly different from a long-memory process or the model with a finite spike at frequency zero described in Theorem 1.

It should be emphasized that long memory (or fractional integration) is an asymptotic concept. In Theorem 2 we establish that when $\rho < 1$ and $\delta < 0$, the rate of decay of the moving-average coefficients of the price–dividend ratio is hyperbolic. Yet as we saw in Theorem 1, the spectral density at frequency zero is finite and bounded away from zero. It follows that the price–dividend ratio will be (asymptotically) stationary, but due to its structure as integrated expectations, in a finite sample it can be well described as a non-stationary long-memory process.

9 See Appendix B.

Theorem 2. *Let Assumptions 1 and 2 hold. Then, the moving-average coefficients of the log-linearized price–dividend ratio in (6), ψ_k, decay asymptotically at the same rate as the moving-average coefficients of the underlying process:*

$$\psi_k \sim c_1 k^{\delta-1}, \tag{12}$$

where $c_1 < 0$ is a constant.

Proof. See Appendix B. □

3.2. State-space ARFIMA model of the price–dividend ratio

The analysis presented in Section 2 indicates that taking the fractional integration of returns and dividend growth into account should result in improved statistical inference and better ability of the present-value model to account for different aspects of the data. Therefore, in this section we present a present-value model with a fractional integration component.

Van Binsbergen and Koijen (2010) specified expected returns ($m_t \equiv E_t[r_{t+1}]$) and expected dividend growth ($g_t \equiv E_t[\Delta d_{t+1}]$) as AR(1) processes. We consider instead the more general ARFIMA process. In particular, we model expected returns and expected dividend growth as ARFIMA$(1, \delta_m, 0)$ and ARFIMA$(1, \delta_g, 0)$ processes, respectively:

$$(1-\phi_m L)(1-L)^{\delta_m}(m_t - \mu_m) = \varepsilon_{m,t}, \tag{13a}$$

$$(1-\phi_g L)(1-L)^{\delta_g}(g_t - \mu_g) = \varepsilon_{g,t}, \tag{13b}$$

where L is the lag operator, δ_m and δ_g are fractional integration coefficients, and $\varepsilon_{m,t}$ and $\varepsilon_{g,t}$ are zero-mean i.i.d. series. To satisfy Assumption 1, the processes m_t and g_t are assumed to be stationary, which holds when $|\phi_i| < 1$ and $\delta_i < 1/2$, for $i \in \{m, g\}$. Theorems 1 and 2, however, require a stronger assumption regarding the rate of decay of the moving-average terms (see Assumption 2), which is satisfied when both δ_m and δ_g are smaller than zero (see Assumption 2'). Whether or not these conditions are satisfied is an empirical issue and, as such, it remains our hypothesis of interest in the empirical implementation of the model. Finally, we note that the moving-average coefficients of the process $g_t - m_t$ (ϕ in Eq. (5)) are a sum of moving-average coefficients of g_t and the negative of m_t, i.e. $\varphi_j = \varphi_{g,j} - \varphi_{m,j}$. It follows that the rate of decay of the coefficients, and thus the order of integration of $g_t - m_t$, is determined by the series with the higher order of integration, as mentioned in Remark 1. When $\delta_m = \delta_g = 0$, our model corresponds to the one in Van Binsbergen and Koijen (2010).

We specify the expectation series m_t and g_t as:

$$m_t = \mu_m + \mathbf{w}'\mathbf{C}_{m,t}, \tag{14a}$$

$$g_t = \mu_g + \mathbf{w}'\mathbf{C}_{g,t}, \tag{14b}$$

where $\mathbf{w} = [1, 0, 0, \cdots]'$, and $\mathbf{C}_{m,t}$ and $\mathbf{C}_{g,t}$ are infinite-dimensional state vectors that can be expressed as $\mathbf{C}_{k,t} = \left[x_{k,t},\ E_t(x_{k,t+1}),\ E_t(x_{k,t+2}), \ldots\right]'$, for $k = \{m, g\}$, where $E_t(x_{k,t+j}) = \sum_{i=j}^{\infty} \varphi_{k,i}\varepsilon_{k,t+j-i}$ and $\varphi_{k,i}$ are functions of ARFIMA parameters in (13a)–(13b). The transition equations are:

$$\mathbf{C}_{m,t+1} = \mathbf{F}\mathbf{C}_{m,t} + \mathbf{h}_m\varepsilon_{m,t+1}, \tag{15a}$$

$$\mathbf{C}_{g,t+1} = \mathbf{F}\mathbf{C}_{g,t} + \mathbf{h}_g\varepsilon_{g,t+1}, \tag{15b}$$

with $\mathbf{F}$, $\mathbf{h}_m$, and $\mathbf{h}_g$ given by[10]:

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \\ \vdots & & & \ddots \end{bmatrix}, \quad \mathbf{h}_m = \begin{bmatrix} 1 \\ \varphi_{m,1} \\ \varphi_{m,2} \\ \vdots \end{bmatrix},$$

$$\mathbf{h}_g = \begin{bmatrix} 1 \\ \varphi_{g,1} \\ \varphi_{g,2} \\ \vdots \end{bmatrix}.$$

The vectors $\mathbf{h}_m$ and $\mathbf{h}_g$ contain coefficients of the moving-average representation of the expected returns and expected dividend growth series, as in (13a) and (13b), respectively. In particular, denote the moving-average coefficients of the pure fractionally integrated process of order δ by a_j, such that $a_j = \frac{j+\delta-1}{j}a_{j-1}$ (starting from $a_0 = 1$), and the moving average of a stationary ARMA process by b_j (in the AR(1) case which we consider, $b_j = \phi^j$). Then, the moving average of ARFIMA$(1, \delta, 0)$ is $\varphi_j = \sum_{k=0}^{j} a_k b_{j-k}$.

The realized returns and dividend growth rate are equal to the expectation series plus an orthogonal shock:

$$r_{t+1} = m_t + \varepsilon_{r,t+1}, \tag{16a}$$

$$\Delta d_{t+1} = g_t + \varepsilon_{d,t+1}, \tag{16b}$$

where $\varepsilon_{r,t+1}$ and $\varepsilon_{d,t+1}$ are Gaussian white noise processes. Eqs. (16a)–(16b) constitute a signal plus noise model (see e.g. Sun & Phillips, 2003, Dalla, Giraitis, & Hidalgo, 2006). The results presented in Section 2.2 suggest that the expected series (signal) may be antipersistent. However, since the order of integration of a series is equal to the highest integration order of its components, the additive Gaussian noise makes the realized series formally an $I(0)$ process. In consequence, when the antipersistent signal is sufficiently strong, in small samples it could be detected by standard estimators, but distorted and biased towards zero by the presence of noise.

As demonstrated by (Cochrane, 2008), shocks to either realized returns $\varepsilon_{r,t+1}$ or realized dividend growth $\varepsilon_{d,t+1}$ can be expressed as a function of other innovations. Thus, effectively we need only two observation equations. Following Cochrane we substitute out the shocks to realized returns (see Appendix C) and thus, we obtain the following measurement equations:

$$\Delta d_{t+1} = \mu_g + \mathbf{w}'\mathbf{C}_{g,t} + \varepsilon_{d,t+1}, \tag{17a}$$

$$pd_t = A + \mathbf{b}'\mathbf{C}_{g,t} - \mathbf{b}'\mathbf{C}_{m,t}, \tag{17b}$$

where $A = (\kappa + \mu_g - \mu_m)/(1-\rho)$ and $\mathbf{b} = [1, \rho, \rho^2, \ldots]'$.

10 A similar state-space long-memory model was proposed by Goliński and Zaffaroni (2016) in an application to the term structure of interest rates.

To complete the model we also need to specify the covariance matrix of the structural shocks, which we assume to be homoscedastic[11]:

$$\mathbf{\Sigma} = var\left(\begin{bmatrix} \varepsilon_{m,t+1} \\ \varepsilon_{g,t+1} \\ \varepsilon_{d,t+1} \end{bmatrix}\right) = \begin{bmatrix} \sigma_m^2 & \sigma_{mg} & \sigma_{md} \\ \sigma_{mg} & \sigma_g^2 & \sigma_{gd} \\ \sigma_{md} & \sigma_{gd} & \sigma_d^2 \end{bmatrix}.$$

4. Estimation results

4.1. Identification and estimation methodology

Since log linearization is based on the first-order approximation, it does not hold exactly for the observed data. Following Cochrane (2008a) and Van Binsbergen and Koijen (2010), in the model estimation we impose the identity structure given in (3) by using the observed log returns and generating the dividend growth rates from the identity.[12]

In the preliminary estimation of the model, we considered different autoregressive and moving-average orders of ARFIMA processes, as well as different orders of integration δ for expected returns and expected dividend growth. We found the ARFIMA$(1, \delta, 0)$ model of expected returns and expected dividend growth with $\delta_m = \delta_g = \delta$ to be the most favorable specification in terms of the Bayesian information criterion and likelihood-ratio test. In particular, relaxing the constraint $\delta_m = \delta_g$ has negligible effect on the value of the likelihood function. As such, we maintain this assumption in the rest of the paper. As a benchmark, we also estimate the model with expected returns and expected dividend growth parametrized as AR(1). We refer to these two specifications as the PV-ARFIMA and PV-AR models, respectively.

Following Rytchkov (2012), we parametrize the model using correlations rather than covariances: $\rho_{mg} = \sigma_{mg}/(\sigma_m\sigma_g)$, $\rho_{gd} = \sigma_{gd}/(\sigma_g\sigma_d)$ and $\rho_{md} = \sigma_{md}/(\sigma_m\sigma_d)$. As pointed out by Rytchkov (2012) and Cochrane (2008), the dimension of the covariance matrix of shocks is not identified in a system with expected returns and expected dividend growth following an AR(1) process. For other dynamic systems, however, the identification status of the covariance parameters has not been clear, which posed a gap in the literature. We examine this issue in Appendix C (in which we use results from Abadir and Magnus (2005), ch. 11). We show that all $\mathbf{\Sigma}$ coefficients are identified, for present-value models with richer short-memory dynamics (e.g. AR(p) with $p \geq 2$) or with long-memory dynamics. Nonetheless, following Rytchkov (2012) and Van Binsbergen and Koijen (2010), we decided to use the same specification of $\mathbf{\Sigma}$ in both PV-AR and PV-ARFIMA and set the correlation between the expected dividend growth and the realized dividend shock to zero ($\rho_{gd} = 0$) for two reasons: (a) it allows us to attribute the difference in performance of the two models solely to the fractional integration feature, and (b) in the preliminary analysis we estimated the PV-ARFIMA with all $\mathbf{\Sigma}$ parameters and found that ρ_{gd} is close to zero and statistically insignificant ($\hat{\rho}_{gd} = 0.09$ with a p-value equal to 0.87).

Thus, the set of parameters to estimate is:

$$\mathbf{\Theta} \equiv (\mu_m, \phi_m, \delta, \mu_g, \phi_g, \sigma_m, \sigma_g, \sigma_d, \rho_{mg}, \rho_{gd}, \rho_{md})$$

for the PV-ARFIMA model, and the same excluding δ for the PV-AR model. The log-linearization parameters (κ, ρ) are defined by the sample mean (see Section 2.2) and as such are not subject to estimation.

The model is estimated by means of maximum likelihood estimation (MLE). We assume that the error terms have a multivariate Gaussian distribution, which, since the measurement and transition equations consist of a linear dynamic system, allows us to compute the likelihood using the Kalman filter (Hamilton, 1994).[13] The transition equations are given by (15a) and (15b) and the measurement equations by (17a) and (17b). Despite the fact that the state vectors are infinitely dimensional, Chan and Palma (1998) showed that the consistent estimator of an ARFIMA process is obtained when the state vector is truncated at a lag $l \geq \sqrt{T}$. In their Monte Carlo simulation, Chan and Palma (1998) showed that the approximate MLE works well in sample sizes as small as 100 observations, which is close to our sample of 89 observations. In the estimation we set the truncation lag at $l = 40$.[14]

4.2. Results

The estimates of the models are reported in Table 3. Expected returns exhibit strong and positive autoregressive dynamics. Allowing for fractional integration in the model increases the autoregressive coefficient from 0.84

Table 3
Estimation results of the present-value models.
The models of expected returns and expected dividend growth are specified as AR(1) for the PV-AR model and ARFIMA$(1, \delta, 0)$ for the PV-ARFIMA model. The asymptotic standard errors are reported in small font. *, ** and, *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is 1934–2022.

	Model	
	PV-AR	PV-ARFIMA
μ_m	0.0698***	0.0682***
std.err	0.0111	0.0052
ϕ_m	0.8442***	0.9233***
std.err	0.0634	0.0342
μ_g	0.0265***	0.0244***
std.err	0.0099	0.0044
ϕ_g	−0.1813	0.3027**
std.err	0.1880	0.1405
δ	–	−0.3109***
std.err		0.1014
σ_g	0.0491***	0.0948***
std.err	0.0095	0.0071
σ_m	0.0287***	0.0460***
std.err	0.0103	0.0114
σ_d	0.0839***	0.0002
std.err	0.0062	0.0094
ρ_{mg}	−0.8639***	−0.2324***
std.err	0.0001	0.0001
ρ_{md}	0.5036***	0.9725***
std.err	0.0001	0.0001

11 The present-value model with time-varying risk was analyzed by Piatti and Trojani (2017).

12 In our sample the average approximation error amounts to 36 basis points, and the correlation between these two series amounts to 0.9989.

13 See Appendix D for details.

14 We checked that the results are robust to other choices of the truncation lag.

Table 4
Estimation statistics of the present-value models.
In Panel A we report: the likelihood ratio test performed relative to the PV-AR model with associated p-values, asymptotic and bootstrap, reported in small font; the sample standard deviation of expected returns; the sample standard deviation of expected dividend growth; the model-implied correlation between expected returns and expected dividend growth, one lag for realized returns, dividend growth, expected returns, and expected dividend growth; and the R^2 coefficient of returns and dividend growth. Panel B of the table presents the variance decomposition of the price–dividend ratio. The sample period is 1934–2022.

	Model	
	PV-AR	PV-ARFIMA
Panel A: Model statistics		
LR	–	5.70
$p-value(asymptotic)$		(0.0169)
$p-value(bootstrap)$		[0.0638]
$\sigma(m_t)$	5.36%	6.53%
$\sigma(g_t)$	4.99%	9.64%
$\widehat{corr}(r_t, r_{t-1})$	−0.05	−0.12
$\widehat{corr}(\Delta d_t, \Delta d_{t-1})$	−0.05	0.02
$\widehat{corr}(\mu_t, \mu_{t-1})$	0.84	0.70
$\widehat{corr}(g_t, g_{t-1})$	−0.18	0.02
$\widehat{corr}(m_t, g_t)$	−0.39	−0.14
R^2_r	0.14	0.20
$R^2_{\Delta d}$	0.24	0.25
Panel B: *pd* variance decomposition		
exp. returns	41.71%	44.05%
exp. div. growth	0.87%	7.82%
covariance	57.42%	48.13%

to 0.92. High autocorrelation of expected returns is consistent with the findings of others in the literature (see Fama & French, 1988, Ferson, Sarkissian, & Simin, 2003, and Pástor & Stambaugh, 2009). The autoregressive coefficient for expected dividend growth for the PV-AR is close to zero and not statistically significant, while for the PV-ARFIMA model it is positive (0.30) and significant at the 5% level.

The estimate of the fractional integration parameter is negative at −0.31 and statistically significant. The point estimate is more negative than the semiparametric estimates reported in Table 1, but the difference is not statistically significant. The smaller semiparametric estimates (in absolute terms) of the realized series are also consistent with the signal plus noise model, as in (16a)–(16b), since the antipersistent signal is biased towards zero in the observed series.

The estimate of the volatility shocks to the expected dividend growth σ_g in the PV-ARFIMA model is almost twice as big as in the PV-AR model (0.09 and 0.05 for the two models, respectively). Similarly, the estimate of the volatility of shocks to expected returns σ_m is higher for the PV-ARFIMA model, about 0.05, while the PV-AR model it amounts to 0.03. The correlation between the innovations of expected returns and expected dividend growth ρ_{mg} is higher in the PV-AR model at −0.86, while for the PV-ARFIMA it is −0.23. On the other hand, the correlation between shocks to expected returns and realized dividends ρ_{md} is lower for the PV-AR model, 0.50, while it amounts to 0.97 for the PV-ARFIMA model.

The statistics of the two estimated models are presented in Table 4. The first line shows the likelihood-ratio (LR) test of equal fit to the data of the two models, with PV-AR being the nested model. The LR test favors the PV-ARFIMA model with the asymptotic and the bootstrap p-values equal to 1.7% and 6.38%. The next two lines of Table 4 report the sample standard deviations of filtered expected returns and expected dividend growth. The variability of the implied time series increases when we allow for fractional integration.

In the next two lines we report the model-implied first-order autocorrelation of returns and dividend growth implied by the model parameters. In our sample, the autocorrelation of returns amounts to about −0.09 (not reported in any table). The autocorrelation based on the PV-ARFIMA estimates is slightly more negative, −0.12, while the PV-AR model implies −0.05. The sample first-order autocorrelation in the dividend growth series is slightly positive and amounts to 0.08 (not reported in any table). The value implied by the PV-ARFIMA model is smaller (0.02), but closer than the autocorrelation implied by the PV-AR model (−0.05). It is interesting to compare these numbers to the figures reported in the next two lines for expected returns and expected dividend growth. The model-implied first-order autocorrelations for expected returns for the PV-ARFIMA and PV-AR models are 0.70 and 0.84, respectively, and are comparable to others reported in the literature.[15] It should be noted that the pattern of short-run positive and long-run negative autocorrelation implied by the PV-ARFIMA model is consistent with the well-known phenomenon of long-run reversal in stock returns (see e.g. Fama & French, 1988, and Cutler, Poterba, & Summers, 1988). The implied autocorrelation of expected dividend growth is close to zero for the PV-ARFIMA model and negative (−0.18) for the PV-AR model.

In the following row of Table 4 we report the model-implied correlation between expected returns and expected dividend growth. The correlation between the two series amounts to about −0.39 and −0.14 for the PV-AR and PV-ARFIMA models, respectively. The negative correlation between the two expected series is in line with the findings of Van Binsbergen and Koijen (2010), but goes against the conjecture of positive correlation made by Lettau and Ludvigson (2005).

In the last two lines of Panel A in Table 4 we report the R^2 statistics calculated as:

$$R^2_r = 1 - \frac{\mathrm{var}(r_t - m^F_{t-1})}{\mathrm{var}(r_t)}, \tag{18a}$$

$$R^2_{\Delta d} = 1 - \frac{\mathrm{var}(\Delta d_t - g^F_{t-1})}{\mathrm{var}(\Delta d_t)}, \tag{18b}$$

where m^F_{t-1} and g^F_{t-1} are filtered series of expected returns and expected dividend growth rates, respectively. As in Van Binsbergen and Koijen (2010) the filtered series

15 Van Binsbergen and Koijen (2010) found the autoregressive parameter in the state space to range from 0.932 to 0.956, while Rytchkov (2012) found that the autoregressive coefficient was 0.78–0.85. Cochrane (2008) simulated the persistence of expected returns from 0.91–0.96 under an AR(1) specification.

Table 5
Bootstrap p-values for data simulated under the null and the alternative hypotheses.
Given the estimated parameters from Table 3, we use the stationary bootstrap procedure proposed in Politis and Romano (1994) to generate 5000 artificial samples of data of the same length as the original sample size (89 time series observations) for both models, PV-AR and PV-ARFIMA. For each series we calculate the statistics reported in Table 1. We calculate the bootstrap p-values by the frequency with which the statistic calculated in the simulated data ($\widehat{M}^b$) is higher or lower (whichever is lower) than the original statistic ($\widehat{M}$): $\min\left[\Pr(\widehat{M}^b > \widehat{M}),\ \Pr(\widehat{M}^b < \widehat{M})\right]$. The p-values for the data generated under the PV-AR model and under the PV-ARFIMA model are reported in round brackets and square brackets, respectively.

	r	Δd	pd
Panel A: $I(1)$ tests			
ADF	(0.4176)	(0.0120)**	(0.1870)
	[0.2832]	[0.0350]**	[0.2104]
Phillips–Perron	(0.4176)	(0.0950)*	(0.1870)
	[0.2832]	[0.1624]	[0.2104]
Fractional ADF			(0.2610)
			[0.1514]
Panel B: $I(0)$ test			
KPSS	(0.2702)	(0.3228)	(0.1774)
	[0.4528]	[0.3060]	[0.1534]
Lobato–Robinson	(0.4130)	(0.4546)	(0.0664)*
	[0.4024]	[0.4234]	[0.0650]*
Panel C: Fractional difference estimates			
GPH	(0.0746)*	(0.1020)	(0.1536)
	[0.2070]	[0.2734]	[0.2048]
Robinson	(0.0878)*	(0.1680)	(0.1190)
	[0.3210]	[0.4040]	[0.1418]
Shimotsu	(0.0552)*	(0.3278)	(0.0670)*
	[0.2418]	[0.4254]	[0.1000]

are updated one step ahead and are readily obtained from the Kalman filter. For the PV-AR model, the R_r^2 value amounts to 14%. When we add the fractional integration component, the R_r^2 value increases to 20%. The dividend growth process also seems significantly predictable in-sample. For the PV-AR model, the $R_{\Delta d}^2$ for dividend growth is 24% and for the PV-ARFIMA model, it reaches 25%. This is contrary to some results reported in the literature (e.g. Cochrane, 2008a) but is on the other hand in line with, for example, Van Binsbergen and Koijen (2010) and Koijen and Van Nieuwerburgh (2011).

Finally, in Panel B of Table 4 we report the results of the price–dividend ratio variance decomposition. Using the estimates of the model, we calculate the unconditional variance of the state vectors $\mathbf{C}_{g,t}$ and $\mathbf{C}_{m,t}$ using truncation at the $l = 2000$ lag. The portion of the variance attributed to the discount rates (expected returns) is thus $\mathbf{b}'\,\mathrm{var}(\mathbf{C}_{m,t})\mathbf{b}/\,\mathrm{var}(pd_t)$, and the part corresponding to expected dividend growth is $\mathbf{b}'\,\mathrm{var}(\mathbf{C}_{g,t})\mathbf{b}/\,\mathrm{var}(pd_t)$. The covariance part is calculated as the remainder of the total variance. Both the PV-AR and PV-ARFIMA models attribute a substantial part of the variation in the price–dividend ratio to fluctuations in expected returns, about 42% and 44%, respectively, and a negligible portion to fluctuations in the expected dividend growth. However, due to significant correlation of both components, a big part of the price–dividend ratio variance is attributed to the covariance term, which accounts for about 57% and 48% of the variation for the PV-AR and PV-ARFIMA models, respectively.

4.3. Bootstrap comparison

In this section we address the question of how likely it is that the observed features of the data arise under the short-memory PV-AR generating process and whether the proposed PV-ARFIMA model offers any improvement in this respect. To this end, based on the parameters reported in Table 3, we use the Politis and Romano (1994) stationary bootstrap to simulate 5000 samples of the data of the same length as the original sample (89 time series observations), under both models.[16] For each simulated sample, we calculate the same statistics that we reported in Table 1. Based on the distribution of each statistic, we examine how likely it is to observe the original estimate. If the simulated data are close to the true data-generating process, we should expect that the observed statistic (denoted generically by $\widehat{M}$) is close to the center of the distribution. On the other hand, if the assumed process is not closely aligned with the true process, the observed statistic would be far in one of the tails of the distribution. Specifically, we calculate the bootstrap p-values as:

$$\min\left[\Pr\left(\widehat{M}^b > \widehat{M}\right),\ \Pr\left(\widehat{M}^b < \widehat{M}\right)\right], \tag{19}$$

where $\widehat{M}^b$ is the generic statistic calculated on the simulated sample. In Table 5 we report the bootstrap p-values for the simulations under the PV-AR model (round brackets) and the PV-ARFIMA model (square brackets) for each statistic.

The results suggest that in terms of the $I(1)$ and $I(0)$ tests (Table 5, Panel A and Panel B, respectively), there is not sufficient evidence to discriminate between the two competing models, although the p-values are generally smaller for the short memory model, at least for such a short sample size. In particular, both models are rejected at the 5% level for the ADF test for dividend growth; the ACF test strongly rejects the unit-root hypothesis on the data generated under both models, but the value of the test is typically not smaller than the original statistic. A similar situation occurs with the (Lobato & Robinson, 1998) test, where both models are rejected at the 10% level. The only difference is for the (Phillips & Perron, 1988) test for dividend growth, where the PV-AR model is rejected at the 10% level, while the p-value for the PV-ARFIMA model amounts to 16.24%.

The discriminatory power of the bootstrap exercise, however, is stronger for the estimates of the fractional difference parameters (Panel C). The PV-AR model is rejected for all three estimators for the return series (see Figure A–6 in Appendix E), and for the (Shimotsu, 2010) estimator for the price–dividend ratio. In other cases, the test does not reject either model, but the p-values are always smaller for the PV-AR model.

The presented results allow us to conclude that although the PV-AR model is able to reproduce some features of the data, such as the in-sample values of the

16 More details on the bootstrap results are reported in Appendix E.

$I(1)$ and $I(0)$ tests, it is unable to reproduce other features of the data, such as the estimates of the fractional integration parameter for returns. In this respect, the PV-ARFIMA model fares much better: although it reproduces the unit root and stationarity statistics similarly to the PV-AR model, at the same time it delivers the estimates of δ for each series centered around the estimates found in the data. Thus, our results indicate that the PV-ARFIMA model is more compelling. This is in line with the bootstrap p-value rejecting the PV-AR model in favor of the PV-ARFIMA model in the present-value framework, as reported in Table 4 (see also Figure A–7 in Appendix E).[17]

However, the ability to fit a model to descriptive statistics is not the only way to compare models. Indeed, further evidence in favor of the PV-ARFIMA model is due to its superior forecasting performance and its closer relation with the macro-variables. We discuss these issues in detail in Section 5.

4.4. Model-implied persistence of the price–dividend ratio

In light of our theoretical discussion in Section 3.1, it is interesting to check the implications of different assumptions about the dynamics of expected returns and expected dividend growth for the price–dividend ratio within the present-value framework.

The kth autocovariance of the price–dividend ratio is given by:

$$\gamma(k) = \sum_{j=0}^{\infty}[\sigma_g^2 \mathbf{b}'\mathbf{F}^j\mathbf{h}_g\mathbf{b}'\mathbf{F}^{j+|k|}\mathbf{h}_g + \sigma_m^2\mathbf{b}'\mathbf{F}^j\mathbf{h}_m\mathbf{b}'\mathbf{F}^{j+|k|}\mathbf{h}_m - \rho_{mg}\sigma_g\sigma_m\left(\mathbf{b}'\mathbf{F}^j\mathbf{h}_g\mathbf{b}'\mathbf{F}^{j+|k|}\mathbf{h}_m + \mathbf{b}'\mathbf{F}^{j+|k|}\mathbf{h}_g\mathbf{b}'\mathbf{F}^j\mathbf{h}_m\right)], \tag{20}$$

where the terms are as defined in Section 3. Since the vectors and matrices are infinitely dimensional, we use the approximation by truncating their dimension at 2000.

In Fig. 4(a) we plot the autocorrelation function (calculated as $\gamma(k)/\gamma(0)$) of the price–dividend series implied by the PV-AR and PV-ARFIMA models calibrated with the parameter estimates from Table 3. In the same figure, we also plot the empirical autocorrelations for the price–dividend ratio. The difference in the decay implied by the two models is striking. The PV-AR autocorrelations decrease rapidly, and at the 30-th lag they are virtually zero, while the autocorrelations implied by the PV-ARFIMA model exhibit slow, hyperbolic decay. We also note that although the PV-ARFIMA model does not replicate some remaining seasonal patterns present in the data, it successfully captures high and slowly decaying autocorrelation in the price–dividend ratio at long lags.

Furthermore, we examine the spectral density defined as:

$$f(\lambda) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)e^{-ij\lambda}. \tag{21}$$

In Fig. 4(b) we plot the spectral densities of the price–dividend series for the PV-AR and PV-ARFIMA models calculated based on the first 1500 autocovariances. In the figure we also superimpose a periodogram for the price–dividend ratio. In the neighborhood of frequency zero, the spectral density of the PV-AR model becomes flat, while the PV-ARFIMA model exhibits a spike resembling a long-memory series but finite. We can also observe that although the PV-ARFIMA model does not fit the data points perfectly, it provides a much better approximation to real data than the PV-AR model.

Since our model of returns and dividend growth in (16a)–(16b) is a signal plus noise model, the order of integration of the realized series is formally $I(0)$. Thus, although realized returns and dividend growth are close to serially uncorrelated, the price–dividend ratio might appear to have long-memory features.

Table 6
Mincer and Zarnowitz (1969) regressions for returns and dividend growth for the present-value models.
In Panel A we regress returns on a constant and the filtered values of expected returns. In the first two lines we report the estimated coefficients with their standard errors, and in the following two lines the t-statistic for the null hypothesis of unbiased forecasts, that is $H_0 : \alpha = 0$ and $H_0 : \beta = 1$. In the next line we report the F-test of the joint null hypothesis $H_0 : \alpha = 0$ *and* $\beta = 1$ with the p-values. In Panel B we report the corresponding results for dividend growth. The data sample is 1934–2022.

	Model	
	PV-AR	PV-ARFIMA
Panel A: $r_{t+1} = \alpha + \beta \times m_t^F + u_t$		
α	−0.0029	0.0010
std.err.	0.0422	0.0315
β	1.0210	1.0188
std.err.	0.5591	0.3970
$t-val.\ (H_0 : \alpha = 0)$	−0.0680	0.0313
$t-val.\ (H_0 : \beta = 1)$	0.0375	0.0475
$F\ (H_0 : \alpha = 0, \beta = 1)$	0.0036	0.0080
p−value	0.9864	0.9921
Panel B: $\Delta d_{t+1} = \alpha + \beta \times g_t^F + u_t$		
α	−0.0032	−0.0040
std.err.	0.0104	0.0103
β	1.0024	1.0681
std.err.	0.1903	0.1951
$t-val.\ (H_0 : \alpha = 0)$	−0.3064	−0.3893
$t-val.\ (H_0 : \beta = 1)$	0.0129	0.3489
$F\ (H_0 : \alpha = 0, \beta = 1)$	0.0596	0.0920
p−value	0.9422	0.9122

5. Analysis of forecasting power

5.1. Forecast diagnostics

In this section we evaluate the predictions given by the present-value models. We estimate the (Mincer & Zarnowitz, 1969) regressions, where the filtered series of expected returns and expected dividend growth are used as predictors:

$$r_{t+1} = \alpha + \beta \times m_t^F + u_{t+1}, \tag{22a}$$

$$\Delta d_{t+1} = \alpha + \beta \times g_t^F + u_{t+1}. \tag{22b}$$

Unbiased predictors should yield $\beta = 1$ and $\alpha = 0$. In Table 6 the regression results for returns are reported in Panel A and for dividend growth in Panel B. The estimates for returns for both the PV-AR and the PV-ARFIMA models do not deviate from their hypothesized values

[17] The bootstrap p-value for the PV-ARFIMA model amounts to 44.04% (not reported in any Table).

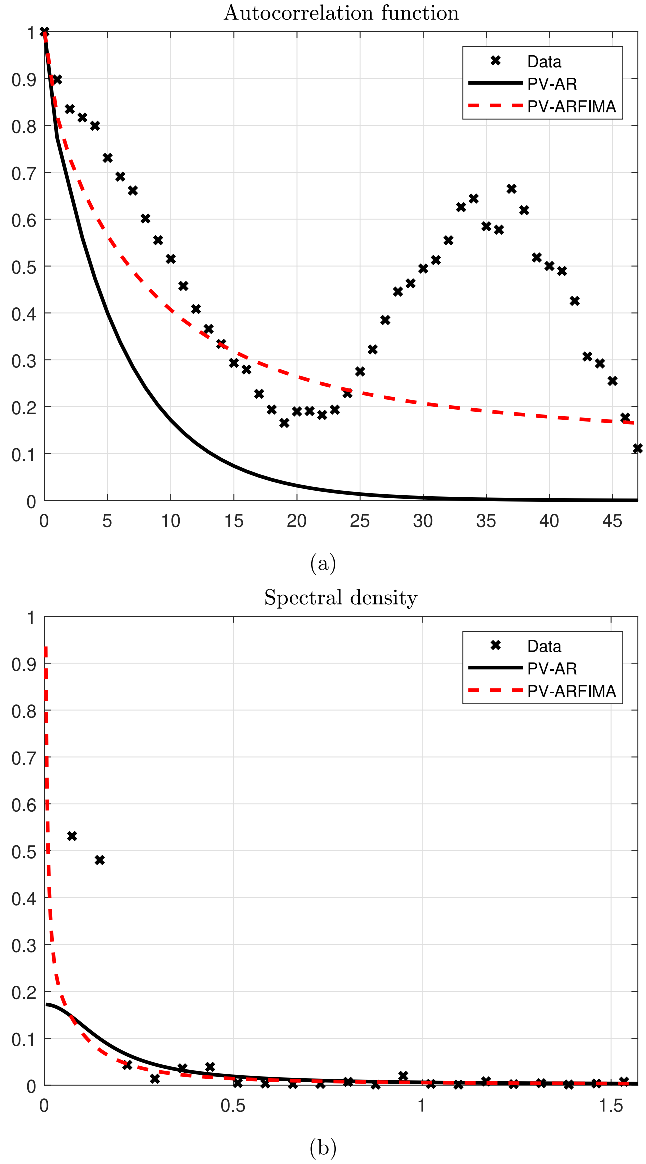

Fig. 4. Autocorrelation function and spectral density of the price–dividend ratio.
The subfigures show the values of autocorrelations (a) and spectral density (b) implied by the PV-AR (black line) model and the PV-ARFIMA (red dashed line) model together with the quantities estimated from the data ('×' markers). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

significantly, as evidenced by small t-test values (in absolute terms) and the F-test not rejecting the joint null hypothesis ($H_0 : \alpha = 0$ and $\beta = 1$) at any conventional significance level.

In Panel B of Table 6 we report the results for the dividend growth predictions. The β parameters for both the PV-AR and PV-ARFIMA models are estimated at 1.002 and 1.068, respectively, and are not statistically different from one. The intercept (α) is close to zero and not statistically significant for both models. The joint test fails to reject the null hypothesis that the forecasts are unbiased for both specifications. In summary, both models seem to yield unbiased forecasts of returns and dividend growth.

5.2. Present-value model versus forecasting regressions

The fact that stock prices are forward-looking as in (4) motivates the use of the price–dividend ratio in classical forecasting regressions:

$$r_{t+1} = \alpha_r + \beta_r pd_t + u_{r,t+1}, \quad (23a)$$

$$\Delta d_{t+1} = \alpha_d + \beta_d pd_t + u_{d,t+1}. \quad (23b)$$

Forecasting regressions ((23a)–(23b)) are not the optimal method of inference for several reasons. First, despite the fact that they are based on the present-value identity, the regressions are considered individually. As shown by Pástor and Stambaugh (2009), there are potential gains

Table 7
Root mean square forecasting errors for the present-value models and predictive regressions.
In Panel A we report the RMSFEs for in-sample forecasts calculated by using the estimates obtained from the whole sample and evaluated on the subsample. In Panel B we estimate the models for 1934–2002 and use these estimates to generate forecasts for the rest of the sample. In Panel C we start by estimating the models on the sample from 1934–2002 and then make the prediction for 2003. In the next step, we extend the estimation sample by one observation and make a prediction for the next year, and so on. In Panel D we estimate the models recursively, keeping the same number of observations in the estimation sample. For example, to make the prediction in the second step for 2004, we estimate the models on the sample from 1935–2003 and so on. The last column reports the RMSFEs from predicting with the sample mean calculated over the relevant estimation period. All forecasts are evaluated on the sub-sample from 2003–2022. The RMSFEs are reported in percentage points.

	PV model PV-ARFIMA	PV-AR	Forecasting regressions pd	Δpd	Sample mean
	Panel A: In-sample forecast				
r_t	**16.54**	17.65	17.79	18.05	17.92
Δd_t	**7.90**	8.32	9.55	8.94	9.33
	Panel B: Fixed point estimation forecast				
r_t	**16.88**	18.05	18.58	18.05	17.92
Δd_t	9.16	**8.84**	8.93	9.33	9.58
	Panel C: Recursive estimation forecast				
r_t	**16.85**	17.99	18.43	18.17	18.03
Δd_t	**8.65**	8.78	8.91	9.25	9.50
	Panel D: Rolling window estimation forecast				
r_t	**16.94**	18.07	18.36	18.02	18.01
Δd_t	9.08	**8.85**	9.03	9.36	9.50

Table 8
Diebold and Mariano (1995) test of equal forecasting ability.
The tests are performed out-of-sample for the same evaluation methods as reported in Table 7. The two-sided test compares uses a quadratic loss function against the benchmark forecasts from the PV-ARFIMA model, so that a positive statistic indicates the superior performance of the long-memory model. The test corrects for a small sample and the autocorrelation of forecast errors by using fixed-b asymptotics with the long-run variance estimated with the Bartlett kernel proposed by Coroneo and Iacone (2020). * and ** denote significance at the 10% and 5% levels, respectively.

	PV-AR	pd	Δpd	Sample mean
	Panel A: Fixed point estimation forecast			
r_t	2.21*	1.21	3.22**	3.60**
Δd_t	−0.91	−0.17	0.31	0.35
	Panel B: Recursive estimation forecast			
r_t	1.80	1.66	3.44**	3.70**
Δd_t	0.36	0.20	1.41	0.27
	Panel C: Rolling window estimation forecast			
r_t	2.17*	1.38	3.62**	3.69**
Δd_t	−0.83	−0.08	0.84	0.27

in considering jointly the system of predictive regressions. The second drawback of the predictive regressions stems from the fact that they ignore information in the dynamics of the underlying time series. These issues are addressed in our present-value model with latent variables.

We examine the predictive ability of the estimated present-value models and draw comparisons to the forecasting regressions (23a)–(23b) using the price–dividend ratio not only in levels but also in its first difference. As an additional benchmark, we include the forecasts made with a historical mean, which is known to be hard to beat (see e.g. Goyal & Welch, 2003, and Welch & Goyal, 2008). We examine both the in-sample and out-of-sample forecasting ability of the models on the last 20 years of data, that is 2003–2022. In Table 7 we present the root mean square forecasting error (RMSFE) for four forecasting exercises. In bold font we highlight the lowest RMSFE across all models.

Panel A of Table 7 reports the in-sample forecast results obtained by using the parameters estimated on the whole sample. The model with the fractional integration component exhibits consistent forecasting power for both returns and dividend growth, with the lowest RMSFE (16.5% and 7.9%, respectively). The performance of the PV-AR model is worse for both series, but particularly worse for returns (17.65%), yielding an RMSFE over 1% larger than that produced by the PV-ARFIMA model. The forecasting regressions with the level of the price–dividend ratio and its first difference exhibit worse in-sample forecasting performance than the present-value models, and struggle to beat the sample mean, which echoes the findings in Welch and Goyal (2008).

In Panels B, C, and D of Table 7 we report forecasting results produced by three out-of-sample schemes. In Panel B we report the results obtained by estimating the models only once on the sample period from 1934–2002 and then using these estimates to compute the subsequent point forecasts. The results in Panel C are obtained by expanding the data used in estimation recursively by one observation each time and making the prediction for the next year. In Panel D we use the rolling-window method. That is, for each prediction we use the parameters estimated on the preceding 69 observations.

Similar to the in-sample results from Panel A, in terms of return predictability, the PV-ARFIMA model performs much better than other models regardless of the forecasting strategy. For dividend growth, both present-value models have similar performance and the ranking here depends on the forecasting scheme. Predictive regressions, with both the level and first difference of the price–dividend ratio, perform worse than the present-value models in terms of return forecasts. Regarding the dividend growth forecast, there is some evidence that in predictive regressions the level of the price–dividend ratio can perform about as well as the present-value models. Finally, we note that both present-value models, with one exception for the PV-AR model, beat the forecasts made with the sample mean.

To assess the statistical significance of the difference in forecasting ability of different models, we perform the Diebold and Mariano (1995) test for the out-of-sample forecasts. We use the standard quadratic loss function with fixed-b asymptotics, with the long-run variance estimated with the Bartlett kernel proposed by Coroneo and Iacone (2020) that corrects for a small sample and the autocorrelation of forecast errors.[18, 19] In Table 8 we

[18] We are grateful to Laura Coroneo for providing us with the Matlab code.

[19] We note that, despite the common practice, the Diebold and Mariano (1995) test for the recursive scheme in Panel C should be

report the test values comparing forecasts from the PV-ARFIMA model to all other models. The positive entry denotes a superior performance of the PV-ARFIMA model, and vice versa. The stars denote the level of significance based on critical values for a two-sided test. Despite the small sample, we find that return forecasts made with the PV-ARFIMA model are generally statistically better than those produced by other methods, although the test does not reject the null hypothesis for forecasts with the level of the price–dividend ratio. On the other hand, we do not find any statistically significant difference in predictive performance for the dividend growth, even against the sample mean. This reverberates the argument for the lack of predictability for dividend growth in Cochrane (2008a).

5.3. Expected returns and dividend growth over the business cycle

Aggregate stock prices have long been viewed as leading indicators of the business cycle (see Mitchell & Burns, 1938, Zarnowitz, 1992, and Stock & Watson, 1999). This is consistent with the PV model that considers stock prices to reflect expectations of future outcomes. The risk premium does appear to be countercyclical, rising in recessions and falling in expansions. Di Tella and Hall (2022) show that a higher risk premium can lead to recessions, decreasing investment and employment. Lettau and Ludvigson (2005) showed that expected dividend growth also varies with the business cycle. This can happen because managers smooth dividends imperfectly over the business cycle. Consistent with that hypothesis, Gertler and Hubbard (1993) found that firm dividend payouts appear to be procyclical (lower in periods of low growth than in economic expansions). In this section we contribute to the literature by examining the relation between macroeconomic variables and the series of expected returns and expected dividend growth filtered from our model.

In Fig. 5(a) we plot the time series of realized and expected returns as implied by the models. The grey areas denote the NBER recession periods. Since our data are annual, we plot only recessions that lasted at least nine months. We can see that the higher variability of expected returns implied by the PV-ARFIMA model in comparison to PV-AR is prominent. The expected returns series seem to have a strong countercyclical pattern: they fall in the period prior to and at the start of economic downturns, and then increase as the period of expansion approaches. Consistent with the results reported in Table 4, the variability of expected dividend growth plotted in Fig. 5(b) is significantly higher for the PV-ARFIMA model and exhibits a clear procyclical pattern with low expected dividend growth at the outset of an economic slowdown.

In order to examine properties related to business cyclicality we regress a set of macro-variables on the filtered series of expected returns and expected dividend growth. We proceed in the spirit of Liew and Vassalou (2000), who showed that risk premium correlates with future GDP growth. As such, in a predictive fashion, we regress the log growth of real personal consumption expenditures ($\Delta Cons$) and the log growth of industrial production of consumption goods (ΔIP) on expected returns and expected dividend growth.[20] We chose these variables because they are meaningful indicators of the business cycle. Since the time series of industrial production growth is available only from 1940, the regressions with this series are therefore run on a shorter sample.[21] In Table 9 we report the slope coefficients with the t-statistics (reported in small font) calculated using the heteroscedasticity and autocorrelation consistent standard errors and the adjusted R^2 coefficients. In Panel A we report the regressions with $\Delta Cons$ as the dependent variable. The regressions with ΔIP as the dependent variable are reported in Panel B.

We find that the coefficient associated with expected returns for both present-value models is negative and statistically significant (consistent with theory and evidence that the risk premium is countercyclical). We also find that the coefficient associated with expected dividend growth for both present-value models is positive and statistically significant (consistent with the hypothesis of imperfect dividend-smoothing and evidence of the procyclicality of dividend growth). The expected returns and dividend growth series obtained from the PV-ARFIMA model seem to retain predictive ability even after adding to the regressions the filtered series from the PV-AR model or the realized series.

The results allow us to make a few observations. First, in a simple regression setting, although both expected returns and expected dividend growth have a cyclical nature, the latter is a stronger predictor of the business cycle: the R^2 for regressions with the expected dividend growth is about two to three times higher than with expected returns. Second, although the t-statistics of the predictors do not differ much between models, we can observe that the model with fractional integration predicts the macro-variables better than the PV-AR model in basically every regression, as can be observed from higher adjusted R^2 values.

Taken together, these results suggest that implied expected returns and dividend growth series can have a potential application as leading economic indicators—particularly more so if the present-value model includes a fractional integration component.

6. Conclusion

The present-value identity shows that the price–dividend ratio is the sum of discounted expected returns and expected dividend growth. Since in the data the price–dividend ratio series appears to be non-stationary—while the return and dividend growth series are sta-

interpreted with caution, since it does not account for the uncertainty of estimated parameters.

[20] In the preliminary analysis we also used GDP growth but did not find a statistically significant relation with expected returns and expected dividend growth in our sample.

[21] The macro-data were obtained from the Federal Reserve Economic Data (FRED) freely available on the website of the Federal Reserve Bank of St. Louis.

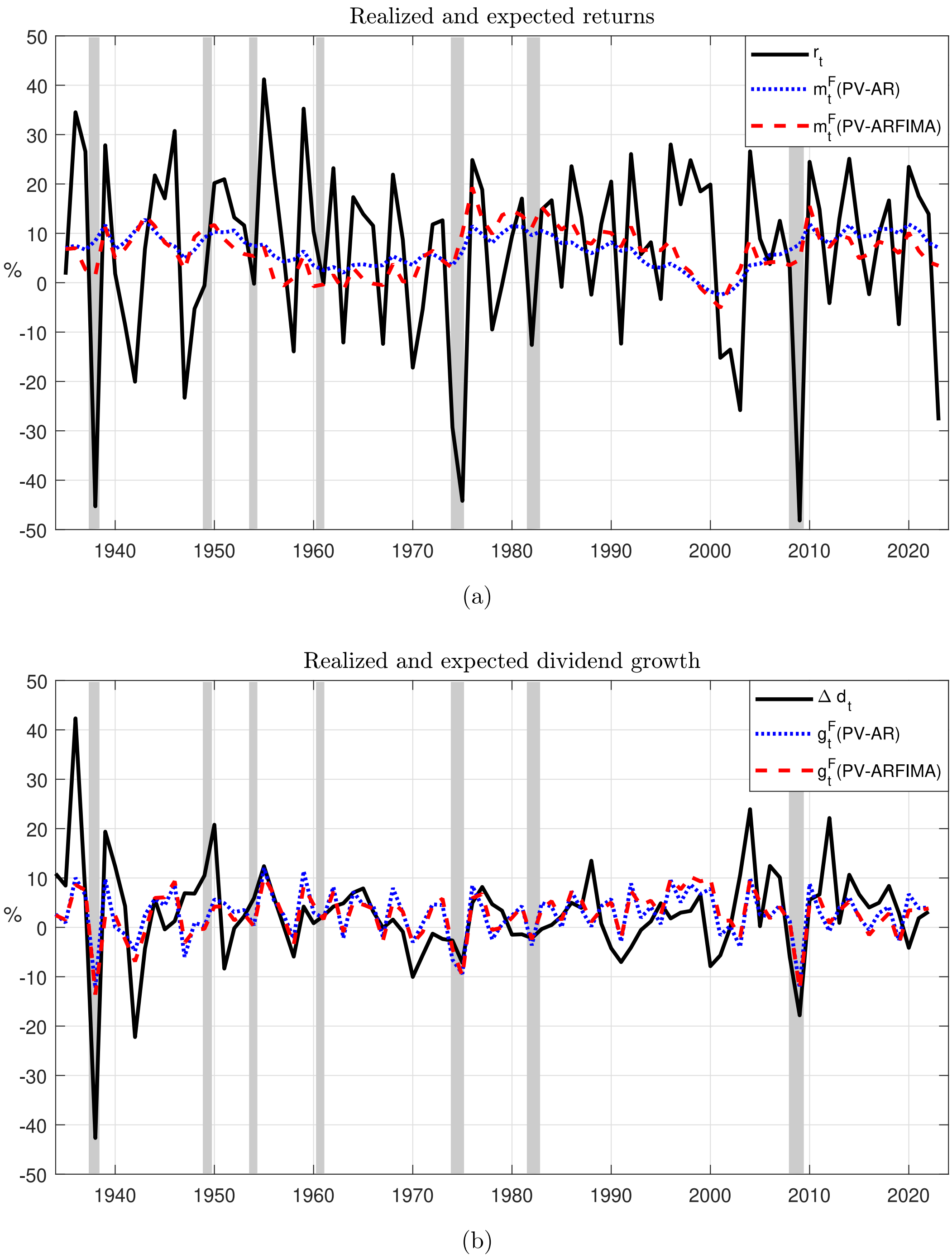

Fig. 5. Realized and expected returns and dividend growth.
The subfigures show expected returns (a) and expected dividend growth (b) as implied by the PV-AR (blue dotted lines) and *PV − ARFIMA* (red dashed lines) models, together with their realized values (black line). The grey areas denote the NBER identified recession periods (only those longer than nine months). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tionary, even antipersistent—the difference in the order of integration of these series has been a puzzle. We showed that through the aggregation of antipersistent expectations, the price–dividend ratio can appear non-stationary in finite samples. Despite being asymptotically stationary, it will exhibit slowly decaying (but summable)

Table 9
Results from the regression of macro-variables on returns, dividend growth, model-filtered expected returns, and expected dividend growth. The macro-variables are real consumption growth ($\Delta Cons$) and growth of industrial production of consumer goods (ΔIP). The intercept is omitted from the table. The t-statistics calculated from heteroscedasticity and autocorrelation consistent standard errors are reported in small font. The implied series are obtained from the whole sample (1934–2022) and the regressions are run on available samples of macro-variables, that is 1934–2022 for consumption and 1940–2022 for industrial production growth.

	Dependent variable: $\Delta Cons_{t+1}$							
	i	ii	iii	iv	v	vi	vii	viii
m_t^{ARFIMA}	−0.1932***		−0.1542**	−0.0908*				
	−3.9997		−2.0959	−1.8231				
m_t^{AR}		−0.2368***	−0.0684					
		−3.4712	−0.6789					
g_t^{ARFIMA}					0.2855***		0.2289**	0.2910***
					7.9000		2.0322	7.9309
g_t^{AR}						0.2644***	0.0588	
						7.3085	0.5432	
r_t				0.0555***				
				4.4205				
Δd_t								−0.0226
								−1.5525
$\bar{R}^2$	0.1406	0.1078	0.1314	0.2674	0.3065	0.2808	0.3004	0.3073
	Dependent variable: ΔIP_{t+1}							
	i	ii	iii	iv	v	vi	vii	viii
m_t^{ARFIMA}	−0.3029***		−0.3068**	−0.1473				
	−3.3893		−2.5440	−1.6377				
m_t^{AR}		−0.3333**	0.0065					
		−2.1635	0.0308					
g_t^{ARFIMA}					0.4694***		0.3412	0.4738***
					4.9608		1.3753	4.7263
g_t^{AR}						0.4400***	0.1325	
						5.2366	0.5900	
r_t				0.0943***				
				3.3116				
Δd_t								−0.0232
								−0.3803
$\bar{R}^2$	0.1026	0.0606	0.0914	0.1997	0.2135	0.2006	0.2062	0.2053

autocorrelations and a (finite) spike of the spectral density at frequency zero.

We included the fractional integration feature in the present-value model using an ARFIMA specification. The fractional integration parameter for expected returns and expected dividend growth is negative and statistically significant, and close to the non-parametric estimates. The benchmark model based on AR(1) processes is decisively rejected by the likelihood-ratio test. Although in the univariate analysis, the dividend growth provides the stronger motivation for a model with fractional integration, the economic significance of the fractional model is more strongly manifested in the expected return series, as evidenced by the enhanced predictability of the return series. We also showed that the PV-ARFIMA model filtered series of expected returns and dividend growth are aligned with the business cycle, as they help to predict macroeconomic variables such as consumption and industrial production growth. This is important, since, as emphasized by Cochrane (2011), a correct understanding of the risk premium is vital for macro-prudential regulation and monetary policy.

We demonstrated that the fractionally integrated model closely replicates the apparent long memory behavior in the price–dividend ratio, as predicted by our theory. To have a full understanding of the mechanism that generates persistent dynamics of the price–dividend ratio, however, we should understand the mechanism that leads to the formation of antipersistent expected returns and expected dividend growth. Such a mechanism can be due to investors' overreaction to news and biased expectations, as documented by Lakonishok et al. (1994) and La Porta (1996). However, the design of a formal micro-founded model of this type we leave for future research.

Declaration of competing interest

We declare that we do not have any undisclosed personal or financial interests that may be perceived as influencing our work.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2024.03.005.

References

Abadir, K. M., Heijmans, R. D. H., & Magnus, J. R. (2018). *Statistics*. Cambridge University Press.
Abadir, K. M., & Magnus, J. R. (2005). *Matrix algebra*. Cambridge University Press.
Ang, A., & Bekaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies*, *20*, 651–707.
Balke, N. S., & Wohar, M. E. (2009). Market fundamentals versus rational bubbles in stock prices: A Bayesian perspective. *Journal of Applied Econometrics*, *24*, 35–75.
Campbell, S. D., & Diebold, F. X. (2009). Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business & Economic Statistics*, *27*, 266–278.
Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, *1*, 195–228.

Chan, N. H., & Palma, W. (1998). State space modeling of long-memory processes. *The Annals of Statistics*, *26*, 719–740.
Chevillon, G., & Mavroeidis, S. (2017). Learning can generate long memory. *Journal of Econometrics*, *198*, 1–9.
Cochrane, J. H. (1999). New facts in finance. *Economic Perspectives*, *23*, 36–58.
Cochrane, J. H. (2008). State-space vs. VAR models for stock returns. Unpublished working paper.
Cochrane, J. H. (2008a). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, *21*, 1533–1575.
Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, *66*, 1047–1108.
Coroneo, L., & Iacone, F. (2020). Comparing predictive accuracy in small samples using fixed-smoothing asymptotics. *Journal of Applied Econometrics*, *35*, 391–409.
Cunado, J., Gil-Alana, L., & de Gracia, F. P. (2005). A test for rational bubbles in the NASDAQ stock index: A fractionally integrated approach. *Journal of Banking & Finance*, *29*, 2633–2654.
Cutler, D. M., Poterba, J. M., & Summers, L. H. (1988). What moves stock prices?. NBER Working Papers 2538, National Bureau of Economic Research, Inc.
Dalla, V., Giraitis, L., & Hidalgo, J. (2006). Consistent estimation of the memory parameter for nonlinear time series. *Journal of Time Series Analysis*, *27*, 211–251.
Di Tella, S., & Hall, R. (2022). Risk premium shocks can create inefficient recessions. *Review of Economic Studies*, *89*, 1335–1369.
Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427–431.
Diebold, F. X., & Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, *105*, 131–159.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*, 253–263.
Diebold, F. X., & Rudebusch, G. D. (1991). On the power of Dickey–Fuller tests against fractional alternatives. *Economics Letters*, *35*, 155–160.
Dolado, J. J., Gonzalo, J., & Mayoral, L. (2002). A fractional Dickey–Fuller test for unit roots. *Econometrica*, *70*, 1963–2006.
Engsted, T., Pedersen, T. Q., & Tanggaard, C. (2012). The log-linear return approximation, bubbles, and predictability. *Journal of Financial and Quantitative Analysis*, *47*, 643–665.
Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*, 3–25.
Ferson, W. E., Sarkissian, S., & Simin, T. T. (2003). Spurious regressions in financial economics? *The Journal of Finance*, *58*, 1393–1414.
Gertler, M., & Hubbard, R. G. (1993). Corporate financial policy, taxation, and macroeconomic risk. *Rand Journal of Economics*, *24*, 286–303.
Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, *4*, 221–238.
Goliński, A., & Zaffaroni, P. (2016). Long memory affine term structure models. *Journal of Econometrics*, *191*, 33–56.
Goyal, A., & Welch, I. (2003). Predicting the equity premium with dividend ratios. *Management Science*, *49*, 639–654.
Granger, C. W. J. (1966). The typical spectral shape of an economic variable. *Econometrica*, *34*, 150–161.
Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, *14*, 227–238.
Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, *1*, 15–29.
Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.
Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, *68*, 165–176.
Koijen, R. S., & Van Nieuwerburgh, S. (2011). Predictability of returns and cash flows. *Annual Review of Financial Economics*, *3*, 467–491.
Koustas, Z., & Serletis, A. (2005). Rational bubbles or persistent deviations from market fundamentals? *Journal of Banking & Finance*, *29*, 2523–2539.
Kramer, W. (1998). Fractional integration and the augmented Dickey–Fuller test. *Economics Letters*, *61*, 269–272.
Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*, 159–178.
La Porta, R. (1996). Expectations and the cross-section of stock returns. *The Journal of Finance*, *51*, 1715–1742.
Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance*, *49*, 1541–1578.
Lettau, M., & Ludvigson, S. C. (2005). Expected returns and expected dividend growth. *Journal of Financial Economics*, *76*, 583–626.
Liew, J., & Vassalou, M. (2000). Can book-to-market, size and momentum be risk factors that predict economic growth?. *Journal of Financial Economics*, *57*, 221–245.
Lobato, I. N., & Robinson, P. M. (1998). A nonparametric test for I(0). *Review of Economic Studies*, *65*, 475–495.
Mankiw, G. N., & Shapiro, M. D. (1986). Do we reject too often?: Small sample properties of tests of rational expectations models. *Economics Letters*, *20*, 139–145.
Maynard, A., & Phillips, P. C. B. (2001). Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics*, *16*, 671–708.
Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*. NBER.
Mitchell, W. C., & Burns, A. F. (1938). Statistical indicators of cyclical revivals. In *Statistical Indicators of Cyclical Revivals* (pp. 1–12). National Bureau of Economic Research, Inc, NBER Chapters.
Nielsen, M. Ø. (2010). Nonparametric cointegration analysis of fractional systems with unknown integration orders. *Journal of Econometrics*, *155*, 170–187.
Pástor, L., & Stambaugh, R. F. (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance*, *64*, 1583–1628.
Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*, 335–346.
Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review*, *52*, 201–226.
Piatti, I., & Trojani, F. (2017). Predictable risks and predictive regression in present-value models. Unpublished working paper.
Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*, 1303–1313.
Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, *22*, 27–59.
Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, *23*, 1630–1661.
Rytchkov, O. (2012). Filtering out expected dividends and expected returns. *Quarterly Journal of Finance*, *2*, 1–56.
Shimotsu, K. (2010). Exact local Whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory*, *26*, 501–540.
Sowell, F. (1990). The fractional unit root distribution. *Econometrica*, *58*, 495–505.
Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, *54*, 375–421.
Stock, J. H. (1986). Unit roots, structural breaks and trends. In R. F. Engle, & D. McFadden (Eds.), *Handbook of econometrics*: *vol. 4*, (pp. 2739–2841). Elsevier, chap. 46.
Stock, J. H., & Watson, M. W. (1999). Business cycle fluctuations in US macroeconomic time series. In J. B. Taylor, & M. Woodford (Eds.), *Handbook of macroeconomics*: *vol. 1*, (pp. 3–64). Elsevier, chap. 1.
Sun, Y., & Phillips, P. C. B. (2003). Nonlinear log-periodogram regression for perturbed fractional processes. *Journal of Econometrics*, *115*, 355–389.
Van Binsbergen, J. H., & Koijen, R. S. J. (2010). Predictive regressions: A present-value approach. *The Journal of Finance*, *65*, 1439–1471.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, *21*, 1455–1508.

Wu, Y. (1997). Rational bubbles in the stock market: Accounting for the U.S. stock-price volatility. *Economic Inquiry*, *35*, 309–319.

Zarnowitz, V. (1992). Composite indexes of leading, coincident, and lagging indicators. In V. Zarnowitz (Ed.), *Business Cycles: Theory, History, Indicators, and Forecasting* (pp. 316–356). University of Chicago Press.