Fractional Integration of the Price-Dividend Ratio in a Present-Value Model.

Golinski, Adam and Madeira, Joao and Rambaccussing, Dooruj

University of York, University of York, University of Dundee

13 September 2014

Fractional Integration of the Price-Dividend Ratio in a Present-Value Model of Stock Prices

Adam Goliński, João Madeira and Dooruj Rambaccussing[*]

12 September 2014

Abstract

We re-examine the dynamics of returns and dividend growth within the present-value framework of stock prices. We find that the finite sample order of integration of returns is approximately equal to the order of integration of the first-differenced price-dividend ratio. As such, the traditional return forecasting regressions based on the price-dividend ratio are invalid. Moreover, the nonstationary long memory behaviour of the price-dividend ratio induces antipersistence in returns. This suggests that expected returns should be modelled as an $ARFIMA$ process and we show this improves the forecast ability of the present-value model in-sample and out-of-sample.

JEL Classification: G12, C32, C58

Keywords:price-dividend ratio, persistence, fractional integration, return predictability, present-value model.

1 Introduction

The price-dividend ratio has been shown to have strong forecasting power of future returns at long horizons (see Fama and French (1988a) and Cochrane (1999)). The prediction properties of the price-dividend ratio have a strong theoretical foundation grounded in the present-value (PV) identity, popularized in the log linear form by Campbell and Shiller (1988). As shown in Cochrane (2005), price-dividend ratios "can only move at all if they forecast future returns, if they forecast future dividend growth, or if there is a bubble – if the price-dividend ratio is nonstationary and is expected to grow explosively". Cochrane (2008a) argues that the lack of predictability of dividend growth reinforces the evidence for forecastability of stock returns. However, many research studies have pointed out that return predictability has been overstated due to the highly persistent price-dividend ratio (e.g. Stambaugh (1986), Stambaugh (1999), Mankiw and Shapiro (1986), Hodrick (1992) and Goyal and Welch (2003)).

Many (see Cochrane (2005)) however, find it implausible that the price-dividend ratio has unit root properties (i.e. are integrated of order one), since this would mean that the series is unbounded (implying that it can achieve negative values and go to infinity). As such, it is tempting to think that the price-dividend ratio process is, in fact, a persistent but, nonetheless, a mean reverting process (long memory in price-dividend ratio was earlier suggested by Andersson and Nydahl (1998) in the context of a test of rational bubbles). Our paper re-examines the hypothesis of time variation in expected returns from the perspective of long range persistence in the price-dividend ratio (that is, fractionally integrated of order higher than zero but smaller than one). Indeed, we find the price-dividend ratio to exhibit long memory and estimate it to have an order of integration of about 0.8.[1] We show that the

relation between returns, dividend growth and price-dividend ratio implies that the order of integration of returns is (in finite sample) approximately equal to the order of integration of the first differenced price-dividend ratio. We find the time series of returns to be integrated of order -0.2, confirming this conjecture.[2]

The fact that the variables in simple linear regression have different orders of integration invalidates statistical inference (see Maynard and Phillips (2001)). The negative fractional order of integration in returns and dividend growth in the data must be taken into account when estimating the PV model of stock prices. This motivates specifying the expected return and expected dividend growth series in a PV model as autoregressive fractionally integrated moving average ($ARFIMA$) processes. We derive the unobserved series for expected returns and expected dividend growth through a structural state-space approach. The state-space (or latent-variables) representation has shown to be a 'useful structure for understanding and interpreting forecasting relations' as stated by Cochrane (2008b). Recent important examples of this include Van Binsbergen and Koijen (2010) and Rytchkov (2012), who found the state-space methodology to be able to increase the return forecast R^2 over price-dividend ratio regressions. As in these papers, we specify expected returns and expected dividend growth as latent variables defined within a PV model of the aggregate stock market to which we subsequently apply the Kalman filter and obtain parameter estimates through maximum likelihood.

The fractional integration parameter in expected returns is found to be statistically significant and negative. Using model selection criteria we find the $ARFIMA(1,\delta,0)$ model for expected returns and $ARMA(1,1)$ for expected dividend growth to be the preferred specification. Our results suggest that allowing for an autoregressive fractionally integrated process in expected returns leads not only to a better in-sample fit to the data but also to a better out-of-sample forecast. Assuming expected re-

turns follow an autoregressive process of order one results in R^2 values ranging from 13 to 14 percent for returns and about 32 percent for dividend growth rates in the 1926-2011 sample. The use of an autoregressive fractionally integrated process in expected returns results in R^2 values for returns of about 20 percent and a range of 36 to 38 percent for dividend growth rates in the 1926-2011 sample. Several prediction exercises on the last 40 years of data $(1971 - 2011)$ confirm the relevance of using a model which accounts for fractional integration in improving the forecast ability of the PV model both in-sample and out-of-sample. Assuming a first order autoregressive process of expected returns results in R^2 values of about 3 percent for returns and a range of 11 to 15 percent for dividend growth rates in-sample and negative R^2 for returns and about 7 to 12 percent for dividend growth rates out-of-sample. On the other hand, the use of an $ARFIMA$ process in expected returns results in R^2 values for returns of about 9 percent and about 13 to 17 percent for dividend growth rates in-sample, and out-of-sample R^2 values of about 1 to 4 percent for returns and 9 to 13 percent for dividend growth.

Using Mincer-Zarnowitz style regressions (Mincer and Zarnowitz (1969)) we check that our model produces latent counterparts that jointly match the time series properties of returns and dividends very well. The expected returns and expected dividend growth series forecast observed series better than the $AR(1)$ model by Van Binsbergen and Koijen (2010) both in-sample and out-of-sample. Our filtered series of expected returns and expected dividend growth are clearly countercyclical, which is in line with many other studies (e.g. Chen, Roll and Ross (1986), Fama and French (1989), Fama (1990), Barro (1990)).[3]

This paper is related to the research stream focused on testing for long memory in stock returns and volatility. While long memory has been well documented in the volatility literature, the evidence of long memory in returns is rather weak.[4] Our

results support to the view that fractional integration processes are relevant in asset pricing.

The remainder of the paper is organized as follows. Section 2 explores the potential imbalance in the return forecasting regression due to the high persistence in the price-dividend ratio. In section 3 we set out the PV model with $ARFIMA$ dynamics. In section 4.1 we describe the data. Section 4.2 describes the estimation methodology and results while in section 4.3 we present a series of diagnostics of the model. Section 4.4 presents the out-of-sample performance of particular models. We examine the business cycle fluctuations of the model implied expected returns and dividend growth in Section 4.5. Section 5 concludes.

2   The implications of persistence in the price-dividend ratio

We start by defining the aggregate stock market's total log return (r_{t+1}) and log dividend growth rate (Δd_{t+1}) as:

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

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

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 then 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 and Shiller (1988)).

We now proceed to the analysis of the data (see section 4.1 for details). Figure 1 shows the time series for the log of the price-dividend ratio. The price-dividend ratio is lower preceding economic booms and high values preceding recessions, suggesting it could be relevant for forecasting returns. One can also observe that it is a very persistent variable. In Table 1 we report the results of the following regression:

$$pd_{t+1} = \alpha + \beta \times pd_t + u_{t+1}. \tag{4}$$

The estimated coefficient on the lagged price-dividend value is very close to one (0.9417). Table 1 also reports the results of the Augmented Dickey-Fuller unit root test (Dickey and Fuller (1979)). The $t-$statistic (-1.52) is much smaller than the 5% critical value (-2.9), which means that we cannot reject the null of a unit root in the series.[5] Thus, in principle the nonstationarity of the price-dividend ratio invalidates a return forecasting regression (see Granger and Newbold (1974)) of the type considered by Fama and French (1988a):

$$r_{t+1} = \alpha + \beta \times pd_t + u_{t+1}. \tag{5}$$

However, as discussed in the previous section it is likely that the price-dividend ratio is integrated of order less than one but we fail to reject the null due to the small power of the test. In Table 2 we explore further the possibility that the price-dividend ratio is a variable with long memory or long range persistence (i.e. fractionally integrated of order higher than zero but smaller than one). Table 2 reports the estimates of the order of integration of the price-dividend ratio, returns and dividend

growth series obtained using three different semi-parametric estimators, proposed by Geweke and Porter-Hudak (1983), Robinson (1995) and Shimotsu (2010). The Shimotsu estimator, as opposed to the other two, has been designed to deal with a nonstationary time series. As can be seen, the estimates of the fractional parameter (δ) revolve around 0.8, meaning that the series is a nonstationary but mean-reverting process. On the other hand, the time series of returns and dividend growth seem to exhibit antipersistence (i.e. are integrated of order smaller than zero). Inference and forecasting based on estimates of (5) is therefore invalid (as shown by Maynard and Phillips (2001)) due to the different order of integration of the variables included in it.

At this point, we should consider as well the balance in the order of integration in the log-linearized return equation, (3). A closer inspection of (3) reveals that, due to the discount parameter (ρ) being very close to unity, the return series is (almost) over-differenced. Indeed, as reported in Table 2, the estimate of the fractional integration parameter for the return series is about -0.2, which is exactly the expected order of integration of the price-dividend ratio after taking first difference. Naturally, this also applies to the dividend growth process (the point estimates of δ for dividend growth are more negative than those of returns, but well within a confidence interval of two standard deviations of -0.2).

The PV model (see Campbell and Shiller (1988) and Van Binsbergen and Koijen (2010)) is derived from the return accounting identity shown in (3). The preceding analysis indicates that taking into account the fractional integration of returns and dividend growth should allow for improved statistical inference.

3 The Present-Value Model

In this section we present a log-linearized PV model of stock prices similar to that of Van Binsbergen and Koijen (2010). The crucial difference is the way we model the persistence in expected returns. Van Binsbergen and Koijen (2010) specified expected returns as an $ARFIMA$ process while we will consider the more general $ARFIMA$ process. The discussion in the previous section indicates that the $ARFIMA$ process should improve the ability of the PV model in accounting for the data. In our empirical application we consider two different specifications for expected returns $(m_t \equiv E_t[r_{t+1}])$ and expected dividend growth $(g_t \equiv E_t[\Delta d_{t+1}])$; we model expected returns as an $ARFIMA$ $(1, \delta_m, 0)$ process and expected dividend growth as an $ARMA(1,1)$ process:[6]

$$(1 - \phi_m L)(1 - L)^{\delta_m}(m_t - \mu_m) = \varepsilon_{m,t}, \tag{6}$$

$$(1 - \phi_g L)(g_t - \mu_g) = (1 + \theta_g L)\varepsilon_{g,t}, \tag{7}$$

where L is the the lag operator and $|\delta_m| < 1/2$ is a fractional integration parameter. For the sake of notational simplicity, since the fractional integration term (δ_m) and moving average term (θ_g) are unique, from now on we will call them simply δ and θ, respectively. When $\delta = \theta = 0$ then our model becomes identical to that in Van Binsbergen and Koijen (2010).

It is often convenient (see Cochrane (2008b)) to re-write the model as an infinite moving average:[7]

$$m_t = \mu_m + \varepsilon_{m,t} + \varphi_{m,1}\varepsilon_{m,t-1} + \varphi_{m,2}\varepsilon_{m,t-2} + \dots, \tag{8}$$

$$g_t = \mu_g + \varepsilon_{g,t} + \varphi_{g,1}\varepsilon_{g,t-1} + \varphi_{g,2}\varepsilon_{g,t-2} + \dots. \tag{9}$$

An $AR(1)$ specification of the expected returns time series process, imposes a tight restriction on the moving average coefficients, such that $\varphi_{m,j} = \phi_m^j$. The addition of the fractionally integrated component and the extension to the $ARFIMA(p, \delta, q)$ series allows for additional flexibility in modelling the series dynamics. If the fractional integration parameter (δ) is larger than zero the series is characterized by slow decay of autocorrelations, at an hyperbolic rate. On the other hand, if $\delta < 0$, we say that the series is antipersistent. For $\delta = 0$ the series is a simple short memory process and the model reduces to an $ARMA(p, q)$. Moreover, the series is stationary if $\delta < 1/2$ and invertible if $\delta > -1/2$.[8]

For estimation purposes we will use a state space representation of m_t and g_t. Thus, we specify the state space equations:

$$m_t = \mu_m + \mathbf{w}' \mathbf{C}_{m,t}, \tag{10}$$

$$g_t = \mu_g + \mathbf{w}' \mathbf{C}_{g,t}, \tag{11}$$

where $\mathbf{w} = [1\ 0\ 0\ \cdots\ ]'$, and $\mathbf{C}_{r,t}$ and $\mathbf{C}_{d,t}$ are infinite dimensional state vectors. The transition equations are

$$\mathbf{C}_{m,t+1} = \mathbf{F}\mathbf{C}_{m,t} + \mathbf{h}_m \varepsilon_{m,t+1}, \tag{12}$$

$$\mathbf{C}_{g,t+1} = \mathbf{F}\mathbf{C}_{g,t} + \mathbf{h}_g \varepsilon_{g,t+1}, \tag{13}$$

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

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

As in Van Binsbergen and Koijen (2010), the realized dividend growth rate is equal to the expected dividend growth rate plus an orthogonal shock:

$$\Delta d_{t+1} = g_t + \varepsilon_{d,t+1}. \tag{14}$$

Rearranging (3) for the price-dividend ratio and iterating forward we obtain the PV identity

$$pd_t = \frac{\kappa}{1-\rho} + E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, \tag{15}$$

which relates the log price-dividend ratio to expected future dividend growth and returns.

Using (12) and (13) in (14) and (15) we obtain the following measurement equations:

$$\Delta d_{t+1} = \mu_g + \mathbf{w}'\mathbf{C}_{g,t} + \varepsilon_{d,t+1}, \tag{16}$$

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

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

We now need only specify the covariance matrix of the "structural" shocks (which we assume to have mean zero and to be independent and identically distributed over

time) to complete the specification of this model:

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

The vector of parameters to be estimated is:

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

where ρ_{mg_r}, ρ_{gd} and ρ_{md} are correlation coefficients defined as: $\rho_{mg} = \sigma_{mg}/(\sigma_m \sigma_g), \rho_{gd} = \sigma_{gd}/(\sigma_g \sigma_d)$ and $\rho_{md} = \sigma_{md}/(\sigma_m \sigma_d)$.

4   Estimation

4.1   Data and Methodology

In our empirical investigation we use value-weighted NYSE/Amex/Nasdaq index data available from the Center for Research in Securities Prices (CRSP). We downloaded monthly data from January 1926 to December 2011 with and without dividends to construct series of annual returns.[9] These are the same series as in Van Binsbergen and Koijen (2010) and Rytchkov (2012) but updated to include observations for more recent years. We then obtain real returns and real dividend growth series by using the CPI index from the US Bureau of Labor Statistics.[10] Since the PV model is a first order approximation, it does not hold exactly for the observed data. Following Cochrane (2008a) and Van Binsbergen and Koijen (2010) we use exact measures of returns to find the dividend growth rates from the PV model and use it in subsequent analysis.[11]

We estimate four different specifications: $AR(1) - AR(1)$, $AR(1) - ARMA(1,1)$,

$ARFIMA(1, \delta, 0) - AR(1)$ and $ARFIMA(1, \delta, 0) - ARMA(1, 1)$, where the time series specifications refer to expected returns and expected dividend growth, respectively.

As pointed out by Rytchkov (2012) and Cochrane (2008b), the dimension of the covariance matrix of shocks is not identified in our system. Following Rytchkov (2012) and Van Binsbergen and Koijen (2010) we set the correlation between the expected dividend growth and unexpected dividend innovation to zero ($\rho_{gd} = 0$). Additionally, we found that the correlation between expected returns and unexpected dividend growth (ρ_{md}) in all our estimated models to be close to zero and statistically insignificant. As such, we decided to set it to zero as well, in order to reduce the number of parameters and increase the power of the estimation.

The remaining parameter values are obtained 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)).[12] The transition equations are given by (12), (13) and the measurement equations (16) and (17). 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 the lag $l \geq \sqrt{T}$. In estimation of the model we use the truncation at $l = 30$, but the results are robust to other choices.

4.2 Results

We found that, according to the $t-$test, the four model specifications to have all the parameters as statistically significant (except for the intercept of expected dividend growth). The estimates are reported in Table 4.

The dynamics of expected returns have a strong positive autoregressive component and a negative fractional integration component. Allowing for fractional integration in the model increases the estimate of the autoregressive coefficient from about 0.83 to about 0.89. The high persistence of expected returns is consistent with the findings in the literature (see Fama and French (1988a), Campbell and Cochrane (1999), Ferson, Sarkissian and Simin (2003), and Pástor and Stambaugh (2009)).

The autoregressive coefficient for expected dividend growth is strongly negative (this is in line with the results in Van Binsbergen and Koijen (2010), who also found similar results in the case of market-invested dividends), in the range -0.53 to -0.61 when assuming $AR(1)$ dynamics. Extending the model of expected dividend growth from $AR(1)$ to $ARMA(1,1)$ renders an even more negative autoregressive coefficient (around -0.88). The moving average component in dividend growth is about 0.6 and has relatively large standard errors (about 0.24).

As in Van Binsbergen and Koijen (2010) we find the correlation between expected dividend growth rates and expected returns to be positive, high and statistically significant (this is consistent with the work of Menzly, Santos and Veronesi (2004), and Lettau and Ludvigson (2005)).

The descriptive statistics of the estimated models are reported in Table 5. The first line shows the estimated number of parameters of particular models, which is between 8 for the most parsimonious $AR(1) - AR(1)$ to 10 for the $ARFIMA(1, \delta, 0) - ARMA(1,1)$ model. The likelihood ratio test examines the null hypothesis of equal fit to the data by particular models in relation to the most restricted model, which is the $AR(1) - AR(1)$. The test rejects the null hypothesis at the 10% significance level for the $AR(1) - ARMA(1,1)$, albeit marginally, and $ARFIMA(1, \delta, 0) - ARMA(1,1)$ models. The 10% significance level might seem rather high as a standard testing level, but since our annual time series have only 86 observations the power of the test must

be appropriately adjusted.[13]

For model selection criteria we calculated the Akaike (AIC) and Bayesian Information Criteria (BIC):

$$AIC = -2\ln(\mathcal{L}) + 2k, \tag{18}$$

$$BIC = -2\ln(\mathcal{L}) + k\ln(TN), \tag{19}$$

where $\mathcal{L}$ is the likelihood function evaluated at the maximum, k is the number of parameters, and T and N is the sample size in the temporal and cross-sectional dimensions, respectively. The two criteria select different models as the preferred one: the AIC favours the $ARFIMA(1,\delta,0) - ARMA(1,1)$ model, while the most parsimonious $AR(1) - AR(1)$ is preferred by the BIC. As can be seen from the information criteria formulae (18-19), the difference in model selection stems from the fact that while the BIC penalizes for the number of observations, the AIC does not.

The next two lines of Table 5 report the sample standard deviations of expected dividend growth and expected returns. As can be seen, the variability of the implied time series increases when we allow for a more flexible model specification than $AR(1) - AR(1)$. The variability of the expected excess returns almost doubles when we move from the short memory models to the models that include the fractional integration component. The standard deviation of the expected returns implied by the $AR(1) - AR(1)$ model is 3.82% and it goes up to 7.23% for the $ARFIMA(1,\delta,0) - ARMA(1,1)$ model. The $ARFIMA$ specification offers a higher variability which is more similar to the volatility of realized returns. On the other hand, the increase in the variability of expected dividend growth is rather moderate and it does not exceed 1%.

In the following line we report the sample correlation between the expected returns and expected dividend growth. The correlation between the two series for the $AR(1) - AR(1)$ model amounts to 0.12, but adding only the moving average component to the expected dividend process increases the correlation to 0.30. Adding the fractional integration component to the expected returns series increases the correlation to 0.38 for the $ARFIMA(1, \delta, 0) - AR(1)$ model and to 0.5 for the $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ model. Lettau and Ludvigson (2005). point out that this large positive correlation is consistent with higher variation in expected returns and expected dividend growth than apparent from the price-dividend ratio.

In the last two lines of Table 5 we report the R^2 statistics calculated as:

$$R_r^2 = 1 - \frac{v\hat{a}r(r_t - m_{t-1}^F)}{v\hat{a}r(r_t)},$$

and

$$R_{\Delta d}^2 = 1 - \frac{v\hat{a}r(\Delta d_t - g_{t-1}^F)}{v\hat{a}r(\Delta d_t)}.$$

where $v\hat{a}r$ is the sample variance, m_t^F is the filtered series for expected returns (m_t), and g_t^F is the filtered series for expected dividend growth rates (g_t). The filtered series are easily obtained from the Kalman filter.

Two facts are striking from these numbers. The first important fact is the predictability of returns and its dependence on the assumed time series specification. For the $AR(1) - AR(1)$ model the R^2 amounts to 0.13 and it raises slightly to 0.14 when allowing for $ARMA(1, 1)$ process in expected dividend growth. However, when we add the fractional integration component, the R^2 jumps to 0.21. Van Binsbergen and Koijen (2010) report R^2 values of 8% to 9%, which is in line with our $AR(1) - AR(1)$ model.[14]

Our results show that in a reduced form framework it is important to consider

carefully the underlying time series process. Second, the dividend growth process seems very predictable. For the $AR(1) - AR(1)$ model the R^2 for dividend growth is 0.32 and it is the highest for the $ARFIMA(1, \delta, 0) - AR(1)$ model, where it reaches 0.38. This is contrary to some results reported in related 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).

4.3 Forecasts Diagnostics

In this section we formally evaluate the predictions given by the PV model. To do so, we use the classic Mincer and Zarnowitz (1969) regressions, where the filtered series of expected returns and expected dividend growth are used as predictors:

$$r_{t+1} = \alpha + \beta \times m_t^F + u_t. \tag{20}$$

$$\Delta d_{t+1} = \alpha + \beta \times g_t^F + u_t. \tag{21}$$

Good predictors should be optimal ($\beta = 1$) and unbiased ($\alpha = 0$). In Table 6 we report the results of the regressions for returns in Panel A and for dividend growth in Panel B. From Panel A we can see that all models have a negative bias and the forecasts tend to 'overshoot' the realized values of returns, as evidenced by negative estimates of α and larger than 1 estimates of β. The statistical significance of these deviations seems to be particularly against the short memory models $AR(1) - AR(1)$ and $AR(1) - ARMA(1,1)$, as both the $t-$test values are larger than 2 in absolute values, and the F test of the joint null hypothesis ($H_0 : \alpha = 0$ and $\beta = 1$) rejects the null at the 5% and 10%, respectively. On the other hand, the models that incorporate the long memory component are characterized by the intercept and slope coefficients close to their hypothesized values, and their $t-$statistics and F test don't reject the

null hypothesis at any conventional significance level. Additionally, we also test for serial correlation in the residuals up to the second order. The models show a similar performance. The null hypothesis of no correlation can be rejected at the 5% level for the $AR(1) - ARMA(1,1)$ model, while the remaining models have $p-$values just above 5%.

In Panel B of Table 6 we report the same statistics for the dividend growth series. We can see that all models display a significant and positive bias of forecasts as evidenced by the $t-$test values of the intercept exceeding 2. The slope coefficients, on the other hand are uniformly smaller than 1, but only the $t-$statistic for the $AR(1) - AR(1)$ model rejects the null at the 5% level. The joint hypothesis is rejected for all models at the 5% level, which is the outcome of the statistical bias of the intercept. The only model not rejected jointly at the 1% level is the $ARFIMA(1,\delta,0) - ARMA(1,1)$. The test of the serial correlation in residuals rejects the $AR(1) - AR(1)$ model at the 5% level, and the $ARFIMA(1,\delta,0) - AR(1)$ at the 10% level, while there is no evidence of serial correlation in residuals for the $AR(1) - ARMA(1,1)$ and $ARFIMA(1,\delta,0) - ARMA(1,1)$ models.

In summary, the short memory models exhibit strong departure from the unbiasedness and optimal hypothesis for returns and dividend growth. Especially bad performance is noted for the $AR(1) - AR(1)$ model that fails all the tests, both of simple and joint hypotheses. Models which account for fractional integration seem to yield good forecasts of returns and also improve the forecasts of dividend growth. The $ARFIMA(1,\delta,0) - ARMA(1,1)$ model seems to perform the best overall.

4.4   Out-of-Sample Forecast Exercises

Since it is well known that some models can predict stock returns very well in-sample but perform badly out-of-sample (see e.g. Goyal and Welch (2003), and Welch and

Goyal (2008)), we examine the out-of-sample forecasting ability of our time series models in Table 7. Specifically, we consider prediction of the models on the last 40 years of data, that is $1971 - 2011$. Panel A presents the benchmark results obtained from using the parameters estimated on the whole sample, thus it consists of in-sample forecast. As can be seen, the chosen subperiod is much less predictable than the whole period, since the R^2 coefficients from both dividend growth and expected returns are much smaller than those reported in Table 5. The models with the long memory component show a better fit to both dividend growth and expected returns time series in this subperiod (the $AR(1) - AR(1)$ has lower R^2_r and $R^2_{\Delta d}$ values than the $ARFIMA(1, \delta, 0) - AR(1)$ model, while the $AR(1) - ARMA(1, 1)$ also has lower R^2_r and $R^2_{\Delta d}$ values than the $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ model). In Panels B and C we report the out-of-sample forecast produced by two methodologies. In Panel B we report the results obtained by estimating the models only once on the data $1926 - 1970$ and using these estimates to find the subsequent forecast. The results in Panel C were obtained by expanding the data used in estimation recursively by one observation each time and making the prediction for the next year. As could be expected, the out-of-sample forecasts deteriorate significantly as compared to the in-sample predictions.

The returns predictions generated by the short memory models, $AR(1) - AR(1)$ and $AR(1) - ARMA(1, 1)$, perform worse than the sample mean, as evidenced by the negative R^2 values. On the other hand, the models that include the fractional integration component predicted better than the sample mean. The degree of prediction is very modest, but nevertheless the R^2 statistic is positive for all models, both for the fixed point method and recursive forecast. We emphasize that the out-of-sample results should be taken with caution since the sample period is very small.

The dividend growth, however, remains still strongly predictable by all models

with R^2 ranging between 0.7 to 0.11 for the recursive estimation method. Interestingly, dividend growth is better predicted by the fixed point estimation method than by recursive forecasts. One could expect the opposite relationship, since the recursive forecast should make use of increasing information available to make new forecasts. We interpret this as the effect of small sample uncertainty. Just as with expected returns, the introduction of the long memory component leads to an improvement in the model's out-of-sample forecast performance of dividend growth (the $AR(1) - AR(1)$ has lower R^2 than the $ARFIMA(1, \delta, 0) - AR(1)$ model and the $AR(1) - ARMA(1, 1)$ also has lower R^2 than the $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ model) with either of the two methodologies.

4.5 Expected Returns and Dividend Growth over the Business Cycle

In Figure 2 we plot the time series of realized (blue line) and expected returns as implied by the models $AR(1) - AR(1)$ (green line) and $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ (red line). The time series for dividend growth (blue line) and expected dividend growth for the $AR(1) - AR(1)$ (green line) and $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ (red line) are shown in Figure 3. The grey areas denote the NBER recession periods. Since our data is annual, we plotted only recessions that lasted at least 9 months. We can see that the higher variability of expected returns implied by the $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ model in comparison to the $AR(1) - AR(1)$ is very prominent. The expected returns seem to have a very strong countercyclical pattern: they fall in the period prior and at the start of economic downturns and then increase as the period of expansion approaches. Since the fit of the expected dividend growth is not so sensitive to a choice of a particular model, the two implied series are quite close to each other. From Figure 3 can be seen that the expected dividend growth

series also exhibits a countercyclical pattern.

In order to examine the cyclicality of expected returns and dividend growth we regress a set of macro variables on the filtered series of expected returns and expected dividend growth. The macro series are growth of real consumption ($\Delta Cons$), growth of real GDP (ΔGDP) and growth of industrial production of consumption goods (ΔIP). The growth of the series is defined as the log difference. We chose these variables since they are meaningful indicators of the business cycle. The filtered series are obtained from the whole available sample of returns, that is $1926 - 2011$, however the consumption growth and GDP growth are available only from 1930 and industrial production growth only from 1940, the regressions are therefore run for those respective periods.[15] In Table 8 we report the slope coefficients with the $t-$statistics calculated from the ordinary least squares (OLS) standard errors (reported in small font) and the regression R^2.[16] In Panel A we report the regression on expected returns while the regression on implied dividend growth is reported in Panel B.

The results allow us to make a few observations. First, despite the countercyclical nature of both expected returns and expected dividend growth, the latter is a stronger predictor of the business cycle. It is especially visible for predictive regressions of GDP growth; while the expected returns are not significant and have the slope coefficients close to zero, the expected dividend growth have statistically significant slopes at 5% level for all models.

Second, although the significance of the predictors does not change significantly for different models, we can observe that the models that incorporate the fractional integration component seem to predict better than short memory models. This observation can be made for almost all regressions (the only exception consists of the predictive regressions of GDP growth by the expected dividend growth series), but it is especially evident for prediction of consumption growth by the expected dividends

series, where the R^2 value increases from 0.13 for the $AR(1) - AR(1)$ model to 0.19 for the $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ model.

These results suggest that obtaining implied expected returns and dividend growth series can have an important application as leading economic indicators. Particularly more so if the PV model includes a fractional integration component, since this leads to a larger degree of countercyclicality for both the case of expected returns and expected dividend growth as indicated by a larger R^2 statistic.

5   Conclusion

In this paper we show that the long range persistence of the price-dividend ratio renders the simple return forecasting regression considered by Fama and French (1988a) as invalid. Moreover, we argue that in finite sample the order of integration of the log return series should be approximately the same as that of the first differenced price-dividend ratio, which induces negative memory in the return series. We found evidence confirming this conjecture using semi-parametric estimators; we found that the dividend ratio series is nonstationary but mean reverting with a fractional integration parameter estimate of about 0.8, while the return series is characterized by a fractional integration parameter amounting to about -0.2.

We incorporate the fractional integration feature in the PV model using an $ARFIMA$ time series specification. Using model selection criteria we found that the preferable joint model is the one with $ARFIMA(1, \delta, 0)$ expected returns and $ARMA(1, 1)$ expected dividend growth processes. The fractionally integrated model yields better returns and dividend growth forecasts than the $AR(1)$ model, both in-sample and out-of-sample. Using Mincer-Zarnowitz style regressions we found that our model correctly captures the variation in expected returns and dividend growth.

Our work has important implications for the popular return forecasting literature.

The potential imbalance in the regression can be a reason for the very mixed and hotly debated regression results. Also, using a structural model that takes into account the stylized features of the data can certainly help understand the underlying forces. As emphasized by Cochrane (2011), a correct understanding of the risk premia is vital for macro-prudential regulation and monetary policy.

Notes

[1]Lettau and Van Nieuwerburgh (2008) explained the strong persistence in the price-dividend series as a result of structural breaks (or shifts) in the steady state mean of the economy. They showed that if the shifts are accounted for, then the return forecasting ability of the price-dividend ratio is stable over time. These findings reinforce the long memory argument in the price-dividend ratio. As showed by Diebold and Inoue (2001), rare structural breaks and long memory are really two sides of the same coin and they cannot be distinguished from each other in finite samples. On the other hand, Granger and Hyung (2004) established that, if the true series is a long memory process, it is very likely that spurious breaks will be detected. Conversely, even if the true process was generated by occasional breaks, the long memory process can successfully reproduce many features of the true series and (under some conditions) can yield better forecasts. Indeed, Lettau and Van Nieuwerburgh (2008) reported that difficulties with detecting the breaks in real time makes it hard to forecast stock returns. An alternative explanation of the long memory feature in the aggregate price-dividend ratio could be due to aggregation. Granger (1980) has shown that "integrated series can occur from realistic aggregation situations" (for example: independent series generated by a first order autoregressive process can result in a fractionally integrated series when aggregated).

[2]Negative serial correlation in common stock returns at long horizons has indeed been found in many studies (e.g. De Bondt and Thaler (1985), Fama and French (1988b), Poterba and Summers (1988)). This evidence has been interpreted in many studies (such as De Bondt and Thaler (1985)) as overreaction to past events due to waves of optimism or pessimism among investors. The finding that investors expect

such long-run return reversals, however, supports the idea that mean reversion can be consistent with the efficient functioning of markets as argued by Malkiel (2003).

[3]See also Campbell and Diebold (2009) and references therein.

[4]Willinger, Taqqu and Teverovsky (1999) found evidence of small degree of long range dependence in stock returns. On the other hand, Lo (1991) found no statistical evidence of long memory in stock returns. Lobato and Savin (1998) study rejected the hypothesis of long memory in the levels of returns but found the presence of long memory in squared returns (in line with the findings of Ding, Granger and Engle (1993)). Recent studies have reinforced the view that long memory is important to the understanding of asset prices. Bollerslev, Osterrieder, Sizova and Tauchen (2013) estimate a fractionally cointegrated VAR model for returns, objective and risk-neutral volatilities using high-frequency intraday data. Sizova (2013) demonstrates that accounting for long memory in predictive variables is important when considering long-horizon return regressions.

[5]Since the number of lags in the test selected by AIC is 0, the test collapses to the simple Dickey-Fuller test.

[6]We considered extensions of different autoregressive and moving average orders of $ARFIMA$ processes for both the expected returns and expected dividend growth processes and found the $ARFIMA(1, \delta_m, 0)$ model of expected returns and $ARMA(1, 1)$ model of expected dividend growth to be the most general specification with all coefficients significant at the 5% level.

[7]See Brockwell and Davis (2009) for details on deriving the moving average coefficients.

[8]See Granger and Joyeux (1980) and Hosking (1981). For a textbook treatment

of long memory see Brockwell and Davis (2009) or Palma (2007).

[9]Although monthly or quarterly data would be preferable, we found a strong seasonal pattern in the correlogram of dividend growth series at quarterly frequency, which, if not accounted for, invalidates the time series analysis of the dynamics. See also Ang and Bekaert (2007) and Cochrane (2011), appendix. A.1.

[10]Since firms can pay dividends at different times of a year, as shown by Cochrane (1991), dividends paid early in the year are treated as reinvested at the market rate of return to the end of the year. Van Binsbergen and Koijen (2010) considered the PV model with market reinvested and risk-free rate reinvested strategies and showed that the resulting aggregate dividend growth series are very similar.

[11]The difference between the observed and implied dividend growth is negligible. The correlation between these two series amounts to 0.9997.

[12]See appendix for details.

[13]For a more detailed discussion of this argument see Hendry (1995).

[14]The R^2 values for annual returns reported in the literature for the long sample, starting in 1926 are about $3\% - 9\%$, see e.g. Campbell, Lo and MacKinlay (1997), ch.7, Goyal and Welch (2003). The price-dividend ratio is generally found to forecast returns better in the second half of the twentieth century until the 1990s, as evidenced by Campbell et al. (1997), Goyal and Welch (2003), Lewellen (2004) and Koijen and Van Nieuwerburgh (2011). Lettau and Van Nieuwerburgh (2008) considered a 30 year rolling sample and found R^2 values ranging from close to zero to 30%.

[15]The macro data was obtained from the Federal Reserve Economic Data (FRED), which is freely available at the website of the Federal Reserve Bank of St.Louis.

[16] In the regressions we did not detect neither heteroskedasticity nor autocorrelation in the residuals.

References

Andersson, M. K. and Nydahl, S.: 1998, Rational bubbles and fractional alternatives, *Unpublished working paper* .

Ang, A. and Bekaert, G.: 2007, Stock return predictability: Is it there?, *Review of Financial Studies* **20**(3), 651–707.

Barro, R. J.: 1990, The stock market and investment, *Review of Financial Studies* **3**(1), 115–131.

Bollerslev, T., Osterrieder, D., Sizova, N. and Tauchen, G.: 2013, Risk and return: Long-run relations, fractional cointegration, and return predictability, *Journal of Financial Economics* **108**(2), 409–424.

Brockwell, P. J. and Davis, R. A.: 2009, *Time series: theory and methods*, Springer, second edition.

Campbell, J. Y. and Cochrane, J. H.: 1999, By force of habit: A consumption based explanation of aggregate stock market behavior, *Journal of Political Economy* **107**(2), 205–251.

Campbell, J. Y., Lo, A. W. and MacKinlay, C. A.: 1997, *The Econometrics of Financial Markets*, Princeton University Press Princeton, NJ.

Campbell, J. Y. and Shiller, R. J.: 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* **1**(3), 195–228.

Campbell, S. D. and Diebold, F. X.: 2009, Stock returns and expected business conditions: Half a century of direct evidence, *Journal of Business & Economic Statistics* **27**(2), 266–278.

Chan, N. H. and Palma, W.: 1998, State space modeling of long-memory processes, *Annals of Statistics* **26**(2), 719–740.

Chen, N.-F., Roll, R. and Ross, S. A.: 1986, Economic forces and the stock market, *Journal of Business* **59**(3), 383–403.

Cochrane, J. H.: 1991, Volatility tests and efficient markets: A review essay, *Journal of Monetary Economics* **27**(3), 463–485.

Cochrane, J. H.: 1999, New facts in finance, *Economic Perspectives* (Q III), 36–58.

Cochrane, J. H.: 2005, *Asset pricing*, Princeton University Press, Princeton.

Cochrane, J. H.: 2008a, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* **21**(4), 1533–1575.

Cochrane, J. H.: 2008b, State-space vs. var models for stock returns, *Unpublished working paper* .

Cochrane, J. H.: 2011, Presidential address: Discount rates, *Journal of Finance* **66**(4), 1047–1108.

De Bondt, W. F. M. and Thaler, R.: 1985, Does the stock market overreact?, *Journal of Finance* **40**(3), 793–805.

Dickey, D. A. and Fuller, W. A.: 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* **74**(366a), 427–431.

Diebold, F. X. and Inoue, A.: 2001, Long memory and regime switching, *Journal of Econometrics* **105**(1), 131–159.

Ding, Z., Granger, C. W. J. and Engle, R. F.: 1993, A long memory property of stock market returns and a new model, *Journal of Empirical Finance* **1**(1), 83–106.

Fama, E. F.: 1990, Stock returns, expected returns, and real activity, *Journal of Finance* **45**(4), 1089–1108.

Fama, E. F. and French, K. R.: 1988a, Dividend yields and expected stock returns, *Journal of Financial Economics* **22**(1), 3–25.

Fama, E. F. and French, K. R.: 1988b, Permanent and temporary components of

stock prices, *Journal of Political Economy* **96**(2), 246–273.

Fama, E. F. and French, K. R.: 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* **25**(1), 23–49.

Ferson, W. E., Sarkissian, S. and Simin, T. T.: 2003, Spurious regressions in financial economics?, *Journal of Finance* **58**(4), 1393–1414.

Geweke, J. and Porter-Hudak, S.: 1983, The estimation and application of long memory time series models, *Journal of Time Series Analysis* **4**(4), 221–238.

Goyal, A. and Welch, I.: 2003, Predicting the equity premium with dividend ratios, *Management Science* **49**(5), 639–654.

Granger, C. W. J.: 1980, Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics* **14**(2), 227–238.

Granger, C. W. J. and Hyung, N.: 2004, Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns, *Journal of Empirical Finance* **11**(3), 399–421.

Granger, C. W. J. and Joyeux, R.: 1980, An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis* **1**(1), 15–29.

Granger, C. W. J. and Newbold, P.: 1974, Spurious regressions in econometrics, *Journal of Econometrics* **2**(2), 111–120.

Hamilton, J. D.: 1994, *Time series analysis*, Princeton University Press, Princeton.

Hendry, D. F.: 1995, *Dynamic econometrics*, Oxford University Press.

Hodrick, R. J.: 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial studies* **5**(3), 357–386.

Hosking, J. R. M.: 1981, Fractional differencing, *Biometrika* **68**(1), 165–176.

Koijen, R. S. J. and Van Nieuwerburgh, S.: 2011, Predictability of returns and cash flows, *Unpublished working paper* .

Lettau, M. and Ludvigson, S. C.: 2005, Expected returns and expected dividend

growth, *Journal of Financial Economics* **76**(3), 583–626.

Lettau, M. and Van Nieuwerburgh, S.: 2008, Reconciling the return predictability evidence, *Review of Financial Studies* **21**(4), 1607–1652.

Lewellen, J.: 2004, Predicting returns with financial ratios, *Journal of Financial Economics* **74**(2), 209–235.

Lo, A. W.: 1991, Long-term memory in stock market prices, *Econometrica* **59**(5), 1279–1313.

Lobato, I. N. and Savin, N. E.: 1998, Real and spurious long-memory properties of stock-market data, *Journal of Business & Economic Statistics* **16**(3), 261–268.

Malkiel, B. G.: 2003, The efficient market hypothesis and its critics, *Journal of Economic Perspectives* **17**(1), pp. 59–82.

Mankiw, G. N. and Shapiro, M. D.: 1986, Do we reject too often?: Small sample properties of tests of rational expectations models, *Economics Letters* **20**(2), 139–145.

Maynard, A. and Phillips, P. C. B.: 2001, Rethinking an old empirical puzzle: econometric evidence on the forward discount anomaly, *Journal of Applied Econometrics* **16**(6), 671–708.

Menzly, L., Santos, T. and Veronesi, P.: 2004, Understanding predictability, *Journal of Political Economy* **112**(1), 1–47.

Mincer, J. A. and Zarnowitz, V.: 1969, The evaluation of economic forecasts, *In Mincer, J. (ed.), Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER.

Palma, W.: 2007, *Long-memory time series: theory and methods*, Vol. 662, John Wiley & Sons.

Pástor, L. and Stambaugh, R. F.: 2009, Predictive systems: Living with imperfect predictors, *Journal of Finance* **64**(4), 1583–1628.

Poterba, J. M. and Summers, L. H.: 1988, Mean reversion in stock prices: Evidence

and implications, *Journal of Financial Economics* **22**(1), 27–59.

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

Rytchkov, O.: 2012, Filtering out expected dividends and expected returns, *Quarterly Journal of Finance* **2**(03), 1–56.

Shimotsu, K.: 2010, Exact local whittle estimation of fractional integration with unknown mean and time trend, *Econometric Theory* **26**(02), 501–540.

Sizova, N.: 2013, Long-horizon return regressions with historical volatility and other long-memory variables, *Journal of Business & Economic Statistics* **31**(4), 546–559.

Stambaugh, R. F.: 1986, Bias in regressions with lagged stochastic regressors, *Unpublished working paper* .

Stambaugh, R. F.: 1999, Predictive regressions, *Journal of Financial Economics* **54**(3), 375–421.

Van Binsbergen, J. H. and Koijen, R. S. J.: 2010, Predictive regressions: A present-value approach, *Journal of Finance* **65**(4), 1439–1471.

Welch, I. and Goyal, A.: 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* **21**(4), 1455–1508.

Willinger, W., Taqqu, M. S. and Teverovsky, V.: 1999, Stock market prices and long-range dependence, *Finance and stochastics* **3**(1), 1–13.

Appendix A

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

In order to obtain the Kalman equations it is convenient to write the measurement equations in the form where the shocks are lagged relatively to the state vector (see e.g. Brockwell and Davis (2009)). Therefore we define the new state variables $\mathbf{x}_{m,t+1} = \mathbf{C}_{m,t}$ and $\mathbf{x}_{g,t+1} = \mathbf{C}_{g,t}$, so the transition equations are now

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

$$\mathbf{x}_{g,t+1} = \mathbf{F}\mathbf{x}_{g,t} + \mathbf{h}_g \varepsilon_{g,t}, \tag{A-2}$$

and the measurement equations are:

$$\Delta d_t = \mu_g + \mathbf{w}' \mathbf{x}_{g,t} + \varepsilon_{d,t}, \tag{A-3}$$

$$pd_t = A + \mathbf{b}' \mathbf{F} \mathbf{x}_{g,t} - \mathbf{b}' \mathbf{F} \mathbf{x}_{m,t} + \mathbf{b}' \mathbf{h}_g \varepsilon_{g,t} - \mathbf{b}' \mathbf{h}_m \varepsilon_{m,t}. \tag{A-4}$$

In general notation the transition and measurement equations are

$$\mathbf{x}_{t+1} = \overline{\mathbf{F}} \mathbf{x}_t + \mathbf{v}_t, \tag{A-5}$$

$$\mathbf{y}_t = \mathbf{e} + \overline{\mathbf{W}} \mathbf{x}_t + \mathbf{z}_t, \tag{A-6}$$

with

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

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

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

The Kalman recursive equations of the model are:

$$\left\{\begin{aligned} \boldsymbol{\Delta}_t &= \overline{\mathbf{W}}\boldsymbol{\Omega}_t\overline{\mathbf{W}}' + \mathbf{R} \\ \boldsymbol{\Theta}_t &= \overline{\mathbf{F}}\boldsymbol{\Omega}_t\overline{\mathbf{W}}' + \mathbf{S} \\ \boldsymbol{\Omega}_{t+1} &= \overline{\mathbf{F}}\boldsymbol{\Omega}_t\overline{\mathbf{F}}' + \mathbf{Q} - \boldsymbol{\Theta}_t\boldsymbol{\Delta}_t^{-1}\boldsymbol{\Theta}_t' \\ \widehat{\mathbf{x}}_{t+1} &= \overline{\mathbf{F}}\widehat{\mathbf{x}}_t + \boldsymbol{\Theta}_t\boldsymbol{\Delta}_t^{-1}\left(\mathbf{y}_t - \mathbf{e} - \overline{\mathbf{W}}\widehat{\mathbf{x}}_t\right) \end{aligned}\right\} \qquad \text{(A-7)}$$

where

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

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

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

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

The log likelihood function is then given by:

$$\ell = (2\pi)^{-kT/2}\left(\prod_{t=1}^{T}\det\boldsymbol{\Delta}_t\right)^{-1/2}\exp\left(-\frac{1}{2}\sum_{t=1}^{T}(\mathbf{y}_t - \widehat{\mathbf{y}}_t)'\boldsymbol{\Delta}_j^{-1}(\mathbf{y}_t - \widehat{\mathbf{y}}_t)\right) \qquad \text{(A-8)}$$

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

α	β	$t - adf$
0.2046	0.9417	-1.5199
0.1285	0.0384	

Table 1: ADF test for the price-dividend ratio. The 5 percent critical value is -2.90. The data sample is 1926-2011.

		p/d ratio	returns	div.growth
GPH	δ	1.0763	-0.2313	-0.3738
	std.err.	0.2138	0.2138	0.2138
Robinson	δ	0.8117	-0.2154	-0.3479
	std.err.	0.1091	0.1091	0.1091
Shimotsu	δ	0.8079	-0.1690	-0.2438
	std.err.	0.1078	0.1078	0.1078

Table 2: Semiparametric estimation of the fractional integration parameter δ for real returns and dividend growth series. The estimators applied are proposed by: Geweke and Porter-Hudak (1985), Robinson (1995) and Shimotsu (2010). The standard errors are reported in small font. The data sample is 1926-2011.

	p/d ratio	returns	div.growth
Mean	3.3294	6.1290	1.9898
Std Dev	0.4301	20.0743	14.8599
Skew	0.6154	−0.7672	0.1650
Ex. Kurtosis	−0.1252	0.4177	0.2310
Min	2.6268	−48.8422	−30.8588
Max	4.4991	44.6672	44.0666
1st lag Autocorr.	0.9243	0.0164	−0.1279
10th lag Autocorr.	0.3983	−0.0442	−0.0395

Table 3: Descriptive statistics of the real returns and dividend growth series. The data sample is 1926-2011.

		Model		
m_t :	$AR(1)$	$AR(1)$	$ARFIMA(1,\delta,0)$	$ARFIMA(1,\delta,0)$
g_t :	$AR(1)$	$ARMA(1,1)$	$AR(1)$	$ARMA(1,1)$
μ_m	0.0644	0.0634	0.0665	0.0659
std.err	0.0159	0.0169	0.0159	0.0182
ϕ_m	0.8321	0.8254	0.8917	0.8840
std.err	0.0747	0.0719	0.0498	0.0510
δ			-0.3269	-0.2929
std.err			0.1368	0.1303
μ_g	0.0179	0.0170	0.0180	0.0176
std.err	0.0145	0.0151	0.0157	0.0179
ϕ_g	-0.6050	-0.8861	-0.5348	-0.8714
std.err	0.2392	0.0861	0.2156	0.1009
θ		0.5874		0.5764
std.err		0.2229		0.2360
σ_m	0.0322	0.0369	0.0661	0.0684
std.err	0.0131	0.0141	0.0158	0.0179
σ_g	0.0599	0.0651	0.0719	0.0731
std.err	0.0200	0.0167	0.0191	0.0174
σ_d	0.1293	0.1296	0.1268	0.1286
std.err	0.0145	0.0113	0.0116	0.0115
ρ_{mg}	0.8410	0.9098	0.8664	0.9183
std.err	0.0854	0.0684	0.0677	0.0574

Table 4: Estimation results of different present-value models. The standard errors are reported in small font. The data sample is 1926-2011.

			Model	
m_t :	$AR(1)$	$AR(1)$	$ARFIMA(1,\delta,0)$	$ARFIMA(1,\delta,0)$
g_t :	$AR(1)$	$ARMA(1,1)$	$AR(1)$	$ARMA(1,1)$
k	8	9	9	10
LR	$-$	2.7075	2.6716	4.8418
$p-value$		0.0999	0.1022	0.0888
AIC	-164.2585	-164.9660	-164.9302	-165.1004
BIC	-139.0786	-136.6386	-136.6027	-133.6254
$\sigma(m_t)$	3.8189	4.2736	7.2302	7.4888
$\sigma(g_t)$	5.1369	5.4733	5.8259	6.0329
$corr(m_t, g_t)$	0.1239	0.3043	0.3817	0.4986
R_r^2	0.1321	0.1443	0.2060	0.2058
$R_{\Delta d}^2$	0.3166	0.3200	0.3763	0.3577

Table 5: Estimation statistics ofdifferent present-value models. In rows we report: number of parameters in a model; likelihood ratio test performed relatively to the AR(1)-AR(1) model with associated p-values reported in small font; Akaike Information Criterion; Bayesian Information Criterion; sample standard deviation of expected dividend growth; sample standard deviation of expected stock returns; sample correlation between expected returns and expecte dividend growth; R-squared coefficient of dividend growth and stock returns.

		Model		
$m_t:$	$AR(1)$	$AR(1)$	$ARFIMA(1,\delta,0)$	$ARFIMA(1,\delta,0)$
$g_t:$	$AR(1)$	$ARMA(1,1)$	$AR(1)$	$ARMA(1,1)$

Panel A: $r_{t+1} = \alpha + \beta \times m_t^F + u_t$

	$AR(1)$ / $AR(1)$	$AR(1)$ / $ARMA(1,1)$	$ARFIMA(1,\delta,0)$ / $AR(1)$	$ARFIMA(1,\delta,0)$ / $ARMA(1,1)$
α	-0.0869	-0.0724	-0.0252	-0.0227
std.err.	0.0382	0.0352	0.0263	0.0261
β	2.3248	2.0924	1.2939	1.2394
std.err.	0.5144	0.4589	0.2680	0.2593
$t-val.\ (H_0:\alpha=0)$	-2.2777	-2.0558	-0.9582	-0.8685
$t-val.\ (H_0:\beta=1)$	2.5755	2.3806	1.0964	0.9232
$F\ (H_0:\alpha=0,\beta=1)$	3.3246	2.8425	0.6428	0.4820
$p-value$	0.0408	0.0639	0.5284	0.6193
$F\ test\ [AR(1)-AR(2)]$	2.8533	3.2463	3.1044	3.0333
$p-value$	0.0634	0.0440	0.0502	0.0536

Panel B: $\Delta d_{t+1} = \alpha + \beta \times g_t^F + u_t$

	$AR(1)$ / $AR(1)$	$AR(1)$ / $ARMA(1,1)$	$ARFIMA(1,\delta,0)$ / $AR(1)$	$ARFIMA(1,\delta,0)$ / $ARMA(1,1)$
α	0.0585	0.0544	0.0500	0.0474
std.err.	0.0232	0.0229	0.0227	0.0225
β	0.1468	0.3685	0.5697	0.6855
std.err.	0.4261	0.3982	0.3708	0.3553
$t-val.\ (H_0:\alpha=0)$	2.5202	2.3713	2.2026	2.1077
$t-val.\ (H_0:\beta=1)$	-2.0024	-1.5860	-1.1605	-0.8852
$F\ (H_0:\alpha=0,\beta=1)$	7.2104	5.2610	7.3222	4.7647
$p-value$	0.0013	0.0070	0.0012	0.0110
$F\ test\ [AR(1)-AR(2)]$	3.7835	1.9651	2.5621	1.4814
$p-value$	0.0269	0.1468	0.0834	0.2334

Table 6: Results of the Mincer-Zarnowitz (1969) regressions for real returns (Panel A) and dividend growth (Panel B) for different 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 and consistent 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. The last line shows the value of the $F-$test of no autocorrelation of order 1 and 2 in regression residuals with the corresponding p-values. In Panel B we report the same results for dividend growth.

			Model	
$m_t:$	$AR(1)$	$AR(1)$	$ARFIMA(1,\delta,0)$	$ARFIMA(1,\delta,0)$
$g_t:$	$AR(1)$	$ARMA(1,1)$	$AR(1)$	$ARMA(1,1)$
		Panel A: In-sample forecast		
R^2_r	0.0310	0.0297	0.0948	0.0914
$R^2_{\Delta d}$	0.1099	0.1535	0.1315	0.1659
		Panel B: Fixed point estimation forecast		
R^2_r	-0.0311	-0.0396	0.0433	0.0359
$R^2_{\Delta d}$	0.1015	0.1258	0.1080	0.1304
		Panel C: Recursive estimation forecast		
R^2_r	-0.0209	-0.0295	0.0183	0.0083
$R^2_{\Delta d}$	0.0733	0.1005	0.0883	0.1058

Table 7: Out-of-sample forecast power of different present-value models. The forecasts are evaluated on the subsample 1971-2011. In Panel A we report the 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 1926-1970 and use these estimates to generate forecasts for the rest of the sample. In Panel C we start from estimating the model on the sample 1926-1970 and making the prediction for the next year. In the next step we extend the estimation sample by one observation and make a prediction for the next year, and so on.

	Model
$m_t:$	$AR(1)$	$AR(1)$	$ARFIMA(1,\delta,0)$	$ARFIMA(1,\delta,0)$
$g_t:$	$AR(1)$	$ARMA(1,1)$	$AR(1)$	$ARMA(1,1)$

	$AR(1)$ / $AR(1)$	$AR(1)$ / $ARMA(1,1)$	$ARFIMA(1,\delta,0)$ / $AR(1)$	$ARFIMA(1,\delta,0)$ / $ARMA(1,1)$
Panel A:				
$\Delta Cons_{t+1} = \alpha + \beta \times m_t^F + u$				
β	-0.2666	-0.2460	-0.1502	-0.1439
$t-value$	-3.3171	-3.3988	-3.5959	-3.5393
R^2	0.1236	0.1290	0.1422	0.1384
$\Delta GDP_{t+1} = \alpha + \beta \times m_t^F + u$				
β	-0.0062	-0.0122	-0.0239	-0.0222
$t-value$	-0.0437	-0.0956	-0.3225	-0.3088
R^2	0.0000	0.0001	0.0013	0.0012
$\Delta IP_{t+1} = \alpha + \beta \times m_t^F + u$				
β	-0.2785	-0.2592	-0.1535	-0.1474
$t-value$	-2.0031	-2.0630	-2.1268	-2.0990
R^2	0.0557	0.0589	0.0624	0.0608
Panel B:				
$\Delta Cons_{t+1} = \alpha + \beta \times g_t^F + u$				
β	-0.2063	-0.2205	-0.2082	-0.2156
$t-value$	-3.4067	-3.9450	-4.0264	-4.3383
R^2	0.1295	0.1663	0.1721	0.1944
$\Delta GDP_{t+1} = \alpha + \beta \times g_t^F + u$				
β	-0.2108	-0.2084	-0.1843	-0.1785
$t-value$	-2.0274	-2.1308	-2.0250	-2.0118
R^2	0.0501	0.0550	0.0499	0.0493
$\Delta IP_{t+1} = \alpha + \beta \times g_t^F + u$				
β	-0.3330	-0.3437	-0.3158	-0.3196
$t-value$	-3.1928	-3.4493	-3.5102	-3.6055
R^2	0.1304	0.1489	0.1534	0.1605

Table 8: Results from the regression of macro variables on model filtered expected returns(Panel A) and expected dividend growth (panel B). The macro variables are real consumption growth ($\Delta Cons$), real GDP growth (ΔGDP) and growth of industrial production of consumer goods (ΔIP). The intercept is ommited from the table. The t-statistics calculate from OLS standard errors are reported in small font. The implied series are obtained from the whole sample 1926-2011 and the regressions are run on available samples of macro variables, that is 1930-2011 for consumption and GDP growth and 1940-2011 for industrial production growth.
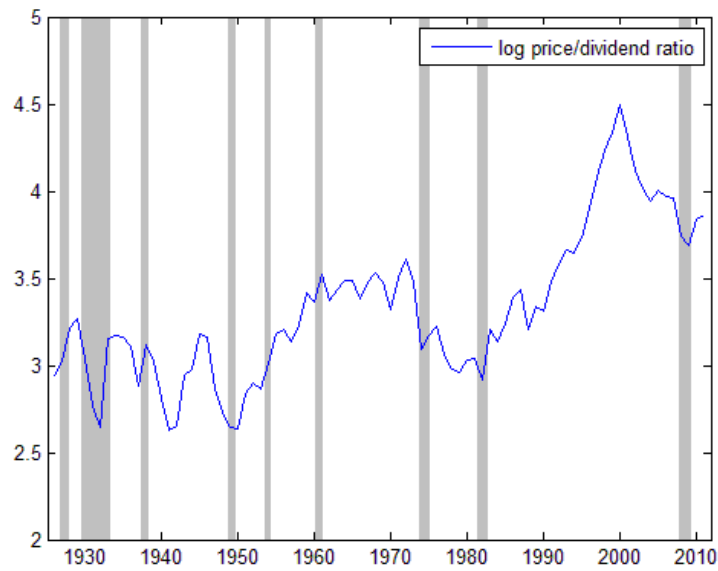
Figure 1: Time series of log price/dividend ratio. The grey areas denote the recession periods (only those longer than 9 months).
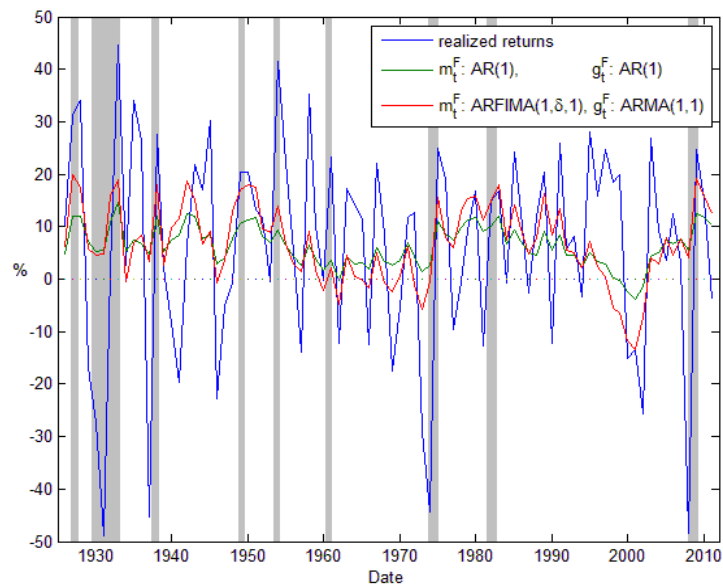
Figure 2: Realized returns (blue line) and expected returns as implied by the models $AR(1) - AR(1)$ (green line) and $ARFIMA(1, \delta, 0) - ARMA(1, 1)$ (red line). The grey areas denote the recession periods (only those longer than 9 months).
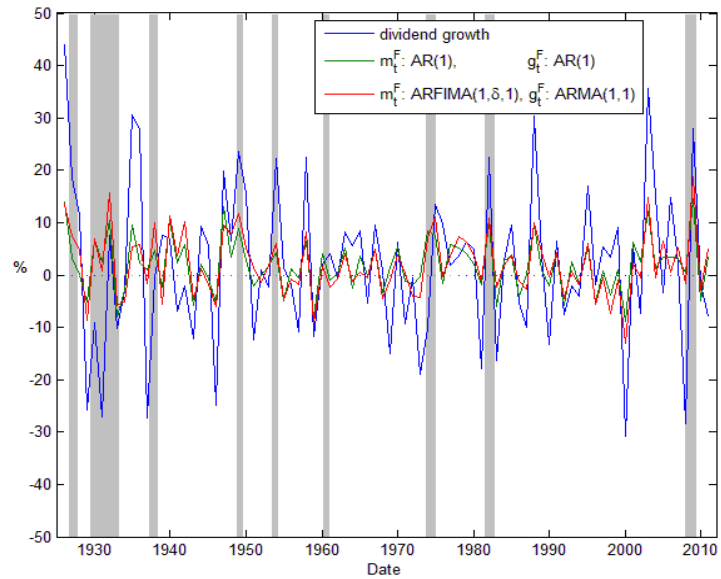
Figure 3: Realized dividend growth (blue line) and expected dividend growth as implied by the models $AR(1)-AR(1)$ (green line) and $ARFIMA(1,\delta,0)-ARMA(1,1)$ (red line). The grey areas denote the recession periods (only those longer than 9 months).