# “Superstitious” Investors*

Hongye Guo† Jessica A. Wachter‡

September 6, 2019

**Abstract**

We consider an economy in which investors believe dividend growth is predictable, when in reality it is not. We show that these beliefs lead to excess volatility and return predictability. We also show that these beliefs are reasonable in the face of evidence on dividend growth. We apply this framework to explaining the value premium, predictability of bond returns, and the violation of uncovered interest rate parity.

Keywords: Excess volatility, Extrapolative expectations, Rare events, Overconfidence

JEL codes: G12, G15, G41

*We thank Nicholas Barberis, Simcha Barkai, John Campbell, Anna Cieslak, Marco Grotteria, Campbell Harvey, David Hirshleifer, Michael Kahana, Karen Lewis, Andrei Shleifer, Adrien Verdelhan, and seminar participants at Boston University, Duke, EPFL/University of Lausanne, London Business School, MIT, and at Wharton for helpful comments.

†Department of Finance, The Wharton School, University of Pennsylvania. Email: hoguo@wharton.upenn.edu

‡Department of Finance, The Wharton School, University of Pennsylvania and NBER; Email: jwachter@wharton.upenn.edu;

# 1 Introduction

Why is aggregate stock price volatility so high? Starting with Shiller (1981) and Campbell and Shiller (1988), an influential literature shows that stock market volatility is too large to arise from rational expectations of future dividends. In response, the literature has proposed several explanations that maintain the notion of the rational investor. The "excess" volatility could arise from time-varying discount rates, which could in turn be driven by time-varying volatility of dividends (Bansal and Yaron, 2004; Calvet and Fisher, 2007; Lettau et al., 2008), or time-varying risk aversion Campbell and Cochrane (1999). Stock price volatility could also arise from time-varying forecasts of the occurrence or impact of rare events (Gabaix, 2012; Wachter, 2013).

These models of fully rational investors have considerable appeal. They hold out the promise that asset pricing puzzles can be solved in a world that is complicated and unpredictable, yet well-understood by investors. The careful development of these models have led to specific tests, which have not always worked in the models' favor. Problems include counterfactual predictions for the term structure of dividend claims (Binsbergen et al., 2012; Lettau and Wachter, 2007), interest rates (Backus et al., 2014) and variance risk (Dew-Becker et al., 2017). Some models do not extend well to economies where agents in aggregate can transfer resources across states and time (Lettau and Uhlig, 2000; Kaltenbrunner and Lochstoer, 2010). Finally, because these models are rational, risk premia must ultimately represent a return for bearing risk. Empirical studies have looked for this relation and failed to find it (Duffee, 2005; Moreira and Muir, 2017).[1]

This paper proposes a model for stock return volatility that does not assume rational investors with full information. This is not the same as assuming investors are irrational, it simply means that investors have a biased prior on the data generating process. Inspired by

[1]Models with time-varying rare events would seem to hold out the best hope, among rational models, for disentangling the relation between risk and return. However, requiring infinitely precise knowledge of a difficult-to-measure time-varying quantity does not seem like a victory.

the literature on behavioral finance (Barberis et al., 2003; Shiller, 2003; Hirshleifer, 2015), we motivate beliefs based on psychological studies. We take as motivation the classic animal learning study of Skinner (1948). In Skinner's study, hungry pigeons were presented food at regular intervals. Most of the pigeons developed bizarre habits of behavior, the reason for which is that they happened to have displayed that specific behavior when the food was offered.

What do these pigeons have to do with investors? While the pigeons' associations between behavior and food may seem ridiculous, their behavior illustrates a tendency to create structure out of randomness. The strong tendency to find structure where none exists characterizes human subjects as well, both in the laboratory and real-world situations (Bar-Hillel and Wagenaar, 1991). It persists even when subjects are trained to know what is random and what is not (Neuringer, 1986).

In our base case, we assume (for simplicity) that investors are risk-neutral. They believe they can forecast dividend growth using a persistent signal, though dividend growth is in fact iid. Like the pigeons they believe events can be forecasted (dividend growth, as opposed to food) even when they are completely random. We show that this condition itself is sufficient to generate excess volatility and return predictability seen in the data. Prices embed the incorrect beliefs about dividend growth, and thus are excessively volatile. Moreover, prices revert to more correct values as the expected growth fails to materialize, generating excess returns that appear to vary over time. However, in this risk-neutral environment, true risk premia are always equal to zero.

A slightly generalized model can produce an unconditional equity premium assuming investors have time-additive CRRA utility and rare disasters that occur with constant probability(Barro, 2006; Rietz, 1988). In such a model, there is no time series relation between risk and return. Moreover, time-additive CRRA utility implies flat average term structures of equity and interest rates, rather than a counterfactual upward-sloping term structure of equities and a downward-sloping term structure of interest rates.

Finally, we extend the model to address other asset pricing puzzles. A longstanding puzzle is the high abnormal returns on value stocks (Fama and French, 1992). We show that the same mechanism that explains excess volatility in the time series can explain this abnormal performance. We show that a belief in an excessive amount of interest rate predictability can explain the ability of the yield spread to forecast excess returns (Campbell and Shiller, 1991). We apply the model to forecastability of exchange rates, and shows it accounts for the failure of uncovered interest rate parity and the forward premium puzzle.

Our paper relates to a literature on subjective expectations and asset pricing. This literature has proposed a number of interesting mechanisms by which investors' beliefs may fail to match reality. Early work on this subject includes Barsky and De Long (1993) and Cecchetti et al. (2000). A recent literature focuses on extrapolative expectations: agents incorrectly believe the future will look like the past. However, the difficulty with "price extrapolation" – namely the tendency, as measured in survey data, to believe positive stock market outcomes will be followed by further outcomes – is that it is hard to incorporate into a quantitative asset pricing model. One strand of the literature focuses on counterfactual CARA utility with normally distributed returns (Fuster et al., 2010; Barberis et al., 2015), or lead to excess consumption volatility (Adam et al., 2017).[2] A second strand assumes "fundamentals extrapolation" (Alti and Tetlock, 2014; Hirshleifer et al., 2015), which is easier to incorporate but does not have a basis in survey evidence and appears to require a long-run-risk type mechanism (Bansal and Yaron, 2004) for quantitative significance.

Related to the idea of fundamentals extrapolation is the difficulty that agents have in learning about the true data generating process.[3] Learning on its own can produce excess volatility, but effects eventually dissipate (Timmermann, 1993; Veronesi, 1999; Lewellen and Shanken, 2002). The literature has proposed several ways in which learning might be persistent or impaired. These include ambiguity aversion (Hansen and Sargent, 2010; Bidder and

[2]See Jin and Sui (2018) for a recent exception.
[3]See Pastor and Veronesi (2009) for a survey on learning and asset prices.

Dew-Becker, 2016), overconfidence (Scheinkman and Xiong, 2003; Dumas et al., 2009; Daniel and Hirshleifer, 2015), deviations from Bayesian updating (Nagel and Xu, 2018), and a persistent failure of the data to match the likelihood that the investors assume (Jagannathan and Liu, 2019). Our paper, which takes as an underlying assumption that investors believe in persistent dividend growth when no such persistence exists, is in principle consistent with a wide range of psychological theories of bias or impaired learning.[4] In fact, the precise form of this belief in persistence need not be specified. In contrast to models of fundamentals-based extrapolation, which have no direct support in survey evidence, the mechanism in our paper is supported by survey evidence, as shown by de la O and Myers (2018). Specifically, de la O and Myers (2018) show that valuations are nearly entirely driven by beliefs, as opposed to discount rates, just as our model implies.

Our further contribution relative to this literature is to quantitatively explain a number of seemingly unrelated asset pricing anomalies (e.g. stock return predictability, stock return volatility, the value premium, the success of the value-minus-growth factor, the failure of the expectations hypothesis and of uncovered interest rate parity) with a single behavioral mechanism that has a long-established psychological foundation. As noted above, our mechanism is consistent with many behavioral models that have already been proposed. What matters is that that the expectation varies over time and over assets, indicating that investors think they know more than they in fact do. Regardless of the form of the expectation, it is embedded into the asset price, creating the myriad of anomalies described above, which together are very difficult to explain in a fully rational model.

The remainder of this paper is organized as follows. In Section 2, we present our benchmark model for the aggregate stock market, which explains stock market volatility. Section 3 explores the quantitative implications of this model. Section 4 shows that the beliefs in

[4]One advantage our framework has over several previous ones is that we do not require a belief in over-persistence of consumption growth. Persistence in consumption growth is harder to find in the data than persistence in dividend growth, and hence this belief is less reasonable. Moreover, the iid nature of consumption growth in our model greatly simplifies the analysis.

the model could arise from a Bayesian statistician computing the predictability in dividend growth. Section 5 extends the model to one of the cross-section of returns, the term structure of interest rates, and the carry trade. Section 6 concludes.

# 2 Model

In this section, we describe our benchmark model in two stages. Section 2.1 focuses on pricing and returns in the case of a risk-neutral investor. Section 2.2 considers the case of a risk averse investor who faces a constant probability of rare disaster.[5] The implications for volatility are virtually the same in the two cases, though the latter case also accounts for the equity premium.

## 2.1 Risk-neutral model

Consider an infinite-horizon discrete time economy with risk-neutral investors. Let $D_t$ denote the aggregate dividend at time $t$, and $d_t = \log D_t$. Assume that investors believe

$$\Delta d_{t+1} = x_t + u_{t+1}, \tag{1}$$

where

$$x_{t+1} = \phi x_t + v_{t+1}, \tag{2}$$

and

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right) \tag{3}$$

[5]While we could also consider the rare disaster case with risk neutrality, the impact on pricing would be minimal.

Assume $0 < \phi < 1$, so that dividend growth is stationary and positively autocorrelated. The assumption that realized and expected dividends are uncorrelated is for convenience.[6]

Under risk neutrality and assuming a discount factor $\delta$, the absence of arbitrage implies that the value today of a dividend paid an integer $n \geq 0$ periods in the future is

$$P_{nt} = E_t^*[\delta^n D_{t+n}], \tag{4}$$

where we use the notation $E^*$ to denote the expectations of investors. The law of iterated expectations then implies the following recursion for (4):

$$P_{nt} = E_t^*[\delta P_{n-1,t+1}], \quad n \geq 1, \tag{5}$$

with boundary condition $P_{0t} = D_t$. The asset priced in (4) is an "equity strip" (see Lettau and Wachter (2007)), analogous to a zero-coupon bond.

Equations (1–3) define a Markov structure for dividend growth, so if we divide both sides of (4) by $D_t$, we obtain a function of $x_t$. Let

$$F_n(x_t) = \frac{P_{nt}}{D_t}. \tag{6}$$

The recursion (5) pins down the functions $F_n(\cdot)$:

$$F_n(x_t) = E_t^*\left[\delta F_{n-1}(x_{t+1})\frac{D_{t+1}}{D_t}\right] \tag{7}$$

with boundary condition $F_0(x_t) = 1$. The solution is

$$F_n(x_t) = e^{a_n + b_n x_t}, \tag{8}$$

[6]Dividend data alone is not sufficient to identify the correlation in (3).

where the coefficients are defined recursively as

$$a_n = a_{n-1} + \frac{1}{2}b_{n-1}^2\sigma_v^2 + \frac{1}{2}\sigma_u^2 + \log\delta$$
$$b_n = b_{n-1}\phi + 1. \tag{9}$$

with boundary conditions $a_0 = b_0 = 0$. The recursion for $b_n$ has the well-known solution

$$b_n = \frac{1-\phi^n}{1-\phi}. \tag{10}$$

The price-dividend ratio on the aggregate stock market is a sum of these claims:

$$\frac{P_t}{D_t} = \sum_{n=1}^{\infty}\frac{P_{nt}}{D_t} = \sum_{n=1}^{\infty}F_n(x_t). \tag{11}$$

The return on the equity strip with maturity $n$ equals

$$\begin{aligned} 1 + R_{n,t+1} &= \frac{P_{n-1,t+1}}{P_{nt}} \\ &= \frac{F_{n-1}(x_{t+1})}{F_n(x_t)}\frac{D_{t+1}}{D_t}. \end{aligned} \tag{12}$$

Defining $R_t^m$ as the net return on the aggregate market and $R_{n,t+1}$ as that on the $n$-period equity strip, it follows from (11) that

$$R_{t+1}^m \equiv \frac{P_{t+1} - P_t + D_{t+1}}{P_t} = \sum_{n=1}^{\infty}\left(\frac{P_{nt}}{\sum_{k=1}^{\infty}P_{kt}}\right)R_{n,t+1}. \tag{13}$$

Namely, the market return is a weighted average of the returns on the equity strips.[7] The

[7] The intermediate steps in this calculation are as follows:

$$R_{t+1}^m = \frac{\sum_{n=1}^{\infty}P_{n,t+1} + D_{t+1}}{\sum_{n=1}^{\infty}P_{nt}} - 1 = \frac{\sum_{n=1}^{\infty}P_{n-1,t+1}}{\sum_{n=1}^{\infty}P_{nt}} - 1 = \sum_{n=1}^{\infty}\left(\frac{P_{nt}}{\sum_{k=1}^{\infty}P_{kt}}\right)\left(\frac{P_{n-1,t+1}}{P_{nt}} - 1\right).$$

weights depend on the value of $x_t$ (an increase in $x_t$ shifts the weight toward high-maturity claims), but this effect is second-order under our distributional assumptions. It will also be useful, in what follows, to note that an alternative characterization of prices and of $F_n(x_t)$, following directly from the recursion (7), is

$$\frac{P_{nt}}{D_t} = F_n(x_t) = E_t^* \left[ \delta^n e^{\sum_{s=1}^{n} \Delta d_{t+s}} \right] \tag{14}$$

In the remainder of this section, we give intuition for why this model can easily describe return patterns. We focus on the returns on an equity strip (12) because it makes the calculations easier and, because of (13), the intuition carries over to the market. However, in the subsequent section, we report quantitative implications for the returns on the aggregate market (11).

Suppose first that the investor's beliefs match reality, so that (1–3) represent the physical process for dividends. Substituting (1) and (8) into (12), we find

$$\begin{aligned} \log(1 + R^*_{n,t+1}) &= a_{n-1} - a_n + b_{n-1}x_{t+1} - b_n x_t + x_t + u_{t+1} \\ &= a_{n-1} - a_n + (b_{n-1}\phi - b_n + 1)x_t + b_{n-1}v_{t+1} + u_{t+1}, \end{aligned}$$

where we use $R^*$ to denote returns when the physical distribution matches the subjective one. Substituting from (9) implies that

$$\log(1 + R^*_{n,t+1}) = a_{n-1} - a_n + b_{n-1}v_{t+1} + u_{t+1}. \tag{15}$$

When dividend growth is in fact predictable, returns are iid. Prices incorporate all available information, and so any innovation to returns must come from an innovation to expected dividend growth represented by $v_{t+1}$, or an innovation to dividend growth itself, represented by $u_{t+1}$. Furthermore, (9) implies $E^*[R^*_t] = \delta^{-1}$, namely there is zero risk premium, as must be the case because investors are risk neutral.

Assume however, that investors' beliefs do not match reality. The physical process for dividends is not (1–3), but rather

$$\Delta d_{t+1} = u_{t+1}. \tag{16}$$

For simplicity, we assume investors are correct about the evolution of the state variable $x_t$, namely (2) represents the physical process. Additional effects could arise from incorrect beliefs concerning the persistence of $x_t$. For simplicity, we do not consider these here.[8] The contrast between (1) and (16) is what we mean by superstition in this paper. Dividend growth is simply white noise. However, investors believe it is partially forecastable. It may be forecastable based on previous prices, previous dividends, or on something else entirely. Regardless of what is driving $x_t$, this forecastability will lead to the observed dynamics in the data.

We can clearly illustrate this intuition by computing returns. First note that prices reflect agents' (incorrect) beliefs and are given by (8) and (9). These prices are identical under both correct and incorrect beliefs and, because they accurately represent some form of beliefs, are arbitrage-free. However, consider returns:

$$\begin{aligned}
\log(1+R_{n,t+1}) &= \log\left(\frac{F_{n-1}(x_{t+1})}{F_n(x_t)}\frac{D_{t+1}}{D_t}\right) && (17)\\
&= a_{n-1} - a_n + b_{n-1}x_{t+1} - b_n x_t + u_{t+1} && (18)\\
&= a_{n-1} - a_n + b_{n-1}(\phi x_t + v_{t+1}) - b_n x_t + u_{t+1}. && (19)
\end{aligned}$$

[8]Cochrane (2008) argues that dividend growth is in fact unpredictable. The strength of the predictability in the data depends on how dividends are measured, a point made by van Binsbergen and Koijen (2010), Larrain and Yogo (2008). Dividend predictability, to the extent it exists, appears to be transient (Lettau and Ludvigson, 2005; Li and Wang, 2018). While we focus, for clarity, on the case in which investors believe there is no persistence in dividend growth, what matters for our mechanism is that investors overestimate the persistence of expected dividend growth.

Thus,

$$\log(1 + R_{n,t+1}) = a_{n-1} - a_n - x_t + b_{n-1}v_{t+1} + u_{t+1}, \tag{20}$$

and, under the physical expectation,

$$\log E_t \left[1 + R_{n,t+1}\right] = -\log \delta - x_t.$$

Unlike the case where investors' beliefs are correct (see Eq.15) excess returns are predictable. When $x_t$ is high, prices are high and future returns are low.

Equation (20) shows that superstition on the part of investors leads to return predictability. It also leads to return volatility. It is again useful to contrast superstition with rational (i.e. correct) beliefs. When the physical and subjective distributions coincide,

$$\mathrm{Var}(\log(1 + R^*_{nt})) = b^2_{n-1}\sigma^2_v + \sigma^2_u, \tag{21}$$

whereas

$$\mathrm{Var}(\log(1 + R_{nt})) = \sigma^2_x + b^2_{n-1}\sigma^2_v + \sigma^2_u, \tag{22}$$

where

$$\sigma^2_x \equiv \frac{\sigma^2_v}{1 - \phi^2}.$$

At first glance, it appears that return volatility arises from the term $\sigma^2_x$, because this is the source of predictability. Also, this is missing in the case of rationality. However, the link between superstition and volatility is more subtle. In fact, almost all of the volatility arises, in both cases, from the $\sigma^2_v$ term: as discussed in the next paragraph, this term is an order of magnitude bigger than the others. It appears in both the rational and superstition cases, and in both cases it represents changes in investors subjective expectations about dividend growth. In the rational case, however, these expectations coincide with the true distribution. In the case with superstition, it will appear, ex post, as a time-varying discount rate. Volatility is similar in both cases; in one case it is accompanied by predictable dividends (counterfactually) whereas in the other it is accompanied by predictable returns.

We now return to the question of the volatility decomposition in (22). In the paragraph above, we claimed that nearly all the volatility in returns arises from the volatility in expected dividends, as represented by $b_{n-1}^2\sigma_v^2$. We now explain why this is so. First note that $\sigma_u^2$ is the volatility of realized dividends. This $0.07^2$ per annum in postwar data. On the other hand, the volatility of shocks to $x_t$, $\sigma_v$, and the unconditional volatility of $x_t$, $\sigma_x$, are unobserved. To understand the magnitude of the remaining terms, we turn to the prices of dividend claims, normalized by current dividends. These are denoted by $F_n(x_t)$ and given in (8) and (9).

Recall that the price-dividend ratio on the market is a sum of these component price-dividend ratios. Furthermore, even if the persistence $\phi$ is high, decay is geometric, and so for $n$ sufficiently large, $b_n \approx (1-\phi)^{-1}$. If we let $\sigma_{pd}^2$ be the variance of the log price-dividend ratio on the market, roughly speaking,[9]

$$\sigma_{pd}^2 \equiv \lim_{n\to\infty} \mathrm{Var}(\log F_n(x_t)) = \frac{\sigma_x^2}{(1-\phi)^2}$$

Then, for long-maturity equity strips (which, due to the properties of geometric decay, best represents the return on the market) the decomposition (22) takes the form

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Var}(\log(1+R_{nt})) &= \sigma_x^2 + \frac{\sigma_v^2}{(1-\phi)^2} + \sigma_u^2 \\ &\approx (1-\phi)^2\sigma_{pd}^2 + (1-\phi^2)\sigma_{pd}^2 + \sigma_u^2. \end{aligned} \tag{23}$$

[9]Note that the log price-dividend ratio equals

$$pd = \log \sum_{n=1}^{\infty} F_n(x_t) \approx \sum_{n=1}^{\infty} a_n + b_n x_t = a^* + b^* x_t.$$

Because of geometric decay, $b^* \approx (1-\phi)^{-1}$.

While $\sigma_u \approx 0.07$, $\sigma_{pd} \approx 0.42$. The persistence $\phi$ will equal the persistence of the price-dividend ratio. At $\phi = 0.92$, the first term in (23) equals $(0.08 \times 0.42)^2$, whereas the second term equals $(0.39 \times 0.42)^2$. The second term, representing the effect of innovations to $x_t$ is thus roughly 25 times larger than the term representing $x_t$ itself, and roughly 5 times larger than the term representing dividend volatility.[10] Finally note that these terms add up to $(0.18)^2$, thus (roughly) accounting for the annual volatility in stock returns.

This accounting exercise suggests that this simple model can explain return volatility, predictability in excess returns, together with the lack of predictability in dividends. As yet, it has nothing to say about the equity premium. Below, we address this lack, and perform a more formal calibration exercise.

## 2.2 Model with IID Disasters

We now show that a realistic equity premium can be incorporated into the model above. Assume a representative agent who maximizes a time-additive utility function with constant relative risk aversion:

$$E \sum_{t=0}^{\infty} \delta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma},$$

where $\gamma$ is relative risk aversion and $\delta$ remains the time discount factor. The agent holds the following beliefs about the consumption and dividend growth processes:

$$\Delta c_{t+1} = \mu + u_{t+1} + w_{t+1}, \tag{24}$$

$$\Delta d_{t+1} = \mu + x_t + u_{t+1} + w_{t+1}, \tag{25}$$

[10] This will also be true in a rational model with prices driven by discount rate variation. Most of the variation in realized returns comes from *innovations* in the discount rate, which are unpredictable. Very little comes from the variation in the discount rate itself.

where $x_t$ is as in (2) above, with shocks $u_{t+1}$ and $v_{t+1}$ distributed as in (3). We further assume, following Barro (2006), that

$$w_t \overset{iid}{\sim} \begin{cases} \xi & \text{probability } = p \\ 0 & \text{probability } = 1-p \end{cases} \tag{26}$$

where $\xi$ is a constant and $w_t$ is independent of $u_t$ and $v_t$.[11]

In equilibrium, the aggregate market and the riskfree rate are priced using the representative investor's Euler equation. That is, if we let $P_{nt}$ be the price of an $n$ period ahead equity strip, then $P_{nt}$ satisfies the recursion

$$P_{nt} = E_t^* \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} P_{n-1,t+1} \right],$$

where $E^*$ denote expectations taken with respect to the subjective distribution, and where $P_{0t} = D_t$. Defining $F_n(x_t) = P_{nt}/D_t$, as in the previous section, we have

$$F_n(x_t) = E_t^* \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} F_{n-1}(x_{t+1}) \frac{D_{t+1}}{D_t} \right] \tag{27}$$

with boundary condition $F_0(x_t) = 1$. The solution is again

$$F_n(x_t) = e^{a_n + b_n x_t}, \tag{28}$$

where $a_n$ follows the modified recursion

$$a_n = a_{n-1} + \log\delta + (1-\gamma)\mu + \frac{1}{2} b_{n-1}^2 \sigma_v^2 + \frac{1}{2}(1-\gamma)^2 \sigma_u^2 + \log(p e^{(1-\gamma)\xi} + (1-p)) \tag{29}$$

[11] Given the process for consumption and dividends, the agent should be able to back out $x_t$. We assume that the agent does not do this; alternatively we could make the standard assumption that dividends contain an additional shock relative to consumption so that one cannot be perfectly inferred from the other.

with $a_0 = 0$. The recursion for $b_n$ is the same, and so $b_n = (1-\phi^n)/(1-\phi)$ still holds.

The riskfree asset is also priced using the investor's Euler equation. Let $R_f$ be the one-period riskfree rate. Then:

$$E^* \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + R_f) \right] = 1,$$

implying

$$\log(1 + R^f) = -\log\delta + \gamma\mu - \frac{1}{2}\gamma^2\sigma_u^2 - \log(pe^{-\gamma\xi} + (1-p)). \tag{30}$$

We assume that the investor has correct beliefs about the consumption distribution (24). Moreover, the investor correctly assumes that dividends are equally subject to disasters as are consumption. However, the investor believes that dividends are predictable, when in reality they are not. We parsimoniously capture these assumptions by setting the physical distribution of $\Delta d_{t+1}$ equal to $\Delta c_{t+1}$.

Defining $R_{n,t+1}$, as in the previous section, as the return on the $n$-period dividend claim:

$$\begin{aligned}
\log(1 + R_{n,t+1}) &= \log\left( \frac{F_{n-1}(x_{t+1})}{F_n(x_t)} \frac{D_{t+1}}{D_t} \right) \\
&= a_{n-1} - a_n + b_{n-1}x_{t+1} - b_n x_t + \mu + u_{t+1} + w_{t+1} \\
&= a_{n-1} - a_n + \mu - x_t + b_{n-1}v_{t+1} + u_{t+1} + w_{t+1}.
\end{aligned}$$

We therefore have, under the physical measure,

$$\log E_t\left[1 + R_{n,t+1}\right] = a_{n-1} - a_n + \mu - x_t + \frac{1}{2}b_{n-1}^2\sigma_v^2 + \frac{1}{2}\sigma_u^2 + \log(pe^{\xi} + (1-p)),$$

and, for the expected excess return under the physical measure:

$$\log E_t\left[(1+R_{n,t+1})/(1+R^f)\right] = -x_t + \gamma\sigma_u^2 + \\ \log(pe^{\xi} + (1-p)) + \log(pe^{-\gamma\xi} + (1-p)) - \log(pe^{(1-\gamma)\xi} + (1-p)).$$

For small $p$ (or, as the time interval shrinks):

$$\log E_t\left[(1+R_{n,t+1})/(1+R^f)\right] \approx -x_t + \gamma\sigma_u^2 - p(1-e^{-\gamma\xi})(1-e^{\xi}), \tag{31}$$

where we have used, e.g., $\log(pe^{\xi} + (1-p)) = \log(1 + p(e^{\xi} - 1)) \approx p(e^{\xi} - 1)$. The expected excess return has its usual unconditional component, $\gamma\sigma_u^2 - p(1-e^{-\gamma\xi})(1-e^{\xi})$, the first term of which represents the normal risk, and the second term of which represents the risk of disasters. This term captures the negative covariance between returns and marginal utility during disaster periods. These components represent a risk premium, namely a return to bearing the risk of equity, which might go down during a disaster. The first term, $x_t$, does not represent a return to bearing risk, but rather is mispricing.[12]

Note that our assumption that the agent correctly assesses disaster risk is to discipline the model. We would find nearly the same equity premium if the agent overly assessed disaster risk; i.e. was pessimistic. If the values here represent an optimistic assessment of disaster risk (namely, disasters should have occurred with probability greater than 2%), then that simply implies that we were lucky and that the equity premium is not as much of a puzzle as believed. Also, allowing the agent to believe consumption growth is forecastable would also not affect our results; however we believe this is less of a plausible assumption. As discussed above, the literature shows less predictability in consumption growth than in dividend growth. As we show below, beliefs in favor of dividend growth predictability are

[12]As described in the previous section, the variance of $x_t$ is relatively small. Thus the wedge between the unconditional expectation of (31) and the true unconditional equity premium is small as well.

reasonable (though not required) given the data.

# 3 Data and Calibration Results

## 3.1 Data

We use the value-weighted CRSP index to represent the market. We compute an annual dividend by taking monthly dividends and summing. The dividend-price-ratio of the market is the trailing one year aggregate dividend divided by the ex-dividend price of the market. We use 3-month Treasury bill returns to proxy for the riskfree rate. We use the CPI index to go from nominal returns and dividend growth to real returns and dividend growth. The full sample for this study ranges from 1927 to 2017, and the post-war subsample ranges from 1948 to 2017. All data are annual.

## 3.2 Parameter Values

Table 1 shows the parameter choices for our simulations, done at an annual frequency. We choose $\sigma_u$ to be the volatility of log real dividend growth in the data. We choose $\phi = 0.95$ to match the observed first-order autocorrelation in the log dividend price ratio in postwar data. The discount factor $\delta$, provided it is within a reasonable range and high enough to ensure convergence, has a second-order effect on the results. We choose $\delta = 0.97$, which is consistent with a low riskfree rate, and still allows for convergence of the infinite sum (11). For the risk-neutral model, the remaining parameter is $\sigma_v$, which we choose to be 0.01. This generates the correct volatility of the price-dividend ratio under risk neutrality.

For the model with disaster risk, we follow Barro (2006) and choose risk aversion $\gamma$ to be 3, the average growth rate of consumption $\mu$ to be 2%, the annual disaster probability $p$ to be 2%, and the size of the disaster to be 33%. We set the time-discount factor $\delta$ to match the average return on the riskfree asset, which we set at the average annual (real) return on three-month Treasury bills.

### 3.3 Results

We simulate 4000 samples of either 91 years of data (to represent the 1927–2017 sample) or 70 years of data (to represent the 1948 to 2017 sample). We report three types of results: the results for the risk neutral model with the longer simulation, the results for the disaster model, with the longer simulation, and the results from the disaster model for the shorter simulation, in which we consider only samples with no disasters. Reporting the risk-neutral results for the shorter simulations would be repetitive, as the only difference is in the degree of small-sample bias of some of the statistics.

Table 2 show the results for the 1927–2017 sample, and compare these to the model. This comparison confirms the informal analysis in Section 2: the model can simultaneously match the standard-deviation of returns, of the price-dividend ratio, of dividend growth. The model fits the slight negative autocorrelation of annual returns. However, the model, by construction does not fit the slight autocorrelation in dividend growth, which is 20% at an annual horizon.

Including rare disasters in the model, which account for a high equity premium and low riskfree rate, have little impact on the second moments. While there is a slight reduction in the standard deviation of the divided-price ratio (due to the duration effect; the equity premium causes a down-weighting of long-horizon claims which are the most sensitive to changes in expectations), the data value remains well-within the 10% confidence bounds. Table 3 show similar results for the postwar sample.

Table 4 reports reports from regressions of excess returns on the price-dividend ratio. The first panel replicates the well-known result that excess returns are indeed predictable by the dividend-price ratio. This result also holds in post-war data (Table 5). Coefficients are statistically significant at nearly all horizons, with $R^2$ statistics increasing from 2% to 28%.

Table 6 shows that, in contrast, dividend growth is significantly forecastable only at the 1-year horizon.[13] This forecastability is transient, in that the $R^2$ statistics do not increase with horizon. Even this short-horizon effect becomes insignificant in post-war data (Table 7).

These tables also show simulations from the model. Note that by simulating the series of the correct length under a model that captures the correlational structure of the data, we capture the source of bias in the model that is also in the data (Stambaugh, 1999). The superstitious investor model captures the correct magnitude of return predictability, and the lack of dividend growth predictability. The amount of dividend growth predictability in the data (with the exception of the shortest horizon in the 1927–2017 series) can easily be accounted for by finite-sample noise. The superstitious investor model captures the economically and statistically significant predictability. However, returns are not *too* predictable in the model; the data coefficients lie within the confidence intervals. It is not easy to take advantage of the superstitious agent because there is a sense in which he is correct.

Finally, Figure 1 shows a time series plot of the level of prices and the level of dividends, post-1926. On the figure, the level of dividends is multiplied by a constant so the average level of the series are the same. Consistent with the superstition model, but inconsistent with a model in which investors have correct beliefs, deviations from the mean of the price-dividend ratio are usually followed by adjustments in prices, rather than adjustments in dividends. This is a graphical illustration of the results in Table 4. Figure 2 shows that this effect holds more dramatically in postwar data.

# 4 A Bayesian view of dividend predictability

A possible objection to the model in Section 2 is that, over time, investors would learn that dividends are in fact unpredictable. If investors did learn the correct distribution, prices

[13]This and earlier statements of significance are under the assumption of a single test. Accounting for multiple comparisons would likely further decrease the significance of dividend growth predictability.

would remain volatile, but return predictability would dissipate. In this section, we confront the hypothesized beliefs with data. We consider an investor whose prior beliefs include the possibility of dividend growth predictability. The agent updates these beliefs given the historical time series, seen through the lens of the likelihood implied by (1–3). Our evidence speaks to the difficulty of learning the true process for dividend growth.

We assume, as in Section 2, the agent believes that dividend growth contains a predictable component. Should this predictable component exist, it follows from the reasoning in Section 2 that it should be captured by the price-dividend ratio.[14] The agent therefore considers the predictive system:

$$\Delta d_{t+1} = \beta \hat{x}_t + u_{t+1} \tag{32}$$

$$\hat{x}_{t+1} = \hat{\phi} \hat{x}_t + \hat{v}_{t+1}, \tag{33}$$

where $\hat{x}_t = p_t - d_t$, the log price-dividend ratio, and where

$$\begin{bmatrix} u_t \\ \hat{v}_t \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \hat{\sigma}_v^2 \end{bmatrix}\right). \tag{34}$$

We refer to the predictor variable as $\hat{x}_t$ in contrast to $x_t$. Up to linearization error, the assumptions in Section 2 imply that $\hat{x}$ and $x$ differ only by a scale factor, approximately equal to $1/(1-\phi)$. For convenience, we de-mean both variables.[15]

Under conditions described in Appendix A, it suffices to consider a prior on the parameters of the dividend process and the marginal likelihood for the dividend process, taking observations on $\hat{x}_t$ as given. That is, the time-series regression (32) for dividend growth is, in this case, equivalent to standard OLS in which the regressor is strictly exogenous.

[14]To the extent that the price-dividend ratio fails to capture this component, we are biased against finding dividend growth predictability, and therefore proving the beliefs to be less justifiable than otherwise.

[15]De-meaning the variables simplifies the analysis, and only affects the conclusions through a degree-of-freedom adjustment that becomes negligible as the same size grows.

We assume a prior inverse-gamma distribution for $\sigma_u^2$ and, conditional on $\sigma_u^2$, a normal distribution for the predictive coefficient $\beta$:

$$\beta \mid \sigma_u \sim N(\beta_0, g^{-1}\sigma_u^2\Lambda_0^{-1}) \quad (35)$$

$$\sigma_u^2 \sim IG(a_0, b_0). \quad (36)$$

We set parameters $a_0$ and $b_0$ so that the prior on $\sigma_u^2$ is diffuse.[16] Equation 36 implies a conjugate prior on $\beta$ (Zellner, 1996). As explained below, $\Lambda_0$ is a scale factor that will allow us to interpet $g$ as indexing the strength of the prior.

Given the priors (35) and (36), and the likelihood defined by (32–34), the agent forms a posterior. Let $\hat{\mathbf{x}}_t = \{\hat{x}_0, \ldots, \hat{x}_t\}$, namely the set of observations on $\hat{x}_s$, up to and including time $t$. Let $\mathbf{y}_t = \{\Delta d_1, \ldots \Delta d_t\}$ be the dividend growth observations up to and including time $t$. The agent calculates

$$p(\beta, \sigma_u \mid \hat{\mathbf{x}}_t, \mathbf{y}_t) \propto \mathcal{L}(\mathbf{y}_t \mid \hat{\mathbf{x}}_t, \beta, \sigma_u) p(\beta, \sigma_u), \quad (37)$$

where $p(\beta, \sigma_u)$ is the prior specified in (35) and (36) and $\mathcal{L}(\mathbf{y}_t \mid \hat{\mathbf{x}}_t, \beta, \sigma_u)$ is the likelihood of observing the dividend growth data given the predictor variable and the parameters.

We fix time $T$ as the last data point observed. We stack the observations on $\hat{x}_t$ and $\Delta d_t$ into vectors:

$$Y = \begin{bmatrix} \Delta d_1 \\ \vdots \\ \Delta d_T \end{bmatrix}, \qquad X = \begin{bmatrix} \hat{x}_0 \\ \vdots \\ \hat{x}_{T-1} \end{bmatrix}.$$

Note that the OLS estimate of $\beta$ equals

$$\hat{\beta} = (X^\top X)^{-1} X^\top Y,$$

[16]Because our focus will be on the posterior mean of $\beta$, these play no further role in our analysis.

and that (32) implies

$$Y = \beta X + U,$$

where $U \sim N(0, \sigma_u^2 I)$, and $I$ is the $T \times T$ identity matrix. It follows that the posterior (37) is given by

$$p(\beta, \sigma_u \,|, \hat{\mathbf{x}}_T, \mathbf{y}_T) \propto \sigma_u^{-n} \exp\left\{-\frac{1}{2\sigma_u}(Y - X\beta)^\top (Y - X\beta)\right\} \sigma_u^{-1} \exp\left\{-\frac{g\Lambda_0(\beta - \beta_0)^2}{2\sigma_u^2}\right\}$$

where $\propto$ means up to a proportionality factor that does not depend on $\beta$ and $\sigma_u$. Completing the square implies

$$p(\beta, \sigma_u \,|, \hat{\mathbf{x}}_T, \mathbf{y}_T) \propto \sigma_u^{-1} \exp\left\{-\frac{(X^\top X + g\Lambda_0)(\beta - \bar{\beta})^2}{2\sigma_u^2}\right\} \times p(\sigma_u \,|\, \hat{\mathbf{x}}_T, \mathbf{y}_T), \tag{38}$$

where

$$\begin{aligned} \bar{\beta} &= (g\Lambda_0 + X^\top X)^{-1}(g\Lambda_0 \beta_0 + X^\top Y) \\ &= (g\Lambda_0 + X^\top X)^{-1}(g\Lambda_0 \beta_0 + (X^\top X)\hat{\beta}), \end{aligned}$$

and where $p(\sigma_u \,|\, \hat{\mathbf{x}}_T, \mathbf{y}_T)$ is a term that does not depend on $\beta$ and is therefore the marginal posterior of $\sigma_u$ (see (Zellner, 1996, Chapter 8) for more detail). It is clear from (38) that the posterior of $\beta$ conditional on $\sigma_u$ is normal with posterior mean $\bar{\beta}$. Note also that $\bar{\beta}$ is a weighted average between the prior mean $\beta_0$ and the sample mean $\hat{\beta}$, with the weights determined by the precisions of the prior and of the sample respectively.

If we, ex post, set $\Lambda_0 = X^\top X$, then $g$ corresponds to the weight on $\beta_0$ as a percent of the weight on $\hat{\beta}$, so that $g = 0.1$ implies that the prior receives 1/10 of the weight of the sample, and $g = 0.01$ means it receives 1/100 of the weight. We set the prior mean of $\beta$ to a value consistent with the agent's beliefs in Section 2. For comparability with Tables 4–7, which show regressions on the dividend-price ratio, Figure 3 shows the negative of the posterior mean of $\beta$. We consider an informative prior, with $g = 0.10$, and a diffuse prior, with $g = 0.01$.

Figure 3 shows that the agent does indeed revise her prior beliefs, at least at first. She revises it to imply more, not less predictability of dividend growth. Indeed, from the 1930s to the 1970s, it appears that dividend growth was more predictable than later in the sample.[17] Only when nearly the full sample is used, namely around 2000, does the posterior mean converge to the sample estimate, which happens to be close to, though implying slightly more predictability than, the prior. Note that the convergence implies that the prior does not matter when the full sample is used.

Thus an agent, viewing the evidence on annual dividend growth rates in isolation, would be justified in maintaining a belief that dividend growth rates are predictable. This agent, however, is not fully rational. He incorrectly extrapolates the predictability from the one-year horizon to long horizons. Moreover, he fails to notice that excess returns are also predictable.

# 5 Extensions

Cochrane (2011) notes that predictability, both in the time series and in the cross-section, appears to be ubiquitous. He attributes this predictability to variation in discount rates across time and across assets. He notes that, within a no-arbitrage setting (like the current paper), discount-rate based explanations of phenomena and belief-based explanations are isomorphic (see Harrison and Kreps (1979)). However, the fact that one can be mapped into the other does not necessarily make them equally good explanations, as the discount-rate equivalent of a belief-based model might be complicated (and likewise, for a belief-based equivalent of an discount-rate explanation). Furthermore, time-varying discount rates are

[17] Jagannathan and Liu (2019) also show that dividend growth predictability features striking instability over the sample, declining after 1970.

ideally viewed as an endogenous outcome of an economic model. In most models, time-varying discount rates are tied to time-varying risk, providing testable implications discussed in the introduction. Discount rates might also vary because risk aversion varies, or a rare event probability varies; yet this would suggest a co-movement in measures of discount rates, which is absent in the data Lettau and Wachter (2011).[18]

On the other hand, if investors display superstitious behavior about aggregate market dividends, it is natural to assume that this behavior could be seen in other asset classes, and would produce the kind of (ex post) predictability seen in the data. We give specific parametric examples below.

## 5.1 The value premium

Fama and French (1992) show that stocks with high ratios of book equity to market equity (value stocks) exhibit significantly higher excess returns than those with low ratios (growth stocks). Moreover, market betas line up in roughly the opposite direction of the expected returns, as do standard deviations. Thus standard risk-based stories fail to account for the observed value premium. Here, we show that a simple extension to the model presented in Section 2 naturally accounts for this finding.

As is well known, the value premium result extends to ratios of other fundamentals to price, such as earnings to price (see, e.g. Lettau and Wachter (2007)). What appears to be important is having price in the denominator and a plausible non-price scaling variable in the numerator. For our model, the most natural scaling variable is payouts, namely dividends, though these could be connected, through a standard production framework, to book value. We focus on the earnings-to-price ratio in the data because dividends are to some extent arbitrary.

[18]Differential time-varying exposure to rare events offers another potential route for unifying this evidence (Gabaix, 2012). Moreover, existing models of time-varying risk aversion or rare events do not in practice entirely break the link between first and second moments.

Assume $n$ risky assets. Let $D_{jt}$ denote the time-$t$ dividend, and $\Delta d_{jt}$ log dividend growth, for stock $j$, where $j = 1, ..., n$. Investors believe that dividend growth is predictable, as before. However, besides a component that affects all firms in the same way, there is a second component that affects firms differentially. That is,

$$\Delta d_{j,t+1} = x_t + \beta_{z,j} z_t + u_{j,t+1}, \tag{39}$$

where

$$x_{t+1} = \phi_x x_t + v_{x,t+1} \tag{40}$$

$$z_{t+1} = \phi_z z_t + v_{z,t+1}. \tag{41}$$

We assume the shocks $u_{j,t+1}$ (for $j = 1, \ldots, n$), $v_{x,t+1}$, and $v_{z,t+1}$, are normally distributed, independent of one another, and independent over time, with variances $\sigma_u^2$ $(\forall j)$, $\sigma_{vx}^2$, and $\sigma_{vz}^2$ respectively.

Equation 39 indicates that subjective expectations are driven by $x_t$ and $z_t$. Firms are affected by $x_t$ in the same way, while they are differentially affected by $z_t$. So that $x_t$ has the interpretation of expected dividend growth in the aggregate, we assume that $\sum_j \beta_{z,j} = 0$.

We assume risk-neutral investors with discount rate $\delta$. Let $P_t^j$ denote the price of stock $j$. As in Section 2,

$$P_t^j = \sum_{n=1}^{\infty} P_{n,t}^j,$$

where $P_{n,t}^j$ is the price of the $n$-period dividend strip for stock $j$. Prices $P_{n,t}^j$ satisfy a recursion analogous to (7). Conjecture that the solution takes the form

$$\frac{P_{n,t}^j}{D_{j,t}} = F_n^j(x_t, z_t) = e^{a_{j,n} + b_{x,n} x_t + \beta_{z,j} b_{z,n} z_t}. \tag{42}$$

Using a recursion analogous to (7), we find the difference equations

$$
\begin{aligned}
a_{j,n} &= a_{j,n-1} + \frac{1}{2}b_{x,n-1}^2\sigma_{vx}^2 + \frac{1}{2}\beta_{z,j}^2 b_{z,n-1}^2\sigma_{vz}^2 + \frac{1}{2}\sigma_u^2 + \log\delta \\
b_{x,n} &= b_{x,n-1}\phi_x + 1 \\
b_{z,n} &= b_{z,n-1}\phi_z + 1,
\end{aligned} \tag{43}
$$

with $P_{0,t}^j/D_{jt} = 1$ implying boundary conditions $a_{j,0} = b_{x,0} = b_{z,0} = 0$. Thus

$$
\begin{aligned}
b_{x,n} &= \frac{1-\phi_x^n}{1-\phi_x} \\
b_{z,n} &= \frac{1-\phi_z^n}{1-\phi_z}.
\end{aligned}
$$

It is also useful to note that:

$$
\frac{P_{n,t}^j}{D_{j,t}} = E_t^*\left[\delta^n e^{\sum_{s=1}^n \Delta d_{j,t+s}}\right] = \exp\{a_{j,n} + b_{x,n}x_t + \beta_{z,j}b_{z,n}z_t\}, \tag{44}
$$

with $a_{j,n}$, $b_{x,n}$, $b_{z,n}$ as above.

Equation 42 implies a cross-section of scaled-price ratios as long as there is a cross-section of exposures $\beta_{z,j}$. Following the empirical literature, we refer to stocks with high price ratios as growth and those with low price ratios as value. For example, if $z_t > 0$, then growth stocks will have high $\beta_{zj}$ and value stocks will have low $\beta_{zj}$. On the other hand, if $z_t < 0$, the reverse pattern will be the case.[19] Note that all that is required to produce a spread in price-dividend ratios is variation in the loadings $\beta_{z,j}$. Value stocks need not be pre-assigned some $\beta_{z,j}$.

We assume, as in Section 2, that dividend growth is in fact unpredictable. We define the market portfolio to be the weighted average of the individual assets. We take this

[19] We disregard for the moment the Jensen's inequality adjustments in the $a_{j,n}$ terms. In our calibration, these are small.

simple model to the data.[20] Table 8 reports means of portfolios formed on earnings-to-price ratios in postwar data, and in simulations from the model. In historical data, firms are sorted into quintiles based on earnings-to-price ratios (details can be found on Kenneth French's website). In the data, value firms (those with high earnings-to-price ratios) have high expected returns relative to growth firms. Except for the extreme value quintile, they have lower standard deviations and lower betas with respect to the market. Thus the Capital Asset Pricing Model does not explain the spread in expected returns, and abnormal returns are large.

Table 8 shows that the model can replicate the high returns on value stocks and low returns on growth stocks. True risk premia are zero in the model, however, measured risk premia are not. Thus the entire value-minus-growth return is attributable to alpha.

To understand the intuition behind this result, consider, as in Section 2, the realized returns on a dividend strip:

$$\begin{aligned}
\log(1+R^j_{n,t+1}) &= \log P^j_{n-1,t+1} - \log P^j_{n,t} \\
&= a_{j,n-1} - a_{j,n} + b_{x,n-1}x_{t+1} - b_{x,n}x_t + \beta_{z,j}(b_{z,n-1}z_{t+1} - b_{z,n}z_t) + u_{j,t+1} \\
&= a_{j,n-1} - a_{j,n} - x_t - \beta_{z,j}z_t + b_{x,n-1}v_{x,t+1} + \beta_{z,j}b_{z,n-1}v_{z,t+1} + u_{j,t+1}. \quad (45)
\end{aligned}$$

Thus the differential return between a stock $j$ and $k$ equals:

$$\begin{aligned}
&\log(1+R^j_{n,t+1}) - \log(1+R^k_{n,t+1}) \\
&\quad = (a_{j,n-1} - a_{j,n} - (a_{k,n-1} - a_{k,n})) \underbrace{-(\beta_{z,j} - \beta_{z,k})z_t}_{\text{value premium}} + \\
&\quad\quad (\beta_{z,j} - \beta_{z,k})b_{z,n-1}v_{z,t+1} + u_{j,t+1} - u_{k,t+1}. \quad (46)
\end{aligned}$$

[20] The persistences $\phi_x = \phi_z = 0.85$. The volatilities $\sigma_{vx} = \sigma_{vz} = 2.5\%$, while $\sigma_u = 20\%$. $\log\delta = -5.7\%$ so that prices converge. The loading on $z_t$, $\beta_{jz}$ ranges from -1 to 1.

To fix ideas, assume $z_t > 0$ and $\beta_{z,k} > \beta_{z,j}$. Stock $k$ is now overpriced (investors forecast high future dividend growth), relative to stock $j$. It follows from (44) that $k$ has a high price-earnings ratio (and will be identified as a growth stock), whereas $j$ has a low one (and will be identified as a value stock). When dividends are realized – and they are in fact an iid shock rather than predictable as investors believe – the growth stock experiences a negative return relative to the value stock.

In fact, one can compute the expected excess return on value-minus-growth. It is equal to

$$\log E_t\left[1 + R^j_{n,t+1}\right] - \log E_t\left[1 + R^k_{n,t+1}\right] = (\beta_{z,k} - \beta_{z,j})z_t \tag{47}$$

Because $z_t > 0$, value stocks appear to offer a premium over growth stocks. It follows that the value factor always has a positive average return.

The question remains: why might we think that differences in beliefs drive the difference in value and growth returns? Evidence in favor of a belief-driven model comes directly from analysts expectations in IBES and subsequent realizations.[21] Figure 4 compares realizations of earnings relative to earnings expectations. Also shown are previous years earnings (again, scaled by expectations). Over our sample, expectations were systematically inflated for all stocks, perhaps reflecting agency problems on the part of the analysts.[22] However, whereas value stocks were did not underperform by much (realized earnings were more than 80% of expected earnings), the realization of earnings on growth stocks was less than 60% of what analysts forecast. While a full analysis of IBES data is outside the scope of this paper, this result suggests that disappointment in growth stocks drives at least part of the value premium.

A second question is how our findings could be differentiated from alternative belief-driven

[21]Earnings realization data come from Compustat. IBES provides EPS forecasts for fiscal year 1 and fiscal year 2. We interpolate the two EPS forecasts to form a forward one-year EPS expectation. We multiple this EPS expectation by shares outstanding in IBES to form the forward one-year earnings expectation.

[22]See Bordalo et al. (2019) for a discussion.

explanations of differences in returns in the cross-section. A previous literature considered over-reaction as an explanation of the value premium (Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999). Most recently, Alti and Tetlock (2014) write down an extrapolative expectations model. Building on earlier work by La Porta (1996), Bordalo et al. (2019) propose a diagnostic-expectations model for why firms with high long-term analysists' growth forecasts in IBES subsequently underperform. They do not explicitly address the value premium, but their idea could be potentially applied to do so. Along a different line, Tsai and Wachter (2016) consider a rational expectations model of rare booms; when these fail to realize, a value premium is observed.

With the exception of Tsai and Wachter (2016), these models imply over or under-optimism for individual stocks. They are not models of a common component in expected dividend growth that investors believe is there but may not be.[23] It is this common time-series component that connects our model in this section to the model in Section 2. This common time-series component appears necessary to explain the data. As Cochrane (2011) emphasizes, the value premium is "explained" by the HML, the return on a value-minus-growth portfolio (Fama and French, 1993). On its face, this suggests a risk-based explanation. However, our model provides a natural explanation. Again consider the return differential (46) in the case of $j$ a value stock and $k$ a growth stock. We can interpret (46) as a time-series HML factor. From (45), we see that stocks such that $\beta_{z,j}$ is negative will have a positive loading on the factor, whereas stocks for which it is positive will have have a negative loading. Moreover, including the factor in a cross-sectional regression will soak up the excess return. Table 9 shows that indeed, including HML in the regression leads to zero abnormal returns in model, just as in the data.

To summarize, the model offers a simple explanation of the finding that value stocks

[23]Perhaps the models could be reconfigured in this way. As in the previous section, it is not our purpose to argue that our mechanism is substantially different from others that have been proposed, but rather it is perhaps the simplest possible implementation of a variety of mechanisms.

outperform growth stocks. Similarly to the model in Section 2, the underlying cause is that investors believe that they can predict dividend growth when in fact it is unpredictable. Data on analyst forecast errors offer direct evidence in favor of a belief-driven explanation, while time-series regressions suggest a belief-driven explanation of the form we describe.

## 5.2 Violations of the expectations hypothesis of interest rates

We now apply these ideas to the pricing of Treasury bonds. Assume that investors believe that the continuously-compounded short-term interest rate $r_t$ follows a first-order autoregressive process, so that

$$\Delta r_{t+1} = (\phi - 1)(r_t - \bar{r}) + v_{t+1} \tag{48}$$

where $\Delta r_{t+1} = r_{t+1} - r_t$, $|\phi| < 1$, $\bar{r}$ is the unconditional mean of $r_t$, and $v_{t+1} \overset{iid}{\sim} N(0, \sigma_v^2)$. Note that $\phi$ is the first-order autocorrelation of $r_t$.[24]

As with dividend growth, investors believe that changes in interest rates are more forecastable than they are in reality. That is, while (48) represent beliefs, the true process is governed by

$$\Delta r_{t+1} = (\zeta - 1)(r_t - \bar{r}) + v_{t+1}, \tag{49}$$

with

$$|\zeta - 1| < |\phi - 1|. \tag{50}$$

[24] The analysis in this section takes the short-term interest rate $r_t$ as a given. Perhaps the simplest way to micro-found variation in this rate is to consider a risk-neutral investor with discount rate $\delta$ and an exogenous inflation process $\Delta\pi_{t+1}$ such that

$$\Delta\pi_{t+1} = \bar{\pi} + z_t + u_{t+1}$$

and

$$z_{t+1} = \phi z_t + v_{t+1},$$

with $u_{t+1}$ and $v_{t+1}$ distributed as in (3). The interest rate $r_t$ then solves

$$E_t\left[\delta e^{-\Delta\pi_{t+1} + r_t}\right] = 1.$$

Under these assumptions, the analysis proceeds exactly as described.

We focus on the case where $\zeta, \phi \in [0,1]$ so that (50) implies $\zeta > \phi$. In forecasting next period's interest rate, (50) implies that investors put more weight on previous values of the interest rate than they should. Alternatively stated, interest rates are closer to a random walk (they mean revert more slowly) in the data than investors believe ($\zeta > \eta$).

We consider risk-neutral pricing for bonds. The dynamics thus far define a discrete-time Vasicek (1977) model.[25] Let $B_n(r_t)$ denote the price of the $n$-period bond as a function of the riskfree rate between periods $t$ and $t+1$. Then bond prices satisfy the recursion

$$B_n(r_t) = E_t^*\left[e^{-r_t} B_{n-1}(r_{t+1})\right], \tag{51}$$

with $B_0(r_t) = 1$ and $B_1(r_t) = e^{-r_t}$. It follows that

$$\log B_n(r_t) = -a_n - b_n r_t \tag{52}$$

with

$$\begin{aligned} a_n &= a_{n-1} + b_{n-1}(1-\phi)\bar{r} - \frac{1}{2} b_{n-1}^2 \sigma_v^2 \\ b_n &= 1 + b_{n-1}\phi \end{aligned} \tag{53}$$

and $a_0 = b_0 = 0$. Note that $a_1 = 0$ and $b_1 = 1$, so that $B_1(r_t) = e^{-r_t}$. The solution for $b_n$ is again

$$b_n = \frac{1-\phi^n}{1-\phi}. \tag{54}$$

Defining the continuously compounded yield on the $n$-period bond as

$$y_{nt} = -\frac{1}{n} \log B_n(r_t)$$

[25] A substantial literature on latent factor models strongly rejects a single-factor model in favor of multi-factor alternatives (Dai and Singleton, 2002; Duffee, 2002). Piazzesi et al. (2015) show how subjective expectations can be incorporated into a model with richer dynamics. For the purpose of illustrating our mechanism, however, this simple model suffices.

It follows from (54) that the yield spread equals

$$y_{nt} - y_{1t} = \text{constant} + \left( \frac{1}{n} \frac{1-\phi^n}{1-\phi} - 1 \right) r_t$$

(recall that $y_{1t} = r_t$). The (continuously compounded) holding period return on the $n$-period bond is given by

$$r_{n,t+1} = \log B_{n-1}(r_{t+1}) - \log B_n(r_t)$$

(note that $r_{1,t+1} = r_t$). Substituting in for (52), (54), and for the physical evolution of $r_t$, (49), we find the following equation for continuously-compounded excess returns:

$$rx_{n,t+1} = r_{n,t+1} - r_{1,t+1} = \text{constant} + (\zeta - \phi) \frac{b_{n-1}}{1-(1/n)b_n} (y_{nt} - y_{1t}) + b_{n-1} v_{t+1}.$$

When $\zeta = \phi$, we recover the equilibrium with correct beliefs in which excess returns are unpredictable. However, when $\zeta > \phi$, the yield spread will predict excess returns with a positive sign, as in the data.

The economic intuition is similar to that of predictability in equity ratios. Yields fluctuate based on forecasts of future interest rates. Relatively high values of long-term yields indicate investor forecasts of rising short-term interest rates. Short-term rates are not as predictable as investors think, and on average, when the yield spread is high, interest rates fall relative to investor's expectations. As a result, an above-average yield spread forecasts positive excess returns on bonds.

The ability of the yield spread to forecast excess bond returns was first noted in the data by Campbell and Shiller (1991). According to the expectations hypothesis of interest rates, yields on long-term bonds should reflect forecasts of future short-term interest rates.[26]

[26] There are slight differences depending on whether this hypothesis is articulated in logs or levels (Campbell, 1986).

Indeed, the recursion (51) implies

$$y_{nt} = -\frac{1}{n} \log E_t^* \left[ e^{\sum_{\tau=0}^{n-1} r_{t+\tau}} \right].$$

If investors correctly anticipate yields, then bond returns will be unpredictable. However, Campbell and Shiller (1991), Fama and Bliss (1987) and a large subsequent literature show that excess bond returns are strongly forecastable. We replicate this finding in Table 10, which reports coefficients from regressing bond returns on yield spreads using the Fama-Bliss data set for zero-coupon bonds.

As an illustrative calculation, we calibrate $\sigma_v$ and $\phi$ to jointly match the volatility and first-order autocorrelations of yields. This implies $\sigma_v = 1.5\%$ per annum and an annual autocorrelation $\zeta$ of (roughly) 0.90. Given these parameters, $\phi = 0.45$ gives us roughly the amount of predictability in the data.

Table 10 shows results from historical data and from simulating 1000 samples of length 70 years. We run the regression

$$rx_{n,t+1} = \alpha_n + \beta_n (y_{nt} - y_{1t}) + \epsilon_{t+1}$$

for zero-coupon bonds for maturities ranging from 2 to 5 years. Bond excess returns are strongly predictable in both data and model.

Though there are aspects of the data that this one factor model cannot match (for example, yield spreads are less persistent then yields themselves), it offers a simple explanation for a difficult feature of the data: namely, why investors appear to require, at some points in time, very different term premia for long-term bonds. In this model, the answer is that they do not require such premia, but rather, they do not know the correct process for interest rates. Recent empirical work has assembled direct evidence in favor for this hypothesis. Cieslak (2018) shows that survey errors are forecastable, and that the forecastable component predicts excess return on bonds, in the decreasing pattern shown in Table 10. Moreover, Cieslak (2018) and Piazzesi et al. (2015) show that, over the 1980–2010 period, which featured a decline in interest rates, survey expectations were systematically above expectations formed based on an econometric model. Consistent with (50), it appears that investors kept expecting a reversion to the mean, and were surprised, time and again, that such a reversion failed to occur.

### 5.3 Uncovered interest rate parity

While a full account of the behavior of currencies and international interest rates is far outside the scope of this paper, we offer a simple extension of the previous ideas to the forward premium anomaly, otherwise known as the failure of uncovered interest rate parity.[27]

Let $S_t$ be the exchange rate in units of foreign currency per U.S. dollar. Let $R_{t+1}$ be the nominal interest rate in the U.S. available between times $t$ and $t+1$, and $\tilde{R}_{t+1}$ be the nominal interest rate in the foreign country, denominated in that country's currency. Let $F_t$ be the forward price of the foreign currency. That is, at time $t$, one dollar can be converted into $F_t$ units of the foreign currency at time $t+1$. We consider nominal rates, and assume no risk of sovereign default, so that $R_{t+1}$ and $\tilde{R}_{t+1}$ are known at time $t$.

Risk-neutral pricing for the U.S. investor requires that expected rates of return be equal when computed with respect to the investor's probability distribution:

$$E_t^* \left[ 1 + R_{t+1} - \frac{S_t}{S_{t+1}} (1 + \tilde{R}_{t+1}) \right] = 0. \tag{55}$$

That is, returns from investing risk-free in the U.S. should be equal, in expectation, to investing in the foreign country's riskfree rate. Of course, one first needs to convert U.S.

[27] The potential for distorted beliefs to resolve exchange rate puzzles has also been noted by Gourinchas and Tornell (2004) and Burnside et al. (2011). The field of international finance offers a rich array of puzzles in which mis-specified beliefs could play a role, as shown in Dumas et al. (2017). Frankel and Froot (1987) offer survey evidence that is consistent with the explanation we propose here.

dollars into the foreign currency, and then back again in the following period.[28] Equation 55 can be rewritten as:

$$\frac{1+\tilde{R}_{t+1}}{1+R_{t+1}} = \left(E_t^*\left[\frac{S_t}{S_{t+1}}\right]\right)^{-1}. \tag{56}$$

Furthermore, no-arbitrage implies *covered* interest rate parity. That is, investing at the U.S. interest rate must equal buying the foreign currency today, investing at the foreign country's interest rate, and then converting back via a forward contract:

$$1+R_{t+1} = \frac{S_t}{F_t}(1+\tilde{R}_{t+1}). \tag{57}$$

Combining (56) and (57) implies that the so-called *forward discount* $F_t/S_t$ is related to appreciation (or depreciation) in the exchange rate via the expectation:

$$\frac{F_t}{S_t} = \left(E_t^*\left[\frac{S_t}{S_{t+1}}\right]\right)^{-1}. \tag{58}$$

That is, high forward discounts indicate investors expect appreciation of the currency.

Equations 56 and 58, each constituting uncovered interest rate parity, have been extensively tested and found to fail in the time series and in the cross section of currencies. For example, Lustig et al. (2011) sort currencies on the basis of the left-hand-side of (58) and then compute subsequent changes in exchange rates. Contrary to (58), they find no relation between a high forward discount and future appreciation of the currency. Nor is there a relation between the interest rate differential (56) and future appreciation. What they do find is a relation between the forward discount and excess returns on the foreign currency.

[28] The risk-neutral investor cares about expected returns being equated in real terms. Thus, if $\Delta\pi_{t+1}$ is log inflation between $t$ and $t+1$ in the U.S., (55) should have, inside the square brackets, $e^{\Delta-\pi_{t+1}}$. By using (55), we effectively assume (for simplicity) that U.S. inflation is uncorrelated with innovations in the foreign exchange rate.

Specifically, define the continuously-compounded excess return on the currency as

$$rx_{t+1} = \log\left(\frac{S_t}{S_{t+1}}(1+\tilde{R}_{t+1})\right) - \log(1+R_{t+1}). \tag{59}$$

Lustig et al. (2011) show sorting on the forward discount produces a large spread in excess currency returns in the next period. We show their results in the data row of Table 11. The currencies with the lowest forward discount have a subsequent excess currency return of -3%, while those with the highest have a return of 6%. However, their volatilities are approximately the same. This result parallels a time-series finding that high forward discounts (equivalently, high interest rate differentials), predict high excess returns on the currency (Backus et al., 2001).

To understand these results, we consider a very simple model for the exchange rate. Consider a set of countries indexed by $j$, $j = 1, \ldots, n$. Let $s_{jt} = \log S_{jt}$, and $\Delta s_{j,t+1} = s_{j,t+1} - s_{jt}$. Assume that

$$\Delta s_{j,t+1} = x_{jt} + \sigma_j u_{j,t+1}, \tag{60}$$

for some random variable $x_{jt}$, with $u_{j,t+1} \overset{iid}{\sim} N(0,1)$. It follows from (58) that the log of the forward discount $f_{jt} - s_{jt} = \log(F_{jt}/S_{jt})$ equals

$$f_{jt} - s_{jt} = x_{jt} - \frac{1}{2}\sigma_j^2 \tag{61}$$

Define continuously-compounded returns $\tilde{r}_{j,t+1} = \log(1+\tilde{R}_{j,t+1})$ and $r_{t+1} = \log(1+R_{t+1})$.

Consider the excess return on the foreign currency, defined in (59).

$$
\begin{aligned}
rx_{j,t+1} &= s_{jt} - s_{j,t+1} + \tilde{r}_{j,t+1} - r_{j,t+1} \\
&= -\Delta s_{j,t+1} + f_{jt} - s_{jt} && (62) \\
&= -x_{jt} - \sigma_j u_{j,t+1} + x_{jt} - \frac{1}{2}\sigma_j^2 && (63) \\
&= -\sigma_j u_{j,t+1} - \frac{1}{2}\sigma_j^2, && (64)
\end{aligned}
$$

where (62) follows from (57), and (63) imposes equality between the subjective and physical distribution, and uses (60). Thus continuously compounded excess returns are unpredictable. In the cross-section, $E[rx_{j,t+1}]$ depend only $\sigma_j^2$, a Jensen's inequality effect. On the other hand, exchange rates should be predictable by the forward discount, as is clear from comparing (61) and (60). This predictability is absent in the data.

Suppose instead that the true process for exchange rates is a random walk:

$$
\Delta s_{j,t+1} = \sigma_j u_{j,t+1}.
$$

In this case,

$$
\begin{aligned}
rx_{j,t+1} &= -\Delta s_{j,t+1} + f_{jt} - s_{jt} && (65) \\
&= -\sigma_j u_{j,t+1} + f_{jt} - s_{jt} && (66)
\end{aligned}
$$

where $u_{j,t+1}$ is an iid shock. Clearly excess returns on currencies will be forecastable, both in the time series and the cross-section, by the forward discount. Table 11 shows that indeed this is the case, and that the model replicates the magnitude of the cross-sectional relation.[29]

Note that, unlike previous results, the one-period nature of the forward-premium re-

[29] The results of Lustig et al. (2011) indicate an HML-type factor in currency premia. To capture this common factor, one could proceed as in Section 5.1 and model differential loadings on a common forecast.

gressions implies that we need not specify a process for $x_{jt}$. Bringing in term structure information, such as in Lustig et al. (2018), would help pin down such a process. We leave such extensions to future work.

# 6 Conclusion

Like the pigeons in Skinner's classic (1948) experiment, investors discover meaning in randomness. In this paper, we have shown that this simple insight has far-reaching consequences for asset pricing. An asset price is today's forecast of the future outcome of a random process, such as a company's dividend, or a country's exchange rate. Any information investors *think* they have about this future outcome will be in today's price. And yet if the process in question is not in fact forecastable, the price will adjust to meet reality, rather than reality adjusting to meet the price. We have shown, in four distinct settings, that the former is what occurs. For the aggregate stock market, prices have adjusted to meet dividends. For the cross-section of stocks, those with high prices relative to earnings see their prices fall. Long-term bond prices adjust to meet stable short-term interest rates, rather than the other way around. Forward prices of currencies adjust to meet spot prices.

A difficult and interesting question is how investors form their expectations. We have shown that, regardless of the specifics of this process, a tendency to find structure in randomness leaves a signature pattern in asset prices, one which we can observe in a strikingly consistent way.

# Appendix

## A Bayesian analysis of predictive regressions

Consider the predictive system

$$y_{t+1} = \beta_0 + \beta x_t + u_{t+1} \tag{A.1}$$

$$x_{t+1} = \phi_0 + \phi x_t + v_{t+1} \tag{A.2}$$

The agent observes $x_0, \ldots, x_T$ and $y_1, \ldots, y_T$, perhaps because $y_t$ represents a return or a growth rate (and so one observation is lost relative to $x_t$). We assume

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} \overset{iid}{\sim} N(0, \Sigma) \tag{A.3}$$

for a positive-semidefinite matrix $\Sigma$, representing the variance-covariance matrix.

Define

$$B = \begin{bmatrix} \beta_0 & \phi_0 \\ \beta & \phi \end{bmatrix}$$

and $\mathbf{x}_t, \mathbf{y}_t$, analogously to Section 4. Let $\mathcal{L}$ denote the joint likelihood of the data. The agent forms the posterior

$$p(B, \Sigma \,|\, \mathbf{x}_T, \mathbf{y}_T) \propto \mathcal{L}(\mathbf{x}_T, \mathbf{y}_T | B, \Sigma) p(B, \Sigma), \tag{A.4}$$

where $\propto$ denotes up to a factor that does not depend on $B$ and $\Sigma$.

We make the following assumptions, to reduce the problem to the one considered in Section 4.

**Assumption 1.** *The matrix $\Sigma$ is diagonal.*

**Assumption 2.** *The parameters $\beta_0, \beta_1$ and $\sigma_u$ are independent, under the prior, of $\phi_0, \phi_1$ and $\sigma_v$, where $\sigma_u^2$ is the first, and $\sigma_v^2$ the second, diagonal element of $\Sigma$.*

As shown below, these assumptions guarantee strict exogeneity of $x_t$ in relation to $y_t$. If these assumptions hold approximately, i.e. if the contemporaneous correlation between $y_{t+1}$ and $x_{t+1}$ is small, then it is likely that inference will not be strongly effected. See Wachter and Warusawitharana (2015) for the analysis when these don't hold, as is the case when $y_t$ represents stock returns.

We now show the marginal posterior for $\beta_0$, $\beta$ and $\sigma_u$ reduces to (37). Define $l(x_{t+1}, y_{t+1} \,|\, x_t, B, \Sigma)$ as the likelihood of the time-$t$ observation. Note that (A.1–A.3) imply

$$l(x_{t+1}, y_{t+1} \,|\, x_t, B, \Sigma) = l(x_{t+1}, y_{t+1} \,|\, \mathbf{x}_t, \mathbf{y}_t B, \Sigma).$$

Conditional probability calculations imply

$$\mathcal{L}(\mathbf{x}_T, \mathbf{y}_T | B, \Sigma) = \prod_{t=0}^{T-1} l(x_{t+1}, y_{t+1} \,|\, x_t, B, \Sigma) l(x_0 \,|\, B, \Sigma),$$

where we use $l(x_0 \,|\, B, \Sigma)$ to denote the likelihood of the initial observation.

Assumption 1 implies that, conditional on $x_t$ and on the parameters, $y_{t+1}$ is independent of $x_{t+1}$. We can factor $l$ as follows:

$$\begin{aligned} l(x_{t+1}, y_{t+1} \,|\, x_t, B, \Sigma) &= l(y_{t+1} \,|\, x_t, x_{t+1}, B, \Sigma) l(x_{t+1} \,|\, x_t, B, \Sigma) \\ &= l(y_{t+1} \,|\, x_t, B, \Sigma) l(x_{t+1} \,|\, x_t, B, \Sigma) \qquad \text{(A.5)} \\ &= l(y_{t+1} \,|\, x_t, \beta_0, \beta, \sigma_u) l(x_{t+1} \,|\, x_t, \phi_0, \phi_1, \sigma_v) \qquad \text{(A.6)} \end{aligned}$$

Note that (A.5) does indeed require Assumption 1. If this assumption does not hold, then realizations of $x_{t+1}$ give additional information about the shocks $u_{t+1}$. Given (A.5), (A.6) follows from the form of (A.1) and the definition of $\Sigma$.

We apply Assumption 2 and (A.6) to find the following form of the posterior:

$$p(B, \Sigma \,|\, \mathbf{x}_T, \mathbf{y}_T) \propto \prod_{t=0}^{T-1} l(y_{t+1} \,|\, x_t, \beta_0, \beta, \sigma_u) \prod_{t=0}^{T-1} l(x_{t+1} \,|\, x_t, \phi_0, \phi, \sigma_v) l(x_0 \,|\, \phi_0, \phi_1, \sigma_v) \\ \times p(\beta_0, \beta, \sigma_u) p(\phi_0, \phi, \sigma_v). \quad \text{(A.7)}$$

Furthermore,

$$p(B, \Sigma \,|\, \mathbf{x}_T, \mathbf{y}_T) = p(\beta_0, \beta, \sigma_u \,|\, \phi_0, \phi, \sigma_v, \mathbf{x}_T, \mathbf{y}_T) p(\phi_0, \phi, \sigma_v \,|\, \mathbf{x}_T, \mathbf{y}_T). \quad \text{(A.8)}$$

The right hand side of (A.7) factors into two terms, one of which depends on $(\beta_0, \beta, \sigma_u)$, and one of which depends on $(\phi_0, \phi, \sigma_v)$. Thus we can write:

$$p(\phi_0, \phi, \sigma_v \,|\, \mathbf{x}_T, \mathbf{y}_T) \propto \prod_{t=0}^{T-1} l(x_{t+1} \,|\, x_t, \phi_0, \phi, \sigma_v) l(x_0 \,|\, \phi_0, \phi_1, \sigma_v) p(\phi_0, \phi, \sigma_v),$$

and, from (A.8),

$$p(\beta_0, \beta, \sigma_u \,|\, \mathbf{x}_T, \mathbf{y}_T) \propto \prod_{t=0}^{T-1} l(y_{t+1} \,|\, x_t, \beta_0, \beta, \sigma_u) p(\beta_0, \beta, \sigma_u).$$

This proves that (37) is the correct posterior.

## References


Adam, K., Marcet, A., and Beutel, J. (2017). Stock price booms and expected capital gains. *American Economic Review*, 107(8):2352–2408.

Alti, A. and Tetlock, P. C. (2014). Biased beliefs, asset prices, and investment: A structural approach. *The Journal of Finance*, 69(1):325–361.

Backus, D., Chernov, M., and Zin, S. (2014). Sources of entropy in representative agent models. *The Journal of Finance*, 69(1):51–99.

Backus, D. K., Foresi, S., and Telmer, C. I. (2001). Affine term structure models and the forward premium anomaly. *The Journal of Finance*, 56(1):279–304.

Bansal, R. and Yaron, A. (2004). Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509.

Bar-Hillel, M. and Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics*, 12(4):428–454.

Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115(1):1–24.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, page 37.

Barberis, N., Thaler, R., G. M. Constantinides, M. H., and Stulz, R. M. (2003). *A survey of behavioral finance*, volume Volume 1, Part B, chapter 18, pages 1053–1128. Elsevier.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3):823–866.

Barsky, R. B. and De Long, J. B. (1993). Why does the stock market fluctuate? *The Quarterly Journal of Economics*, 108(2):291–311.

Bidder, R. and Dew-Becker, I. (2016). Long-run risk is the worst-case scenario. *American Economic Review*, 106(9):2494–2527.

Binsbergen, J. H. v., Brandt, M. W., and Koijen, R. S. (2012). On the timing and pricing of dividends. *American Economic Review*, 102(4):1596–1618.

Bordalo, P., Gennaioli, N., La Porta, R., and Shleifer, A. (2019). Diagnostic expectations and stock returns. Forthcoming, *Journal of Finance*.

Burnside, C., Han, B., Hirshleifer, D., and Wang, T. Y. (2011). Investor overconfidence and the forward premium puzzle. *Review of Economic Studies*, 78(2):523–558.

Calvet, L. E. and Fisher, A. J. (2007). Multifrequency news and stock returns. *Journal of Financial Economics*, 86(1):178–212.

Campbell, J. Y. (1986). A defense of traditional hypotheses about the term structure of interest rates. *Journal of Finance*, 41:183–193.

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*, 107:205–251.

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

Campbell, J. Y. and Shiller, R. J. (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies*, 58:495–514.

Cecchetti, S. G., Lam, P.-s., and Mark, N. C. (2000). Asset pricing with distorted beliefs: Are equity returns too good to be true? *American Economic Review*, 90(4):787–805.

Cieslak, A. (2018). Short-rate expectations and unexpected returns in treasury bonds. *The Review of Financial Studies*, 31(9):3265–3306.

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

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

Dai, Q. and Singleton, K. (2002). Expectations puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics*, 63:415–442.

Daniel, K. and Hirshleifer, D. (2015). Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives*, 29(4):61–88.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53(6):1839–1885.

de la O, R. and Myers, S. (2018). Subjective cash flow and discount rate expectations. Working paper, Stanford University.

Dew-Becker, I., Giglio, S., Le, A., and Rodriguez, M. (2017). The price of variance risk. *Journal of Financial Economics*, 123(2):225–250.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57:369–443.

Duffee, G. R. (2005). Time-variation in the covariance between stock returns and consumption growth. *Journal of Finance*, 60(4):1673–1712.

Dumas, B., Kurshev, A., and Uppal, R. (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *Journal of Finance*, 64(2):579 – 629.

Dumas, B., Lewis, K. K., and Osambela, E. (2017). Differences of opinion and international equity markets. *The Review of Financial Studies*, 30(3):750–800.

Fama, E. F. and Bliss, R. R. (1987). The information in long-maturity forward rates. *American Economic Review*, 77:680–692.

Fama, E. F. and French, K. R. (1992). The cross-section of expected returns. *Journal of Finance*, 47:427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on bonds and stocks. *Journal of Financial Economics*, 33:3–56.

Frankel, J. A. and Froot, K. A. (1987). Using survey data to test standard propositions regarding exchange rate expectations. *The American Economic Review*, 77(1):133.

Fuster, A., Laibson, D., and Mendel, B. (2010). Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives*, 24(4):67–84.

Gabaix, X. (2012). An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics*, 127(2):645–700.

Gourinchas, P.-O. and Tornell, A. (2004). Exchange rate puzzles and distorted beliefs. *Journal of International Economics*, 64(2):303 – 333.

Hansen, L. P. and Sargent, T. J. (2010). Fragile beliefs and the price of uncertainty. *Quantitative Economics*, 1(1):129–162.

Harrison, M. and Kreps, D. (1979). Martingales and multiperiod securities markets. *Journal of Economic Theory*, 20:381–408.

Hirshleifer, D. (2015). Behavioral finance. *Annual Review of Financial Economics*, 7(1):133–159.

Hirshleifer, D., Li, J., and Yu, J. (2015). Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics*, 76(C):87–106.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54(6):2143–2184.

Jagannathan, R. and Liu, B. (2019). Dividend dynamics, learning, and expected stock index returns. *The Journal of Finance*, 74(1):401–448.

Jin, L. J. and Sui, P. (2018). Asset pricing with return extrapolation. Working paper, California Institute of Technology.

Kaltenbrunner, G. and Lochstoer, L. A. (2010). Long-run risk through consumption smoothing. *Review of Financial Studies*, 23(8):3190–3224.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *The Journal of Finance*, 51(5):1715–1742.

Larrain, B. and Yogo, M. (2008). Does firm value move too much to be justified by subsequent changes in cash flow? *Journal of Financial Economics*, 87(1):200–226.

Lettau, M. and Ludvigson, S. C. (2005). Expected returns and expected dividend growth. *Journal of Financial Economics*, 76:583–626.

Lettau, M., Ludvigson, S. C., and Wachter, J. A. (2008). The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies*, 21:1653–1687.

Lettau, M. and Uhlig, H. (2000). Can habit formation be reconciled with business cycle facts? *Review of Economic Dynamics*, 3:79–99.

Lettau, M. and Wachter, J. A. (2007). Why is long-horizon equity less risky? a duration-based explanation of the value premium. *Journal of Finance*, 62:55–92.

Lettau, M. and Wachter, J. A. (2011). The term structures of equity and interest rates. *Journal of Financial Economics*, 101(1):90 – 113.

Lewellen, J. and Shanken, J. (2002). Learning, asset-pricing tests, and market efficiency. *Journal of Finance*, 57:1113–1145.

Li, Y. and Wang, C. (2018). Rediscover predictability: Information from the relative prices of long-term and short-term dividends. Working paper, Ohio State University.

Lustig, H., Roussanov, N., and Verdelhan, A. (2011). Common risk factors in currency markets. *Review of Financial Studies*, 24(11):3731–3777.

Lustig, H., Stathopoulos, A., and Verdelhan, A. (2018). The term structure of currency carry trade risk premia. Available at SSRN: https://ssrn.com/abstract=2340547 or http://dx.doi.org/10.2139/ssrn.2340547.

Moreira, A. and Muir, T. (2017). Volatility-managed portfolios. *The Journal of Finance*, 72(4):1611–1644.

Nagel, S. and Xu, Z. (2018). Asset pricing with fading memory. Working paper, University of Chicago and University of Michigan.

Neuringer, A. (1986). Can people behave "randomly?": The role of feedback. *Journal of Experimental Psychology*, 115(1):62–75.

Pastor, L. and Veronesi, P. (2009). Learning in financial markets. *Annual Review of Financial Economics*, 1(1):361–381.

Piazzesi, M., Salomao, J., and Schneider, M. (2015). Trend and cycle in bond premia. Working Paper, Stanford University and University of Minnesota.

Rietz, T. A. (1988). The equity risk premium: A solution. *Journal of Monetary Economics*, 22:117–131.

Scheinkman, J. and Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6):1183–1220.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436.

Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of Economic Perspectives*, 17(1):83–104.

Skinner, B. F. (1948). 'Superstition' in the pigeon. *Journal of Experimental Psychology*, 38(2):168.

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

Timmermann, A. G. (1993). How learning in financial markets generates excess volatility and predictability in stock prices. *The Quarterly Journal of Economics*, 108(4):1135–1145.

Tsai, J. and Wachter, J. A. (2016). Rare booms and disasters in a multisector endowment economy. *Review of Financial Studies*, 29(5):1113–1169.

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

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188.

Veronesi, P. (1999). Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *The Review of Financial Studies*, 12(5):33.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance*, 68(3):987–1035.

Wachter, J. A. and Warusawitharana, M. (2015). What is the chance that the equity premium varies over time? Evidence from regressions on the dividend-price ratio. *Journal of Econometrics*, 186(1):74–93.

Zellner, A. (1996). *An introduction to Bayesian inference in econometrics*. John Wiley and Sons, Inc., New York, NY.

Table 1: Parameters Used in Simulations

| Parameter | Risk Neutral | Disaster |
|---|---|---|
| Shock to realized log dividend growth $\sigma_u$ | 0.11 | 0.11 |
| Shock to expected log dividend growth $\sigma_v$ | 0.01 | 0.01 |
| Subjective persistence in expected log dividend growth $\phi$ | 0.95 | 0.95 |
| Time-discount factor$\delta$ | 0.97 | 0.95 |
| Expected dividend growth $\mu$ | 0 | 0.02 |
| Relative risk aversion $\gamma$ | 0 | 3.00 |
| Disaster probability $p$ | 0 | 0.02 |
| Disaster size $1 - e^{\xi}$ | - | 0.33 |

The table shows parameters used in the simulations. For the model with disasters, the agent has constant relative risk aversion with parameter $\gamma$. The physical distribution of aggregate consumption growth is the same as that of dividends growth and is not subject to bias. The model is simulated at an annual frequency.

Table 2: Empirical and Simulated Moments for the Aggregate Market, Full Sample

| | Data | Model: Risk Neutral | | | Model: Disaster | | |
|---|---|---|---|---|---|---|---|
| | 1927-2017 | 5 | 50 | 95 | 5 | 50 | 95 |
| $\sigma(R^m)$ | 0.20 | 0.20 | 0.23 | 0.26 | 0.17 | 0.19 | 0.22 |
| AC of $R^m$ | -0.01 | -0.19 | -0.02 | 0.14 | -0.18 | -0.01 | 0.16 |
| $\sigma(d-p)$ | 0.45 | 0.30 | 0.46 | 0.71 | 0.21 | 0.33 | 0.52 |
| AC of $d-p$ | 0.88 | 0.80 | 0.91 | 0.96 | 0.80 | 0.91 | 0.96 |
| $\sigma(\Delta d)$ | 0.11 | 0.10 | 0.11 | 0.12 | 0.10 | 0.12 | 0.14 |
| AC of $\Delta d$ | 0.19 | -0.18 | -0.01 | 0.16 | -0.19 | -0.01 | 0.16 |
| $E[R^m]$ | 0.09 | 0.00 | 0.03 | 0.06 | 0.03 | 0.06 | 0.09 |
| $E[R^f]$ | 0.01 | 0.03 | 0.03 | 0.03 | 0.01 | 0.01 | 0.01 |

We simulate 4000 samples each consisting of 91 years of data from the model with risk-neutral investors, and the model with risk-averse investors and rare disasters. The table reports moments from the 1927–2017 sample (second column), and medians, 5th percentile values, and 95th percentile values (remaining columns). $R^m$ denotes the net return on the market, $d-p$ the log dividend-price ratio,$\Delta d$ log dividend growth, and $R^f$ the riskfree rate. AC refers to the first-order autocorrelation and $\sigma(\cdot)$ the standard deviation. The model is simulated at an annual frequency.

Table 3: Empirical and Simulated Moments for the Aggregate Market, Post-war

| | Data | Model: Disaster, No Realization | | |
|---|---|---|---|---|
| | 1948-2017 | 5 | 50 | 95 |
| $\sigma(R^m)$ | 0.17 | 0.16 | 0.19 | 0.22 |
| AC of $R^m$ | -0.07 | -0.22 | -0.02 | 0.17 |
| $\sigma(d-p)$ | 0.42 | 0.19 | 0.31 | 0.50 |
| AC of $d-p$ | 0.92 | 0.76 | 0.90 | 0.96 |
| $\sigma(\Delta d)$ | 0.07 | 0.10 | 0.11 | 0.12 |
| AC of $\Delta d$ | 0.24 | -0.21 | -0.01 | 0.18 |
| $E[R^m]$ | 0.09 | 0.04 | 0.07 | 0.10 |
| $E[r_f]$ | 0.01 | 0.01 | 0.01 | 0.01 |

We simulate 4000 samples each consisting of 70 years of data from the model with rare disasters. We remove samples that contain disaster realizations. The table reports moments from the 1947–2017 sample (second column), and medians, 5th percentile values, and 95th percentile values (remaining three columns). $R^m$ denotes the net return on the market, $d-p$ the log dividend-price ratio,$\Delta d$ log dividend growth, and $R^f$ the riskfree rate. AC refers to the first-order autocorrelation and $\sigma(\cdot)$ the standard deviation. The model is simulated at an annual frequency.

Table 4: Predictability of Stock Market Excess Return, Full Sample

| | Horizon in Years | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 4 | 6 | 8 | 10 |
| Panel A: Data 1927-2017 | | | | | | |
| $\beta$ | 0.07 | 0.16 | 0.26 | 0.36 | 0.49 | 0.59 |
| $t$-stat | [1.39] | [1.95] | [2.72] | [2.71] | [2.86] | [2.75] |
| $R^2$ | 0.02 | 0.06 | 0.10 | 0.16 | 0.25 | 0.28 |
| Panel B: Risk Neutral Model | | | | | | |
| $\beta$ | 0.09 | 0.18 | 0.34 | 0.48 | 0.60 | 0.71 |
| 5th percentile | 0.02 | 0.05 | 0.09 | 0.12 | 0.15 | 0.17 |
| 95th percentile | 0.22 | 0.41 | 0.72 | 0.97 | 1.16 | 1.30 |
| $R^2$ | 0.04 | 0.07 | 0.14 | 0.20 | 0.25 | 0.29 |
| Panel C: Disaster Model | | | | | | |
| $\beta$ | 0.11 | 0.22 | 0.41 | 0.58 | 0.73 | 0.86 |
| 5th percentile | 0.03 | 0.07 | 0.12 | 0.16 | 0.20 | 0.21 |
| 95th percentile | 0.25 | 0.46 | 0.85 | 1.13 | 1.39 | 1.60 |
| $R^2$ | 0.04 | 0.08 | 0.15 | 0.21 | 0.26 | 0.30 |

This table reports predictive coefficients and $R^2$-statistics from regressions of the form

$$\sum_{i=1}^{H} r^m_{t+i} - r^f_{t+i} = \beta_0 + \beta(d_t - p_t) + \epsilon_{t+H},$$

where $r^m_{t+i} = \log(1 + R^m_{t+i})$ is the continuously-compounded aggregate market return between $t+i-1$ and $t+i$, $r^f_{t+i} = \log(1 + R^f_{t+i})$ is the continuously-compounded Treasury Bill return between $t+i-1$ and $t+i$, and $d_t - p_t = \log D_t/P_t$ is the aggregate dividend-price ratio. Panel A reports results from the 1927–2017 sample. Panel B and Panel C report medians and 5th and 95th percentile values from simulated data for predictive regressions, and medians for $R^2$-statistics as described in Table 2. For the data panel, $t$-statistics are adjusted for heteroskedasticity and autocorrelation.

Table 5: Predictability of Stock Market Excess Return, Post-war

| | Horizon in Years | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 4 | 6 | 8 | 10 |
| Panel A: Data 1948-2017 | | | | | | |
| $\beta$ | 0.10 | 0.20 | 0.28 | 0.40 | 0.51 | 0.59 |
| $t$-stat | [2.27] | [2.59] | [2.90] | [2.91] | [2.87] | [2.71] |
| $R^2$ | 0.07 | 0.13 | 0.16 | 0.22 | 0.27 | 0.31 |
| Panel B: Disaster Model No Realization | | | | | | |
| $\beta$ | 0.12 | 0.24 | 0.44 | 0.62 | 0.77 | 0.89 |
| 5th percentile | 0.03 | 0.06 | 0.11 | 0.13 | 0.16 | 0.16 |
| 95th percentile | 0.30 | 0.55 | 0.95 | 1.28 | 1.53 | 1.74 |
| $R^2$ | 0.05 | 0.09 | 0.18 | 0.24 | 0.30 | 0.34 |

This table reports predictive coefficients and $R^2$-statistics from regressions of the form

$$\sum_{i=1}^{H} r^m_{t+i} - r^f_{t+i} = \beta_0 + \beta(d_t - p_t) + \epsilon_{t+H},$$

where $r^m_{t+i} = \log(1 + R^m_{t+i})$ is the continuously-compounded aggregate market return between $t+i-1$ and $t+i$, $r^f_{t+i} = \log(1 + R^f_{t+i})$ is the continuously-compounded Treasury Bill return between $t+i-1$ and $t+i$, and $d_t - p_t = \log D_t/P_t$ is the aggregate dividend-price ratio. Panel A reports results from the 1947–2017 sample. Panel B and Panel C report medians and 5th and 95th percentile values from simulated data for predictive regressions, and medians for $R^2$-statistics as described in Table 3. For the data panel, $t$-statistics are adjusted for heteroskedasticity and autocorrelation.

Table 6: Predictability of Aggregate Dividend Growth, Full Sample

| | Horizon in Years | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 4 | 6 | 8 | 10 |
| Panel A: Data 1927-2017 | | | | | | |
| $\beta$ | -0.07 | -0.10 | -0.12 | -0.14 | -0.14 | -0.13 |
| $t$-stat | [-2.14] | [-1.54] | [-1.72] | [-1.85] | [-1.45] | [-1.18] |
| $R^2$ | 0.09 | 0.07 | 0.06 | 0.07 | 0.08 | 0.07 |
| Panel B: Risk Neutral Model | | | | | | |
| $\beta$ | 0.00 | 0.00 | 0.00 | 0.00 | -0.00 | -0.00 |
| 5th percentile | -0.04 | -0.09 | -0.17 | -0.24 | -0.31 | -0.37 |
| 95th percentile | 0.05 | 0.09 | 0.17 | 0.24 | 0.32 | 0.39 |
| $R^2$ | 0.01 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 |
| Panel C: Disaster Model | | | | | | |
| $\beta$ | -0.00 | -0.00 | -0.01 | -0.00 | -0.00 | -0.00 |
| 5th percentile | -0.07 | -0.13 | -0.26 | -0.38 | -0.48 | -0.59 |
| 95th percentile | 0.07 | 0.14 | 0.26 | 0.38 | 0.49 | 0.59 |
| $R^2$ | 0.01 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 |

This table reports predictive coefficients and $R^2$-statistics from regressions of the form

$$\sum_{i=1}^{H} \Delta d_{t+i} = \beta_0 + \beta(d_t - p_t) + \epsilon_{t+H},$$

where $\Delta d_{t+i}$ is the change in log aggregate dividends between $t+i-1$ and $t+i$ and $d_t - p_t = \log D_t/P_t$ is the aggregate dividend-price ratio. Panel A reports results from the 1927–2017 sample. Panel B and Panel C report medians and 5th and 95th percentile values from simulated data for predictive regressions, and medians for $R^2$-statistics as described in Table 2. For the data panel, $t$-statistics are adjusted for heteroskedasticity and autocorrelation.

Table 7: Predictability of Aggregate Dividend Growth, Post-war

| | Horizon in Years | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 4 | 6 | 8 | 10 |
| | Panel A: Data 1948-2017 | | | | | |
| $\beta$ | -0.01 | -0.01 | -0.04 | -0.08 | -0.09 | -0.12 |
| $t$-stat | [-0.59] | [-0.29] | [-0.72] | [-1.00] | [-0.83] | [-0.86] |
| $R^2$ | 0.01 | 0.00 | 0.01 | 0.04 | 0.05 | 0.06 |
| | Panel B: Disaster Model No Realization | | | | | |
| $\beta$ | -0.00 | -0.00 | -0.00 | 0.00 | -0.01 | -0.01 |
| 5th percentile | -0.07 | -0.15 | -0.29 | -0.41 | -0.52 | -0.61 |
| 95th percentile | 0.08 | 0.15 | 0.28 | 0.42 | 0.52 | 0.63 |
| $R^2$ | 0.01 | 0.01 | 0.03 | 0.04 | 0.05 | 0.06 |

This table reports predictive coefficients and $R^2$-statistics from regressions of the form

$$\sum_{i=1}^{H} \Delta d_{t+i} = \beta_0 + \beta(d_t - p_t) + \epsilon_{t+H},$$

where $\Delta d_{t+i}$ is the change in log aggregate dividends between $t + i - 1$ and $t + i$ and $d_t - p_t = \log D_t/P_t$ is the aggregate dividend-price ratio. Panel A reports results from the 1947–2017 sample. Panel B and Panel C report medians and 5th and 95th percentile values from simulated data for predictive regressions, and medians for $R^2$-statistics as described in Table 3. For the data panel, $t$-statistics are adjusted for heteroskedasticity and autocorrelation.

Table 8: Return Statistics for Value and Growth Portfolios

| | 1 (Low) | 2 | 3 | 4 | 5 (High) | 5 - 1 |
|---|---|---|---|---|---|---|
| | | | Panel A: Data 1952-2017 | | | |
| $E[R]$ | 6.46 | 7.61 | 8.96 | 11.34 | 13.65 | 7.19 |
| $t$-stat | [2.72] | [3.73] | [4.25] | [4.86] | [4.79] | [3.46] |
| $\sigma(R)$ | 19.29 | 16.60 | 17.13 | 18.97 | 23.17 | 16.87 |
| $\alpha$ | -2.05 | -0.05 | 1.20 | 2.96 | 3.77 | 5.82 |
| $t$-stat | [-1.99] | [-0.09] | [1.59] | [2.74] | [2.72] | [2.58] |
| $\beta_{mkt}$ | 1.03 | 0.93 | 0.94 | 1.01 | 1.19 | 0.17 |
| | | | Panel B: Model | | | |
| $E[R]$ | -0.14 | -0.14 | 0.39 | 1.37 | 2.67 | 2.83 |
| $\sigma(R)$ | 21.63 | 17.65 | 16.19 | 17.00 | 19.51 | 25.18 |
| $\alpha$ | -1.01 | -1.01 | -0.42 | 0.57 | 1.89 | 2.93 |
| $\beta_{mkt}$ | 1.07 | 1.02 | 0.99 | 0.97 | 0.95 | -0.12 |

Each year we form portfolios based on the earnings-to-price ratio and compute value-weighted portfolio returns over the subsequent year. Panel A reports the mean, standard deviation, CAPM alpha and beta with respect to the market in annual data from 1952 to 2017. Panel B reports the 50th percentiles of these statistics over 1000 simulations of length designed to match the data, each with 1000 stocks in the cross section.

Table 9: Value and Growth Performance Relative to a Two-Factor Model

| | 1 (Low) | 2 | 3 | 4 | 5 (High) |
|---|---|---|---|---|---|
| | | | Panel A: Data 1952-2017 | | |
| $\alpha$ | 0.27 | 0.08 | -0.05 | 0.95 | 0.27 |
| $t$-stat | [0.57] | [0.12] | [-0.09] | [1.47] | [0.57] |
| $\beta_{mkt}$ | 1.10 | 0.93 | 0.90 | 0.96 | 1.10 |
| $\beta_{hml}$ | -0.40 | -0.02 | 0.22 | 0.35 | 0.60 |
| | | | Panel B: Model | | |
| $\alpha$ | 0.49 | -0.30 | -0.48 | -0.18 | 0.49 |
| $\beta_{mkt}$ | 1.01 | 1.00 | 1.00 | 1.00 | 1.01 |
| $\beta_{hml}$ | -0.52 | -0.24 | 0.02 | 0.26 | 0.48 |

Each year we form portfolios based on the earnings-to-price ratio and compute value-weighted portfolio returns over the subsequent year. Panel A reports coefficients generated from the regression $r_{i,t} = \alpha + \beta_{hml} hml_t + \beta_{mkt} mkt_t$ where $r_{i,t}$ is the portfolio return in excess of the riskfree rate, $hml_t$ is the return on the 5th quintile (high) minus that of the 1st quintile (low), and $mkt_t$ is the average excess return on all 5 portfolios. Panel B reports the 50th percentiles of those coefficients over 1000 simulated samples of length designed to match the data. Data are annual, from 1952 to 2017.

Table 10: Moments of Bond Yields

| | Maturity in Years | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Panel A: Data 1952-2017 | | | | | |
| $\beta_n$ | | 1.61 | 2.13 | 2.39 | 2.50 |
| $t$-stat | | [2.92] | [3.51] | [3.81] | [3.60] |
| $\sigma(y_n)$ | 3.10 | 3.05 | 2.97 | 2.92 | 2.85 |
| $AC(y_n)$ | 0.88 | 0.90 | 0.90 | 0.91 | 0.91 |
| $\sigma(y_n - y_1)$ | | 0.33 | 0.54 | 0.69 | 0.81 |
| $AC(y_n - y_1)$ | | 0.40 | 0.46 | 0.52 | 0.55 |
| Panel B: Model | | | | | |
| $\beta_n$ | | 1.45 | 1.29 | 1.17 | 1.08 |
| $\sigma(y_n)$ | 2.80 | 2.03 | 1.54 | 1.22 | 1.00 |
| $AC(y_n)$ | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 |
| $\sigma(y_n - y_1)$ | | 0.77 | 1.26 | 1.58 | 1.80 |
| $AC(y_n - y_1)$ | | 0.85 | 0.85 | 0.85 | 0.85 |

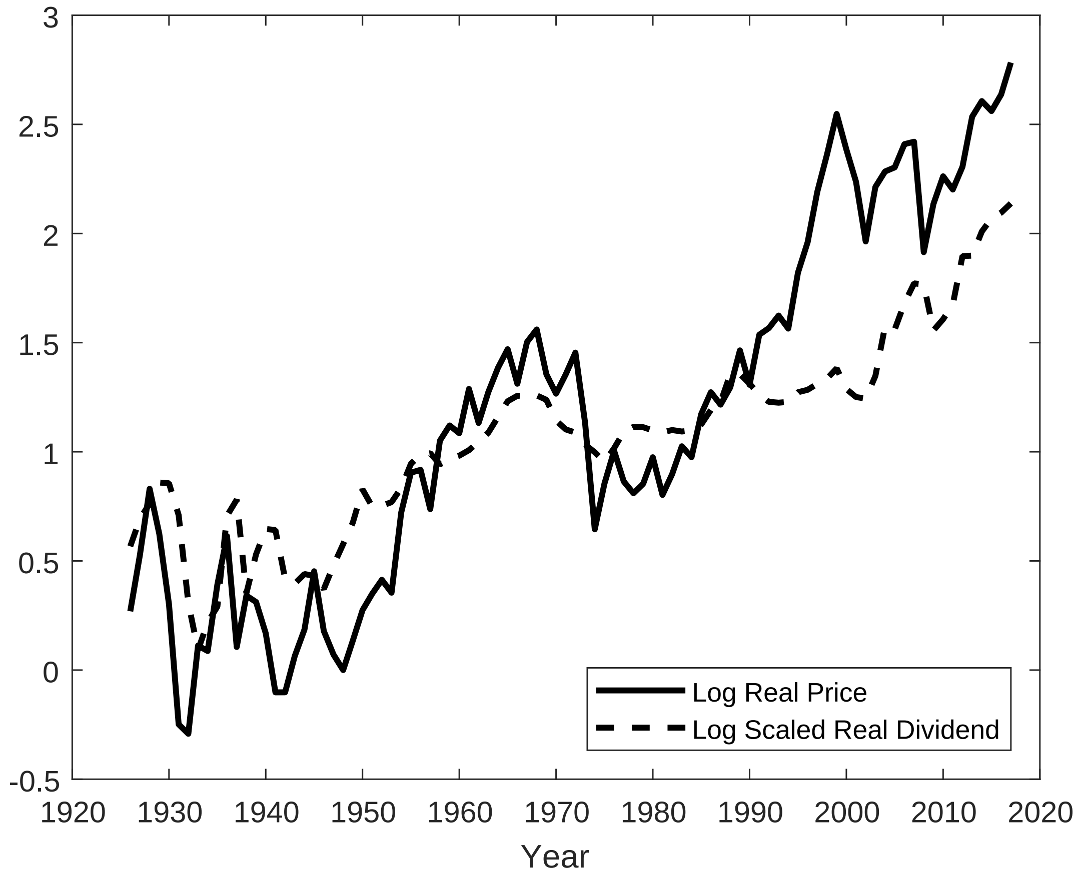
Panel A of the table reports the volatility and the first-order autocorrelation of zero-coupon bond yields and yields spread, as well as the regression coefficients $\beta_n$ as in $rx_{n,t+1} = \alpha_n + \beta_n(y_{nt} - y_{1t}) + \epsilon_{t+1}$, where $rx_{n,t+1}$ is the return of n-year bond in excess of $y_1$ over period $t+1$. The $t$-statistics adjust for heteroskedasticity. Panel B report the percentiles of those moments computed over 1000 simulations, each with 66 years of length. Data are from 1952 to 2017.

Table 11: Moments of Portfolios Sorted on Forward Discount

| | 1 (Low) | 2 | 3 | 4 | 5 | 6 (High) |
|---|---|---|---|---|---|---|
| Panel A: Data 1983-2008 | | | | | | |
| $\mu(rx)$ | -2.92 | 0.02 | 1.40 | 3.66 | 3.54 | 5.90 |
| $\sigma(rx)$ | 8.22 | 7.36 | 7.46 | 7.53 | 7.85 | 9.26 |
| $\mu(f-s)$ | -3.90 | -1.30 | -0.15 | 0.94 | 2.55 | 7.78 |
| $\sigma(f-s)$ | 1.57 | 0.49 | 0.48 | 0.53 | 0.59 | 2.09 |
| Panel B: Model | | | | | | |
| $\mu(rx)$ | -7.80 | -4.66 | -2.86 | -1.15 | 0.71 | 3.80 |
| $\sigma(rx)$ | 8.24 | 8.22 | 8.19 | 8.20 | 8.20 | 8.23 |
| $\mu(f-s)$ | -7.82 | -4.71 | -2.84 | -1.16 | 0.71 | 3.81 |
| $\sigma(f-s)$ | 1.07 | 0.85 | 0.79 | 0.79 | 0.85 | 1.07 |

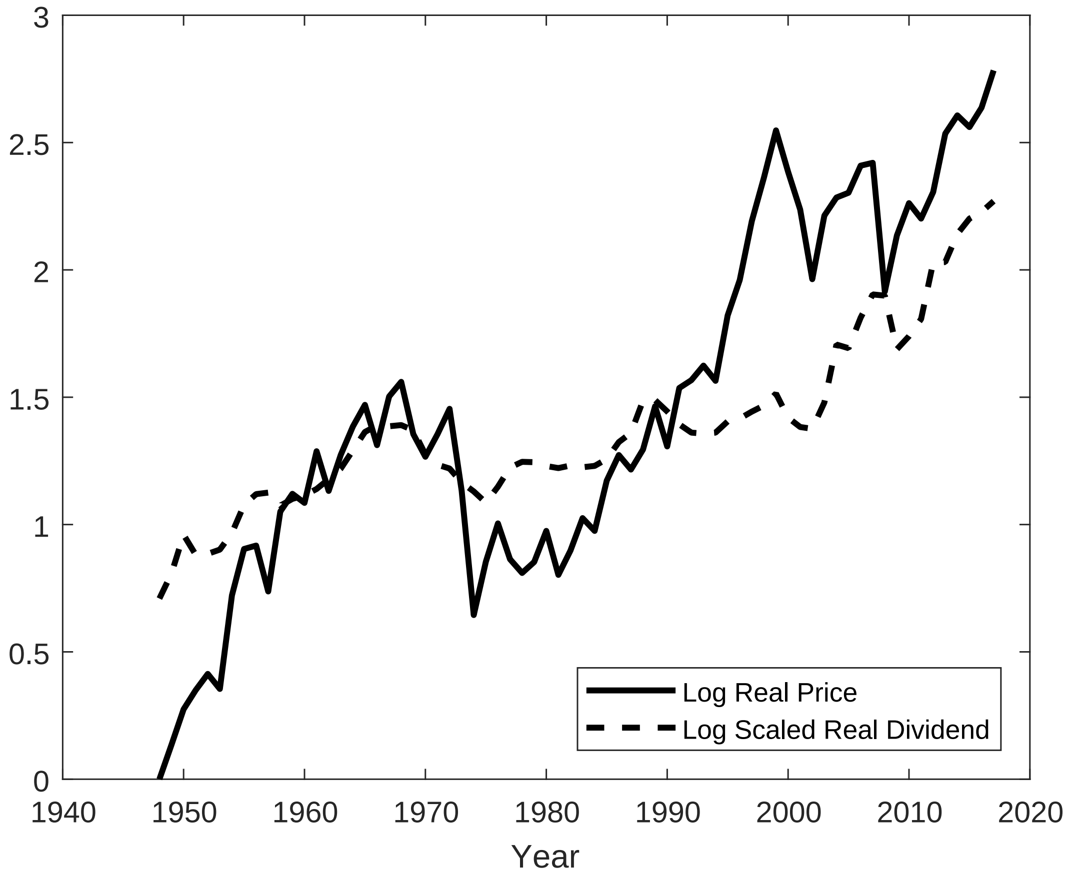
Panel A of the table reports means and standard deviations of average log excess currency returns $rx$ and log forward discount $f - s$ within each of 6 currency portfolios formed on the forward discount. Data, from Lustig et al. (2011), are monthly, from 1983–2018. Panel B reports the 50th percentiles of those moments over 1000 simulations of the model, each with 293 monthly observations.

Figure 1: Log Real Dividend and Log Real Price Level of the Aggregate Market, Full Sample



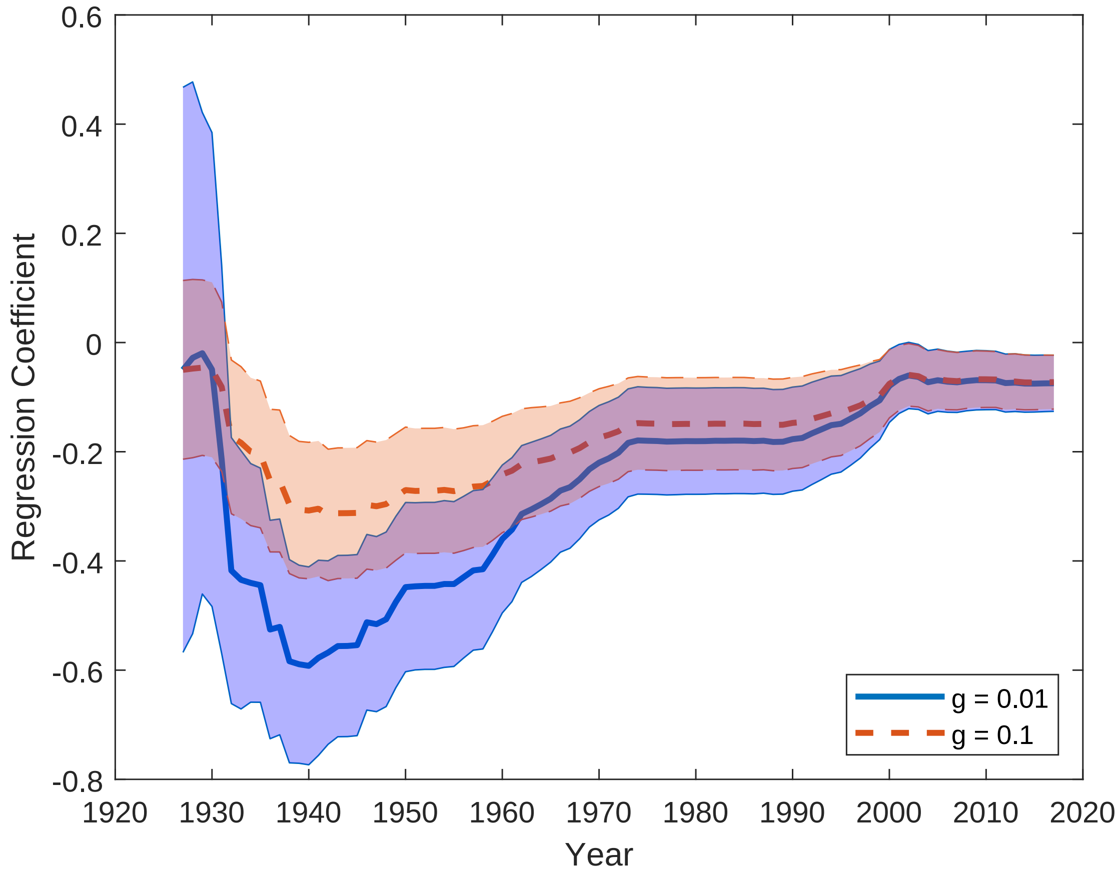

This figure plots the annual frequency log price level and log dividend of the US stock market in the post-1926 era. The log real dividend is multiplied by 27.43, the mean P/D ratio post-1926. The dividend and the price are adjusted for inflation.

Figure 2: Log Real Dividends and Log Real Price Level of the Aggregate market, Post-War



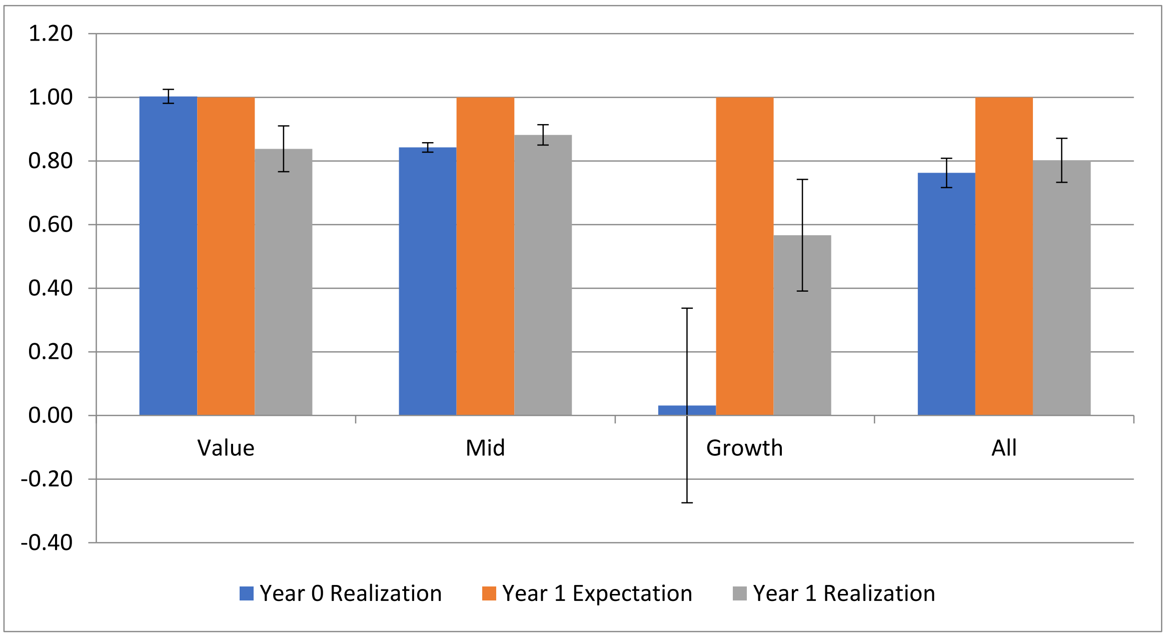

This figure plots the annual frequency log price level and log dividend of the US stock market in the post-1948 era. The log real dividend is multiplied by 27.43, the mean P/D ratio post-1948. The dividend and the price are adjusted for inflation.

Figure 3: Predicting dividend growth using the dividend-price ratio




This figure shows the posterior mean of the predictive coefficient in a regression of one-year ahead dividend growth on the dividend-price ratio. The posterior mean is calculated using Bayesian methods, assuming an informative prior, where $g$ indexes the degree of informativeness. For each year in the sample, the agent uses all available data to form a posterior for the predictive coefficient. Data begin in 1927. A prior parameter of $g = 0.1$ implies that the prior mean of the coefficient receives a weight of 10% relative to the sample estimate, whereas a prior parameter of $g = 0.01$ implies that the prior mean receives a weight of 1%. Shaded areas denote plus and minus 2 posterior standard deviations.

Figure 4: Earnings Expectations and Realizations




For each month, we sort stocks in the S&P 500 index into 3 bins based on their trailing one-year earnings-to-price ratio. For each bin, we compute trailing one-year earnings, forward one-year earnings forecasts from IBES, and earnings realization over the following year. We then compute two ratios: trailing one-year earnings over forward one-year earnings forecasts and forward one-year earnings over forward one-year earnings forecasts. We report the average ratios and standard error bars representing 2 standard errors (adjusted for heteroskedasticity and autocorrelation). Data are from 1985 to 2015.