Supplemental Appendix for “Superstitious” Investors

May 28, 2024

Contents

A Discussion of related models

A.1 Bordalo et al. (2023)

Bordalo et al. (2023) feature the following data generating process of dividend growth:

$$g_{t+1} = \mu g_t + \nu_{t+1}$$

$$\nu_{t+1} = \eta_t + \tau_{t+1}$$

The dividend growth rates' AR(1) residuals consist of η_t, the intangible news, and τ_{t+1}, the tangible news. Notice η_t is revealed one period before τ_{t+1}. The stock market excess return is[1]:

$$r_{t+1} = -\frac{1-\alpha\rho}{1-\alpha\mu}\epsilon_t + \frac{1+\alpha\mu\theta}{1-\alpha\mu}\tau_{t+1} + \frac{\alpha+\alpha\theta}{1-\alpha\mu}\eta_{t+1}$$

Here α is the AR constant in Campell-Shiller decomposition (around 0.97); θ is the parameter between 0 and 1 capturing the degree of diagnostic belief; ρ is the persistence of subjective growth; ϵ_t is an exponentially weighted moving average of a linear combination of past η and τ, multiplied by θ. The surprise in cash flow growth relative to survey expectation is:

$$\tau_{t+1} - \epsilon_t$$

In this setting, η is a key component that determines the contemporaneous correlation stock market return and cash flow surprise. If η is not disperse relative to τ, then r_{t+1}'s movements are mainly driven by $-\frac{1-\alpha\rho}{1-\alpha\mu}\epsilon_t + \frac{1+\alpha\mu\theta}{1-\alpha\mu}\tau_{t+1}$, which appears to correlate highly with survey based cash flow surprise $\tau_{t+1} - \epsilon_t$ and/or AR based cash flow surprise $g_{t+1} - \mu g_t$.

If η is very disperse relative to τ— the view of Bordalo et al. (2023) —then i) r_{t+1} is mainly driven by $-\frac{1-\alpha\rho}{1-\alpha\mu}\epsilon_t + \frac{\alpha+\alpha\theta}{1-\alpha\mu}\eta_{t+1}$, and ii) $g_{t+2} - \mu g_{t+1}$, the next period cash flow growth's AR

[1] This comes from equation (8) of Bordalo et al. (2023). There appears to be a minor typo in equation (8) itself which we correct here.

surprise, is mainly driven by η_{t+1}. At the first glance, this η component appears to function similarly to the v component in our model, in that they both drive contemporaneous return and valuation without appearing in contemporaneous cash flow growth. However, the key difference is that η is still part of the contemporaneous cash flow *shock* that drives the cash flow growth next period. Therefore, in the event that η is very disperse relative to τ, Bordalo et al. (2023) would predict a high correlation between this period's return and next period's AR-based cash glow growth surprise. In contrast, in our model, v captures random shocks that are not attached to realized cash flow. Consequently, the correlation between this period's return and next period's cash flow growth surprise is close to 0.

To summarize, it seems that Bordalo et al. (2023)'s model features a high correlation between either i) contemporaneous market return and AR/survey based cash flow surprise or ii) market return this period and AR-based cash flow surprise next period, or both. In the data, these correlations are positive, lending qualitative support to Bordalo et al. (2023). However, the economic scale and statistical strength of these correlations are not high. It is around 0.3 contemporaneously and 0.27 between this year's market return and next year's AR1 residual in cash flow growth.[2] Even with a correlation of 0.3, only 9% of the variation in market return is explained. Since Bordalo et al. (2023) do not calibrate their model's parameters, it is hard to determine these correlations in their model. But given the economic scale of these correlations in data, at least quantitatively, it seems that Bordalo et al. (2023) could really benefit from the addition of a component like our random shock v.

A.2 Nagel and Xu (2022)

The same distinction applies to Nagel and Xu (2022) as well. In the illustrative model of Nagel and Xu (2022), the realized market return in period $t+1$ is:

[2]The correlation is slightly negative between this year's market return and the AR1 residual in cash flow growth 2 years later.

$$r_{t+1} = (1 + \frac{\rho\nu}{1-\rho})(\Delta d_{t+1} - \tilde{\mu}_{d,t}) + \theta + r_f$$

Here ρ is as in the Campbell-Shiller decomposition, ν is a constant gain parameter, and $\Delta d_{t+1} - \tilde{\mu}_{d,t}$ is surprising dividend growth. This comes from equation (2). In the formal model, the intuition appears very similar (e.g. equation B.3 and D.25). Again, the weakly positive correlation between contemporaneous cash flow growth surprise and market return lends directional support to their model, but since the correlation is quantitatively small, there is likely something else going on in the variation of valuation and returns. Our view is that much of such variation comes from our random shock v.

At some level, if variation in valuation ratio is tied to cash flow shocks too tightly—whether this tie comes from diagnostic expectation or experience effect—the variation is likely insufficiently high relative to data. Notice that in Table 5 of Nagel and Xu (2022), row $\sigma(p-d)$ the number in the "Data" column is about 2 times as large as those in the "Model" columns. It seems that our model can help Nagel and Xu (2022) better match the stock market volatility.

A.3 Bordalo et al. (2019)

Bordalo et al. (2019) model individual stocks rather than the aggregate market. In their model, the data generating process for individual firm i's EPS $x_{i,t}$ is:

$$x_{i,t} = bx_{i,t-1} + f_{i,t} + \epsilon_{i,t}$$

Here the key term $f_{i,t}$ represents the firm's "fundamental", which follows:

$$f_{i,t} = \alpha f_{i,t-1} + \eta_{i,t}$$

Notice $f_{i,t}$ is a persistent component in the firm's cash flow growth, which should therefore drive valuation. The authors show that under diagnostic expectation, the investors' subjective belief of $f_{i,t}$ is:

$$\hat{f}^{\theta}_{i,t} = \alpha \hat{f}_{i,t-1} + K(1+\theta)(x_{i,t} - bx_{i,t-1} - \alpha \hat{f}_{i,t-1})$$

Here $\hat{f}_{i,t-1}$ is the rational expectation of $f_{i,t-1}$, K is rational Kalman gain parameter, and θ represents the degree of diagnostic belief.

Since $\hat{f}^{\theta}_{i,t}$ drives valuation, it seems that this framework, if applied to the aggregate market, would also imply a high correlation between contemporaneous return and surprising cash flow growth $x_t - bx_{t-1} - \alpha \hat{f}_{t-1}$. The implication on this particular aspect is then quite similar to that in Nagel and Xu (2022).

B Analysis of the source of volatility

We address the question of the volatility decomposition in (17). In the main text, we claimed that nearly all the volatility in returns arises from the volatility in expected dividends, as represented by $b^2_{n-1}\sigma^2_v$. Here we explain why this is so. First note that σ^2_u 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, σ_v, and the unconditional volatility of x_t, σ_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 $\Phi_n(x_t)$ and given in (7) and (8).

Recall that the price-dividend ratio on the market is a sum of these component price-dividend ratios. Furthermore, even if the persistence ϕ is high, decay is geometric, and so for n sufficiently large, $b_n \approx (1-\phi)^{-1}$. If we let σ^2_{pd} be the variance of the log price-dividend ratio on the market, roughly speaking,[3]

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

Then, for long-maturity equity strips (which, due to the properties of geometric decay, best represents the return on the market) the decomposition (17) takes the form

$$\begin{aligned} \lim_{n\to\infty} \mathrm{Var}(\log(1+R_{nt})) &= \sigma_x^2 + \frac{\sigma_v^2}{(1-\phi)^2} + \sigma_u^2 \\ &\approx (1-\phi)^2\sigma_{pd}^2 + (1-\phi^2)\sigma_{pd}^2 + \sigma_u^2. \end{aligned} \tag{B1}$$

While $\sigma_u \approx 0.07$, $\sigma_{pd} \approx 0.42$. The persistence ϕ will equal the persistence of the price-dividend ratio. At $\phi = 0.92$, the first term in (B1) 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.[4] Finally note that these terms add up to $(0.18)^2$, thus (roughly) accounting for the annual volatility in stock returns.

C Disaster risk model

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

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

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

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

[4]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.

relative risk aversion:

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

where γ is relative risk aversion and δ remains the time discount factor. The agent holds the following beliefs about the consumption and dividend growth processes:

$$\Delta c_{t+1}=\mu+u_{t+1}+w_{t+1}, \tag{C2}$$

$$\Delta d_{t+1}=\mu+x_t+u_{t+1}+w_{t+1}, \tag{C3}$$

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

$$w_t \overset{iid}{\sim} \begin{cases} \xi & \text{with probability } p \\ 0 & \text{with probability } = 1-p \end{cases} \tag{C4}$$

where ξ is a constant and w_t is independent of u_t and v_t.[5]

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

[5] 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.

$P_{0t} = D_t$. Defining $\Phi_n(x_t) = P_{nt}/D_t$, as in the previous section, we have

$$\Phi_n(x_t) = E_t^* \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \Phi_{n-1}(x_{t+1}) \frac{D_{t+1}}{D_t} \right] \tag{C5}$$

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

$$\Phi_n(x_t) = e^{a_n + b_n x_t}, \tag{C6}$$

where a_n follows the modified recursion

$$a_n = a_{n-1} + \log\delta + (1-\gamma)\mu + \frac{1}{2} b_{n-1}^2 \sigma_v^2 + \frac{1}{2}(1-\gamma)^2\sigma_u^2 + \log(pe^{(1-\gamma)\xi} + (1-p)) \tag{C7}$$

with $a_0 = 0$. The recursion for b_n is the same.

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

$$\log(1 + R^f) = -\log\delta + \gamma\mu - \frac{1}{2}\gamma^2\sigma_u^2 - \log(pe^{-\gamma\xi} + (1-p)). \tag{C8}$$

We assume that the investor has correct beliefs about the consumption distribution (C2). 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 Δd_{t+1} equal to Δc_{t+1}.

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

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

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):

$$\log E_t \left[(1 + R_{n,t+1})/(1 + R^f)\right] \approx -x_t + \gamma\sigma_u^2 - p(1 - e^{-\gamma\xi})(1 - e^{\xi}), \tag{C9}$$

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.[6]

Our assumption that the agent correctly assesses disaster risk is purely for parsimony. We would find nearly the same equity premium if the agent were pessimistic. If the agent

[6] As described in the previous section, the variance of x_t is relatively small. Thus the wedge between the unconditional expectation of (C9) and the true unconditional equity premium is small as well.

were optimistic (namely, disasters should have occurred with probability greater than 2%), then we were lucky and that the equity premium is not as much of a puzzle as believed. 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, in this disaster risk model, we assume that agents do not believe they can forecast consumption growth. Empirical work suggests that growth in consumption of non-durables and services is less forecastable than growth in other types of cash flows such as dividends and GDP (Cochrane, 1994; Lettau and Ludvigson, 2005; Beeler and Campbell, 2012). However, we acknowledge that it would not be unreasonable for agents who believed they could forecast dividend growth to also believe they could forecast consumption growth. This raises a possible problem. If, in the disaster risk model, agents were to believe that consumption growth was forecastable with the same mean as dividend growth, then for $\gamma > 1$, expected growth would lead to lower valuations, whereas for $\gamma = 1$ it would have no effect. While $\gamma < 1$ would produce the required effect, such knife-edge dependence on an unknown parameter (combined with the fact that $\gamma < 1$ makes the equity premium very hard to explain) is undesirable.[7]

Agents' first-order conditions states that consumption growth forecasts enter directly into the interest rate. As described in Appendix E, we obtain data on interest rate forecasts from Blue Chip Financial Forecasts. We test for a relation between earnings growth forecasts and interest rate forecasts; if we were to find a positive relation then it would be natural to assume that x_t entered into consumption growth. We would want to specify the relation between consumption and dividend growth (and perhaps preferences), to match the finding shown in

[7]The tradeoff between expected growth and the interest rate is modulated by the elasticity of intertemporal substitution (EIS). This is the inverse of γ in the model of Appendix C, a more general model would allow for a separation and thus could explain the equity premium while allowing the cash flow effect to dominate. One is still left, however, with the undesirable dependence on whether the EIS is above or below one.

Figure 5, namely that price-earnings ratios are very highly correlated with expectations.

Asset pricing models predict a link between contemporaneous consumption growth expectations and the interest rates that prevail over the interval that the consumption growth occurs (see, e.g. Campbell (2003)). To make sure that the results are not sensitive to timing, we correlate earnings forecasts each period both with the lagged interest rate forecast (what the model says we should do), and the contemporaneous interest rate forecast. Both correlations are slightly negative, which is the opposite direction from what theory would predict (Figure E1 in the Supplemental Appendix), though the result is insignificant. It appears that whatever is the primary driver of earnings forecasts is significantly separated from the aggregate economy that the assumptions of Appendix C are accurate; or alternatively whatever drives interest rates is separated from the aggregate economy (this could be understood as the model of Section 2.1 in which the interest rate is exogenous). One might view either disconnect as its own puzzle and an interesting direction for future work.

We also simulate the disaster model and match the data. We follow Barro (2006) and choose risk aversion γ to be 3, the average growth rate of consumption μ 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 δ to match the average return on the riskfree asset, which we set at the average annual (real) return on three-month Treasury bills. These parameter values are listed in Table C1. When computing statistics across simulated samples, we consider only those with no realized disasters.

Table C2 shows that 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. Similarly, Tables C3 and C4 show that the disaster model generates return and dividend growth predictability comparable to those in the risk-neutral model.

Table C1: Parameter values for the aggregate market simulation

Parameter	Disaster
Shock to realized log dividend growth σ_u	0.07
Shock to expected log dividend growth σ_v	0.01
Subjective persistence in expected log dividend growth ϕ	0.95
Time-discount factor δ	0.95
Expected dividend growth μ	0.02
Relative risk aversion γ	3.00
Disaster probability p	0.02
Disaster size $1 - e^{\xi}$	0.33

The table shows parameters used in the simulation of the disaster model for the aggregate market. The agent has constant relative risk aversion with parameter γ. 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 C2: Empirical and simulated moments for the aggregate market

	Data	Model: Disaster		
	1948-2019	5	50	95
$\sigma(R^m)$	0.17	0.14	0.16	0.19
AC of R^m	-0.09	-0.21	-0.02	0.17
$\sigma(d-p)$	0.42	0.20	0.31	0.51
AC of $d-p$	0.92	0.76	0.90	0.97
$\sigma(\Delta d)$	0.07	0.06	0.07	0.08
AC of Δd	0.24	-0.22	-0.01	0.17
$E(R^m)$	0.09	0.04	0.06	0.09
$E(r_f)$	0.01	0.03	0.03	0.03

We simulate 4000 samples each consisting of 72 years of data from the model with risk-averse investors and rare disasters. The table reports moments from the 1948–2019 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, Δd 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 C3: Predictability of stock market excess return

	Horizon in Years					
	1	2	4	6	8	10
Panel A: Data 1948-2019						
β	0.10	0.19	0.27	0.38	0.47	0.56
t-stat	[2.30]	[2.60]	[2.75]	[2.73]	[2.60]	[2.67]
R^2	0.06	0.12	0.15	0.20	0.23	0.27
Panel B: Disaster Model						
β	0.12	0.24	0.45	0.64	0.78	0.90
5th percentile	0.05	0.10	0.19	0.26	0.33	0.36
95th percentile	0.28	0.51	0.88	1.13	1.34	1.49
R^2	0.07	0.13	0.24	0.33	0.40	0.45

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 1948–2019 sample. Panel B reports 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 C4: Predictability of aggregate dividend growth

	Horizon in Years					
	1	2	4	6	8	10
			Panel A: Data 1948-2019			
β	-0.02	-0.02	-0.05	-0.09	-0.12	-0.13
t-stat	[-0.68]	[-0.38]	[-0.82]	[-1.17]	[-1.11]	[-1.03]
R^2	0.01	0.00	0.02	0.05	0.07	0.07
			Panel B: Disaster Model			
β	0.00	0.00	0.00	-0.00	-0.01	-0.01
5th percentile	-0.05	-0.09	-0.16	-0.23	-0.30	-0.38
95th percentile	0.05	0.10	0.18	0.26	0.33	0.40
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 Δd_{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 1948–2019 sample. Panel B reports 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.

D Could investors believe dividends were predictable?

A possible objection to our model is that, over time, investors would learn that dividends are in fact unpredictable. If investors did learn the correct distribution, prices 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 equation (1)-(3) in the main paper. Our evidence speaks to the difficulty of learning the true process for dividend growth.

We assume, as in our model, the agent believes that dividend growth contains a predictable component. Should this predictable component exist, it follows from the reasoning in our model that it should be captured by the price-dividend ratio.[8] The agent therefore considers the predictive system:

$$\Delta d_{t+1} = \beta \hat{x}_t + u_{t+1} \tag{D1}$$

$$\hat{x}_{t+1} = \hat{\phi} \hat{x}_t + \hat{v}_{t+1}, \tag{D2}$$

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

$$\begin{bmatrix} u_t \\ \hat{v}_t \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \hat{\sigma}_v^2 \end{bmatrix}\right). \tag{D3}$$

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.[9]

[8] 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.

[9] 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.

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 (D1) for dividend growth is, in this case, equivalent to standard OLS in which the regressor is strictly exogenous.

We assume a prior inverse-gamma distribution for σ_u^2 and, conditional on σ_u^2, a normal distribution for the predictive coefficient β:

$$\beta \,|\, \sigma_u \sim N(\beta_0, g^{-1}\sigma_u^2\Lambda_0^{-1}) \tag{D4}$$

$$\sigma_u^2 \sim IG(a_0, b_0). \tag{D5}$$

We set parameters a_0 and b_0 so that the prior on σ_u^2 is diffuse.[10] Equation D5 implies a conjugate prior on β (**?**). As explained below, Λ_0 is a scale factor that will allow us to interpet g as indexing the strength of the prior.

Given the priors (D4) and (D5), and the likelihood defined by (D1–D3), 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

$$p(\beta, \sigma_u \,|\, \hat{\mathbf{x}}_t, \mathbf{y}_t) \propto \mathcal{L}(\mathbf{y}_t \,|\, \hat{\mathbf{x}}_t, \beta, \sigma_u) p(\beta, \sigma_u), \tag{D6}$$

where $p(\beta, \sigma_u)$ is the prior specified in (D4) and (D5) and $\mathcal{L}(\mathbf{y}_t \,|\, \hat{\mathbf{x}}_t, \beta, \sigma_u)$ 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 Δd_t

[10] Because our focus will be on the posterior mean of β, these play no further role in our analysis.

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 β equals

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

and that (D1) 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 (D6) 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 β and σ_u. Completing the square implies

$$p(\beta, \sigma_u \,|, \hat{\mathbf{x}}_T, \mathbf{y}_T) \propto \sigma_u^{-1} \exp\left\{ -\frac{(X^\top X + g\Lambda_0)(\beta - \bar{\beta})^2}{2\sigma_u^2} \right\} \times p(\sigma_u \,|\, \hat{\mathbf{x}}_T, \mathbf{y}_T), \tag{D7}$$

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 β and is therefore the marginal posterior of σ_u (see (**?**, Chapter 8) for more detail). It is clear from (D7) that the posterior of β conditional on σ_u is normal with posterior mean $\bar{\beta}$. Note also that $\bar{\beta}$ is a weighted average between the prior mean β_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 β_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 β 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 D1 shows the negative of the posterior mean of β. We consider an informative prior, with $g = 0.10$, and a diffuse prior, with $g = 0.01$.

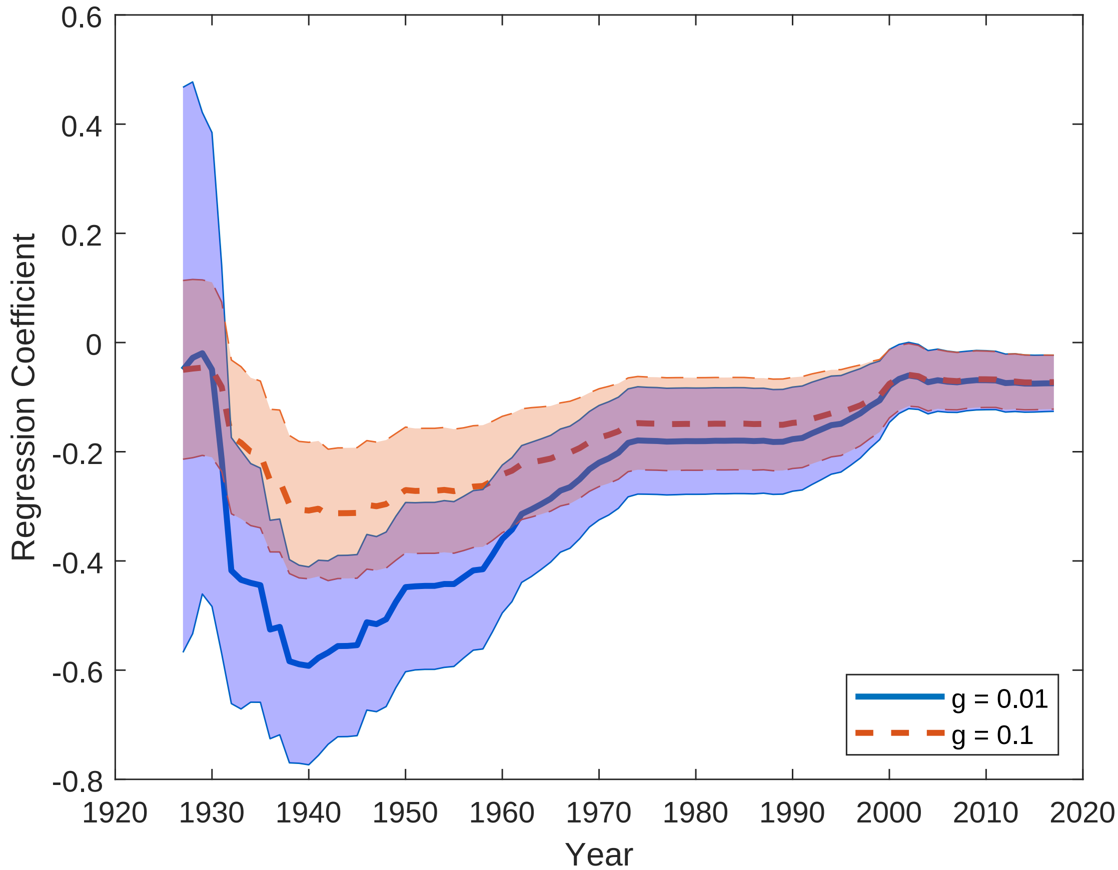
Figure D1 shows that the agent does indeed revise her prior beliefs, at least at first. She revises it to imply more, not less predictability of dividend growth. Indeed, from the 1930s to the 1970s, it appears that dividend growth was more predictable than later in the sample.[11] Only when nearly the 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. In fact, as Table D1 shows, dividend growth is predictable at short horizons. This agent, however, is not fully rational, incorrectly extrapolating the predictability from the one-year horizon to long horizons – note that Table D1 shows that R^2 statistics flatten rather than grow, as would be expected from an ARMA model. Moreover, the agent fails to notice that excess returns are also predictable.[12]

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

[12] While we do not model reinforcement learning (which is a feature of Skinner (1948)), these results suggest that the agent would have received positive reinforcement, throughout the sample, in the sense that

Figure D1: 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.

Table D1: Predictability of aggregate dividend growth, full sample

	Horizon in Years					
	1	2	4	6	8	10
Panel A: Data 1927-2019						
β	-0.08	-0.11	-0.12	-0.15	-0.17	-0.15
t-stat	[-2.18]	[-1.60]	[-1.79]	[-1.98]	[-1.69]	[-1.35]
R^2	0.10	0.08	0.06	0.08	0.10	0.08
Panel B: Risk Neutral Model						
β	0.00	0.00	0.00	0.00	0.00	0.00
5th percentile	-0.04	-0.09	-0.17	-0.24	-0.31	-0.38
95th percentile	0.04	0.09	0.17	0.25	0.32	0.39
R^2	0.01	0.01	0.02	0.03	0.04	0.05
Panel C: Disaster Model						
β	0.00	0.00	0.00	0.00	0.00	-0.00
5th percentile	-0.07	-0.13	-0.25	-0.36	-0.46	-0.56
95th percentile	0.07	0.14	0.27	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 Δd_{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 1927–2019 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.

E Predictability of Treasury bond excess returns and survey data

Campbell and Shiller (1991) show that periods of high term spreads are followed by high excess returns on Treasury bonds. The question is: are investors anticipating these high expected returns? Or, do they avoid Treasury bonds during times of high spreads because they believe (incorrectly) that interest rates are likely to rise? When such an increase fails to occur, bonds exhibit positive returns.

In this section, we first write down a model that formalizes this intuition. Investors believe that changes in interest rates are more forecastable than they are in reality. This model matches the data in that, in the model, term spreads predict returns with a positive sign.

Next, we test the model in survey data. Data from both Blue Chip Financial Forecasts, and from the Survey of Professional Forecasters confirm that high term spreads forecast rising interest rates. Moreover, as the model predicts, high term spreads correlate significantly with analyst forecasts of rising interest rates. The difficulty is that, for a given increase in the term spread, the analyst forecasts rise *less* than the realization. The model predicts that they should rise by more.

This one-factor model has an additional testable implication: high interest rates today forecast declining interest rates in the future. This again, is true across datasets, and for both analyst forecasts, and realizations. However, once again, a given change in the interest rates has a much greater impact on the realization than on the forecast. The model predicts the opposite: that when interest rates are high, investors believe that interest rates will mean revert but in fact they follow a process that is closer to a random walk. The survey data

he or she would have predicted cash flow growth with relative accuracy. While returns would have been different than expect, the low R^2 in return predictability regressions suggests that reinforcement learning through this channel would not have been significant.

therefore reject the model.

E.1 Model and simulation

Assume that investors believe that the continuously-compounded short-term interest rate r_t follows a first-order autoregressive process, so that

$$\Delta r_{t+1} = (\phi - 1)(r_t - \bar{r}) + v_{t+1} \tag{E1}$$

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 ϕ is the first-order autocorrelation of r_t.[13]

As with dividend growth, investors believe that changes in interest rates are more forecastable than they are in reality. That is, while (E1) represent beliefs, the true process is governed by

$$\Delta r_{t+1} = (\zeta - 1)(r_t - \bar{r}) + v_{t+1}, \tag{E2}$$

with

$$|\zeta - 1| < |\phi - 1|. \tag{E3}$$

We focus on the case where $\zeta, \phi \in [0, 1]$ so that (E3) implies $\zeta > \phi$. In forecasting next period's interest rate, (E3) 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

[13]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 δ 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.

walk (they mean revert more slowly) in the data than investors believe ($\zeta > \phi$).

We consider risk-neutral pricing for bonds. The dynamics thus far define a discrete-time ? model.[14] 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

$$B_n(r_t) = E_t^* \left[e^{-r_t} B_{n-1}(r_{t+1}) \right], \tag{E4}$$

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

$$\log B_n(r_t) = -a_n - b_n r_t \tag{E5}$$

with

$$\begin{aligned} a_n &= a_{n-1} + b_{n-1}(1-\phi)\bar{r} - \frac{1}{2} b_{n-1}^2 \sigma_v^2 \\ b_n &= 1 + b_{n-1}\phi \end{aligned} \tag{E6}$$

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

$$b_n = \frac{1-\phi^n}{1-\phi}. \tag{E7}$$

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

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

It follows from (E7) that the yield spread equals

$$y_{nt} - y_{1t} = \text{constant} + \left(\frac{1}{n} \frac{1-\phi^n}{1-\phi} - 1 \right) r_t \tag{E8}$$

[14]A substantial literature on latent factor models strongly rejects a single-factor model in favor of multi-factor alternatives (**??**). **?** 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.

(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 (E5), (E7), and for the physical evolution of r_t, (E2), 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 of stock returns by the price-dividend ratio. Long-term bond yields fluctuate due to changing forecasts of future short-term interest rates. When long-term yields are high relative to short-term yields, it is because (in this model), investors expect short-term yields to rise. However, short-term yields are not as predictable as investors believe, and thus on average, short-term yields will rise less than anticipated (or even fall). This leads to a positive excess return on the long-term bond.

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.[15] Indeed, the recursion (E4) 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), **?** and a large subsequent literature show that excess bond

[15] There are slight differences depending on whether this hypothesis is articulated in logs or levels (**?**).

returns are strongly forecastable. We replicate this finding in Table E1, which reports coefficients from regressing bond returns on yield spreads using the Fama-Bliss data set for zero-coupon bonds.

Table E1: Moments of Bond Yields

	Maturity in Years				
	1	2	3	4	5
Panel A: Data 1952-2019					
β_n		1.58	2.09	2.33	2.44
t-stat		[2.90]	[3.46]	[3.73]	[3.52]
$\sigma(y_n)$	3.08	3.04	2.97	2.93	2.87
$AC(y_n)$	0.88	0.90	0.91	0.91	0.92
$\sigma(y_n - y_1)$		0.33	0.53	0.69	0.81
$AC(y_n - y_1)$		0.41	0.46	0.52	0.55
Panel B: Model					
β_n		1.48	1.31	1.19	1.10
$\sigma(y_n)$	2.83	2.05	1.56	1.23	1.01
$AC(y_n)$	0.86	0.86	0.86	0.86	0.86
$\sigma(y_n - y_1)$		0.78	1.27	1.60	1.82
$AC(y_n - y_1)$		0.86	0.86	0.86	0.86

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 β_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 2019.

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

Table E1 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.

In addition to the moments in Table E1, the model makes predictions that can be tested with survey data. In the subsection below we describe the survey data and then perform the tests.

E.2 Survey data

Our primary survey data soruce is interest rate forecasts from Blue Chip Financial Forecasts (BCFF). This data source contains survey forecasts for a variety of interest rates in the US, in particular the Treasury rates. Behind each Treasury rate consensus are forecasts provided by tens of major banks and financial institutions, e.g. J.P. Morgan and S&P Global. This data source goes back to Q4 of 1982.

An alternate data source of interest rate forecasts to Blue Chip Financial Forecasts (BCFF) is Survey of Professional Forecasters (SPF). This is a quarterly survey contains a large number of economic variables, including the 3-month Treasury rates. The contributors to these surveys are economists of a variety of backgrounds. The interest rate forecast data go back to Q3 of 1981, which is similar to BCFF. While SPF is not a specialized interest rate data source and contains only the 3-month Treasury rate forecasts, it is useful as a robustness check on top of the BCFF data. Here, the correlations between interest rate forecasts and earning growths forecasts computed with BCFF data, as shown in Figure E1 are very similar to those computed with the SPF data in Figure E2.

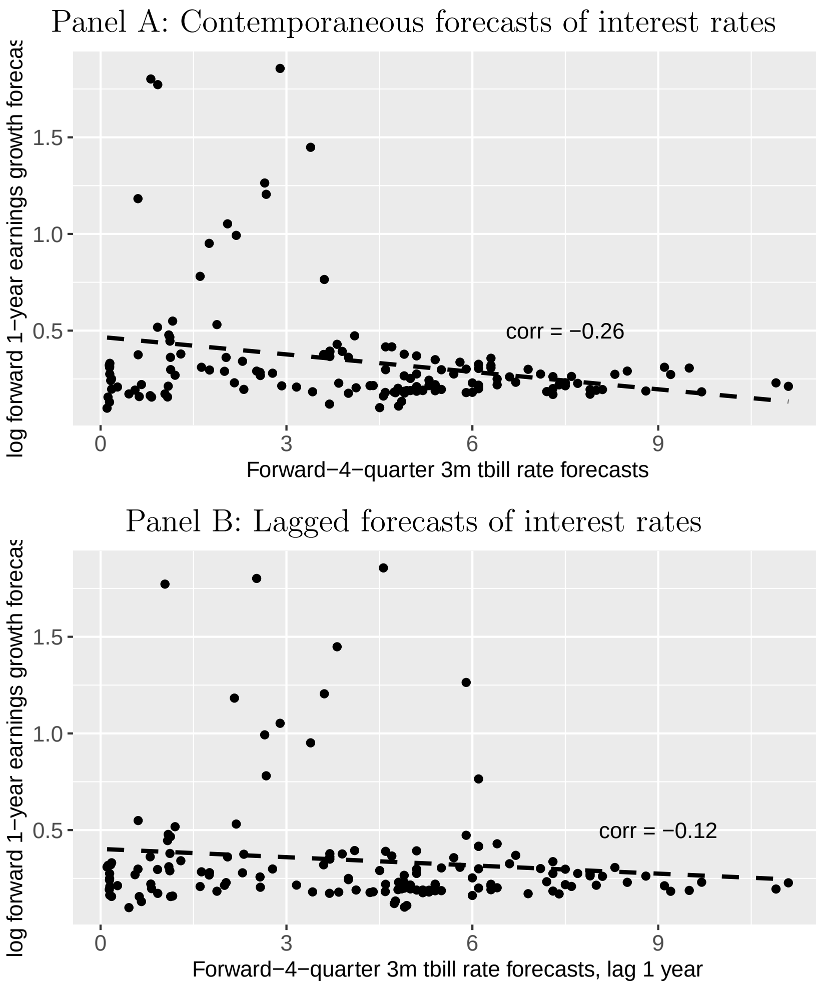
E.3 Results

Equation E1, E2, and E8 predict that the term spread be very positively associated with forecasted change of interest rate and less positively associated with its realization. Column 1-2 and 4-5 of Table E2 show that this does not hold in the survey data. Contrary to the model's prediction, we see that the term spreads are more positively associated with realized interest rate changes in the future. Similar results are seen also in the SPF data, shown in Table E3.

A counterfactual aspect of the model is that the level of the interest rate and the term spread are perfectly negatively correlated. It therefore also predicts that the level of the interest rate should be very negatively associated with forecasted interest rate change and less negatively associated with its realization. Column 3 and 6 of Table E2 and E3 show that this is also not the case either. The model is therefore soundly rejected by the survey data.

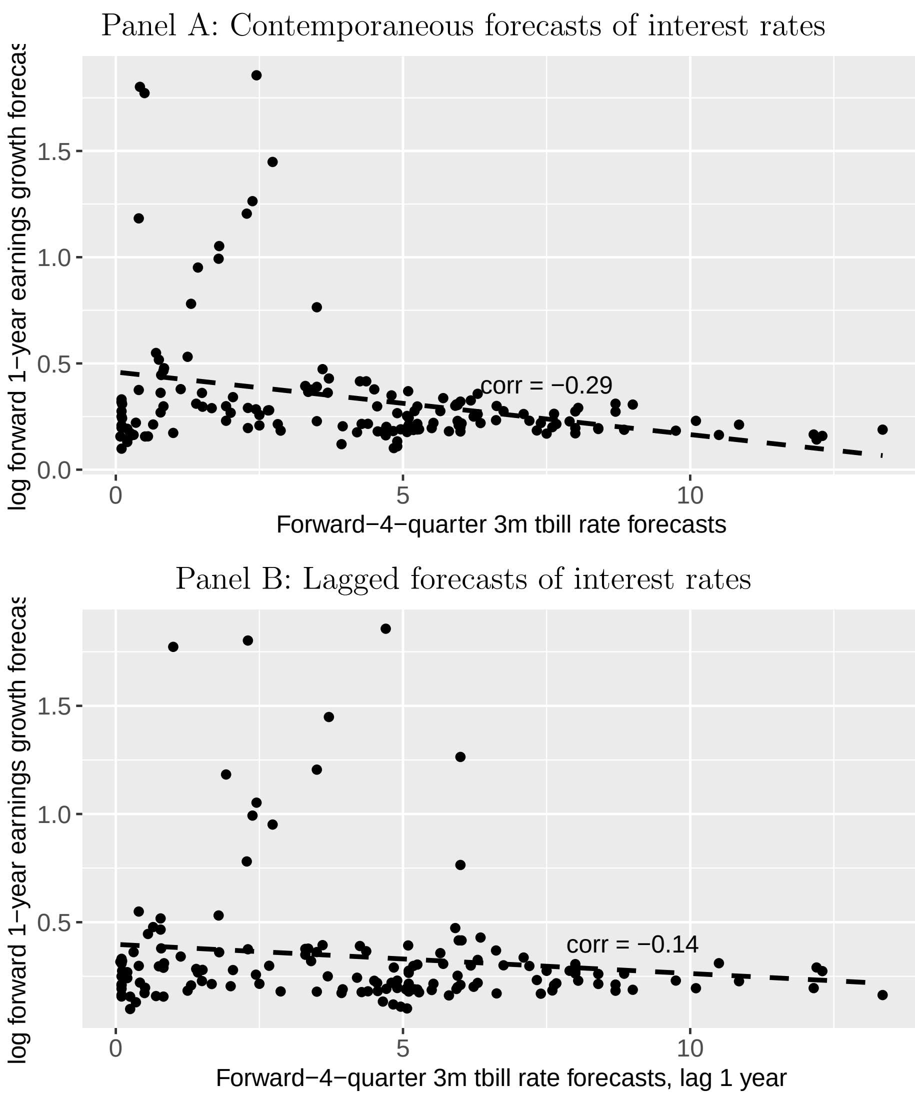
While one could, for example, consider more complicated models for the term structure (Campbell et al. (2020) consider a model in which the interest rate follows an ARMA(1,1) rather than an AR(1) process), it is not clear how model complexity would reconcile the model with the data. One would want such a model to generate predictability of excess bond returns by the term spread. Times of high term spreads should forecast higher than average excess bond returns, in the model, to be consistent with the data. Assuming that the model is in the same spirit as ours (namely, risk premia are constant or zero), then excess returns could only be higher because investors are surprised by lower interest rates. However, this is exactly what is counterfactual in the data: while a term spread does predict rising

Figure E1: Forecasted earnings growth versus forecasted interest rates (BCFF data)



This figure plots log forecasted 1-year earnings growth against forecasted 3-month Treasury bill rate 4 quarters away. Data are quarterly from 1982–2018.

Figure E2: Forecasted earnings growth versus forecasted interest rates (SPF data)



This figure plots log forecasted 1-year earnings growth against forecasted 3-month Treasury bill rate 4 quarters away. Data are quarterly from 1981–2019.

interest rates, and is correlated with higher forecasts, the forecasts respond less (investors are surprised by higher rates, not lower).

It is important to note what these results do not demonstrate. They do not reject a role for expectational errors in bond return volatility. It is possible that bond returns can be predicted by expectational errors (Cieslak, 2018; Wang, 2020). However, they do suggest that that documented predictability by the yield spread is not driven by investor's mis-specified beliefs. The fact that neither the yield spread nor the short rate forecasts errors in forecasts is evidence against the model. Rather, investors do not shift portfolios into long-term bonds when yield spreads are high, not because they predict (incorrectly) that rates will rise. They understand the higher returns offered by the long-term bonds, but do not take advantage of them because they correspond to greater risk.

Table E2: Term Spreads, Short Rates, and Forecasted Changes in Short Rates—BCFF

	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{E}_t[r_{3,t+4}] - r_{3,t}$			$r_{3,t+4} - r_{3,t}$		
$y_{60,t} - r_{3,t}$	0.315***			0.425*		
	[4.36]			[1.68]		
$y_{120,t} - r_{3,t}$		0.269***			0.357**	
		[5.10]			[2.12]	
$r_{3,t}$			-0.105***			-0.139**
			[-3.81]			[-2.47]
constant	-0.076	-0.153	0.726***	-0.743*	-0.836**	0.327
	[-0.60]	[-1.12]	[5.68]	[-1.97]	[-2.18]	[1.37]
N	145	145	145	145	145	145

Column 1 of this table reports results of the follow quarterly time-series regression: $\hat{E}_t[r_{3,t+4}] - r_{3,t} = \alpha + \beta(y_{60,t} - r_{3,t}) + \epsilon_t$. Here $\hat{E}_t[r_{3,t+4}] - r_{3,t}$ is the forecasted change of 3-month Treasury bill rates. $r_{60,t} - r_{3,t}$ is the 5-year Treasury bond rate subtracting 3-month Treasury bill rate in quarter t. Column 2 instead uses the 10-year/3-month term spread as the independent variable. Column 3 instead uses the 3-month Treasury bill rate as the dependent variable. Column 4-6 are analogous regressions with the dependent variables changed to realized short rate changes. Data are quarterly from 1982Q4-2018Q4. T-stats calculated using Newey-West standard errors with 6 lags are reported in the square brackets.

Table E3: Term Spreads, Short Rates, and Forecasted Changes in Short Rates—SPF

	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{E}_t[r_{3,t+4}] - r_{3,t}$			$r_{3,t+4} - r_{3,t}$		
$y_{60,t} - r_{3,t}$	0.329***			0.477**		
	[3.94]			[2.38]		
$y_{120,t} - r_{3,t}$		0.228***			0.463**	
		[3.17]			[2.53]	
$r_{3,t}$			-0.051**			-0.212***
			[-2.50]			[-3.29]
constant	-0.181	-0.161	0.482***	-0.958**	-1.184**	0.552*
	[-1.26]	[-1.00]	[5.10]	[-2.44]	[-2.53]	[1.91]
N	150	150	150	150	150	150

Column 1 of this table reports results of the follow quarterly time-series regression: $\hat{E}_t[r_{3,t+4}] - r_{3,t} = \alpha + \beta(y_{60,t} - r_{3,t}) + \epsilon_t$. Here $\hat{E}_t[r_{3,t+4}] - r_{3,t}$ is the forecasted change of 3-month Treasury bill rates. $y_{60,t} - r_{3,t}$ is the 5-year Treasury bond rate subtracting 3-month Treasury bill rate in quarter t. Column 2 instead uses the 10-year/3-month term spread as the independent variable. Column 3 instead uses the 3-month Treasury bill rate as the dependent variable. Column 4-6 are analogous regressions with the dependent variables changed to realized short rate changes. Data are quarterly from 1981Q4-2018Q4. T-stats calculated using Newey-West standard errors with 6 lags are reported in the square brackets.

F Additional results for equities

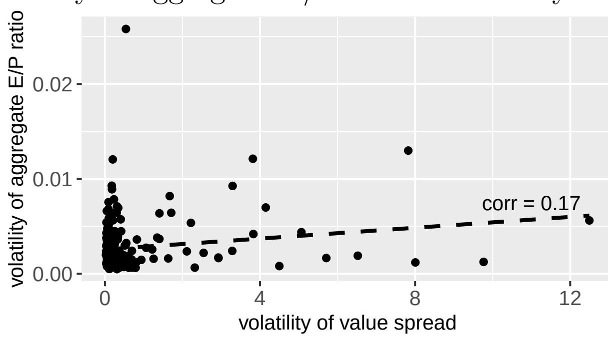
In our model of the cross section of equities, the value spread is the absolute value of z_t. The return to a HML strategy is $-sign(z_t)v_{z,t+1}$. The aggregate valuation ratio is x_t. With the addition of stochastic volatility, because the value spread and HML returns are both driven by the same set of shocks, their volatilities should be highly correlated. However, the aggregate valuation ratio is driven by x_t and thereby $v_{x,t+1}$. Our model therefore predicts weak correlations between the volatility of aggregate valuation and those of the value spread and HML returns.

Panel A of Figure F1 plots quarterly volatility of aggregate E/P ratio versus that of the value spread. They have a low correlation of 0.17. Panel B plots quarterly volatility of aggregate E/P ratio versus that of the HML return. They have a low correlation of 0.22. Panel C plots quarterly volatilites of the HML returns and the value spread we constructed. They have a reasonably high correlation of 0.53. These results are broadly consistent with the model's predictions.

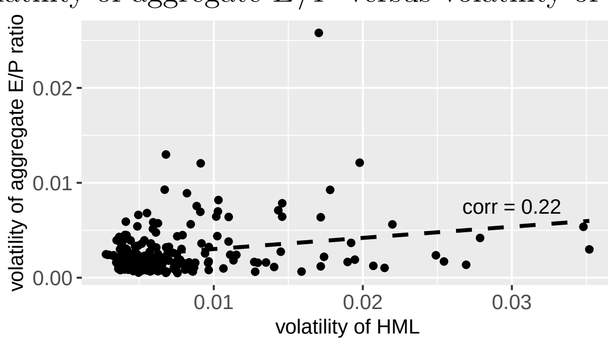
In our model, the value spread is the absolute value of z_t, and aggregate valuation ratio is x_t. Because x_t and z_t are based on difference iid Gaussian shocks, the model predicts a correlation of 0 between these two measures. Also, return to a HML strategy is $-sign(z_t)v_{z,t+1}$. That to a market timing strategy trading on the aggregate valuation ratio is $-x_t v_{x,t+1}$. Again, because they are driven by two different sets of iid Gaussian shocks, our model predicts that they are uncorrelated.

Figure F1: Relation between market risk and risk to a value strategy

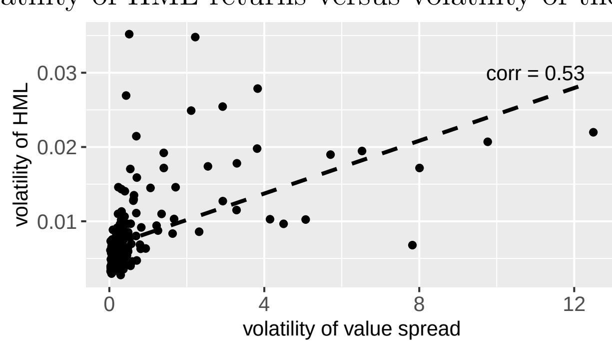
Panel A: Volatility of aggregate E/P versus volatility of value spread



Panel B: Volatility of aggregate E/P versus volatility of HML returns



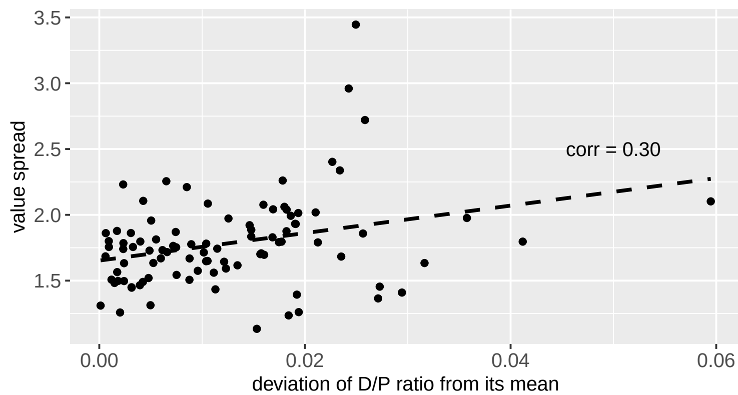
Panel C: Volatility of HML returns versus volatility of the value spread



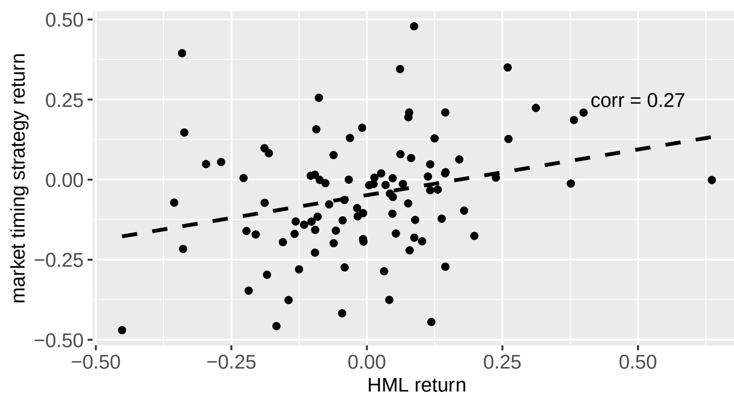
Notes: HML is a portfolio that is long the high E/P ratio quintile and short the low E/P ratio. The value spread is defined as the difference of E/P ratio of bin 5 and bin 1 scaled by the aggregate E/P ratio. Data are quarterly from 1971–2020.

Figure F2: Relation between the value anomaly and aggregate valuations

Panel A: Value spread versus deviation of aggregate dividend-to-price from its mean



Panel B: Returns to timing the market versus HML returns



Notes: The value spread is defined as the difference between the D/P ratio of bin 5 in a value sort and that of bin 1, scaled by the aggregate D/P ratio. Data are monthly from 1926–2020.

References