Distorted Beliefs and Asset Prices

March 17, 2025

Abstract

We examine the impact of distorted investor beliefs on stock market risk premia. By measuring the time-varying bias in investor beliefs, we decompose the objective (statistical) conditional expectation of market returns into two components: investors' beliefs about future returns and their expectational bias. Variation in this bias manifests itself as discount-rate risk in the data but is viewed as cash-flow risk by investors, creating a wedge between investor pricing and the observed risk-return relationship. Thus, accounting for distorted beliefs helps explain the deviations from theoretical predictions observed in the cross-section of expected stock returns.

Keywords: distorted beliefs, return predictability, ICAPM, cross-section of stock returns.

JEL Classification: G12

1 Introduction

There is a growing body of evidence suggesting that economic agents make predictable forecast errors—i.e., have "distorted beliefs"—about macroeconomic and financial outcomes.[1] Various factors, such as rational inattention or behavioral biases, can distort beliefs, leading agents to either underreact or overreact to relevant information. Recognizing and accounting for these distorted beliefs can provide new insights into empirical puzzles that are difficult to explain under the assumption of full-information rational expectations.

In this paper, we investigate the role of distorted beliefs in the stock market, particularly their impact on risk premia. We begin by measuring the conditional bias in investors' beliefs and decompose the predictable component of market returns into two parts: investors' beliefs about future returns and their bias. Variation in this bias manifests itself as discount-rate risk in the data but represents cash-flow risk from investors' perspective. As a result, the bias gives rise to a wedge in observed risk premia, reflecting the difference in the pricing of cash-flow and discount-rate risks. Our findings indicate that distorted beliefs impact both the time series and cross-section of expected returns, helping to explain observed deviations from theoretical predictions, such as those of the Intertemporal Capital Asset Pricing Model (ICAPM).

Our first contribution is to propose a novel way to measure the bias in investors' beliefs—i.e., the difference between beliefs and objective expectations—regarding long-run cash-flow growth and returns. We show that the predictable component of the revisions in investors' cash-flow beliefs identifies the cash-flow bias. In turn, present value identities imply that, in the long-run, cash-flow bias mirrors the bias in expected returns. We estimate this time-varying bias at the market level using revisions in professional analysts' earnings forecasts. By adjusting the statistical expectation of market returns for this bias, we obtain a measure of investors' beliefs about future market returns. The variation in our measure aligns with

[1] Prominent examples include studies by Coibion and Gorodnichenko (2012, 2015), Greenwood and Shleifer (2014), Adam et al. (2017), Bordalo et al. (2020), Nagel and Xu (2021), Bianchi et al. (2022), and Bordalo et al. (2024a).

expectations embedded in index option prices and is consistent with forecasts from industry, banking and government economists, while differing from some other survey-based proxies for expected market returns used in the literature. Overall, our findings suggest that investor beliefs about future market returns are neither constant nor perfectly aligned with the objective variation in expected returns. As a result, it is essential to explicitly allow for the dynamics of these beliefs.

Our second contribution is to show that expectational bias contributes to excess aggregate stock price volatility—i.e., stock price movements unrelated to subsequent movements in cash flows, as hypothesized by Shiller (1981)—but does not explain it completely. Using the decomposition of expected returns into belief and bias components, our analysis reveals that discount-rate beliefs remain a substantial source of stock-market volatility. After accounting for predictability consistent with investor beliefs, we estimate that bias explains approximately 41% of the observed predictability in market returns.

Having estimated the bias in investors' beliefs, our third contribution is to show that this bias can lead to a distorted perception and, consequently, a distorted pricing of risk. A fundamental theoretical insight, originating from Merton (1973) and Campbell (1993), is that financial asset risk premia depend on their return covariances with cash-flow news and expected return—i.e., discount-rate—news. And, non-myopic investors price these two types of risk differently. Under distorted beliefs, investors may misattribute some of the variation in expected returns to the variation in expected cash flows, and vice versa. More specifically, the variation in the bias captures the component of discount-rate news (cash-flow news) that investors see as cash-flow news (discount-rate news). As a result, even if investors price assets according to the ICAPM, the risk-return relationship predicted by the model will not hold in the data, giving rise to an apparent empirical puzzle.

To gauge the gap between pricing under investor's beliefs and the observed risk-return relationship, we derive a bias-adjusted ICAPM that takes the form of a linear factor model that includes cash-flow news, discount-rate news, and bias shocks. We compare this model's performance to the traditional Campbell and Vuolteenaho (2004) two-factor bad beta, good beta ICAPM, which excludes bias shocks. Using 51 test portfolios, our findings reveal that incorporating bias shocks significantly enhances the model's pricing performance, with R^2 values increasing from around 3% to around 36%. Despite this improvement, the model still leaves some pricing errors unexplained, particularly for size-sorted portfolios. The bias shock commands a positive and significant risk price in this cross-section, a result further validated using the supervised PCA approach of Giglio et al. (2024), demonstrating the robustness of our results to the asset selection criteria and to additional factors. Our findings suggest that accounting for distorted beliefs is crucial for understanding the apparent deviations from the risk-return relationship predicted by the theory.

Finally, our analysis also sheds new light on the sources of distorted beliefs. We note that our estimation approach imposes no assumptions on the belief formation mechanism. Instead, we employ VAR forecasts along with present value restrictions to measure the bias. In particular, this multivariate approach accommodates both overreaction and underreaction to different information. We find evidence consistent with overreaction, in particular with theories of diagnostic expectations, as cash flow forecasts tend to negatively predict future forecast revisions and the bias tends to be pro-cyclical. However, we also observe that investors can have an optimistic bias not only during market booms but also during downturns, suggesting that they may underreact to some information and not adequately revise their expectations downward. Examining the drivers of the estimated bias, we find that biases in stock market and macroeconomic expectations—particularly unemployment expectations—are related. Interestingly, we also find that a more optimistic (pessimistic) bias is associated with lower (higher) risk appetite in the corporate bond market. A potential interpretation of this relationship is that stock market responses lag behind shifts in credit market conditions, resulting in an accumulation of bias.

A growing body of literature studies distorted beliefs and their relation to stock market behavior, including studies by Barberis et al. (1998), Bordalo et al. (2019), Nagel and Xu (2021), Bianchi et al. (2024), and Cassella et al. (2024) among others.[2] Closest to our paper are the studies by De La O and Myers (2021), which explore excess stock price volatility, and Bordalo et al. (2024a), which investigate market return predictability. The novelty of our work lies in its focus on the pricing of risk. Our findings differ from these authors in other important ways. For instance, contrary to De La O and Myers (2021), our results suggest that discount-rate variation remains an important source of stock-market volatility under investors' beliefs. Motivated by this finding, we extend the analysis of Bordalo et al. (2024a) to incorporate time-varying beliefs about future market returns and disentangle return predictability arising from belief distortions from that consistent with investors' beliefs.

The argument that distorted beliefs can drive variation in expected returns has been advanced by Kozak et al. (2018). The authors propose a model in which sentiment-driven investors induce fluctuations in expected returns, which creates an inter-temporal risk for rational investors.[3] Our work builds on a related yet distinct premise: a representative non-myopic investor with distorted beliefs mistakes some of the variation in expected returns (expected cash flow) for variation in expected cash-flow (expected returns), and mis-prices this component of risk. We take this framework to the data and find empirical support for its predictions.

Our work is also related to the literature on the distorted beliefs about the macroeconomy; see Coibion and Gorodnichenko (2015), Bordalo et al. (2020), Bianchi et al. (2022), Bhandari et al. (2024), among others. In particular, we find a significant relationship between agents' biases in market cash flow (and market return) expectations and biases in unemployment expectations.

Belief distortions are implicit in the market timing theory of capital structure as proposed by Baker and Wurgler (2002). We find a positive albeit moderate correlation between the

[2] While the literature, including our paper, has primarily focused on distorted beliefs at the market level, recent studies have also examine beliefs regarding firm-level earnings; see Bordalo et al. (2024b).

[3] See Section IV in Kozak et al. (2018).

stock market sentiment index proposed by Baker and Wurgler (2006) that builds on this capital structure theory and our measure of the bias in investors' beliefs. Overall, our findings indicate that our measure effectively captures both previously studied dimensions of belief distortions as well as additional dimensions not identified in prior research.

Our analysis builds on the present value framework developed by Campbell and Shiller (1988), Campbell (1991), Campbell (1993), and applied by Campbell and Vuolteenaho (2004) and Campbell et al. (2013), among others. Recent work has extended this framework in several directions. For instance, Campbell et al. (2018) incorporate stochastic volatility, while Gao and Martin (2021) consider higher-order approximation terms and use options data. In our paper, we explore a different extension of this framework and examine the implications of present value relationships under distorted beliefs.

The rest of the paper is organized as follows. Section 2 presents our analytical framework. Section 3 describes our data. Section 4 presents our main empirical results. Section 5 concludes. An Internet Appendix contains proofs and additional empirical results.

2 Analytical Framework

2.1 Present value relationships and ICAPM

To analyze the impact of distorted beliefs on the baseline predictions of asset pricing theory, we consider the Campbell (1993) version of the ICAPM derived in the Campbell and Shiller (1988) and Campbell (1991) homoscedastic log-linear present value framework. For analytical tractability, we further assume log-normality. To establish our notation, we provide a brief overview of this standard framework and refer to the original papers for a more detailed discussion of the underlying assumptions.

The market log return r_t can be approximated as

$$r_{t+1} \approx k + \Delta d_{t+1} + \rho \left(p_{t+1} - d_{t+1}\right) - \left(p_t - d_t\right), \tag{1}$$

where Δd_t is the market dividend growth, $p_t - d_t$ is the market log price-to-dividend ratio, and k and ρ are constants. Iterating equation (1) forward, assuming $\lim_{h\to\infty} \rho^h \mathbb{E}_t(p_{t+h} - d_{t+h}) = 0$, and taking expectations, the market valuation depends on the difference between expected cash-flow growth $\mathcal{F}_{d,t}$ and expected discount rates $\mathcal{F}_{r,t}$

$$p_t - d_t \approx \frac{k}{1-\rho} + \underbrace{\mathbb{E}_t \sum_{h=0}^{\infty} \rho^h \Delta d_{t+1+h}}_{\equiv \mathcal{F}_{d,t}} - \underbrace{\mathbb{E}_t \sum_{h=0}^{\infty} \rho^h r_{t+1+h}}_{\equiv \mathcal{F}_{r,t}}, \tag{2}$$

Combining equations (1) and (2), we obtain the decomposition of aggregate market return innovations $r_{t+1} - \mathbb{E}_t r_{t+1}$ into cash-flow news $\varepsilon_{d,t+1}$ and discount-rate news $\varepsilon_{r,t+1}$

$$r_{t+1} - \mathbb{E}_t r_{t+1} \approx \underbrace{\left(\mathbb{E}_{t+1} - \mathbb{E}_t\right) \sum_{h=0}^{\infty} \rho^h \Delta d_{t+1+h}}_{\equiv \varepsilon_{d,t+1}} - \underbrace{\left(\mathbb{E}_{t+1} - \mathbb{E}_t\right) \sum_{h=1}^{\infty} \rho^h r_{t+1+h}}_{\equiv \varepsilon_{r,t+1}}. \tag{3}$$

Under Campbell (1993) ICAPM, for any asset return $r_{i,t}$ and corresponding excess return $r^e_{i,t}$

$$\mathbb{E}_t r^e_{i,t+1} + \frac{\mathbb{V}\mathrm{ar}_t r_{i,t+1}}{2} \approx \gamma \times \mathbb{C}\mathrm{ov}_t \left[r_{i,t+1}, \varepsilon_{d,t+1}\right] - 1 \times \mathbb{C}\mathrm{ov}_t \left[r_{i,t+1}, \varepsilon_{r,t+1}\right], \tag{4}$$

where γ is the representative investor's coefficient of relative risk aversion.

2.2 Distorted beliefs

We now assume that, in the framework described above, the representative investor uses an equivalent probability measure $\mathbb{S}$ that differs from the objective probability measure $\mathbb{P}$. For instance, investor might hold diagnostic expectations as in Bordalo et al. (2024a). That said, our framework is intentionally general and we remain agnostic about the belief formation mechanism that gives rise to $\mathbb{S}$.

That said, we assume that the investor is rational otherwise. In particular, her beliefs are constrained by the present value relationships: investor's beliefs about cash-flow growth and expected returns may differ from their objective expectations, but these two components of beliefs have to add up to prevailing market valuations. In other words, from investor's perspective, the present value identity (2) holds but with expectations taken under $\mathbb{S}$.[4]

The discrepancy between $\mathbb{S}$ and $\mathbb{P}$ introduces a bias in investor's beliefs. Our primary focus is on the conditional bias—the difference between conditional expectations under $\mathbb{S}$ and $\mathbb{P}$—and the impact that its time-variation has on asset price fluctuations.[5] Comparing equation (2) with expectations taken, respectively, under $\mathbb{S}$ and under $\mathbb{P}$ we note that, the bias in long-run cash-flow expectations corresponds to an equal bias in long-run market return expectations, and vice versa. Indeed, since both investor's and statistical beliefs are consistent with observed $p_t - d_t$, we have:

$$\mathbb{E}_t^{\mathbb{S}} \sum_{h=0}^{\infty} \rho^h \left(\Delta d_{t+1+h} - r_{t+1+h}\right) = \mathbb{E}_t^{\mathbb{P}} \sum_{h=0}^{\infty} \rho^h \left(\Delta d_{t+1+h} - r_{t+1+h}\right),$$

which after rearranging gives

$$\underbrace{\left(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}\right) \sum_{h=0}^{\infty} \rho^h \Delta d_{t+1+h}}_{= \mathcal{F}_{d,t}^{\mathbb{S}} - \mathcal{F}_{d,t}^{\mathbb{P}} \equiv \mathcal{B}_t^{\mathbb{S}-\mathbb{P}}} = \underbrace{\left(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}\right) \sum_{h=0}^{\infty} \rho^h r_{t+1+h}}_{\mathcal{F}_{r,t}^{\mathbb{S}} - \mathcal{F}_{r,t}^{\mathbb{P}}}. \tag{5}$$

[4]This requires that several technical conditions are satisfied. First, we assume that $\lim_{h\to\infty} \rho^h \mathbb{E}_t^{\mathbb{S}}(p_{t+h} - d_{t+h}) = 0$, ruling out bubbles under distorted beliefs. Second, we assume that the linearization parameter ρ is the same under both $\mathbb{P}$ and $\mathbb{S}$. This can be viewed as a simplifying approximation. Alternatively, it can be justified by assuming that the investor knows the objective long-run mean of the log price-to-dividend ratio, which determines ρ, even though her conditional beliefs may deviate from objective expectations.

[5]We allow for but do not emphasize a potential unconditional bias, to the extent that it is not driving asset price fluctuations. In fact, we do not require that the bias converges. Later in this section, we impose a no-bubble-type condition on the extent to which beliefs can diverge from objective expectations in the limit.

For example, consider overly optimistic beliefs about future cash-flow growth. These beliefs inflate market valuations, resulting in lower long-run returns relative to investor's beliefs.[6]

We will refer to the term $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ in equation (5), which captures the discrepancy between investor's belief and objective expectation of both aggregate cash flows and market returns, as simply the *bias*. A positive (negative) value of $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ indicates that the investor holds more optimistic (pessimistic) beliefs compared to the objective forecast. For convenience, we also refer to the negation $\mathcal{B}_t^{\mathbb{P}-\mathbb{S}} = -\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ as bias. Multiplying equation (5) evaluated at $t+1$ by ρ and subtracting the same equation at t, we obtain a relationship between the bias in cash-flow and discount-rate news defined in equation (3)

$$\underbrace{\varepsilon_{d,t+1}^{\mathbb{S}} - \varepsilon_{d,t+1}^{\mathbb{P}}}_{\rho\mathcal{B}_{t+1}^{\mathbb{S}-\mathbb{P}} - \mathcal{B}_t^{\mathbb{S}-\mathbb{P}}} = \varepsilon_{r,t+1}^{\mathbb{S}} - \varepsilon_{r,t+1}^{\mathbb{P}} - \underbrace{(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}})r_{t+1}}_{\text{time-}t\text{ term}}. \tag{6}$$

Importantly, the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ can be recovered from the predictability of how investors revise their beliefs. Taking conditional expectations under $\mathbb{P}$ of the definition in equation (6), we have

$$\rho\mathbb{E}_t^{\mathbb{P}}\mathcal{B}_{t+1}^{\mathbb{S}-\mathbb{P}} - \mathcal{B}_t^{\mathbb{S}-\mathbb{P}} = \mathbb{E}_t^{\mathbb{P}}\varepsilon_{d,t+1}^{\mathbb{S}} - \underbrace{\mathbb{E}_t^{\mathbb{P}}\varepsilon_{d,t+1}^{\mathbb{P}}}_{=0}.$$

By iterating forward and imposing a no-bubble-type condition, $\lim_{h\to\infty}\rho^h\mathbb{E}_t^{\mathbb{P}}\mathcal{B}_{t+h}^{\mathbb{S}-\mathbb{P}} = 0$, which allows beliefs to deviate from objective expectations while ensuring that such divergences do not become explosive, we obtain

$$\mathcal{B}_t^{\mathbb{S}-\mathbb{P}} = -\sum_{h=0}^{\infty}\rho^h\mathbb{E}_t^{\mathbb{P}}\varepsilon_{d,t+1+h}^{\mathbb{S}}. \tag{7}$$

From equation (7), the bias is equal to the predictable component of the future revisions in investor's forecast of cash-flow growth. In our empirical implementation, we use statistical

[6] This holds true if investor is not marginal: low market valuations that do not reflect her optimism about future cash-flow growth imply, from investor's perspective, high expected returns, which ultimately fail to materialize.

expectations conditional on a set of state variables to proxy for the expectation under $\mathbb{P}$. Note that equation (7) quantifies the difference between expectations under $\mathbb{S}$ and $\mathbb{P}$ conditional on the same variables. Thus, the bias captured by equation (7) does not stem from an informational advanatge that the investor may have over the econometrician but rather reflects a belief distortion conditional on observing the same information.[7]

2.3 Bias in the expected market return

Distorted beliefs explain part of the objective predictability in market returns. Appendix A.1 shows that, using identity (3), we can decompose the H-period expected return into investor's belief and the effect of distorted beliefs

$$\underbrace{\overbrace{\mathbb{E}_t^{\mathbb{P}} \sum_{h=0}^{H-1} \rho^h r_{t+1+h}}^{\text{objective}}}_{\to \mathcal{F}_{r,t}^{\mathbb{P}} \text{ as } H\to\infty} = \underbrace{\overbrace{\mathbb{E}_t^{\mathbb{S}} \sum_{h=0}^{H-1} \rho^k r_{t+1+h}}^{\text{belief}}}_{\to \mathcal{F}_{r,t}^{\mathbb{S}} \text{ as } H\to\infty} + \overbrace{\underbrace{\mathbb{E}_t^{\mathbb{P}} \sum_{h=0}^{H-1} \rho^h \varepsilon_{d,t+1+h}^{\mathbb{S}}}_{\to \mathcal{B}_t^{\mathbb{P}-\mathbb{S}} \text{ as } H\to\infty} + \underbrace{(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}})\mathbb{E}_{t+H}^{\mathbb{S}} \sum_{h=H}^{\infty} \rho^k r_{t+h+1}}_{\to 0 \text{ as } H\to\infty}}^{\text{belief distortion}} .$$

As $H \to \infty$, this expression reduces to the definition in equation (5), with the effect of distorted beliefs converging to $\mathcal{B}_t^{\mathbb{P}-\mathbb{S}}$.

2.4 Bias in the cross-section of expected returns

Distorted beliefs can also help influence expected returns in the cross-section. We start from an intuitive benchmark suggested by finance theory and assume that, from investor's perspective, excess returns compensate for assets' exposure to cash-flow and discount-rate risks. More formally, we assume that the ICAPM in equation (4) holds under $\mathbb{S}$.

The discrepancy between $\mathbb{S}$ and $\mathbb{P}$ creates a wedge between the risk-return relationship that holds under investor's beliefs and the relationship observed in the data. To characterize the

[7] In our empirical implementation, we use a parsimonious set of publicly observed state variables, making it also unlikely that the economtrician has an informational advantage over the investor.

risk-return relationship under $\mathbb{P}$, we express expected returns similarly to Jensen (2024) as the sum of a stock-level bias and the required compensation for perceived risk

$$\mathbb{E}_t^{\mathbb{P}} r_{i,t+1}^e + \frac{\mathbb{V}\mathrm{ar}_t^{\mathbb{S}} r_{i,t+1}}{2} \approx -\left(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}\right) r_{i,t+1}^e + \mathbb{C}\mathrm{ov}_t^{\mathbb{S}}\left[r_{i,t+1}, \gamma\varepsilon_{d,t+1}^{\mathbb{S}} - \varepsilon_{r,t+1}^{\mathbb{S}}\right]. \tag{8}$$

Focusing first on the covariance term in Equation (8), the investor misjudges risk, evaluating asset return covariance with cash-flow news $\varepsilon_{d,t}^{\mathbb{S}}$ and discount-rate news $\varepsilon_{r,t}^{\mathbb{S}}$ rather than their objective counterparts. Using equation (6), we can quantify the resulting difference as

$$\gamma\varepsilon_{d,t+1}^{\mathbb{S}} - \varepsilon_{r,t+1}^{\mathbb{S}} = \gamma\varepsilon_{d,t+1}^{\mathbb{P}} - \varepsilon_{r,t+1}^{\mathbb{P}} + (\gamma - 1)\underbrace{\left(\varepsilon_{d,t+1}^{\mathbb{S}} - \varepsilon_{d,t+1}^{\mathbb{P}}\right)}_{\rho\mathcal{B}_{t+1}^{\mathbb{S}-\mathbb{P}} - \mathcal{B}_t^{\mathbb{S}-\mathbb{P}}} - \underbrace{(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}})r_{t+1}}_{\text{time-}t\text{ term}}. \tag{9}$$

The term in $\rho\mathcal{B}_{t+1}^{\mathbb{S}-\mathbb{P}} - \mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ in equation (9) arises because shocks to the bias manifest themselves as discount-rate risk in the data but represent cash-flow risk from investor's perspective. Since cash-flow and discount-rate risks are priced differently when $\gamma \neq 1$, this distinction affects expected returns. Note that this term does not introduce an additional risk factor—i.e., the investor does not price the risk of her own bias—but instead measures the wedge between perceived and objective risk.

Next, note that the second-moment operators, $\mathbb{V}\mathrm{ar}_t$ and $\mathbb{C}\mathrm{ov}_t$, in equation (8) remain approximately invariant under the change of measure. This result is intuitively related to Girsanov's theorem, which preserves the volatility of a stochastic process under a measure change. In our discrete-time log-normal setting this property holds approximately, as shown in Appendix A.1.[8]

Finally, the asset-specific bias $\left(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}\right) r_{i,t+1}^e$ in equation (8) simplifies considerably in two plausible cases. First, this bias will tend to average out in cross-sectional asset pricing tests

[8] This property does not imply that investors assess risk correctly, as evident from equation (9). Rather, in our setting, the change of measure does not introduce bias into the second moments of a given shock. Extending our framework to a stochastic volatility environment, where investors hold biased beliefs about volatility, represents a promising avenue for future research.

$\mathbb{E}^{\mathbb{P}}\left(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}\right) r_{i,t+1}^e = 0$, unless investors are unconditionally either overly optimistic or overly pessimistic about asset-i returns. Second, if asset-specific bias stems from distorted beliefs about aggregate risk factors, it will be proportional to the return covariance with these factors. For instance, consider an investor with distorted beliefs about aggregate cash flow: $\frac{d\mathbb{S}}{d\mathbb{P}} = S_t$ with $\frac{S_{t+1}}{S_t} \approx 1 + s_t \varepsilon_{d,t+1}^{\mathbb{P}}$, where the process s_t determines the distortion of cash-flow news $\varepsilon_{d,t+1}^{\mathbb{P}}$.[9] Under this assumption, we have

$$(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}) r_{i,t+1} \approx s_t \times \mathbb{C}\text{ov}_t\left[r_{i,t+1}, \varepsilon_{d,t+1}^{\mathbb{P}}\right] \text{ and } \mathcal{B}_t^{\mathbb{S}-\mathbb{P}} \approx s_t \times \mathbb{V}\text{ar}_t\left[\varepsilon_{d,t+1}^{\mathbb{P}}\right].$$

Combining together, we obtain

$$(\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}) r_{i,t+1} \approx \frac{\mathbb{C}\text{ov}_t\left[r_{i,t+1}, \varepsilon_{d,t+1}^{\mathbb{P}}\right]}{\mathbb{V}\text{ar}_t\left[\varepsilon_{d,t+1}^{\mathbb{P}}\right]} \times \mathcal{B}_t^{\mathbb{S}-\mathbb{P}}.$$

In either of these two cases, the unconditional version of the risk-return relationship (8) simplifies to

$$\mathbb{E}^{\mathbb{P}} R_{i,t}^e \approx \lambda_d \mathbb{C}\text{ov}^{\mathbb{P}}\left[r_{i,t}, \varepsilon_{d,t}^{\mathbb{P}}\right] + \lambda_r \mathbb{C}\text{ov}^{\mathbb{P}}\left[r_{i,t}, \varepsilon_{r,t}^{\mathbb{P}}\right] + \lambda_{\mathcal{B}} \mathbb{C}\text{ov}^{\mathbb{P}}\left[r_{i,t}, \varepsilon_{\mathcal{B},t}\right], \tag{10}$$

where $R_{i,t}^e$ is the simple excess return on asset i, $\varepsilon_{\mathcal{B},t+1}$—the bias shock—represents the innovation component of $\rho \mathcal{B}_{t+1}^{\mathbb{S}-\mathbb{P}} - \mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$, and $\lambda_d, \lambda_r, \lambda_{\mathcal{B}}$ are factor loadings.[10] We will consider equation (10) as the baseline for our empirical anaysis.

Alternatively, if investors exhibit a purely idiosyncratic cross-sectional pattern in their beliefs—stemming from unconditional optimism or pessimism about certain stocks—this would introduce significant noise into tests of Equation (8). Accurately testing this relationship would therefore require direct measures of investor beliefs regarding individual asset returns.

[9] This expression provides a discrete-time approximation of a stochastic exponential.

[10] For instance, in the case above, the loadings are $\lambda_d = \gamma - \frac{\mathbb{E}\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}}{\mathbb{V}\text{ar}\varepsilon_{d,t+1}^{\mathbb{P}}}$, $\lambda_r = 1, \lambda_{\mathcal{B}} = \gamma - 1$. If, in additon, the investor is not unconditionally biased, we have $\mathbb{E}\left[\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}\right] = 0$ and the asset-specific bias term vanishes altogether.

2.5 VAR identification

Finally, to measure the key terms in equations (5) and (10), we consider the VAR model

$$\mathbf{x}_{t+1} = \mathbf{B}\mathbf{x}_t + \mathbf{u}_{t+1} \tag{11}$$

where $\mathbf{x}_t = (r_t^e \, \varepsilon_{d,t}^{\mathbb{S}} \, \mathbf{z}_t)^\intercal$ is a vector that includes excess market returns r_t^e, revisions in investor's forecasts of cash-flow growth $\varepsilon_{d,t}^{\mathbb{S}}$, and a set of predictive variables $\mathbf{z}_t$. As discussed in the next section, the forecast revisions $\varepsilon_{d,t}^{\mathbb{S}}$ can be constructed from analysts' earnings forecasts. All variables are demeaned to simplify notation.

The VAR model (11) serves a dual purpose of uncovering statistical predictability in both market returns and forecast revisions, ensuring that both forecasts are based on the same set of information. First, we follow Campbell and Vuolteenaho (2004) and exploit market return predictability captured by the VAR model to obtain discount-rate news

$$\varepsilon_{r,t+1}^{\mathbb{P}} = \mathbf{e}_1^\intercal \rho \mathbf{B} \left(\mathbf{I} - \rho \mathbf{B}\right)^{-1} \mathbf{u}_{t+1}, \tag{12}$$

where $\mathbf{e}_1 = (1\,0\,0...0)^\intercal$ is a selection vector. From equation (3), discount-rate and cash-flow news add up to the innovation in market returns, implying $\varepsilon_{d,t+1}^{\mathbb{P}} = \varepsilon_{r,t+1}^{\mathbb{P}} + \mathbf{e}_1^\intercal \mathbf{u}_{t+1}$.

Next, from equation (7), the bias is equal to the predictable component of forecast revisions and therefore we have

$$\mathcal{B}_t^{\mathbb{S}-\mathbb{P}} = -\mathbf{e}_2^\intercal \mathbf{B}(\mathbf{I} - \rho \mathbf{B})^{-1} \mathbf{x_t}, \tag{13}$$

where $\mathbf{e}_2 = (0\,1\,0...0)^\intercal$ is a selection vector. Additionally, for convenience, we will refer to the the innovation component of forecast revisions $\varepsilon_{d,t+1}^{\mathbb{S}}$ as the belief shock equal to $\mathbf{e}_2^\intercal \mathbf{u}_{t+1}$, and to the innovation component of $\rho \mathcal{B}_{t+1}^{\mathbb{S}-\mathbb{P}} - \mathcal{B}_t^{\mathbb{S}-\mathbb{P}} = \varepsilon_{d,t+1}^{\mathbb{S}} - \varepsilon_{d,t+1}^{\mathbb{P}}$ as the bias shock denoted by $\varepsilon_{\mathcal{B},t+1}$ and equal to

$$\varepsilon_{\mathcal{B},t+1} = \mathbf{e}_2^\intercal \mathbf{u}_{t+1} - \varepsilon_{d,t+1}^{\mathbb{P}}. \tag{14}$$

Finally, we note an alternative way to identify cash-flow and discount-rate news in the VAR model (11). From equation (6) we can obtain cash-flow news as

$$\varepsilon^{\mathbb{P}}_{d,t+1} = \varepsilon^{\mathbb{S}}_{d,t+1} - \rho\mathcal{B}^{\mathbb{S}-\mathbb{P}}_{t+1} + \mathcal{B}^{\mathbb{S}-\mathbb{P}}_{t}, \tag{15}$$

which, through relationship (3), implies that discount-rate news are equal to $\varepsilon^{\mathbb{P}}_{r,t+1} = \varepsilon^{\mathbb{P}}_{d,t+1} - \mathbf{e}_1^{\intercal}\mathbf{u}_{t+1}$. Intuitively, equation (15) shows that (changes in) objective cash-flow growth forecast can be recovered from investor's beliefs corrected for their bias. Comparing the shocks identified through the Campbell and Vuolteenaho (2004) approach with those identified using this alternative approach can serve as a consistency test for our framework.

3 Data

3.1 Earnings forecasts

To construct empirical proxies at the monthly frequency for investors' forecast of market earnings growth ($\mathcal{F}^{\mathbb{S}}_{d,t}$) and for the revisions in this forecast ($\varepsilon^{\mathbb{S}}_{d,t}$), we use the analysts' median forecast of earnings per share and long-run earnings growth from the Thomson Reuters I/B/E/S Estimates Database available from January 1982 onwards. The I/B/E/S data have been extensively used in the literature on distorted beliefs. In particular, De La O and Myers (2021) use short-term earnings per share forecasts whereas Bordalo et al. (2024a) examine long-term growth forecasts. In our analysis, we incorporate both short- and long-term forecasts, as explained below.

As a first step, we obtain market-level forecast by aggregating the earnings per share forecast at one- and two-year horizons as well the long-term growth forecast across all S&P 500 stocks. In a second step, we consider two alternative ways to aggregate forecasts at different horizons to obtain a measure of $\mathcal{F}^{\mathbb{S}}_{d,t} = \mathbb{E}^{\mathbb{S}}_t \sum_{h=0}^{\infty} \rho^h \Delta d_{t+1+h}$. In each approach, we assume a constant payout ratio, implying that dividend growth equals earnings growth.

In our baseline approach, we measure $\mathcal{F}^{\mathbb{S}}_{d,t}$ by combining one- and two-year expected growth rates implied by earnings per share forecasts at one-year (EPS^{1y}_t) and two-year (EPS^{2y}_t) horizons; assuming the growth rate between years two and ten equal to the long-term growth forecast (LTG); and assuming that the growth rate g^{∞} beyond the analysts' long-term forecast horizon, which we set at ten years, is constant:

$$\mathcal{F}^{\mathbb{S}}_{d,t} = \ln \frac{EPS^{1y}_t}{EPS_t} + \rho_a \ln \frac{EPS^{2y}_t}{EPS^{1y}_t} + \frac{\rho_a^2(1-\rho_a^8)}{1-\rho_a} LTG_t + \frac{\rho_a^{10}}{1-\rho_a} g^{\infty}, \tag{16}$$

where $\rho_a = \rho^{12}$. [11] Further, we calibrate g^{∞} such that the unconditional mean of $\mathcal{F}^{\mathbb{S}}_{d,t}$ matches the unconditional mean of the price-to-earnings ratio.[12] Since g^{∞} does not contribute to the variation in $\mathcal{F}^{\mathbb{S}}_{d,t}$, it plays a minimal role in our analysis: while it determines the unconditional mean of the bias, it does not influence its time-varying component, which is the primary focus of our study. As can be seen from equation (16), our baseline approach considers the information contained in both short-term as well as long-term analyst forecasts.

As an alternative, we follow De La O and Myers (2021) and posit a "decay" model for cash-flow expectations:

$$\mathbb{E}^{\mathbb{S}}_t \Delta d_{t+1+h} - g^{\infty} = \phi^h \left(\mathbb{E}^{\mathbb{S}}_t \Delta d_{t+1} - g^{\infty} \right),$$

where the decay parameter ϕ (= 0.445) is estimated by regressing the two-year expected growth rate implied by earnings per share forecasts on the one-year expected growth rate. We then

[11]Following Bordalo et al. (2024a), we calibrate $\rho_a = (1 + e^{-pd})^{-1} \approx 0.9777$ where pd equals the unconditional sample mean of the log price-to-dividend ratio.

[12]That is, g is the average of the growth rate g_t obtained by solving, at each time t, the equation $pe_t = \frac{\kappa - r}{1-\rho_a} + \rho_a \ln \frac{EPS^{2y}_t}{EPS^{1y}_t} + \frac{\rho_a^2(1-\rho_a^8)}{1-\rho_a} LTG_t + \frac{\rho_a^{10}}{1-\rho_a} g_t$, where κ is defined as in Bordalo et al. (2024a) and r equals the unconditional mean return in our sample.

measure $\mathcal{F}^{\mathbb{S}}_{d,t}$ as

$$\mathcal{F}^{\mathbb{S}}_{d,t} = \frac{1}{1-\rho_a\phi} \ln \frac{EPS^{1y}_t}{EPS_t} + \left(\frac{1}{1-\rho_a} - \frac{1}{1-\rho_a\phi}\right) g^{\infty}. \tag{17}$$

As can be seen from equation (17), this second approach relies only on information contained in short-term analyst forecasts. As above, we calibrate g^{∞} such that the unconditional mean of $\mathcal{F}^{\mathbb{S}}_{d,t}$ matches the unconditional mean of the price-to-earnings ratio.

Finally, for each measure of $\mathcal{F}^{\mathbb{S}}_{d,t}$ we calculate cash-flow forecast revisions $\varepsilon^{\mathbb{S}}_{d,t+1} = (\mathbb{E}^{\mathbb{S}}_{t+1} - \mathbb{E}^{\mathbb{S}}_{t}) \sum_{h=0}^{\infty} \rho^h \Delta d_{t+1+h}$ as

$$\varepsilon^{\mathbb{S}}_{d,t} = \frac{\ln EPS_t - \ln EPS_{t-12}}{12} + \rho\mathcal{F}^{\mathbb{S}}_{d,t} - \mathcal{F}^{\mathbb{S}}_{d,t-1},$$

where we approximate the growth in realised monthly earnings per share with the average monthly growth rate over the last year. Note that simply calculating monthly first differences in $\mathcal{F}^{\mathbb{S}}_{d,t}$ would yield series that are highly correlated with the revisions $\varepsilon^{\mathbb{S}}_{d,t}$, with correlation coefficient equal to 0.96.

To construct investors' forecast of market earnings growth ($\mathcal{F}^{\mathbb{S}}_{d,t}$) and the associated cash-flow forecast revisions $\varepsilon^{\mathbb{S}}_{d,t+1}$ at the monthly frequency, we need to interpolate realized quarterly earnings from Compustat as well as earning per share forecasts at different horizons from Thomson Reuters I/B/E/S. From quarterly balance sheet data, we first compute the year-over-year cumulative earnings for each quarter t. Second, we perform a linear interpolation of cumulative earnings from two adjacent quarters to construct monthly realised cumulative earnings and earnings per share, respectively.[13] Further, we construct monthly analyst forecasts of earnings using the methodology on horizon interpolation outlined in the Appendix B of De La O and Myers (2021).

One might be concerned that analysts' earnings forecasts are distorted due to agency con-

[13] For more details on the measurement of earnings consistent with I/B/E/S's definition of earnings see A.2

flicts. However, Bordalo et al. (2019, 2024a) argue that such conflicts are unlikely to affect the time-series variation of forecasts, which is the focus of our analysis. Moreover, brokerage houses typically cover all S&P 500 firms. Therefore, investment banking relationships and analyst sentiment which could incentivize analysts to cover specific firms are not a concern in our sample. In addition, it is important to note that Thomson Reuters gathers their forecasts from hundreds of brokerage houses and independent analysts who track companies as part of their investment research work. Each individual estimate is identified by the name of the analyst or brokerage firm. Because the forecasts are not anonymous, analysts have a strong incentive to accurately report their expectations. In line with this conjecture, previous literature finds evidence that accuracy in earnings forecasts is important for tenure and compensation of analysts; see Mikhail et al. (1999) and Cooper et al. (2001). Furthermore, it has been shown that financial firms' behaviour correlates significantly with their own analysts' forecasts further substantiating the validity of the forecasts. For example, Bradshaw (2004) shows that individual earnings forecasts are correlated with Buy/Sell recommendations, while Chan et al. (2009) show that financial firms' own stock holding changes are significantly positively related to recommendation changes.

Another possible concern regarding the analysts forecasts is that respondents may be using risk-neutral probabilities in forming their expectations. This would imply that analysts overweight negative outcomes and, as a result, consistently under-predict future cash flows. However, that is not what we observe in the data. In fact, analysts forecast errors have zero or negative means, implying a slightly optimistic average bias, the opposite of what we would expect if analysts would indeed be applying risk-neutral probabilities.

3.2 Other data

VAR variables. We use returns of the S&P 500 index in excess of Treasury Bills as a proxy for excess returns on the market (r_t^e). Our market index choice is motivated by the coverage of analysts' earnings forecasts, as discussed above. To be consistent, we construct the price-to-dividend ratio (pd_t) for the S&P 500 by taking the difference between the logarithm of the price and the logarithm of the twelve-months smoothed dividends paid on S&P 500 stocks; see also Goyal and Welch (2008).

Following Campbell and Vuolteenaho (2004), the yield spread (ty_t) is defined as the difference between the long-term yield on government bonds and the short-term taxable Treasury bill yield; the small-stock value spread (vs_t) is measured as the difference between the log book-to-market ratios of small value and small growth stocks; and the default spread (def_t) is defined as the difference between the log yield on Moody's BAA and AAA bonds which are obtained from Amit Goyal's website.

Portfolios. For our cross-sectional asset pricing tests, we use various standard portfolios as test assets. That is, we first rely on a cross-section which consists out of the following 51 test assets: 30 portfolios sorted by market capitalization (ME), book-to-market ratio (BE/ME), and past returns, available from Professor Kenneth French's website; 20 portfolios sorted on past risk loadings as measured by our VAR state variables (following Campbell and Vuolteenaho (2004)); and the return on the S&P 500. Second, we rely on the data from Chen and Zimmermann (2022) which includes a large number of equity portfolios sorted by characteristics.

Survey data. The Livingstone survey is conducted by the Federal Reserve Bank of Philadelphia and have been used by Adam et al. (2021) and Dahlquist and Ibert (2024), among others. We obtain the series of expected returns implied by the survey provided by the Adam et al. (2021). The difference between the fractions of bullish and bearish investors from the AAII sentiment survey is obtained from the American Association of Individual Investors website. The median one-year expected return from the Graham-Harvey survey of CFOs is obtained from the Federal Reserve Bank of Richmond. The one-year individual investor stock market confidence index is from Yale SOM website.

Sentiment measures. Finally, we relate our work to various measures of beliefs and sentiment. We construct the time series for inflation and unemployment biases following the methodology of Bhandari et al. (2024). We obtain the series of stock market sentiment index proposed by Baker and Wurgler (2007) from Jeffrey Wurgler's website. The excess bond premium measure of Gilchrist and Zakrajsek (2012) is from the Federal Reserve Board.

Table 1 reports summary statistics for selected variables. For all variables we have monthly data between January 1982 and December 2020.

4 Empirical Results

4.1 Estimated VAR

In order to capture objective predictability of market returns and of analysts' earnings forecast revisions, we opt for a parsimonious VAR model with the following five state variables: the excess market return, the analysts' forecast of market earnings growth, the revision in analysts' forecast of market earnings growth, the market price-to-dividend ratio, and the yield spread between short-term and long-term government bonds. Panel A of Table 2 reveals that all model variables possess some predictive ability. It is noteworthy that the forecast level negatively predicts subsequent forecast revisions, indicating a pattern consistent with overreaction. This pattern is further supported by the negative coefficients estimated for the relationship between market returns and prior forecasts, as well as between forecast revisions and prior returns. At one-month horizon, the R^2 reaches 0.03 for market return and 0.18 for revision of analysts' forecast. Panel B of Table 2 reports the VAR model's predictive performance at longer horizons.[14,15]

[14]Including additional state variables, namely the small-stock value spread and the corporate bond default yield spread used in Campbell and Vuolteenaho (2004) and Campbell et al. (2013), does not considerably improve our VAR model's predictive ability. That said, our results remain robust when we consider these additional predictors.

[15]Appendix Table A-1 presents the results when we replace the forecast of cash flow growth calculated using equation (16) with those obtained solely from analysts' forecast of short-term earnings growth and an estimate of growth persistence, as per equation (17). The results align with those presented in Table 2.

Panel A of Figure 1 plots the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ implied by the estimated VAR, alongside the market price-to-dividend ratio. The correlation between the bias and the price-to-dividend ratio is 0.32 (0.46 for changes). This pro-cyclicality is consistent with the notion that investors overreact to information. That said, while the bias is pro-cyclical on average, at times we observe the opposite behavior. For instance, an optimistic bias can arise during downturns, consistent with underreaction. This phenomenon is most evident during the Global Financial Crisis: our estimates indicate that investors underestimated the crisis' impact on long-term cash-flow growth, attributing most of the drop in market valuation to an increase in discount rates. This finding aligns with Bianchi et al. (2024) who argue that investors were excessively optimistic about catch-up growth in the aftermath of the crisis.

Finally, Figure A-1 shows the smoothed series of cash-flow news $\varepsilon_{d,t}^{\mathbb{P}}$, discount-rate news $\varepsilon_{r,t}^{\mathbb{P}}$, and belief shocks $\mathbf{e}_2^{\intercal}\mathbf{u}_{t+1}$, where cash-flow and discount-rate news are identified following Campbell and Vuolteenaho (2004). We note that the alternative approach described in Section 2.5, where we first identify cash-flow news and derive discount-rate news as the residual, produces shocks that are highly correlated with baseline, with correlations of 0.74 for cash-flow news and 0.93 for discount-rate news.This finding partially addresses the critique by Chen and Zhao (2009), who argue that directly estimating discount-rate news and deriving cash-flow news as the residual can be problematic due to the potential inheritance of misspecification errors by the CF news. Our results suggest that any misspecification in our case is minimal.

4.2 Bias in expected market returns

We now consider to what extent the bias in investors' beliefs can explain the related stock-market volatility and market return predictability puzzles.

A pro-cyclical bias implies that the Campbell and Shiller (1988) variance decomposition will overestimate (underestimate) the fraction of stock-market volatility explained by the variation in discount rates (cash-flow growth) relative to investors' beliefs. In our sample, the decomposition under $\mathbb{P}$ attributes 98% of the variation in the price-to-dividend ratio to the variation in discount rates, aligning with the original findings by Campbell and Shiller (1988):

$$1 = \overbrace{\frac{\mathbb{C}\text{ov}[pd_t, \mathcal{F}^{\mathbb{P}}_{d,t}]}{\mathbb{V}\text{ar}\,[pd_t]}}^{0.02\ (0.48)} + \overbrace{\frac{\mathbb{C}\text{ov}[pd_t, -\mathcal{F}^{\mathbb{P}}_{r,t}]}{\mathbb{V}\text{ar}\,[pd_t]}}^{0.98\ (23.38)}.$$

Using equation (5) which states $\mathcal{B}^{\mathbb{P}-\mathbb{S}}_t = \mathcal{F}^{\mathbb{S}}_{d,t} - \mathcal{F}^{\mathbb{P}}_{d,t} = \mathcal{F}^{\mathbb{S}}_{r,t} - \mathcal{F}^{\mathbb{P}}_{r,t}$, we can re-write the decomposition as

$$1 = \overbrace{\underbrace{\frac{\mathbb{C}\text{ov}[pd_t, \mathcal{F}^{\mathbb{S}}_{d,t}]}{\mathbb{V}\text{ar}\,[pd_t]}}_{0.26\ (3.83)} - \frac{\mathbb{C}\text{ov}[pd_t, \mathcal{B}^{\mathbb{S}-\mathbb{P}}_t]}{\mathbb{V}\text{ar}\,[pd_t]}}^{0.02\ (0.48)} + \overbrace{\underbrace{\frac{\mathbb{C}\text{ov}[pd_t, -\mathcal{F}^{\mathbb{S}}_{r,t}]}{\mathbb{V}\text{ar}\,[pd_t]}}_{0.74\ (10.94)} + \underbrace{\frac{\mathbb{C}\text{ov}[pd_t, \mathcal{B}^{\mathbb{S}-\mathbb{P}}_t]}{\mathbb{V}\text{ar}\,[pd_t]}}_{0.24\ (2.19)}}^{0.98\ (23.38)} \tag{18}$$

$\frac{\mathbb{C}\text{ov}[pd_t, \mathcal{B}^{\mathbb{S}-\mathbb{P}}_t]}{\mathbb{V}\text{ar}[pd_t]} = 0.24$ implies that 24 percentage points of the price-to-dividend ratio variance that the decomposition under $\mathbb{P}$ attributes to discount-rate variation is due to expected cash-flow growth variation under $\mathbb{S}$. Thus, our results indicate that, even after adjusting for bias, discount-rate variation remains a significant driver of stock-market volatility, accounting for 74% of price-to-dividend ratio variance under investors' beliefs. This stands in contrast to the results in De La O and Myers (2021), who argue that beliefs about cash-flow growth explain the majority of stock-market volatility. Unlike these authors, we derive our conclusion from our estimate of the bias rather than directly from surveys of cash flow and market return expectations. However, the difference in our conclusions is not primarily due a different methodology but rather to our definition of earnings per share. Specifically, we exclude extraordinary items from earnings, whereas De La O and Myers (2021) include them. When we include extraordinary items, the decomposition (18) attributes only 17% price-to-dividend ratio variance to discount-rate variation under investor's beliefs. That said, following the discussion in Nagel (2024), we opt for excluding extraordinary items from earnings as our baseline.

Next, it is natural to ask whether the effect that distorted beliefs have on asset valuations can explain objective predictability in market returns. To answer this question, it is necessary to disentangle market return predictability which reflects belief distortions from that which is consistent with investor's beliefs. In doing so, we extend the analysis of Bordalo et al. (2024a) to incorporate time-varying beliefs about future market returns.[16]

Panel B of Figure 1 shows the series of investor beliefs about long-run expected market return $\mathcal{F}^{\mathbb{S}}_{r,t}$ measured using equation (5) as the difference between the VAR forecast of long-run expected returns $\mathcal{F}^{\mathbb{P}}_{r,t}$, and the bias $\mathcal{B}^{\mathbb{P}-\mathbb{S}}_t$: $\mathcal{F}^{\mathbb{S}}_{r,t} = \mathcal{F}^{\mathbb{P}}_{r,t} - \mathcal{B}^{\mathbb{P}-\mathbb{S}}_t$.[17] First, we observe that beliefs about future returns vary over time. This challenges a common simplifying assumption that attributes *all* of excess volatility (variation in objective expected returns) to beliefs about cash-flow growth.[18] Second, we find that representative investor's beliefs are positively correlated with the VAR forecast, with a correlation of 0.69 (0.00 for changes). That said, the two do not align perfectly. Finally, we note that the correlation between our belief measure and the lower bound on expected market return computed by Martin (2016) from index options, also plotted on Panel B of Figure 1, is 0.60 (0.30 for changes). By contrast, the correlation between this bound and the VAR forecast is substantially lower, at 0.17 (0.04 for changes). The comparison between our measures of long-run expected market returns and the option-implied lower bound on the one-year expected market return is arguably imperfect. That said, to the extent that the Martin (2016) bound reflects investors' beliefs embedded in option prices rather than the statistical predictability of market returns, this comparison supports our interpretation of the belief measure.[19]

[16]Specifically, we relax the assumption in equation (5) in Bordalo et al. (2024a); see also their footnote 15.

[17]Note that $\mathcal{F}^{\mathbb{S}}_{r,t}$, $\mathcal{F}^{\mathbb{P}}_{r,t}$,and $\mathcal{B}^{\mathbb{P}-\mathbb{S}}_t$ are measured using the same set of state variables.

[18]Indeed, assuming that beliefs about future returns $\mathcal{F}^{\mathbb{S}}_{r,t}$ and objective cash-flow growth expectations $\mathcal{F}^{\mathbb{P}}_{d,t}$ are constant, simplifies equation (5) to $\mathcal{F}^{\mathbb{P}}_{r,t} = \mathcal{F}^{\mathbb{S}}_{d,t}$; see equation (40) in Nagel (2024). However, we demonstrate that $\mathcal{F}^{\mathbb{S}}_{r,t}$ is, in fact, time-varying.

[19]It is also instructive to compare our belief measure to survey data. Surveys of expected returns vary considerably in their coverage, horizon, and methodology, which may explain their discrepancies regarding the volatility and cyclical properties of agents' beliefs. Our belief measure aligns most closely with expected returns implied by the Livingston survey, which includes a broad panel of forecasters, with a correlation of 0.01 (0.48 for changes). By contrast, the correlations between our measure and some other survey-based indices considered in the literature is, in fact, negative: -0.22 for the the AAII sentiment survey, -0.34 for the Graham-Harvey survey of CFOs, and -0.26 for the Shiller confidence index. See Appendix Figure A-2 for illustration. Thus, our results point to a

Turning to return predictability, Panel B of Table 2 shows that the VAR forecast explains 3% of market excess returns at the one-month horizon, 24% at the one-year horizon, and 51% in the long-run limit. However, the part of this predictability that is due to distorted beliefs cannot be readily gauged from predictive regressions because belief distortions may not be orthogonal to other factors driving expected returns. To disentangle the contribution of distorted beliefs we note that equation (5) implies a decomposition of expected return variance into the covariance with belief and the covariance with bias:

$$1 = \underbrace{\frac{\mathbb{C}\text{ov}[\mathcal{F}^{\mathbb{P}}_{r,t}, \mathcal{F}^{\mathbb{S}}_{r,t}]}{\mathbb{V}\text{ar}\left[\mathcal{F}^{\mathbb{P}}_{r,t}\right]}}_{0.59\ (6.39)} + \underbrace{\frac{\mathbb{C}\text{ov}[\mathcal{F}^{\mathbb{P}}_{r,t}, \mathcal{B}^{\mathbb{P}-\mathbb{S}}_{t}]}{\mathbb{V}\text{ar}\left[\mathcal{F}^{\mathbb{P}}_{r,t}\right]}}_{0.41\ (4.48)}. \tag{19}$$

This decomposition indicates that 41% of long-run market return predictability uncovered by the VAR model can be attributed to distorted beliefs, while the remaining 59% represent the variation in expected market returns that is consistent with the representative investor's beliefs.

Similar to decomposition in (18), the variance decomposition (19) is sensitive to the definition of earnings used. When extraordinary items are included in earnings measure, we find that distorted beliefs account for more than 100% of return predictability. This counterintuitive result occurs because, although beliefs based on earnings with extraordinary items show substantial variation, they predict future market returns very poorly.

4.3 Bias in the cross-section of stock returns

We now investigate whether distorted beliefs influence the pricing of risk in the cross-section of expected stock returns. To this end, we conduct cross-sectional asset pricing tests that incorporate bias shocks ($\varepsilon_{\mathcal{B},t}$). While these shocks manifest themselves as discount-rate risk in the data, they actually represent cash-flow risk from the investor's perspective, and vice versa.

possible tension between professional analysts' earnings forecasts and some of the surveys of expected returns. See also Bianchi et al. (2024), who make a similar point.

We test the significance of these shocks and assess the improvement in the model's pricing performance when these shocks are included as a way to gauge the wedge between pricing under investor's beliefs and the observed risk-return relationship.

We start by testing an ICAPM-type linear factor model that includes cash-flow news ($\varepsilon^{\mathbb{P}}_{d,t}$) and discount-rate news ($\varepsilon^{\mathbb{P}}_{r,t}$) and bias shocks ($\varepsilon_{\mathcal{B},t}$) as factors. Our test assets comprise a standard set of Fama-French portfolios sorted by market capitalization (ME), book-to-market ratio (BE/ME), and past returns. Following Campbell and Vuolteenaho (2004), we also construct a second set of portfolios sorted by past risk loadings using VAR state variables.[20] Using these 51 test assets—30 size-, book-to-market-, and momentum-sorted portfolios, 20 risk-sorted portfolios, and the market excess return—we evaluate the performance of the traditional two-factor ICAPM with cash-flow and discount-rate factors (columns 1-2 and 5-6 in Table 3), comparing it with the model that includes the bias shock. Each model is estimated in two forms: one with an unrestricted zero-beta rate (odd columns) and another with a restricted zero-beta rate equal to the Treasury bill rate (even columns). In both cases, we allow for unrestricted risk prices for cash-flow, discount-rate, and bias shocks. The first set of results (labeled Baseline) combines short- and long-term analysts' earnings forecasts as specified in equation (16), while the second set of results uses solely short-term analysts' earnings forecasts and an estimate of growth persistence, as specified in equation (17).

Table 3 shows that over the 1982-2020 period, the traditional two-factor ICAPM explains only a small fraction of the cross-section of stock returns, with an R^2 of approximately 3%. Introducing bias shocks considerably enhances the pricing performance, with models in columns (3) and (7) achieving R^2 values of 36% and 28%, respectively. Moreover, we cannot reject that null hypothesis that $R^2 = 1$. However, the χ^2 statistic rejects the null hypothesis that all pricing errors are jointly zero, implying that the two-factor model augmented with beliefs does

[20]The risk loadings are estimated by regressing each stock in the CRSP database, excluding the smallest five percent (by market equity), on market returns, as well as changes in term and value spreads. Similarly to Campbell and Vuolteenaho (2004) we exclude the (changes in) price-to-dividend ratio from this regression due to its high collinearity with market returns.

not fully explain the cross-section of returns on our 51 portfolios. Figure 2 visually summarizes these results by plotting the average realized returns of test portfolios against model-fitted values. While the risk-sorted, momentum, and, to a lesser extent, book-to-market portfolios align well along the 45-degree line—indicating that their returns are well-explained by the model—the size portfolios are scattered across the middle of the plot, suggesting that their fitted returns exhibit meaningful variation, whereas their realized returns do not.

The estimated risk price on bias shocks is significant with both a restricted zero-beta rate and an unrestricted Treasury bill rate, irrespective of whether analysts' long-term earnings forecasts are used. By contrast, estimates for discount-rate and cash-flow news exhibit larger standard errors. When using only short-term analyst forecasts, the risk price per unit of standard deviation in bias exposures is about 7% per year with an unrestricted zero-beta rate.[21] The risk price increases when combining long-term forecasts with short-term ones to compute beliefs.

The magnitude of risk prices also provides insight into the representative investor's risk aversion, as can be seen from equation (9). The implied coefficient of relative risk aversion, based on the estimated risk prices of cash-flow and bias shocks, is calculated as $\frac{0.096}{0.096-0.055} = 2.34$ in column (3) and $\frac{0.067}{0.067-0.029} = 1.78$ in column (7), both of which are well within the range of generally accepted values.

Adding 10 Fama-French industry portfolios to the set of tests assets does not diminish the model's pricing performance, with cross-sectional R^2 remaining at 29%. This is illustrated on Panel A of Figure 3, which plots average realized returns against model-fitted values for the industry portfolios, alongside the other 51 portfolios we consider. As shown in Panel B of Figure 3, the pricing of industry portfolios largely reflects their exposure to bias shocks, with business equipments (HiTec) and consumer durables (Durbl) showing the highest exposure, and oil, gas and coal extraction and products (Enrgy) and utilities (Utils) showing the lowest.

Finally, we note that results in Table 3 are robust to including extraordinary items in earn-

[21] The standard deviation of belief betas is 0.2028, leading to a calculation of $0.029 \times 0.2028 \times 1200 = 7.13\%$.

ings. Appendix Table A-2 replicates the results from Table 3 using this alternative measure of earnings. Notably, bias shocks remain statistically significant across specifications. The models in columns (3) and (7) achieve higher R^2 values than in Table 3, supporting the notion that including extraordinary items tends to overestimate belief distortions. In sum, our cross-sectional asset pricing results remain robust when using both short-term earnings forecasts and earnings that include extraordinary items.

Although our model incorporating beliefs is quite successful in explaining this particular cross-section of average returns, an asset pricing model can spuriously explain the average returns of characteristics-sorted portfolios related to average returns if premia are unconstrained by theory; see Daniel and Titman (1997). To alleviate this concern, we follow the supervised-PCA approach of Giglio et al. (2024), which also accommodates weak latent factors.

We present these results for both bias shocks ($\varepsilon_{\mathcal{B},t}$) and belief shocks ($\mathbf{e}_2^\intercal \mathbf{u}_t$). By construction, the difference between these two shocks is equal to cash-flow news. Thus, considering beliefs shocks can be seen a robustness test to our estimate of cash-flow news.

We use data from Chen and Zimmermann (2022), which includes a large number of equity portfolios sorted by characteristics. For each plausible number of latent factors, p, in our analysis, we have a tuning parameter to select: the number of assets chosen. Following Giglio et al. (2024), the first half of the sample (training period) is used to choose the tuning parameters and estimate the risk premium. The second half of the sample (evaluation period) is used to evaluate the out-of-sample performance of the estimator and the choice of the tuning parameter. Motivated by the scree plot, we present results for p equal to 5, 7, and 11, demonstrating the robustness of our findings across a range of model dimensions.

Table 4 presents the results. The columns labeled ”Baseline” report outcomes based on analysts' forecasts of both short-term and long-term earnings growth. The columns labeled ”Short-term forecasts only” show results using analysts' forecasts of short-term earnings growth. The top panel covers bias shocks, while the bottom panel addresses belief shocks. For each choice of latent factors ($p = 5, 7,$ and 11), we report the SPCA risk-premium estimate and the number of assets selected by SPCA. In these estimates, asset selection is driven solely by the respective shock.

The SPCA estimates of the bias (belief) premia for $p = 5, 7, 11$ are, respectively, 36, 56, and 78 (17, 30, and 46) basis points per month, in the baseline case that uses both short- and long-term forecasts. The corresponding premia when using only short-term earnings forecasts are similar in magnitude, and equal to 54, 80, and 109 (79, 60, and 83) basis points. These estimates are stable across specifications and statistically significant. Importantly, these values align closely with the estimates in Table 3. For instance, in the baseline case, the SPCA mimicking portfolio with 11 latent factors has a loading of 0.186 on the non-tradable bias factor. This results in a risk premium of $0.055 \times 0.186 = 102$ basis points (where 0.055 is the risk price from Table 3), which is smaller but close to the average SPCA excess return. Similarly, when considering the bias shock using only short-term forecasts, the SPCA mimicking portfolio with 11 latent factors has a loading on the non-tradable bias factor of 0.396, delivering a risk premium of $0.029 \times 0.396 = 115$ basis points (where 0.029 is the risk price from Table 3) which is slightly smaller but close to the average SPCA excess return.

To better understand the performance of the SPCA estimator, we can examine the out-of-sample R^2. Note that it is well known in the literature that it is difficult to hedge nontradable factors, like consumption or IP growth, in equity markets. Furthermore, we should not expect large R^2 since SPCA attempts to build a hedging portfolio for the target – in our case bias or beliefs – with factors that must also explain covariation among the universe of test assets. Therefore, an advantage of the SPCA approach is that the hedging portfolio is able to avoid fitting the "measurement error" component in the target, which can be thought of measurement error for nontradable factors like ours.

Interestingly, the out-of-sample R^2 for the baseline bias shock, constructed using both long- and short-term earnings forecasts, ranges from about 4% to 6%, while it is 3% to 4% when using only short-term forecasts. This result has both economic and statistical interpretations. Economically, the positive, albeit modest, out-of-sample R^2 indicates the portfolio built by SPCA for beliefs and biases offers good hedging ability. Statistically, it confirms that the asset selection made by SPCA in the training sample remains valid out of sample. However, the large number of assets selected (ranging from 600 to 950) highlights the challenge of hedging beliefs or biases – similar to most macro factors – using equity markets.

In Table 5, we analyze which assets are selected when extracting latent factors. The portfolio names follow Chen and Zimmermann (2022), with numbers indicating the quintile or decile of the characteristic. For clarity, we group portfolios into six economic categories based on the firm characteristics used to construct them: Intangibles, Investment, Momentum, Profitability, Trading Frictions, Value vs. Growth, and Other.[22]

Focusing on Panel A, SPCA selects assets with the highest correlation to the bias shock to extract the first latent factor. The table reveals that these assets are associated with value, trading frictions, and investment. The highest correlation is achieved by Duration_2 (the second-lowest duration portfolio), with a correlation of 0.58. Interestingly, several of the top assets chosen to hedge the bias shock, such as Duration_2, Investment_3, and Investment_4, are also selected to hedge the belief shock.

The remainder of the table shows assets selected in different iterations. In the second iteration, after orthogonalizing beliefs and test assets to the first factor, selected assets are mostly related to profitability and momentum. In the third iteration, portfolios sorted on share issuance and repurchases play an important role in hedging risk. This aligns with the model by Daniel et al. (2020), which proposes using a long-horizon factor that exploits managers' decisions to issue or repurchase equity in response to persistent mispricing. Overall, the SPCA procedure selects similar assets for both belief and bias shocks.

[22]When the factor is included in Hou et al. (2020), we follow their classification.

4.4 Sources of the bias

We now examine the sources of bias in investors' expectations about the stock market. An advantage of our approach, which employs multivariate forecasts of market returns and analysts' forecast revisions, is that it does not assume any particular belief formation mechanism or rely on a specific proxy for such a mechanism. Nonetheless, it is instructive to investigate which measures of beliefs and sentiment suggested in prior literature are associated with our estimate of the bias. Specifically, we consider the differences between survey and VAR forecasts of inflation and unemployment rates, as constructed by Bhandari et al. (2024), the stock market sentiment index proposed by Baker and Wurgler (2006), and a measure of risk appetite in the corporate bond market proposed by Gilchrist and Zakrajsek (2012).

Panel A of Table 6 shows that the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ tends to be positive (negative) when surveys underestimate (overestimate) future unemployment relative to a VAR forecast, which aligns with economic intuition. We do not find a similar statistically significant relationship for inflation, which may be due to structural shifts in the comovement of inflation with the business cycle and stock market within our sample period; see Cieslak and Pflueger (2023). Additionally, we find a positive and statistically significant relationship between our bias measure and the stock market sentiment index proposed by Baker and Wurgler (2006).

Finally, we find a positive and statistically significant relationship between the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ and the excess bond premium computed by Gilchrist and Zakrajsek (2012). The excess bond premium serves a measure of the risk appetite in the corporate bond market, where a higher premium indicates lower risk appetite. Thus, our findings suggest that investors tend to be overly optimistic (pessimistic) about the stock market when the risk appetite in corporate the bond market is low (high).

One possible interpretation of this relationship is that stock-market investors underreact to changes in corporate bond markets conditions. For instance, their expectations could remain overly optimistic compared to a forecast that fully considers worsening credit conditions, resulting in a positive bias. To support this interpretation, we examine the causality between the excess bond premium and our bias measure $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$. Panel B of Table 6 shows that the excess bond premium Granger-causes the bias, but not vice versa. We leave further investigation into the interaction between beliefs and risk appetite across different markets to future research.

The univariate regression results show that each variable explains only a small portion of stock market bias, with the highest univariate R^2 value of 0.137 observed for the unemployment bias. The multivariate regression yields an R^2 value of 0.372, indicating that our bias measure captures additional dimensions of belief distortions that are not fully accounted for by the explanatory variables considered.

5 Conclusion

We show that accounting for investors' distorted beliefs—i.e., the wedge between between their beliefs and objective expectations—provides new insights into apparent empirical puzzles observed in both the time series and the cross-section of expected stock returns. Our findings offer compelling evidence for the important role played by investors' expectational biases in the stock market. Furthermore, our results suggest that investors' beliefs about the stock market are interconnected to their beliefs about the credit market and the macroeconomy. Exploring the interaction between beliefs across different markets is an interesting avenue for future research. The analytical framework developed in this paper can be instrumental in pursuing this line of inquiry.

References

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

Adam, K., D. Matveev, and S. Nagel (2021). Do survey expectations of stock returns reflect risk adjustments? *Journal of Monetary Economics 117*, 723–740.

Baker, M. and J. Wurgler (2002). Market timing and capital structure. *The Journal of Finance 57*(1), 1–32.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance 61*(4), 1645–1680.

Baker, M. and J. Wurgler (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives 21*(2), 129–152.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49*(3), 307–343.

Bhandari, A., J. Borovicka, and P. Ho (2024). Survey data and subjective beliefs in business cycle models. *Review of Economic Studies forthcoming*.

Bianchi, F., S. C. Ludvigson, and S. Ma (2022). Belief distortions and macroeconomic fluctuations. *American Economic Review 112*(7), 2269–2315.

Bianchi, F., S. C. Ludvigson, and S. Ma (2024). What hundreds of economic news events say about belief overreaction in the stock market. *Working Paper*.

Bohrnstedt, G. W. and A. S. Goldberger (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association 64*(328), 1439–1442.

Bordalo, P., N. Gennaioli, R. LaPorta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *The Journal of Finance 74*(6), 2839–2874.

Bordalo, P., N. Gennaioli, R. LaPorta, and A. Shleifer (2024a). Belief overreaction and stock market puzzles. *Journal of Political Economy forthcoming*.

Bordalo, P., N. Gennaioli, R. LaPorta, and A. Shleifer (2024b). Finance without exotic risk. *NBER Working Paper 33004*.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in macroeconomic expectations. *American Economic Review 110*(9), 2748–82.

Bradshaw, M. T. (2004). How do analysts use their earnings forecasts in generating stock recommendations? *Accounting Review 79*(1), 25–50.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The Economic Journal 101*(405), 157–179.

Campbell, J. Y. (1993). Intertemporal asset pricing without consumption data. *The American Economic Review 83*(3), 487–512.

Campbell, J. Y., S. Giglio, and C. Polk (2013). Hard times. *The Review of Asset Pricing Studies 3*(1), 95–132.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2018). An intertemporal capm with stochastic volatility. *Journal of Financial Economics 128*(2), 207–233.

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

Campbell, J. Y. and T. Vuolteenaho (2004). Bad beta, good beta. *American Economic Review 94*(5), 1249–1275.

Cassella, S., T.-F. Chen, H. Gulen, and Y. Liu (2024). Extracting extrapolative beliefs from market prices: An augmented present-value approach. *Journal of Financial Economics forthcoming*.

Chan, K. W., C. Chang, and A. Wang (2009). Put your money where your mouth is: Do financial firms follow their own recommendations? *The Quarterly Review of Economics and Finance 49*(3), 1095–1112.

Chen, A. Y. and T. Zimmermann (2022). Open source cross-sectional asset pricing. *Critical Finance Review 11*(2), 207–264.

Chen, L. and X. Zhao (2009). Return decomposition. *The Review of Financial Studies 22*(12), 5213–5249.

Cieslak, A. and C. Pflueger (2023). Inflation and asset returns. *Annual Review of Financial Economics 15*, 433–448.

Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

Cooper, R. A., T. E. Day, and C. M. Lewis (2001). Following the leader: a study of individual analysts' earnings forecasts. *Journal of Financial Economics 61*(3), 383–416.

Dahlquist, M. and M. Ibert (2024). Equity Return Expectations and Portfolios: Evidence from Large Asset Managers. *The Review of Financial Studies 37*(6), 1887–1928.

Daniel, K., D. Hirshleifer, and L. Sun (2020). Short- and long-horizon behavioral factors. *The Review of Financial Studies 33*(4), 1673–1736.

Daniel, K. and S. Titman (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance 52*(1), 1–33.

De la O, R., X. Han, and S. Myers (2024). The cross-section of subjective expectations: Understanding prices and anomalies. *Mimeo*.

De La O, R. and S. Myers (2021). Subjective cash flow and discount rate expectations. *Journal of Finance 76*(3), 1339–1387.

Gao, C. and I. W. R. Martin (2021). Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *The Journal of Finance 76*(6), 3211–3254.

Giglio, S., D. Xiu, and D. Zhang (2024). Test assets and weak factors. *Journal of Finance forthcoming*.

Gilchrist, S. and E. Zakrajsek (2012). Credit spreads and business cycle fluctuations. *American Economic Review 102*(4), 1692–1720.

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

Greenwood, R. and A. Shleifer (2014). Expectations of returns and expected returns. *Review of Financial Studies 27*(3), 714–746.

Hillenbrand, S. and O. McCarthy (2022). Heterogeneous investors and stock market fluctuations. *Mimeo*.

Hou, K., C. Xue, and L. Zhang (2020, 12). Replicating anomalies. *Review of Financial Studies 33*(5), 2019–2133.

Jensen, T. I. (2024). Subjective risk and return. *Working Paper*.

Kozak, S., S. Nagel, and S. Santosh (2018). Interpreting factor models. *The Journal of Finance 73*(3), 1183–1223.

Martin, I. (2016). What is the expected return on the market? *Quarterly Journal of Economics 132*(1), 367–433.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica 41*(5), 867–887.

Mikhail, M. B., B. R. Walther, and R. H. Willis (1999). Does forecast accuracy matter to security analysts? *Accounting Review 74*(2), 185–200.

Nagel, S. (2024). Stock market valuation: Explanations, non-explanations, and some open questions. *Mimeo*.

Nagel, S. and Z. Xu (2021). Asset pricing with fading memory. *Review of Financial Studies 35*(5), 2190–2245.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review 71*(3), 421–36.

Table 1: **Summary Statistics.** This Table presents summary statistics of our state variables in the VAR estimations. Note that we do not report the mean and the median of the two alternative definitions of $\mathcal{F}^{\mathbb{S}}_{d,t}$ as they are simply calibrated to the unconditional mean and median of the price-to-earnings ratio in the data as discussed in the Section 3.1. Data is monthly from January 1982 through December 2020.

Variable	Mean	Median	Std.	Min	Max
Market return, r^e_t	0.0085	0.0105	0.0435	-0.1985	0.1401
Price-to-dividend ratio, pd_t	3.7803	3.8858	0.3716	2.7533	4.5236
Term spread, ty_t	2.2794	2.3400	1.2681	-0.4100	4.5500
Variables according to Bordalo et al. (2024a)					
Forecast, $\mathcal{F}^{\mathbb{S}}_{d,t}$	-	-	0.1585	2.1399	2.8645
Short-term component of $\mathcal{F}^{\mathbb{S}}_{d,t}$	0.1812	0.1638	0.0790	0.0123	0.4446
Long-term component of $\mathcal{F}^{\mathbb{S}}_{d,t}$	0.1218	0.1174	0.0170	0.0953	0.1828
Forecast revision, $\varepsilon^{\mathbb{S}}_{d,t}$	0.0007	0.0023	0.0454	-0.1786	0.2335
Variables according to De La O and Myers (2021)					
Forecast, $\mathcal{F}^{\mathbb{S}}_{d,t}$	-	-	0.1399	2.0795	2.8447
Forecast revision, $\varepsilon^{\mathbb{S}}_{d,t}$	0.0003	0.0014	0.0666	-0.2895	0.3430
Additional variables					
Value spread, vs_t	1.5250	1.4949	0.1958	1.1691	2.2310
Default spread, def_t	1.0505	0.9300	0.4332	0.5500	3.3800

Table 2: **VAR model.** Panel A of this table reports the coefficients on the explanatory variables in the VAR model. Each row corresponds to a different equation of the model. The variables include the market excess return (r^e_t), the revisions in analysts' forecast of market earnings growth ($\varepsilon^{\mathbb{S}}_{d,t}$), analysts' forecast of market earnings growth ($\mathcal{F}^{\mathbb{S}}_{d,t}$), the market price-to-dividend ratio (pd_t), and the yield spread between short-term and long-term government bonds (ty_t). Ordinary least squares t-statistics are reported in parentheses below the coefficients. The last column reports the R^2. Panel B reports the fraction of the variation in excess market return and in the revisions in analysts' forecast of market earnings growth at H-month horizon explained by the VAR forecast computed from the estimated VAR model. Analysts' forecast combines short-term and long-term earnings growth forecasts as specified in equation (16). Data are monthly from January 1982 through December 2020.

Panel A: model equations						
	r^e_t	$\varepsilon^{\mathbb{S}}_{d,t}$	$\mathcal{F}^{\mathbb{S}}_{d,t}$	pd_t	ty_t	R^2
Market return, r^e_{t+1}	0.031	-0.030	-0.038	-0.009	-0.002	0.029
	(0.596)	(-0.591)	(-3.028)	(-1.647)	(-0.888)	
Forecast revision, $\varepsilon^{\mathbb{S}}_{d,t+1}$	-0.309	-0.073	-0.071	0.005	-0.003	0.180
	(-6.178)	(-1.507)	(-5.886)	(0.963)	(-2.134)	
Forecast, $\mathcal{F}^{\mathbb{S}}_{d,t+1}$	-0.287	-0.141	0.958	0.002	-0.004	0.935
	(-5.799)	(-2.970)	(80.228)	(0.365)	(-2.440)	
Price-to-dividend ratio, pd_{t+1}	0.069	-0.055	-0.017	0.988	-0.001	0.986
	(1.278)	(-1.058)	(-1.269)	(172.035)	(-0.777)	
Term spread, ty_{t+1}	0.690	-0.899	0.201	-0.045	0.964	0.936
	(1.751)	(-2.370)	(2.110)	(-1.085)	(78.720)	

Panel B: predictability at H-month horizon					
	$H = 1$	$H = 3$	$H = 12$	$H = 120$	$H \to \infty$
$\frac{\mathbb{V}\mathrm{ar}[\mathbb{E}^{\mathbb{P}}_t \sum_{h=0}^{H-1} \rho^h r^e_{t+1+h}]}{\mathbb{V}\mathrm{ar}[\sum_{h=0}^{H-1} \rho^h r^e_{t+1+h}]}$	0.029	0.077	0.239	0.536	0.513
$\frac{\mathbb{V}\mathrm{ar}[\mathbb{E}^{\mathbb{P}}_t \sum_{h=0}^{H-1} \rho^h \varepsilon^{\mathbb{S}}_{d,t+1+h}]}{\mathbb{V}\mathrm{ar}[\sum_{h=0}^{H-1} \rho^h \varepsilon^{\mathbb{S}}_{d,t+1+h}]}$	0.180	0.225	0.540	0.535	0.243

Table 3: **ICAPM.** This table reports the results of the cross-sectional regression of mean excess portfolio returns on portfolio betas estimated in the first-pass univariate regressions. The model is estimated using 51 portfolios: 20 risk portfolios, 10 Fama-French book-to-market ranked portfolios, 10 Fama-French size ranked portfolios, 10 momentum ranked portfolios, and the market portfolio. The table reports the estimates of the constant term and the factor risk premia for cash-flow news ($\varepsilon^{\mathbb{P}}_{d,t}$), discount-rate news ($\varepsilon^{\mathbb{P}}_{r,t}$), and bias shock ($\varepsilon_{\mathcal{B},t}$). t-statistics based on the GMM-VARHAC standard errors for these estimates are reported in the parentheses. The table also reports the R^2 of the cross-sectional regression and its p-value under the null that $R^2 = 1$. Finally, the table reports the statistic and the associated p-value of the chi-squared test of the null hypothesis that all pricing errors are jointly zero. Columns (1)-(4) report the results in which we use the analysts' forecast of both short-term and long-term earnings growth as specified in equation (16). Columns (5)-(8) report the results in which we use analysts' forecast of short-term earnings growth and an estimate of growth rate persistence as specified in equation (17). Data are monthly from January 1982 through December 2020.

	Baseline				Short-term forecast only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.009		0.008		0.009		0.007	
	(5.409)		(4.005)		(5.746)		(4.154)	
Cash-flow news, $\varepsilon^{\mathbb{P}}_{d,t}$	0.005	-0.010	0.096	0.135	0.008	-0.035	0.067	0.084
	(0.382)	(-0.697)	(2.499)	(2.669)	(0.487)	(-1.884)	(2.254)	(2.236)
Discount-rate news, $\varepsilon^{\mathbb{P}}_{r,t}$	0.001	-0.005	0.023	0.031	0.001	-0.006	0.019	0.028
	(0.235)	(-2.141)	(2.317)	(1.895)	(0.373)	(-1.955)	(2.255)	(2.243)
Bias shock, $\varepsilon_{\mathcal{B},t}$			0.055	0.087			0.029	0.052
			(2.480)	(2.641)			(2.389)	(3.028)
R^2	0.032		0.356		0.033		0.282	
	(0.018)		(0.304)		(0.031)		(0.113)	
χ^2	81.651	110.073	80.183	108.326	81.658	109.287	79.518	105.090
	(0.002)	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)

Table 4: **Pricing of bias and belief shocks.** This table reports the risk premia for (nontradable) bias shocks ($\varepsilon_{\mathcal{B},t}$) and belief shocks ($\mathbf{e}_2^\intercal \mathbf{u}_t$) estimated using the supervised principal component analysis (SPCA) procedure proposed by Giglio et al. (2024). The columns report, for different number of latent factors, the estimated risk premia (in percent per month, computed in the training period), the associated t-statistic (in parentheses below), the number of assets selected by SPCA, and the out of sample R^2 (OOS R^2). The universe of assets consists of Chen and Zimmermann (2022) portfolios augmented with 49 industry portfolios. Columns denoted Baseline report the results in which we use the analysts' forecast of both short-term and long-term earnings growth as specified in equation (16). Columns denoted Short-term forecasts only report the results in which we use analysts' forecast of short-term earnings growth and an estimate of growth rate persistence as specified in equation (17). Data are monthly from January 1982 through December 2020.

Bias shocks, $\varepsilon_{\mathcal{B},t}$						
	Baseline			Short-term forecasts only		
# factors	Premium	# assets	OOS R^2	Premium	# assets	OOS R^2
5	36	850	0.046	54	750	0.042
	1.565			2.077		
7	56	950	0.064	80	900	0.035
	2.333			3.077		
11	78	950	0.043	109	950	0.030
	3.000			3.759		

Belief shocks, $\mathbf{e}_2^\intercal \mathbf{u}_t$						
	Baseline			Short-term forecasts only		
# factors	Premium	# assets	OOS R^2	Premium	# assets	OOS R^2
5	17	750	0.133	79	600	0.120
	1.063			3.292		
7	30	950	0.145	60	800	0.116
	1.875			2.609		
11	46	950	0.122	83	950	0.101
	2.556			3.320		

Table 5: **Latent factor portfolios.** This table reports for bias shocks ($\varepsilon_{\mathcal{B},t}$) and belief shocks ($\mathbf{e}_2^\intercal \mathbf{u}_t$) the top-10 portfolios selected by the supervised principal component analysis (SPCA) procedure proposed by Giglio et al. (2024) when extracting the latent factors. Portfolios are sorted by the absolute value of the correlation with shocks. For each latent factor from 1 to 3, the table reports the economic category and the names of the portfolios selected, and the absolute value of the correlation. Naming convention for the portfolios follows Chen and Zimmermann (2022). Results use the analysts' forecast of both short-term and long-term earnings growth combined as specified in equation (16). Data are monthly from January 1982 through December 2020.

Panel A: Bias shock, $\varepsilon_{\mathcal{B},t}$

Factor #1			Factor #2			Factor #3		
Category	Asset	|Corr.|	Category	Asset	|Corr.|	Category	Asset	|Corr.|
Value vs. growth	EquityDuration_2	0.5831	Profitability	roaq_8	0.4103	Intangibles	ChEQ_4	0.2831
Profitability	CBOperProf_10	0.5729	Momentum	NumEarnIncrease_1	0.3916	Investment	ShareIss5Y_4	0.2668
Investment	Investment_3	0.5686	Profitability	ChAssetTurnover_4	0.3895	Value vs.growth	BM_4	0.2564
Intangibles	realestate_2	0.5616	Investment	ShareIss5Y_2	0.3861	Investment	TotalAccruals_4	0.2331
Profitability	OperProfRD_10	0.5612	Profitability	RoE_2	0.3831	Investment	DelNetFin_3	0.2315
Profitability	GP_5	0.5606	Intangibles	ChEQ_2	0.3823	Value vs.growth	SP_3	0.2250
Market	Market	0.5596	Other	Price_2	0.3820	Investment	ShareRepurchase_2	0.2228
Other	Spinoff_1	0.5589	Momentum	EarningsSurprise_4	0.3794	Intangibles	GrSaleToGrInv_4	0.2170
Trading frictions	Coskewness_2	0.5528	Value vs. growth	EquityDuration_2	0.3787	Investment	ShareIss5Y_5	0.2168
Investment	Investment_4	0.5502	Other	Price_3	0.3775	Investment	DelCOL_5	0.2162

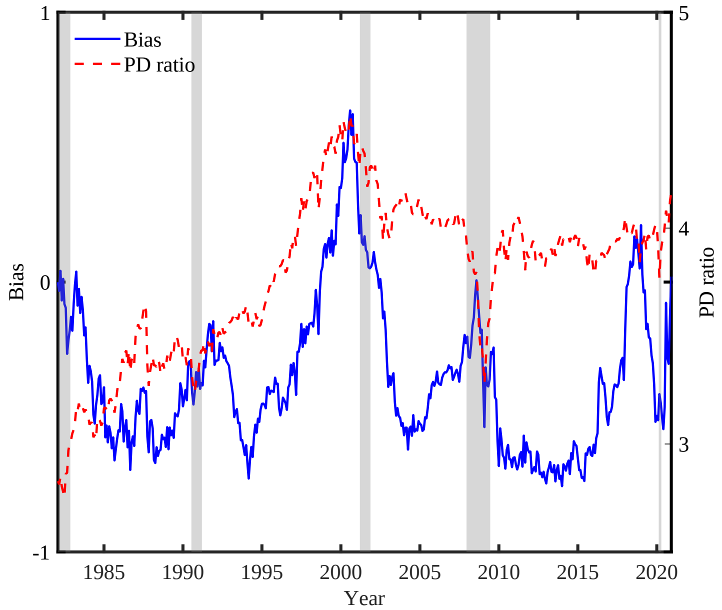
Panel B: Belief shock, $\mathbf{e}_2^\intercal \mathbf{u}_t$

Factor #1			Factor #2			Factor #3		
Category	Asset	|Corr.|	Category	Asset	|Corr.|	Category	Asset	|Corr.|
Value vs. growth	EquityDuration_2	0.6430	Profitability	ChAssetTurnover_4	0.3875	Intangibles	ChEQ_4	0.2813
Trading frictions	MaxRet_6	0.6351	Value vs.growth	BPEBM_2	0.3622	Value vs.growth	BMdec_3	0.2658
Other	Spinoff_1	0.6343	Profitability	OperProf_4	0.3486	Investment	CompositeDebtIssuance_1	0.2455
Market	Market	0.6341	Value vs.growth	EBM_6	0.3453	Investment	ShareIss5Y_4	0.2411
Investment	Investment_3	0.6303	Profitability	Tax_5	0.3430	Value vs.growth	SP_3	0.2391
Trading frictions	Coskewness_2	0.6261	Momentum	EarningsSurprise_8	0.3409	Other	LRreversal_4	0.2277
Investment	Investment_4	0.6259	Momentum	EarningsSurprise_4	0.3406	Momentum	ResidualMomentum_4	0.2276
Intangibles	OrgCap_1	0.6220	Profitability	RoE_4	0.3404	Investment	ShareRepurchase_2	0.2275
Profitability	CBOperProf_10	0.6217	Profitability	roaq_8	0.3389	Investment	DelCOL_4	0.2274
Trading frictions	IdioVol3F_4	0.6206	Momentum	NumEarnIncrease_1	0.3375	Investment	ShareIss1Y_5	0.2265

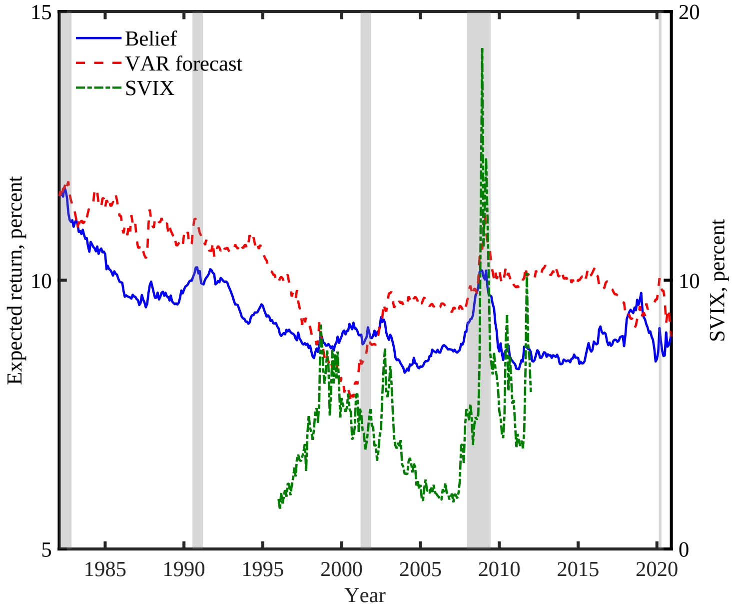
Table 6: **Sources of stock-market bias.** Panel A reports the slope coefficients from regressions of the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ on the differences between survey and VAR forecasts of, respectively, inflation (Inflation bias) and unemployment (Unemployment bias) rates computed following Bhandari et al. (2024), the Baker and Wurgler (2006) stock market sentiment index (Sentiment), and the Gilchrist and Zakrajsek (2012) excess bond premium (Excess bond premium). Newey and West (1987, 1994) t-statistics are reported in parentheses. Panel B reports the F-statistic and the associated p-values from the test of Granger causality between the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ and the excess bond premium (EBP). The VAR for the causality test includes 6 lags chosen by the Akaike information criterion. Results use the analysts' forecast of both short-term and long-term earnings growth combined as specified in equation (16). Data are monthly and run from January 1982 to December 2020.

Panel A: Bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$					
Inflation bias	-0.019				-0.005
	(-0.849)				(-0.183)
Unemployment bias		-0.104			-0.115
		(-3.276)			(-3.783)
Sentiment			0.148		0.143
			(3.305)		(4.745)
Excess bond premium				0.146	0.149
				(2.507)	(4.394)
R^2	0.004	0.137	0.128	0.094	0.372

Panel B: Granger causality	F-stat.	p-val.
Bias does not Granger cause EBP	1.603	0.145
EBP does not Granger cause Bias	3.957	0.001

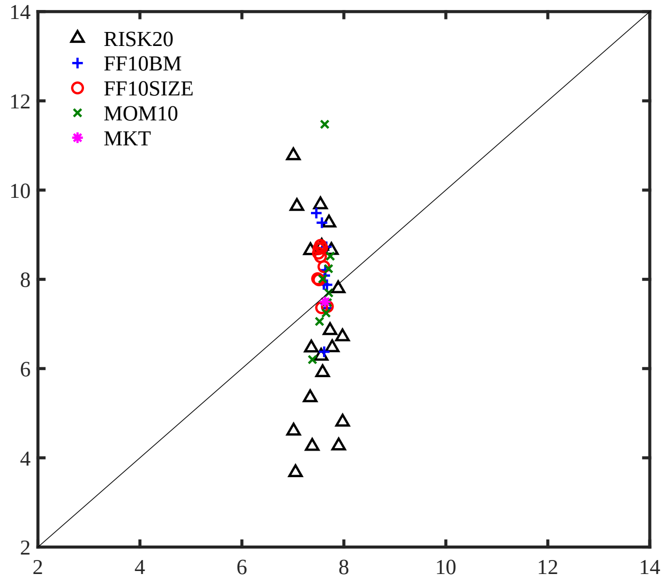


A Bias

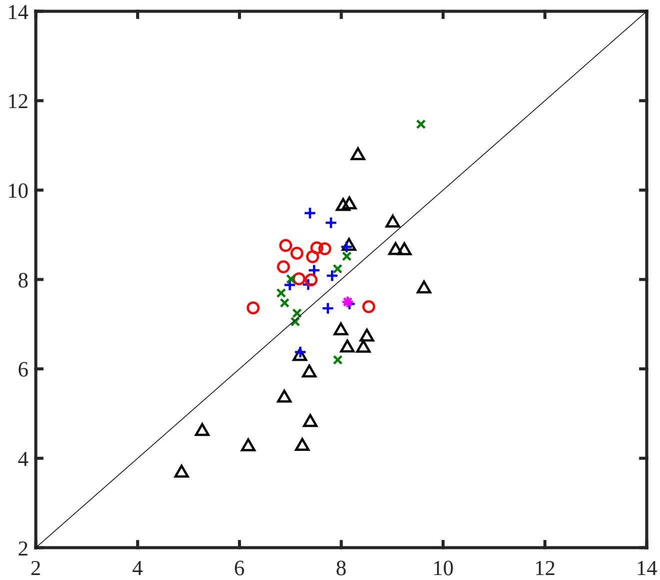


B Expected market returns

Figure 1: **Bias and beliefs**: Panel A plots the bias $\mathcal{B}_t^{\mathbb{S}-\mathbb{P}}$ estimated as the long-run predictable component of analysts' forecast revisions as per equation (7). The panel also shows the market price-to-dividend ratio. Shaded areas correspond to NBER recessions. Panel B plots the VAR forecast of the long-term market returns $\mathcal{F}_{r,t}^{\mathbb{P}}$ and the belief about long-term market returns $\mathcal{F}_{r,t}^{\mathbb{S}}$ obtained as the VAR forecast adjusted for the bias using equation (5). Results use the analysts' forecast of both short-term and long-term earnings growth combined as specified in equation (16). The panel also shows the lower bound on expected market return (at one-year horizon) from Martin (2016).

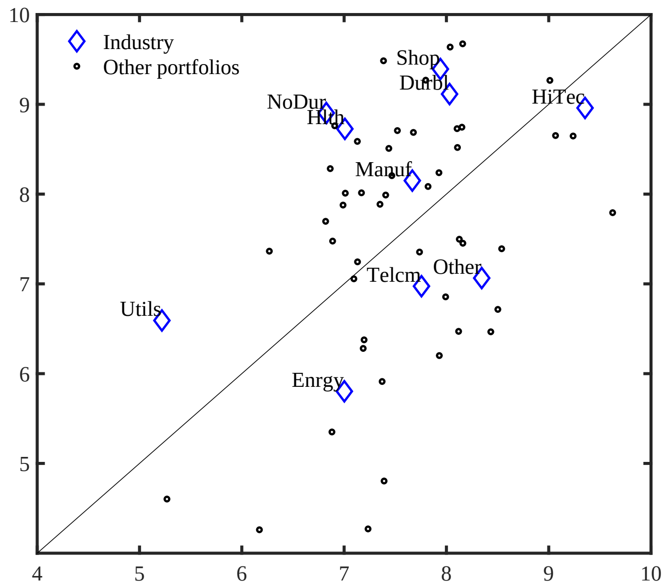


A ICAPM

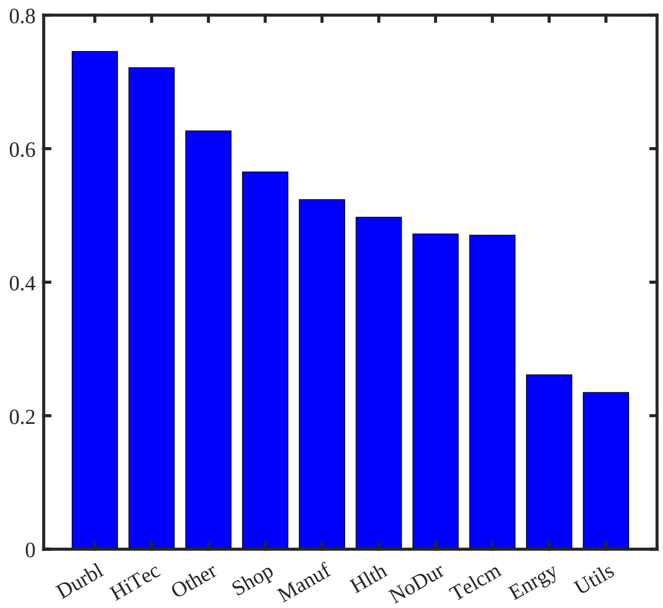


B ICAPM with distorted beliefs

Figure 2: **Performance of ICAPM with distorted beliefs**: The figure plots the predicted mean excess returns (x-axis) against the realized mean excess returns (y-axis) for 51 portfolios including 20 risk portfolios, 10 Fama-French book-to-market ranked portfolios, 10 Fama-French size ranked portfolios, 10 momentum ranked portfolios, and the market portfolio. Predicted returns correspond to regressions (1) and (3) in Table 3. In panel A corresponding to regression (1), the factors are cash-flow news ($\varepsilon^{\mathbb{P}}_{d,t}$) and discount-rate news ($\varepsilon^{\mathbb{P}}_{r,t}$); in panel B corresponding to regression (3), the factors are cash-flow news ($\varepsilon^{\mathbb{P}}_{d,t}$), discount-rate news ($\varepsilon^{\mathbb{P}}_{r,t}$), and the bias shock ($\varepsilon_{\mathcal{B},t}$).



A Predicted and realized returns



B Betas to bias risk

Figure 3: **Pricing of industry portfolios**: Panel A plots the predicted mean excess returns (x-axis) against the realized mean excess returns (y-axis) for 10 Fama-French industry portfolios—namely, consumer nondurables (NoDur), consumer durables (Durbl), manufacturing (Manuf), oil, gas and coal extraction and products (Enrgy), Business Equipment (HiTec), telephone and television transmission (Telcm), wholesale, retail, and some services (Shops), healthcare, medical equipment, and drugs (Hlth), Utilities (Utils), and other (Other)—and 51 portfolios (Other portfolios) we use in our baseine analysis. Predicted returns correspond to a version of the regressions in column (3) of Table 3 that includes industry portfolios. Panel B plots the betas of 10 industry portfolios to the bias shocks. Results use the analysts' forecast of both short-term and long-term earnings growth combined as specified in equation (16).

Appendix

A.1 Proofs

Proof of the result in Section 2.3. Consider identity (3) with expectations taken under $\mathbb{S}$:

$$r_{t+1} - \mathbb{E}_t^{\mathbb{S}} r_{t+1} = \varepsilon_{d,t+1}^{\mathbb{S}} - \varepsilon_{r,t+1}^{\mathbb{S}}.$$

Pre-multiplying by ρ^h and summing over $h = 0, ..., H-1$, we have

$$\sum_{h=0}^{H-1} \rho^h r_{t+1+h} - \sum_{h=0}^{H-1} \rho^h \mathbb{E}_{t+h}^{\mathbb{S}} r_{t+1+h} = \sum_{h=0}^{H-1} \rho^h \varepsilon_{d,t+1+h}^{\mathbb{S}} - \sum_{h=0}^{H-1} \rho^h \varepsilon_{r,t+1+h}^{\mathbb{S}}.$$

Note that

$$\begin{aligned}
\sum_{h=0}^{H-1} \rho^h \mathbb{E}_{t+h}^{\mathbb{S}} r_{t+1+h} - \sum_{h=0}^{H-1} \rho^h \varepsilon_{r,t+1+h}^{\mathbb{S}} &= \sum_{h=0}^{H-1} \left(\rho^h \mathbb{E}_{t+h}^{\mathbb{S}} r_{t+1+h} - (\mathbb{E}_{t+1+h}^{\mathbb{S}} - \mathbb{E}_{t+h}^{\mathbb{S}}) \sum_{k=h+1}^{\infty} \rho^k r_{t+1+k}\right) \\
&= \mathbb{E}_t^{\mathbb{S}} r_{t+1} - \mathbb{E}_{t+1}^{\mathbb{S}} \sum_{k=1}^{\infty} \rho^k r_{t+1+k} + \mathbb{E}_t^{\mathbb{S}} \sum_{k=1}^{\infty} \rho^k r_{t+1+k} \\
&\quad + \rho \mathbb{E}_{t+1}^{\mathbb{S}} r_{t+2} - \mathbb{E}_{t+2}^{\mathbb{S}} \sum_{k=2}^{\infty} \rho^k r_{t+1+k} + \mathbb{E}_{t+1}^{\mathbb{S}} \sum_{k=2}^{\infty} \rho^k r_{t+1+k} \\
&\quad + ... \\
&= \mathbb{E}_t^{\mathbb{S}} \sum_{k=0}^{\infty} \rho^k r_{t+1+k} - \mathbb{E}_{t+H}^{\mathbb{S}} \sum_{k=H}^{\infty} \rho^k r_{t+1+k}.
\end{aligned}$$

Breaking $\mathbb{E}_t^{\mathbb{S}} \sum_{k=0}^{\infty} \rho^k r_{t+1+k}$ into $\mathbb{E}_t^{\mathbb{S}} \sum_{k=0}^{H-1} \rho^k r_{t+1+k} + \mathbb{E}_t^{\mathbb{S}} \sum_{k=H}^{\infty} \rho^k r_{t+1+k}$, taking conditional expectations under $\mathbb{P}$ and re-arranging, one obtains:

$$\mathbb{E}_t^{\mathbb{P}} \sum_{h=0}^{H-1} \rho^h r_{t+1+h} = \mathbb{E}_t^{\mathbb{S}} \sum_{k=0}^{H-1} \rho^k r_{t+1+k} + \mathbb{E}_t^{\mathbb{P}} \sum_{h=0}^{H-1} \rho^h \varepsilon_{d,t+1+h}^{\mathbb{S}} + (\mathbb{E}_t^{\mathbb{S}} - \mathbb{E}_t^{\mathbb{P}}) \mathbb{E}_{t+H}^{\mathbb{S}} \sum_{k=H}^{\infty} \rho^k r_{t+1+k}.$$

Proof of the result in Section 2.4. Denote $s = \frac{d\mathbb{S}}{d\mathbb{P}}$. We have $\mathbb{E}^{\mathbb{P}}(s) = 1$ and $\mathbb{E}^{\mathbb{S}}(x) = \mathbb{E}^{\mathbb{P}}(sx)$. Now,

$$\begin{aligned}
\mathbb{C}\text{ov}^{\mathbb{S}}(x, y) &= \mathbb{E}^{\mathbb{S}}(xy) - \mathbb{E}^{\mathbb{S}}(x)\mathbb{E}^{\mathbb{S}}(y) \\
&= \mathbb{E}^{\mathbb{P}}(sxy) - \mathbb{E}^{\mathbb{P}}(sx)\mathbb{E}^{\mathbb{P}}(sy).
\end{aligned}$$

Using the resultsin Bohrnstedt and Goldberger (1969), we can re-write the first term as

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}(sxy) &= \mathbb{E}^{\mathbb{P}}(s)\mathbb{E}^{\mathbb{P}}(xy) + \mathbb{C}\text{ov}^{\mathbb{P}}(s, xy) \\
&= \mathbb{E}^{\mathbb{P}}(s)\mathbb{E}^{\mathbb{P}}(xy) + \mathbb{E}^{\mathbb{P}}(x)\mathbb{C}\text{ov}^{\mathbb{P}}(s, y) + \mathbb{E}^{\mathbb{P}}(y)\mathbb{C}\text{ov}^{\mathbb{P}}(s, x) + \mathbb{E}^{\mathbb{P}}(\Delta s \Delta x \Delta y),
\end{aligned}$$

where $\Delta. = . - \mathbb{E}^{\mathbb{P}}(.)$. The second term can be rewritten as

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}(sx)\mathbb{E}^{\mathbb{P}}(sy) &= \left(\mathbb{E}^{\mathbb{P}}(s)\mathbb{E}^{\mathbb{P}}(x) + \mathbb{C}\text{ov}^{\mathbb{P}}(s, x)\right) \left(\mathbb{E}^{\mathbb{P}}(s)\mathbb{E}^{\mathbb{P}}(y) + \mathbb{C}\text{ov}^{\mathbb{P}}(s, y)\right) \\
&= \mathbb{E}^{\mathbb{P}}(x)\mathbb{E}^{\mathbb{P}}(y) + \mathbb{E}^{\mathbb{P}}(x)\mathbb{C}\text{ov}^{\mathbb{P}}(s, y) + \mathbb{E}^{\mathbb{P}}(y)\mathbb{C}\text{ov}^{\mathbb{P}}(s, x) + \mathbb{C}\text{ov}^{\mathbb{P}}(s, x)\mathbb{C}\text{ov}^{\mathbb{P}}(s, y).
\end{aligned}$$

Combining together, we obtain

$$\mathbb{C}\text{ov}^{\mathbb{S}}(x, y) = \mathbb{C}\text{ov}^{\mathbb{P}}(x, y) + \underbrace{\mathbb{E}^{\mathbb{P}}(\Delta s \Delta x \Delta y)}_{\text{third-moment term}} - \underbrace{\mathbb{C}\text{ov}^{\mathbb{P}}(s, x)\mathbb{C}\text{ov}^{\mathbb{P}}(s, y)}_{\text{second-order term}}.$$

The third-moment term $\mathbb{E}^{\mathbb{P}}(\Delta s \Delta x \Delta y) = 0$ under joint normality. The second-order terms is approximately zero for variables that represent changes over small time periods and is equal to zero in the continuous time limit.

A.2 Earnings measurement consistent with I/B/E/S

To be consistent with the I/B/E/S's definition of earnings, we follow De la O et al. (2024) and Hillenbrand and McCarthy (2022) and define earnings as Compustat net income (item NIq) excluding non-I/B/E/S items, which comprise extraordinary items and discontinued operations (item XIDOq), special items (item SPIq), and non-recurring income taxes (item NRTXTq) as follows,

$$Earnings_t = NIq_t - XIDOq_t - SPIq_t - NRTXTq_t.$$

We treat missing values in the XIDOq, SPIq, and NRTXTq variables as zeros. That is, if none of these three variables are populated in a given firm-quarter, we do not remove anything from net income in this quarter. Similar to De La O and Myers (2021), De la O et al. (2024), and Bordalo et al. (2024a), we use quarterly updated rather than annual data. This is to match the forecasted earnings from I/B/E/S, which are available at a quarterly frequency. To rule out that our data is affected by the sampling frequency of the balance sheet data, we have compared quarterly net income, extraordinary items and discontinued operations, special items, and non-recurring income taxes with their annual counterparts (i.e., items NI, XIDO, SPI, and NRTXT, respectively). While we identify very small differences when comparing the two time series for net income, we find virtually no differences for all other items.

We then use the quarterly earnings data to construct a monthly sampled time-series of annual earnings at the firm level. To that end, at every month, we calculate annual earnings at the firm level as the sum of quarterly earnings from the most recent four quarters. Finally, following De La O and Myers (2021), the realized aggregate earnings are calculated by summing the total earnings reported by each firm and adjusting by the S&P 500 index divisor.

A.3 Additional tables and figures

Table A-1: **VAR model with short-term earnings forecasts only.** Panel A of this table reports the coefficients on the explanatory variables in the VAR model. Each row corresponds to a different equation of the model. The variables include the market excess return (r_t^e), the revisions in analysts' forecast of market earnings growth ($\varepsilon_{d,t}^{\mathbb{S}}$), analysts' forecast of market earnings growth ($\mathcal{F}_{d,t}^{\mathbb{S}}$), the market price-to-dividend ratio (pd_t), and the yield spread between short-term and long-term government bonds (ty_t). Ordinary least squares t-statistics are reported in parentheses below the coefficients. The last column reports the R^2. Panel B reports the fraction of the variation in excess market return and in the revisions in analysts' forecast of market earnings growth at H-month horizon explained by the VAR forecast computed from the estimated VAR model. Results use analysts' forecast of short-term earnings growth and an estimate of growth rate persistence as specified in equation (17). Data are monthly from January 1982 through December 2020.

Panel A: model equations

	r_t^e	$\varepsilon_{d,t}^{\mathbb{S}}$	$\mathcal{F}_{d,t}^{\mathbb{S}}$	pd_t	ty_t	R^2
Market return, r_{t+1}^e	0.043	-0.003	-0.069	-0.019	0.000	0.049
	(0.814)	(-0.084)	(-4.365)	(-3.317)	(0.245)	
Forecast revision, $\varepsilon_{d,t+1}^{\mathbb{S}}$	-0.423	-0.215	-0.161	-0.018	0.000	0.320
	(-6.158)	(-4.759)	(-7.861)	(-2.351)	(0.075)	
Forecast, $\mathcal{F}_{d,t+1}^{\mathbb{S}}$	-0.431	-0.246	0.899	-0.013	-0.002	0.836
	(-6.103)	(-5.280)	(42.766)	(-1.637)	(-0.730)	
Price-to-dividend ratio, pd_{t+1}	0.072	-0.020	-0.044	0.982	0.000	0.986
	(1.318)	(-0.553)	(-2.697)	(162.853)	(-0.061)	
Term spread, ty_{t+1}	0.434	-0.409	0.345	0.005	0.957	0.914
	(1.082)	(-1.544)	(2.889)	(0.114)	(76.480)	

Panel B: predictability at H-month horizon

	$H=1$	$H=3$	$H=12$	$H=120$	$H\to\infty$
$\frac{\mathbb{V}\text{ar}[\mathbb{E}_t^{\mathbb{P}}\sum_{h=0}^{H-1}\rho^h r_{t+1+h}^e]}{\mathbb{V}\text{ar}[\sum_{h=0}^{H-1}\rho^h r_{t+1+h}^e]}$	0.049	0.126	0.312	0.476	0.518
$\frac{\mathbb{V}\text{ar}[\mathbb{E}_t^{\mathbb{P}}\sum_{h=0}^{H-1}\rho^h \varepsilon_{d,t+1+h}^{\mathbb{S}}]}{\mathbb{V}\text{ar}[\sum_{h=0}^{H-1}\rho^h \varepsilon_{d,t+1+h}^{\mathbb{S}}]}$	0.320	0.347	0.700	0.275	0.106

Table A-2: **ICAPM with EPS that include extraordinary items.** This table reports the results of the cross-sectional regression of mean excess portfolio returns on portfolio betas estimated in the first-pass univariate regressions. The model is estimated using 51 portfolios: 20 risk portfolios, 10 Fama-French book-to-market ranked portfolios, 10 Fama-French size ranked portfolios, 10 momentum ranked portfolios, and the market portfolio. The table reports the estimates of the constant term and the factor risk premia for cash-flow news ($\varepsilon^{\mathbb{P}}_{d,t}$), discount-rate news ($\varepsilon^{\mathbb{P}}_{r,t}$), and bias shock ($\varepsilon_{\mathcal{B},t}$). t-statistics based on the GMM-VARHAC standard errors for these estimates are reported in the parentheses. The table also reports the R^2 of the cross-sectional regression and its p-value under the null that $R^2 = 1$. Finally, the table reports the statistic and the associated p-value of the chi-squared test of the null hypothesis that all pricing errors are jointly zero. Columns (1)-(4) report the results in which we use the analysts' forecast of both short-term and long-term earnings growth as specified in equation (16). Columns (5)-(8) report the results in which we use analysts' forecast of short-term earnings growth and an estimate of growth rate persistence as specified in equation (17). Data are monthly from January 1982 through December 2020.

	Baseline				Short-term forecast only			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.008		0.008		0.008		0.006	
	(5.100)		(4.021)		(4.737)		(3.188)	
Cash-flow news, $\varepsilon^{\mathbb{P}}_{d,t}$	-0.005	-0.025	0.048	0.042	-0.010	-0.048	0.029	0.015
	(-0.421)	(-1.642)	(2.144)	(1.723)	(-0.665)	(-2.268)	(1.305)	(0.488)
Discount-rate news, $\varepsilon^{\mathbb{P}}_{r,t}$	-0.000	0.011	0.019	0.014	-0.002	-0.020	0.014	0.007
	(-0.015)	(-3.551)	(2.509)	(1.638)	(-0.454)	(-3.801)	(1.654)	(0.551)
Bias shock, $\varepsilon_{\mathcal{B},t}$			0.038	0.046			0.029	0.040
			(2.686)	(2.976)			(2.584)	(2.583)
R^2	0.031		0.510		0.042		0.415	
	(0.012)		(0.245)		(0.016)		(0.218)	
χ^2	81.379	110.560	76.297	105.948	81.146	109.781	77.447	105.512
	(0.002)	(0.000)	(0.004)	(0.000)	(0.002)	(0.000)	(0.003)	(0.000)

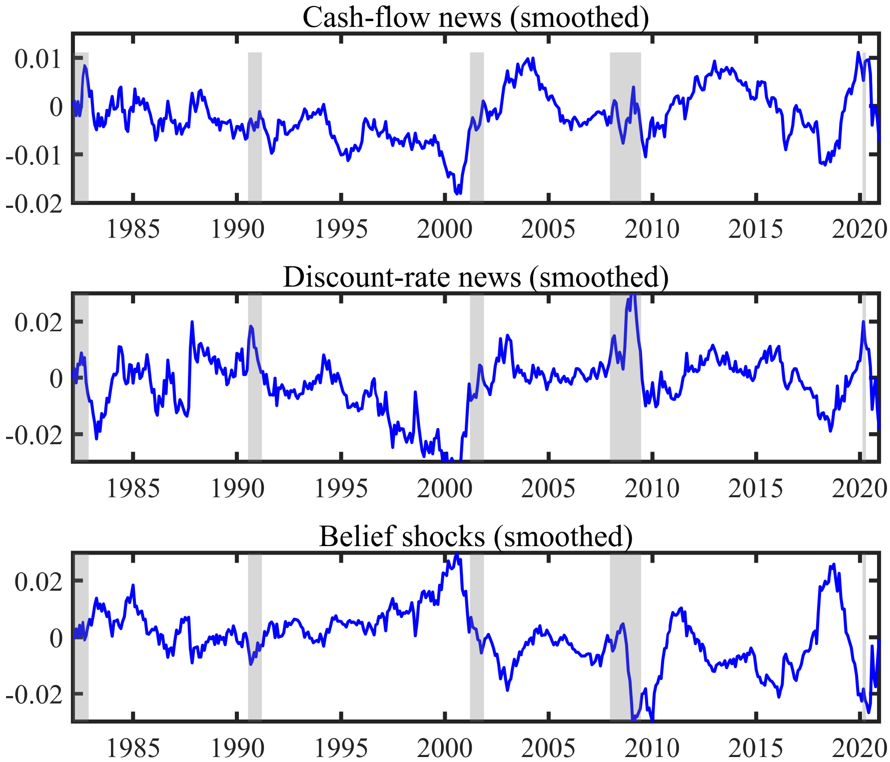


Figure A-1: **Shocks**: This figure plots the ($\varepsilon^{\mathbb{P}}_{d,t}$), discount-rate news ($\varepsilon^{\mathbb{P}}_{r,t}$), and belief shocks ($\mathbf{e}_2^{\intercal}\mathbf{u}_t$) smoothed with a trailing exponentially-weighted moving average. The decay parameter is set to 0.08 and, given the series x_t the smoothed series ma_t are generated as $ma_t = 0.08 \times x_t + (1-0.08) \times ma_{t-1}$. Shaded areas correspond to NBER recessions. Results use the analysts' forecast of both short-term and long-term earnings growth as specified in equation (16). Data are monthly and run from January 1982 to December 2020.

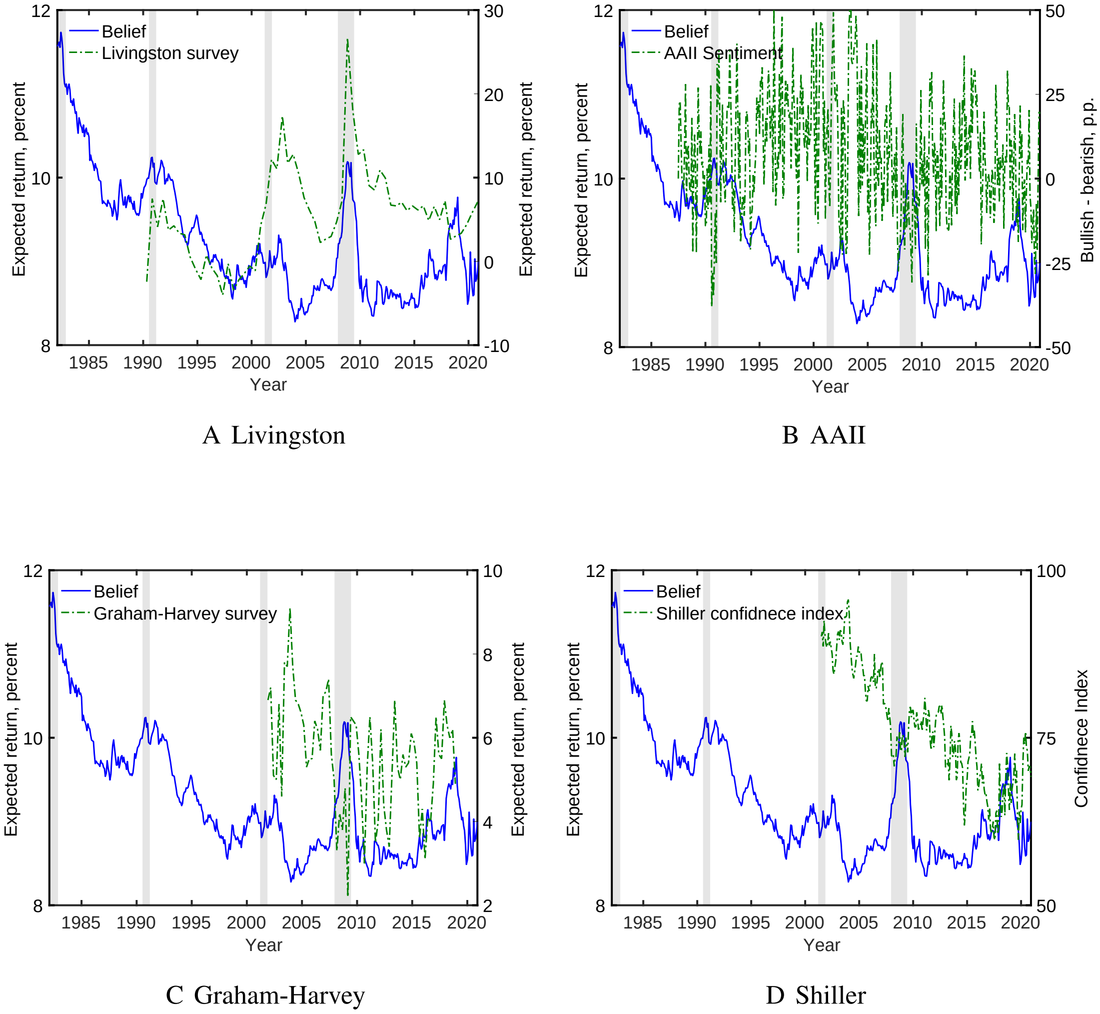


Figure A-2: **VAR beliefs vs survey data**: This figure plots the belief about long-term market returns $\mathcal{F}^{\mathbb{S}}_{r,t}$ obtained as the VAR forecast adjusted for the bias using equation (5) alongside, respectively, the expected market return implied by the Livingston survey from Adam et al. (2021) (Panel A), the difference between the fractions of bullish and bearish investors from the AAII sentiment survey (Panel B), the expected market return from the Graham-Harvey survey of CFOs (Panel C), and the Shiller one-year individual investor stock market confidence index. Results use the analysts' forecast of both short-term and long-term earnings growth as specified in equation (16).