The Cross-section of Subjective Expectations: Understanding Prices and Anomalies

Ricardo Delao, Xiao Han, and Sean Myers*

March 2026

ABSTRACT

We decompose cross-sectional differences in the level of price-earnings ratios using professional forecasts. High price-earnings ratios are accounted for by both low expected returns and overly high expected earnings growth. The magnitudes and timing of the comovements between prices, earnings growth, and returns are consistent with gradual learning rather than expectations being highly sensitive to recent realizations. Earnings growth surprises do not translate 1-1 into one-period returns, but instead are gradually reflected in returns over time. A structural risk-premia model incorporating constant-gain learning about mean earnings growth replicates our findings and generates realistic dispersion and persistence in price-earnings ratios.

*Ricardo Delao is with the Marshall School of Business, University of Southern California. Xiao Han is with the Bayes Business School, City University London. Sean Myers is with the Wharton School at the University of Pennsylvania. We thank Carter Davis, David Hirshleifer, Theis Jensen, Byungwook Kim, Nick Roussanov, Cameron Peng, Petra Sinagl, Stanislav Sokolinski, Michael Ungeheuer, and seminar participants at USC, Wharton, NYU Stern, UT Dallas, UBC, Notre Dame, Vienna Graduate School of Finance, INSEAD, SITE, EFA, SED, Quadrant Behavioral Finance Conference, Frontiers in Quantitative Finance, SQA & CFA Data Science in Finance Conference, the UCI Finance Conference, the Helsinki Finance Summit, the BSE Summer Forum, the Yale Junior Finance Conference, the Junior Valuation Workshop, the Oxford-Man Institute, the Midwest Finance Association, the SFS Cavalcade, and the AFA Annual Meeting for valuable comments. Email addresses: rdelao@marshall.usc.edu, xiao.han.3@city.ac.uk, semyers@wharton.upenn.edu.

It has been known since Basu (1975) and Stattman (1980) that high price ratio stocks (e.g., price-earnings ratios, price-book ratios) earn lower returns than their peers. While one-month differences between Growth and Value stocks have declined over time (Schwert, 2003; Fama and French, 2020), return differences at longer horizons have remained substantial (Delao, Han, and Myers, 2025)[1] and play a large role in accounting for the level of prices (van Binsbergen et al., 2023; Cho and Polk, 2024). Given that a stock's price is the risk-adjusted value of expected future cash flows, these realized return differences imply that high price ratio stocks have low risk exposure, overly high expected cash flows, or a mix of both.

There is a long-standing debate over whether cross-sectional return differences are driven by risk exposure or incorrect cash flow expectations.[2] Our innovation is twofold.

First, using professional forecasts of both returns and cash flows, we decompose cross-sectional differences in the *level* of price ratios. While one-month or one-year returns may be relevant for traders, the level of price ratios is arguably of equal or greater relevance for understanding the allocation of resources across firms (e.g., why are some firms valued at five times earnings while others are valued at fifty times earnings?). To the best of our knowledge, we are the first to quantify the fraction of cross-sectional dispersion in price ratios that is explained by high price ratio stocks having lower subjective return expectations and the fraction that is explained by high price ratio stocks having overly high subjective cash flow expectations. We find that both components play a non-trivial role, however, incorrect cash flow expectations are the quantitatively larger component.

Second, we document that errors in subjective cash flow expectations are "stubborn" in the sense that they are insensitive to recent earnings surprises. On average, high price ratio stocks have high subjective cash flow expectations which are not met by future realized earnings. However, these negative earnings surprises do not lead to a large immediate revision in subjective expectations of future cash flows nor to a large immediate price change.

[1] Table I also confirms that long-term return differences are large even for 1999-2020.

[2] See Fama and French (1995) and Daniel and Titman (1997) for early evidence and Hou, Karolyi, and Kho (2011) and Koijen, Lustig, and Van Nieuwerburgh (2017) for more recent explanations.

Instead, earnings surprises predict a sequence of gradual adjustments in subjective cash flow expectations and prices over the next several years. This aligns with the literature on post earnings announcement drift but emphasizes that these stubborn expectations play an important role in the level of prices, not just one-month to one-year alphas around quarterly earnings announcements.[3] In particular, the presence of stubborn expectations helps to explain how errors in cash flow expectations can have a large impact on the level of prices, as it becomes riskier for informed traders to bet against these expectations. The intuition follows the famous adage, often attributed to Keynes, that "markets can remain irrational longer than you can remain solvent."

For our empirical analysis, we utilize a cross-sectional version of the Campbell-Shiller decomposition.[4] Using professional forecasts, we find that 43.3% of dispersion in price-earnings ratios is accounted for by high price ratio firms having higher expected four-year earnings growth and 12.7% of dispersion is accounted for by high price ratio firms having lower expected four-year returns.[5] Thus, both higher expected earnings growth and lower expected returns help to explain high valuation stocks. Interestingly, while Greenwood and Shleifer (2014) show that expected returns are positively correlated with price ratios in the aggregate time series, in the cross-section investors correctly expect lower returns for high price-ratio firms.[6] The remaining dispersion is explained by expectations of future price-earnings ratios, which reflect expectations of earnings growth and returns beyond four years.

For comparison, realized four-year earnings growth and negative returns account for 9.9% and 32.0% of price-earnings ratio dispersion, respectively, in line with the results of Delao, Han, and Myers (2025). This means that, empirically, high price ratio firms are primarily

[3]For example, Chan, Jegadeesh, and Lakonishok (1996) and Skinner and Sloan (2002) show significant 3- to 12-month return differences across portfolios sorted by earnings surprises.

[4]Because this decomposition is derived from an identity, it holds even if expectations differ from the objective distribution.

[5]For concision, we shorten "subjective expected earnings growth" and "subjective expected returns" to simply "expected earnings growth" and "expected returns." Any time we refer to FIRE (full-information rational expectations) beliefs, we clearly specify that we are using the FIRE-implied distribution.

[6]Dahlquist and Ibert (2024), Bastianello (2025), Büsing and Mohrschladt (2023), and Couts, Gonçalves, and Loudis (2025) also find evidence that expected returns are negatively related to price ratios.

characterized by lower future returns than their peers, rather than by higher future earnings growth. Combining the realized decomposition results with the expected decomposition results, we find that investors overestimate the earnings growth of high price ratio firms, which leads to consistent disappointment in earnings growth for these firms. While investors do expect lower returns for high price ratio firms, they understate the magnitude of this relationship. Consistent with the fact that investors are disappointed by realized earnings growth, the realized returns on high price ratio firms are even lower than investors expected.

Given these errors in expectations, a natural question is how prices and beliefs adjust as future earnings are realized. Under FIRE (full-information rational expectations), earnings surprises should generate immediate, nearly one-for-one revisions in expected future earnings and should not predict any subsequent revisions in earnings expectations. Empirically, however, earnings surprises predict an immediate revision of only 0.11 to 0.14, and predict a sequence of additional revisions of 0.09 to 0.12 in each of the subsequent years. To confirm that this gradual adjustment is not merely an artifact of the subjective expectations data, we show that returns follow a similar path both qualitatively and quantitatively, with earnings surprises predicting a sequence of returns of 0.12 to 0.15 in the current year and each of the subsequent years. Extending this analysis to portfolios sorted on 20 anomalies from Hou, Xue, and Zhang (2015), we consistently find that earnings surprises result in only minor revisions to future earnings expectations and moderate immediate returns, in line with our results for high and low price ratio stocks.

How do these findings fit with FIRE and non-FIRE models? We find that several common behavioral models struggle to match our evidence of stubborn expectations. For example, if agents extrapolate from current earnings growth or have diagnostic expectations of earnings growth, then prices should be highly sensitive to recent realized earnings, and earnings surprises should translate into large immediate returns. Further, as documented previously in Delao, Han, and Myers (2025), standard FIRE models struggle to match the magnitude of the realized relationship between price-earnings ratios and future returns. Thus, while risk premia related to growth options or adjustment costs (Berk, Green, and Naik, 1999; Zhang, 2005) can generate return differences between high and low price-earnings ratio stocks, these models predict a relationship that is an order of magnitude smaller than what we observe in the data.

To explain our empirical findings, we show that constant-gain learning generates volatile prices and reversals (i.e., high price ratios predicting lower future returns) while still matching our evidence of stubborn expectations. In the spirit of Nagel and Xu (2022), we propose a model in which agents learn about mean earnings growth, and we make a small but qualitatively important change by including temporary shocks to the level of earnings. Relative to a FIRE benchmark, constant-gain learning with temporary level shocks increases the volatility of prices and returns substantially due to waves of optimism/pessimism about future earnings. Because of the temporary level shocks, these errors in expectations are stubborn and prices only adjust gradually as future earnings are realized, as earnings surprises are largely attributed to the temporary level shocks.[7] Expanding the model to include risk premia based on cash flow timing à la Lettau and Wachter (2007) allows the model to also match our results on subjective return expectations and improves the model's ability to match the persistence and cross-sectional dispersion of price ratios.

While Bayesian learning can also generate stubborn errors in expectations, we provide two tests that favor constant-gain learning. First, splitting by forecaster tenure, we find that new and old forecasters revise their expectations by similar amounts following earnings surprises and exhibit comparable deviations from the FIRE-implied revision. Second, exploiting cross-sectional heterogeneity in the informativeness of earnings surprises, we find that the magnitude of the earnings revision is largely invariant to signal quality. As a result, the gap between the earnings revision and the FIRE-implied revision is larger for firms with higher signal quality. This second result implies that errors in subjective expectations

[7] In models of learning about mean earning growth without these temporary level shocks, such as Lewellen and Shanken (2002) and Nagel and Xu (2022), the agent believes log earnings follow a random walk. Because of this, prices adjust at least 1-1 to earnings surprises.

should account for a greater share of price–earnings ratio dispersion for firms with higher signal quality, a prediction we confirm empirically. The results for both tests are inconsistent with Bayesian learning, which predicts convergence toward the optimal gain with experience and greater revisions for firms with more informative earnings surprises, but are consistent with a fixed constant-gain learning.

We then estimate and quantitatively test our constant-gain learning model. We set the constant-gain parameter to match previous studies on constant-gain learning (Milani, 2007; Malmendier and Nagel, 2016; Nagel and Xu, 2022) and estimate the remaining 5 parameters solely using realized earnings growth and average aggregate returns. Despite not using any cross-sectional information, the model successfully replicates our decomposition results. Both in terms of magnitudes and timing, the model outperforms standard FIRE models in matching the realized dynamics of price-earnings ratios, earnings growth, and returns and outperforms common behavioral models in matching the dynamics of subjective expectations. Further, the model matches several untargeted aggregate and cross-sectional asset pricing moments.

The quantified structural model allows us to extend our empirical results in two ways. First, we can go beyond the four-year horizon to estimate that expected earnings growth and expected returns for all horizons account for two-thirds (65.7%) and one-third (34.3%) of price-earnings ratio dispersion, respectively. This is largely due to errors in earnings growth expectations, which account for half (50.8%) of all price-earnings ratio dispersion.

Second, we examine how constant-gain learning interacts with risk premia related to cash flow timing to drive expected earnings growth and expected returns. Their combined effect is crucial for generating realistic price volatility, dispersion in price-earnings ratios and persistence in price-earnings ratios. For example, compared to an economy with no learning and no risk premia, introducing only risk premia has little impact on the dispersion in price-earnings ratios and introducing only learning increases the dispersion by a factor of 2.1. However, introducing both increases the dispersion by a factor of 4.5. This highlights the benefit of unifying non-FIRE earnings growth expectations and risk premia related to cash flow timing, as the interaction magnifies the sensitivity of prices to changes in beliefs.

Broadly, this paper contributes to the growing literature using subjective expectations to understand asset prices.[8] In the cross-section, errors in firm-level professional earnings forecasts have been strongly linked to future returns (La Porta, 1996; Frankel and Lee, 1998; Da and Warachka, 2011; So, 2013; Weber, 2018, Bouchaud et al., 2019, van Binsbergen, Han, and Lopez-Lira, 2022).[9] We differ from these studies in two important ways. First and foremost, we focus on explaining cross-sectional differences in the *level* of price ratios, rather than short-term returns. While understanding short-term returns or fluctuations in price ratios is important, understanding the level of a firm's price ratio is of first-order importance for a firm owner or a policy maker interested in allocations. We document stubborn errors in subjective cash flow expectations, which do not translate immediately into large one-period returns, but do play a substantial role in generating large differences between firms in the level of price ratios. Second, by utilizing expectations of both earnings growth and returns, we quantify the relative importance of these two expectations in accounting for cross-sectional dispersion in price-earnings ratios and returns.[10] This decomposition sheds light on the relative importance of risk (discount rates) and mispricing in stock prices.

Our structural model builds on the literature on learning about mean consumption or cash flow growth (Lewellen and Shanken, 2002; Collin-Dufresne, Johannes, and Lochstoer, 2016; Nagel and Xu, 2022) and incorporates risk premia related to cash flow timing, similar in spirit

[8] Amromin and Sharpe (2014); Greenwood and Shleifer (2014); Piazzesi, Salomao, and Schneider (2015); Cassella and Gulen (2018); Delao and Myers (2021); Nagel and Xu (2022); Bordalo et al. (2024); and Gandhi et al. (2025) utilize survey expectations for aggregate outcomes such as returns, cash flows and yields.

[9] Further, Kozak, Nagel, and Santosh (2018) and Engelberg, Mclean, and Pontiff (2018) find that short legs of multiple long-short anomaly strategies comprise stocks with more optimistic earnings forecasts and Engelberg, McLean, and Pontiff (2020) find that anomaly short legs comprise stocks with more optimistic return forecasts. Similarly, Décaire and Graham (2024) use subjective expectations to study fluctuations in price ratios, and Bordalo et al. (2025) use errors in earnings expectations to explain the level and fluctuations in anomaly returns such as HML and RMW. Gómez-Cram and Lawrence (2025) link abnormal returns on software firms to overly pessimistic one- to four-quarter cash flow expectations and show that revenue surprises lead to large one- to four-quarter future returns.

[10] This differs from the implied cost of capital approach (Chen, Da, and Zhao, 2013; Hommel, Landier, and Thesmar, 2025) in which discount rates are inferred using earnings expectations for observable horizons and assumptions about long-term industry growth or GDP growth.

to Lettau and Wachter (2007). We provide new evidence supporting these types of learning models using the cross-sectional dynamics of stocks and show that incorporating learning about temporary shocks to the level of earnings generates distinct qualitative predictions for the timing of earnings growth surprises and returns. We also highlight that learning about cash flows naturally complements risk premia related to cash flow timing. Even if the objective timing of cash flows is relatively similar across all firms (Chen, 2017), these risk premia can still play an important role in stock prices so long as investors *believe* there is a large difference in the timing of cash flows. In other words, as argued in Jensen (2024), once we depart from FIRE, the compensation for risk that investors require should be disciplined by data on investors' believed risks, not the objective risks.

I. Decomposing the cross-section of price ratios

While a large amount of the asset pricing literature has focused on the cross-section of short-term returns, relatively less attention has been paid to the cross-section of prices or price ratios.[11] In particular, we want to understand what can account for the large empirical dispersion in price ratios across stocks, e.g., why do some stocks trade at 50 times earnings while others only trade at 10 times earnings?

To understand dispersion in stock price ratios and how this dispersion relates to subjective cash flow growth expectations and subjective discount rates, we focus on a cross-sectional version of the Campbell-Shiller decomposition. In terms of notation, $E_t^*[\cdot]$ denotes subjective expectations. All other operators use the objective probability distribution. For example, $Var(\cdot)$ and $Cov(\cdot,\cdot)$ denote the observable variance or covariance of variables.

For any stock or portfolio of stocks i, the one-year ahead return $r_{i,t+1}$ can be approximated in terms of the price-earnings ratio $px_{i,t}$, future earnings growth $\Delta x_{i,t+1}$, and the future price-

[11] See Cochrane (2011) for a discussion, "When did our field stop being 'asset pricing' and become 'asset expected returning'?"

earnings ratio, all in logs:

$$r_{i,t+1} \approx \kappa + \Delta x_{i,t+1} + \rho p x_{i,t+1} - p x_{i,t}, \tag{1}$$

where κ and $\rho < 1$ are constants.[12] To understand cross-sectional dispersion in price-earnings ratios, let $\tilde{px}_{i,t}$ be the cross-sectionally demeaned price-earnings ratio of portfolio i and let $\Delta\tilde{x}_{i,t+1}$ and $\tilde{r}_{i,t+1}$ be the cross-sectionally demeaned earnings growth and returns. Rearranging equation (1) and applying subjective expectations $E_t^*[\cdot]$, we see that a higher than average price-earnings ratio must be explained by higher than average expected earnings growth, lower than average expected returns, or a higher than average expected future price-earnings ratio,

$$\tilde{px}_{i,t} \approx \sum_{j=1}^{h} \rho^{j-1} E_t^*\left[\Delta\tilde{x}_{i,t+j}\right] - \sum_{j=1}^{h} \rho^{j-1} E_t^*\left[\tilde{r}_{i,t+j}\right] + \rho^h E_t^*\left[\tilde{px}_{i,t+h}\right]. \tag{2}$$

Importantly, equation (2) does not require that expectations are rational. Because this equation is derived from an identity, it holds under any subjective probability distribution.

To measure the relative contribution of subjective cash flow growth expectations and subjective discount rates to the dispersion in price-earnings ratios, we decompose the variance of $\tilde{px}_{i,t}$ into three components:

$$1 \approx \underbrace{\frac{Cov\left(\sum_{j=1}^{h}\rho^{j-1}E_t^*\left[\Delta\tilde{x}_{i,t+j}\right], \tilde{px}_{i,t}\right)}{Var\left(\tilde{px}_{i,t}\right)}}_{CF_h} + \underbrace{\frac{Cov\left(-\sum_{j=1}^{h}\rho^{j-1}E_t^*\left[\tilde{r}_{i,t+j}\right], \tilde{px}_{i,t}\right)}{Var\left(\tilde{px}_{i,t}\right)}}_{DR_h} + \underbrace{\rho^h\frac{Cov\left(E_t^*\left[\tilde{px}_{i,t+h}\right], \tilde{px}_{i,t}\right)}{Var\left(\tilde{px}_{i,t}\right)}}_{FPX_h}. \tag{3}$$

Note that $Var\left(\tilde{px}_{i,t}\right)$ is the average squared cross-sectionally demeaned price-earnings ratio, which means it measures the average cross-sectional dispersion in price-earnings ratios. The coefficients CF_h and DR_h give a quantitative measure of how much dispersion in price-earnings ratios is accounted for by dispersion in earnings growth expectations and how much

[12] Note that this approximation still holds even for non-dividend paying firms. Appendix E provides a more detailed discussion on the decomposition and its derivation, including the role of the payout ratio.

is accounted for by dispersion in discount rates. Applying the decomposition to multiple horizons h provides information about the timing of expected earnings growth and discount rates. Additionally, the terms in equation (3) can be interpreted as the coefficients from univariate regressions with time fixed effects, e.g., a one unit increase in $px_{i,t}$ is associated with a CF_1 unit increase in expected one-year earnings growth.

When we estimate equation (3) using professional forecasts, we will use expectations of price growth $E_t^*[\Delta p_{i,t+j}]$ as a proxy for expectations of returns $E_t^*[r_{t+j}]$. Empirically, realized price growth and returns are closely related with a correlation of 0.997 to 0.999 for the $j = 1, ..., 4$ horizons that we study in our analysis. However, to ensure that the use of this proxy does not impact the results, we also estimate an exact decomposition based on price growth in Appendix H.1. Because this alternative decomposition is an exact identity, it also addresses any concerns that cross-sectional differences in payout ratios between high and low price-earnings ratio firms may impact the approximation error in equation (3). As shown in Tables I and AV, the results of this exact decomposition closely match the results from equation (3). Further, Delao, Han, and Myers (2025) show that payout ratios do not account for cross-sectional differences in price-earnings ratios, i.e., high price-earnings ratios are not associated with higher or lower dividend-earnings ratios.

A. Errors that impact price ratios versus short-term returns

The key distinction between price-earnings ratios and returns is that price-earnings ratios are tied to the *level* of expectations while short-term returns are tied to *revisions* in expectations. Because of this, the errors in earnings expectations that are most relevant for price-earnings ratios may be distinct from the errors that are most relevant for returns. From equation (2), overly high expectations of earnings growth increase the level of the price-earnings ratio. This over optimism will lead to large forecast errors for future earnings growth. However, whether these errors lead to large short-term returns depends on how agents revise their expectations for subsequent earnings growth after the forecast error is realized.

Combining equations (1) and (2), we have a cross-sectionally demeaned version of the Campbell (1991) formula for unexpected returns,

$$\begin{aligned} \tilde{r}_{i,t+1} - E_t^*\left[\tilde{r}_{i,t+1}\right] &\approx \Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right] + \left(E_{t+1}^* - E_t^*\right)\left[\sum_{j=2}^{\infty}\rho^{j-1}\Delta\tilde{x}_{i,t+j}\right] \\ &\quad - \left(E_{t+1}^* - E_t^*\right)\left[\sum_{j=2}^{\infty}\rho^{j-1}\tilde{r}_{i,t+j}\right]. \qquad (4) \end{aligned}$$

If positive forecast errors $\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]$ lead to investors increasing expected subsequent earnings growth, then these forecast errors will correspond to large immediate return surprises. In fact, the unexpected return will be even larger than the forecast error. However, if positive forecast errors lead to investors *decreasing* expected subsequent earnings growth $\left(E_{t+1}^* - E_t^*\right)\left[\sum_{j=2}^{h}\rho^{j-1}\Delta\tilde{x}_{i,t+j}\right]$ (i.e., they expect earnings to revert back to the previous levels), then the impact on returns will be muted. These changes in expected earnings growth can also be phrased in terms of the earnings level. Equation (4) can be rewritten as

$$\begin{aligned} \tilde{r}_{i,t+1} - E_t^*\left[\tilde{r}_{i,t+1}\right] &\approx w_1\left(\tilde{x}_{i,t+1} - E_t^*\left[\tilde{x}_{i,t+1}\right]\right) + \left(E_{t+1}^* - E_t^*\right)\left[\sum_{j=2}^{\infty}w_j\tilde{x}_{i,t+j}\right] \\ &\quad - \left(E_{t+1}^* - E_t^*\right)\left[\sum_{j=2}^{\infty}\rho^{j-1}\tilde{r}_{i,t+j}\right]. \qquad (5) \end{aligned}$$

where $w_j \equiv (1-\rho)\,\rho^{j-1}$ are weights that sum to 1. Because the weight w_1 is close to 0, the earnings forecast error $\tilde{x}_{i,t+1} - E_t^*\left[\tilde{x}_{i,t+1}\right]$ (which is identical to the earnings growth forecast error $\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]$) will only have a large impact on short-term returns if it is accompanied by a large change in the expected level of future earnings.

Thus, the connection between errors in earnings expectations that matter for price-earnings ratios and errors that matter for short-term returns will depend on how agents revise their expectations after earnings are realized. Because of this, we not only study the decomposition of equation (3), but also study the evolution of earnings expectations after earnings are realized. As discussed more below, we highlight the importance of "stubbornness" in earnings expectations, e.g., investors being overly optimistic about future earnings growth for a firm and then attributing negative forecast errors $\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]$ to temporary shocks that will be reversed rather than revising their expectations for future earnings. These types of stubborn errors in expectations can have a large impact on the level of a firm's price-earnings ratio while only modestly impacting its short-term returns.

II. Empirical decomposition

A. Data

The firm-level realized earnings and prices are collected from Compustat and CRSP. The firm-level expected earnings and prices are collected from I/B/E/S (Institutional Brokers' Estimate System) and Value Line. To perform the decomposition from Section I, we sort these firms into the classic Value and Growth portfolios.[13] Specifically, for each month t, we construct five value-weighted portfolios sorted by book-to-market.[14] For these portfolios, we measure the expectations at time t for earnings growth, price growth, and the future price-earnings ratio over the next four years. We also track the realized buy-and-hold future earnings growth, returns, and price-earnings ratios over the next four years. The *main sample*, which contains expectations of both earnings growth and price growth, ranges from 1999 to 2020. For robustness tests, we also use a *long sample* which ranges from 1982 to 2020 and contains earnings growth expectations. The subsections below provide more detail on the firm-level variable measurements.

A.1. Realized data

The sample of stocks consists of all common stocks (share code 10 and 11) listed on NYSE, AMEX, and NASDAQ. We obtain monthly prices, returns, and shares outstanding from

[13]We focus on these portfolios to connect to a long literature studying Value and Growth portfolios. Additionally, by aggregating firms into portfolios, we can study cash flow growth and discount rates without needing to drop firms with negative future earnings or negative forecasted earnings, whereas studies of firm-level implied cost of capital are generally forced to exclude these firms.

[14]The book-to-market ratio is measured using the market-cap in the portfolio formation month and the total book value from the most recent four quarters. To account for potential data errors, we exclude firms with book-to-market ratios over 100 or below 0.01.

the Center for Research in Security Prices (CRSP). The firm-level accounting variables are constructed from the quarterly Compustat database. Following Davis, Fama, and French (2000) and Cohen, Polk, and Vuolteenaho (2003), we define book value as stockholders' book equity, plus deferred taxes and investment tax credit if available, minus the book value of preferred stock. If stockholders' book equity is not available at Compustat, we define it as the book value of common equity plus the par value of preferred stock, or the book value of assets minus total liabilities in that order. Depending on availability, we use redemption or par value for the book value of preferred stock. To be consistent with the I/B/E/S's definition of earnings, we 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). This aligns with the measure of earnings proposed in Hillenbrand and McCarthy (2026). At every month, annual earnings at the firm level are defined as the sum of quarterly earnings from the most recent four quarters.[15] The main sample includes all firms which have observable returns $r_{i,t+j}$, earnings growth $\Delta x_{i,t+j}$, and price-earnings ratios $px_{i,t+j}$ in future years $j = 1, 2, 3, 4$. We require a future observation so that we can calculate forecast errors for the subjective expectations. However, for robustness, in Appendix H.4, we drop this requirement and estimate a decomposition using delisting returns to reinvest any delisting firms and find similar results.

A.2. Subjective expectations

The subjective earnings and short-term price expectations are extracted from the I/B/E/S Database. The Summary Statistics of the I/B/E/S Database contains the median forecasts for EPS (earnings per share) since 1976 for shorter horizons and 1982 for longer horizons for U.S. publicly traded firms and the median forecasts for prices at the 12-month horizon since 1999. I/B/E/S gathers their forecasts from hundreds of brokerage and independent

[15] To account for possible data errors or extreme outliers, we winsorize annual earnings cross-sectionally at the 1% level.

analysts who track firms as part of their investment research work. Because the forecasts are not anonymous, analysts have a strong incentive to accurately report their expectations.[16] Furthermore, research on I/B/E/S suggests that financial companies' trades are consistent with their own analysts' forecasts and recommendations, which adds to the evidence that reported forecasts genuinely reflect the beliefs of the companies.[17] More importantly, market participants take seriously these analyst forecasts and trade in line with them, with stock prices increasing (decreasing) shortly after upward (downward) revisions in analyst earnings forecasts (Kothari, So, and Verdi 2016).

The long-term price expectations are obtained from the three-to-five-year price targets from the Value Line Investment Survey. Value Line is an independent investment research and financial publishing firm. The price targets cover approximately 1,700 actively traded U.S. companies every period, approximately 90% of the US publicly listed firms market value.[18] Value Line does not have any investment banking relation with the analyzed firms, nor any other obvious reason for providing biased forecasts. To the best of our knowledge, this is the only widely available survey containing firm-level price forecasts at long horizons. In Section VI.B we provide supporting evidence that the Value Line forecasts are consistent with those in I/B/E/S.

We construct monthly earnings expectations for every firm in I/B/E/S at different horizons by using the EPS forecasts for up to three Annual Fiscal Periods (FY1-FY3) and the Long-Term Growth measure (LTG) meant to forecast earnings growth over the next "three-to-five years." For each month, we first interpolate across the different horizons in the annual fiscal periods to estimate an expectation over the next twelve months. We repeat this procedure to calculate two-year expectations. To estimate the three-year expectations, we use the two-year expectations and compound them with the long-term growth forecasts. We repeat

[16] See Mikhail, Walther, and Willis (1999) and Cooper, Day, and Lewis (2001).

[17] Bradshaw (2004) shows that individual earnings forecasts are correlated to Buy/Sell recommendations, while Chan, Chang, and Wang (2009) show that financial companies' own stock holding changes are significantly positively related to recommendation changes.

[18] Value Line is an industry standard to the extent that it has been documented that a large portion of investment newsletters herds towards Value Line recommendations (Graham, 1999).

this procedure to get four-year earnings expectations. We exclude from the main sample the following firms: a) firms without a LTG forecast, b) firms that do not have sufficient forecasts to calculate a 12-month interpolated forecast $E_t^* [\Delta x_{i,t+1}]$, and c) firms that do not have sufficient forecasts in the next year to calculate a 12-month interpolated forecast, $E_{t+1}^* [\Delta x_{i,t+2}]$.[19]

To estimate the price expectations, we obtain the one-year price expectations from the price target in I/B/E/S. We then calculate the four-year price expectation as the three-to-five year price targets from Value Line. We exclude from the main sample those firms missing either a one-year or a three-to-five year price forecast. Since analysts update earnings and price forecasts every month, our expectation data are also in monthly frequency. The main sample covers on average 79.7% of the total market size of firms listed for at least four years in CRSP.

B. Empirical decomposition

Table I and Figure 1 show the results of decomposition (3) using the subjective expectations alongside a FIRE (Full Information Rational Expectations) benchmark based on realized outcomes. The results show the fraction of price-earnings ratio dispersion that is explained by cumulative earnings growth expectations $\sum_{j=1}^{h} \rho^{j-1} E_t^* [\Delta \tilde{x}_{i,t+j}]$ at different horizons as well as the fraction that is explained by cumulative return expectations $\sum_{j=1}^{h} \rho^{j-1} E_t^* [\tilde{r}_{i,t+j}]$. As discussed in Section I, each of the terms in equation (3) can be interpreted as the coefficient from a univariate regression. Concretely, for each horizon h, we estimate three univariate

[19] This last point ensures that for every firm in the main sample we can calculate revisions $E_{t+1}^* [\Delta x_{i,t+2}] - E_t^* [\Delta x_{i,t+2}]$. This allows us to study how expectations of future earnings growth are revised after earnings growth surprises.

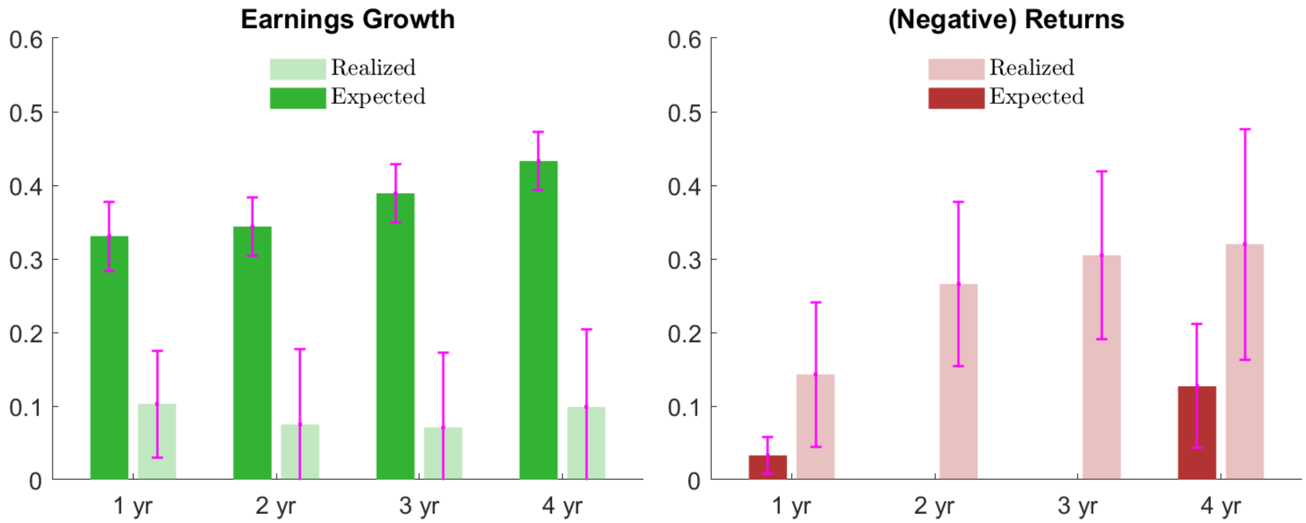

Figure 1. Expected and realized decomposition of price-earnings ratios. This figure illustrates the earnings growth and returns components of the cross-sectional decomposition of $\tilde{px}_{i,t}$ in equation (3). The light bars show the contribution of realized earnings growth and realized returns to the dispersion of price-earnings ratios obtained in the *Realized* columns of Table I. The dark bars show the contribution of expected earnings growth and expected returns to the dispersion of price-earnings ratios obtained in the *Expected* columns of Table I. Each bar shows Driscoll-Kraay 95% confidence intervals.

regressions:

$$\sum_{j=1}^{h} \rho^{j-1} E_t^*[\Delta \tilde{x}_{i,t+j}] = \alpha_h^x + \beta_h^{CF} \tilde{px}_{i,t} + \varepsilon_{i,t}^x \tag{6}$$

$$-\sum_{j=1}^{h} \rho^{j-1} E_t^*[\tilde{r}_{i,t+j}] = \alpha_h^r + \beta_h^{DR} \tilde{px}_{i,t} + \varepsilon_{i,t}^r \tag{7}$$

$$\rho^h E_t^*[\tilde{px}_{i,t+h}] = \alpha_h^{px} + \beta_h^{FPX} \tilde{px}_{i,t} + \varepsilon_{i,t}^{px} \tag{8}$$

using the data for our Value and Growth portfolios from Section II.A. Note that all variables in the regressions are cross-sectionally demeaned, which is equivalent to simply including time fixed effects in all of the regressions. Our estimates of CF_h, DR_h, and FPX_h are then β_h^{CF}, β_h^{DR}, and β_h^{FPX}, respectively. We report the Driscoll-Kraay standard errors, which account for very general forms of spatial and serial correlation, as well as the block-bootstrap standard errors, following the Martin and Wagner (2019) procedure. Section VI provides a number of robustness checks to these main results.

The "Expected" columns of Table I and Figure 1 show the results of the decomposition

Table I

Decomposition of dispersion in price-earnings ratios

This table decomposes the variance of price-earnings ratios using equation (3) at multiple horizons. The *Realized* columns report the elements CF_h, DR_h and FPX_h of the decomposition using future earnings growth, future negative returns and future price-earnings ratios. The *Expected* columns report the elements of the decomposition using expected earnings growth, expected returns and expected price-earnings ratios. The *Error* columns report the contribution of the forecast errors. For instance, for year $h = 1$, $CF_1 = Cov\left(\Delta\tilde{x}_{i,t+1}, \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$ is shown in the *Realized* column. This component can be split into its expected component $Cov\left(E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$ and its error component $Cov\left(\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$. The main sample period is 1999 to 2020. The fourth row shows the element CF_h of the decomposition estimated over the longer sample period of 1982-2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	$h=1$			$h=2$			$h=3$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.103***	0.331***	−0.228***	0.075	0.344***	−0.268***	0.071	0.389***	−0.317***	0.099*	0.433***	−0.335***
	[0.037]	[0.024]	[0.032]	[0.052]	[0.020]	[0.046]	[0.052]	[0.020]	[0.048]	[0.054]	[0.020]	[0.053]
	[0.052]	[0.027]	[0.044]	[0.068]	[0.022]	[0.063]	[0.074]	[0.023]	[0.062]	[0.074]	[0.022]	[0.074]
DR_h	0.143***	0.033***	0.110**	0.266***			0.305***			0.320***	0.127***	0.192**
	[0.050]	[0.013]	[0.053]	[0.057]			[0.058]			[0.080]	[0.043]	[0.082]
	[0.051]	[0.014]	[0.054]	[0.070]			[0.075]			[0.102]	[0.045]	[0.096]
FPX_h	0.746***	0.620***	0.126**	0.642***			0.599***			0.550***	0.385***	0.165***
	[0.050]	[0.019]	[0.056]	[0.034]			[0.039]			[0.057]	[0.027]	[0.062]
	[0.044]	[0.024]	[0.051]	[0.038]			[0.048]			[0.064]	[0.028]	[0.076]
	Long Sample: 1982-2020											
CF_h	0.137***	0.312***	−0.175***	0.102**	0.335***	−0.232***	0.107**	0.399***	−0.291***	0.147***	0.462***	−0.316***
	[0.026]	[0.021]	[0.027]	[0.048]	[0.023]	[0.041]	[0.043]	[0.025]	[0.035]	[0.040]	[0.027]	[0.034]
	[0.026]	[0.021]	[0.027]	[0.049]	[0.022]	[0.042]	[0.042]	[0.026]	[0.036]	[0.040]	[0.027]	[0.033]

using subjective expectations. We find that cash flow growth expectations play a substantial role in explaining the cross-sectional dispersion in price-earnings ratios. Differences in expected one-year earnings growth account for nearly a third (33.1%) of the dispersion in price-earnings ratios, and differences in expected four-year earnings growth account (CF_4) for 43.3% of all price-earnings ratio dispersion. Expected returns explain a much smaller, but positive portion of the observed dispersion, with DR_1 at 3.3% and DR_4 reaching 12.7%. Given that price forecasts are only available at the one-year and four-year horizon, we show DR_h and FPX_h for $h = 1, 4$ and CF_h for all horizons $h = 1, 2, 3, 4$. While we cannot observe expectations out to an infinite horizon, these four-year results demonstrate that over half (56%) of price-earnings ratio dispersion is explained by a combination of expected four-year earnings growth and returns, with earnings growth expectations being the primary component. To estimate the infinite horizon decomposition, Section III introduces a model of belief formation and uses the first four years of CF_h and DR_h as untargeted moments to evaluate the accuracy of the model.

How do these subjective results compare to a FIRE benchmark? Let $E_t^{FIRE}[\cdot]$ denote expectations under FIRE. Because forecast errors $\Delta\tilde{x}_{i,t+j} - E_t^{FIRE}[\Delta\tilde{x}_{i,t+j}]$ are uncorrelated with time t variables under FIRE, we know that $Cov\left(E_t^{FIRE}[\Delta\tilde{x}_{i,t+j}], \tilde{px}_{i,t}\right) = Cov\left(\Delta\tilde{x}_{i,t+j}, \tilde{px}_{i,t}\right)$. The same logic also applies to FIRE expectations of future returns and future price-earnings ratios. Thus, the "Realized" columns of Table I show the estimates of CF_h, DR_h, and FPX_h for each year $h = 1, 2, 3, 4$ using the covariance of $\tilde{px}_{i,t}$ with realized future earnings growth, returns, and price-earnings ratios.[20] High price-earnings ratios are associated with significantly lower future returns and only slightly higher future earnings growth. Empirically, realized one-year returns account for 14.3% of the dispersion, growing to 32.0% at the four-year horizon. In contrast, the contribution of future cash flow growth

[20]For the "Realized" columns, we use the same regressions as equations (6)-(8) but change the dependent variable to the realized future outcome $\sum_{j=1}^{h}\rho^{j-1}\Delta\tilde{x}_{i,t+j}$, $-\sum_{j=1}^{h}\rho^{j-1}\tilde{r}_{i,t+j}$, and $\tilde{px}_{i,t+h}$. The "Error" columns show analogous regressions where the dependent variable is the realized future forecast error.

remains small and relatively flat, around 10% across all horizons. These results are consistent with the findings in Delao, Han, and Myers (2025), who estimate an infinite horizon FIRE decomposition using realized outcomes out to a 15-year horizon and find that future realized cash flow growth accounts for less than 25% of dispersion in price-earnings ratios while future realized returns account for more than 75% of dispersion.

As discussed previously in Delao, Han, and Myers (2025), the magnitudes of the "Realized" decomposition results, such as those shown in Table I, pose a significant challenge for common FIRE asset pricing models. Many common FIRE models (e.g., Berk et al., 1999; Zhang, 2005; Lettau and Wachter, 2007) imply that future realized returns should account for less than 6% of the dispersion in price-earnings ratios. These models struggle to generate a large enough role for discount rates because they tie risk premia directly to the objective timing of cash flows. Since firms differ only modestly in the objective timing of cash flows, as shown by the low values for "Realized" CF_h, the models cannot produce the large return differences observed in the data without assuming implausible levels of risk aversion.

While it is always possible to construct a FIRE model that could match the "Realized" decomposition results, the central challenge would be in generating variation in discount rates that is roughly an order of magnitude larger than what is generated by more common FIRE models. Additionally, this hypothetical model would obviously not align with the direct evidence on subjective expectations shown in Table I. In Sections III and V, we show that introducing stubborn errors in cash flow growth expectations into these types of models can help significantly in terms of matching the "Realized" results as well as the "Expected" results, as risk premia related to cash flow timing are now tied to "Expected" cash flow growth differences between firms, rather than the small "Realized" cash flow timing differences.

Comparing our subjective results to this FIRE benchmark, we find that the differences are statistically significant for all horizons and for all three variables, as shown by the values under the "Error" columns. In particular, this comparison highlights two key results. First, investors substantially overestimate the extent to which high price-earnings ratio stocks will have high future earnings growth. Errors in one-year earnings growth expectations account for 22.8% of price-earnings ratio dispersion, as high price-earnings ratios are consistently associated with disappointment in future earnings growth (as shown in the "Error" columns). Looking at the four-year horizons, more than a third of all dispersion in price-earnings ratios is accounted for by the fact that current price-earnings ratios significantly negatively predict future forecast errors. The final row of Table I shows that our earnings growth results are qualitatively and quantitatively similar over the longer 1982-2020 sample.

Second, investors understand that expensive stocks are associated with lower expected returns, but they underestimate the magnitude of the relationship. As shown in the second row of Table I, lower expected returns for high $\tilde{px}_{i,t}$ firms returns account for 3.3% of dispersion in price-earnings ratios at the one-year horizon and 12.7% at the four-year horizon. The sign of the expected DR_h contrasts sharply with previous findings for aggregate return expectations, which positively comove with aggregate price ratios (Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014; Delao and Myers, 2021) and would imply a negative contribution of return expectations to explaining dispersion in price ratios. While the expected DR_h is positive, it is smaller than the realized DR_h. At the one-year and four-year horizon, differences in realized returns account for 14.3% and 32.0% of price-earnings ratio dispersion. Consistent with the fact that investors overestimate future earnings growth for high $\tilde{px}_{i,t}$ firms, we find that they consistently overestimate the returns for high $\tilde{px}_{i,t}$. In other words, while investors expect lower returns for high $\tilde{px}_{i,t}$ stocks, the realized returns are even worse than expected.

Combined, these two findings emphasize that the mistakes in investors' expectations are about magnitudes, not directions. Investors understand that high price-earnings ratios are associated with higher future earnings growth and lower future returns, but they overestimate the magnitude of the earnings growth relationship and underestimate the magnitude of the return relationship. This underscores the benefit of using a quantitative decomposition which captures magnitudes as well as correlations to study these expectations.

C. Stubbornness in expectations

The results of Table I indicate that errors in earnings growth expectations account for a notable amount of dispersion in price-earnings ratios. Given that we observe earnings expectations at multiple horizons, we now examine how earnings expectations are revised as future earnings are realized. The key finding is that earnings expectations are "stubborn," meaning that positive earnings growth surprises actually lower expected next period earnings growth, as earnings are forecasted to return to their previous levels. This can be more easily understood when we frame the result in levels rather than growth. We show that earnings surprises only lead to small revisions in expected $\tilde{x}_{i,t+2}$, $\tilde{x}_{i,t+3}$, and $\tilde{x}_{i,t+4}$.

As discussed in Section I.A, this stubbornness in expectations drives a wedge between the impact of errors in earnings expectations on the level of price-earnings ratios and the impact on short-term returns. This highlights the importance of directly studying dispersion in price-earnings ratios rather than solely focusing on return predictability. In Section II.D, we discuss how these stubborn expectations differ from expectation-formation models such as Hirshleifer, Li, and Yu (2015), Bordalo et al. (2024), Lewellen and Shanken (2002) and Nagel and Xu (2022), which are primarily designed to generate return predictability and feature earnings expectations that are sensitive to recent realizations.

C.1. Gradual adjustment to surprises

To understand the response to earnings surprises, we regress revisions in expectations on the earnings surprise $\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]$. To give a benchmark, we first consider the regression

$$\tilde{x}_{i,t+h} - E_t^* [\tilde{x}_{i,t+h}] = \alpha_h + \gamma_h (\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]) + \varepsilon_{t+h} \tag{9}$$

for $h = 2, 3, 4$ years. We estimate $\gamma_2 = 0.91^{***}$, $\gamma_3 = 0.89^{***}$, and $\gamma_4 = 0.89^{***}$, and the black dashed line of Figure 2 shows these results for realized future earnings.

We then estimate the revision in expectations after an earnings surprise,

$$E_{t+j}^* [\tilde{x}_{i,t+h}] - E_{t+j-1}^* [\tilde{x}_{i,t+h}] = \alpha_{h,j}^* + \gamma_{h,j}^* (\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]) + \eta_{h,t+j} \tag{10}$$

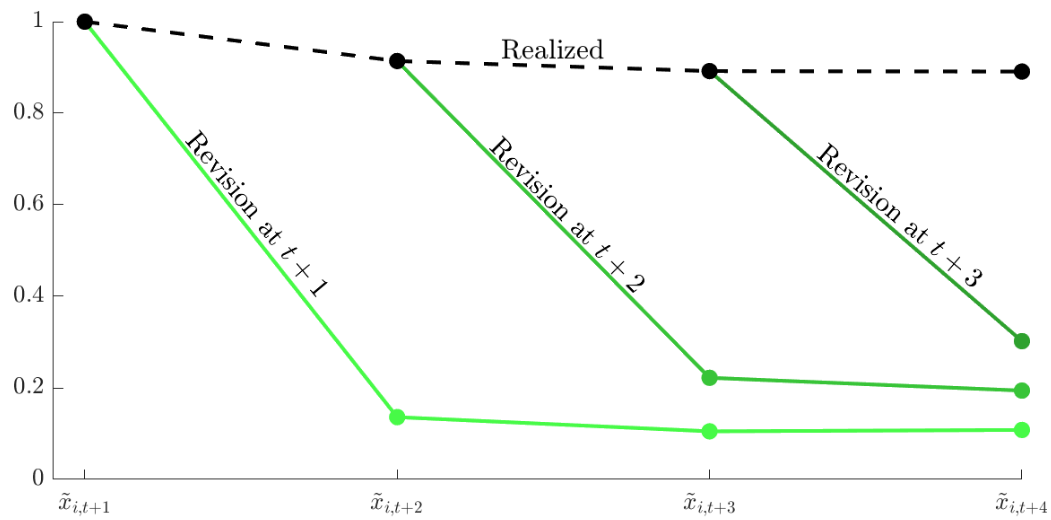

Figure 2. Revisions after an earnings surprise. This figure illustrates the gradual revision in expectations of future earnings following an earnings surprise at $t+1$. Each solid line represents the revised expected path of earnings, $E^*_{t+j}[\tilde{x}_{i,t+h}] - E^*_t[\tilde{x}_{i,t+h}]$, after a one unit earnings surprise, $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$. The line *Revision at t+1* represents the coefficients $\gamma^*_{h,1}$, *Revision at t+2* shows the cumulative sum $\sum_{j=1}^{2} \gamma^*_{h,j}$, and *Revision at t+3* shows the cumulative sum $\sum_{j=1}^{3} \gamma^*_{h,j}$. The dashed line shows the benchmark γ_h from equation (9), representing the impact of earnings surprises on future realized earnings.

for $1 \leq j < h$. The coefficient $\gamma^*_{h,j}$ captures how much earnings surprises at $t+1$ predict forecast revisions at time $t+j$.[21] Under FIRE, these revisions should happen immediately and there should be no further predictable revisions nor predictable forecast errors, i.e., $\gamma^*_{h,1} = \gamma_h$ and $\gamma^*_{h,j} = 0$ for $j > 1$.

Columns 1-3 of Table II Panel A show the results for $\gamma^*_{h,j}$. Figure 2 shows the cumulative sum of the $\gamma^*_{h,j}$ coefficients, i.e., how much $E^*_{t+j}[\tilde{x}_{i,t+h}] - E^*_t[\tilde{x}_{i,t+h}]$ responds to $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$. Rather than an immediate large revision, we find a sequence of gradual revisions over the next three years. After a one unit positive earnings surprise, earnings are expected to largely revert to their previous levels, as shown by the first green line of Figure 2. Expectations of $\tilde{x}_{i,t+2}$, $\tilde{x}_{i,t+3}$, $\tilde{x}_{i,t+4}$ are only revised upwards by 0.14, 0.11, and

[21] This is the complement of Coibion and Gorodnichenko (2015) regressions. Rather than studying how a single revision predicts future forecast errors over time, this regression tests how a single forecast error predicts a sequence of future revisions. To measure these revisions, in Table II and extensions of Table II (i.e., Figures 2 and A5 and Tables III, AXI, and AXIV-AXVI), we restrict our portfolios to firms for which revisions are observed for $h = 2, 3, 4$ which only reduces the total market value in our portfolios by 1.8%. This restriction is not imposed for any other tables.

Table II

Gradual adjustment of expectations

Panel A shows the gradual adjustment of expectations about future earnings $\tilde{x}_{i,t+h}$ after an earnings surprise at $t+1$, i.e., the coefficients $\gamma^*_{h,j}$ estimated using equation (10). For example, the first row shows $\gamma^*_{2,1}$, the effect of an earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ on the revisions to two-year earnings $E^*_{t+1}[\tilde{x}_{i,t+2}] - E^*_t[\tilde{x}_{i,t+2}]$. The second row shows $\gamma^*_{3,1}$ and $\gamma^*_{3,2}$, the effect of an earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ on revisions about $\tilde{x}_{i,t+3}$ occurring in years $t+1$ and $t+2$. Panel B shows the estimated coefficient θ_j from equation (11) which estimates the reaction of returns $\tilde{r}_{i,t+1}, \tilde{r}_{i,t+2}$, and $\tilde{r}_{i,t+3}$ after an earnings surprise at $t+1$. The sample period is 1999 to 2020. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

Panel A: Earnings revisions			
	$j=1$	$j=2$	$j=3$
$\left(E^*_{t+j} - E^*_{t+j-1}\right)[\tilde{x}_{i,t+2}]$	0.14^{*}		
	(0.07)		
$\left(E^*_{t+j} - E^*_{t+j-1}\right)[\tilde{x}_{i,t+3}]$	0.11^{***}	0.12^{***}	
	(0.03)	(0.04)	
$\left(E^*_{t+j} - E^*_{t+j-1}\right)[\tilde{x}_{i,t+4}]$	0.11^{***}	0.09^{**}	0.11^{***}
	(0.03)	(0.03)	(0.03)
Panel B: Returns			
$\tilde{r}_{i,t+j}$	0.13^{**}	0.15^{***}	0.12^{***}
	(0.05)	(0.05)	(0.03)

0.11 respectively. Because of this small revision, in the following year $(t+2)$, investors are on average positively surprised by the $\tilde{x}_{i,t+2}$ earnings and revise their expectations for $\tilde{x}_{i,t+3}$ and $\tilde{x}_{i,t+4}$ by an additional 0.12 and 0.09. This small revision means that investors are again positively surprised in period $t+3$ by $\tilde{x}_{i,t+3}$ and revise their expectations for $\tilde{x}_{i,t+4}$ by another 0.11.

A natural question is whether this stubborn adjustment to new information is unique to analyst forecasts or if it extends to market prices more broadly. To investigate this, we examine how stock returns respond to earnings surprises using the following specification:

$$\tilde{r}_{i,t+j} \quad = \quad \lambda_j + \theta_j\left(\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]\right) + \upsilon_{t+j} \qquad (11)$$

where θ_j captures the extent to which contemporaneous and future returns respond to earnings surprises. Table II Panel B shows the results. Returns do, indeed, respond to earnings surprises, but the reaction is gradual. In line with the gradual adjustment of earnings forecasts by 9 to 14pp each year, we see a gradual return of 12 to 15pp each year after the earnings surprise. This pattern of gradual returns provides strong evidence that the stubborn adjustment to earnings information extends beyond analyst forecasts to market prices themselves, suggesting a broader phenomenon of gradual information incorporation in financial markets.

Another natural question is whether these stubborn expectations could be due to certain analysts not updating their forecasts. The fact that returns follow a similar pattern already points against the idea that this is a mechanical feature of the analyst data. However, to show this point more directly, in Section VI, we repeat the exercise excluding any analyst forecasts that are unchanged from their prior value, and we find nearly identical results. This is because our tests deal with one-year revisions in expectations, $E^*_{t+j}\left[\tilde{x}_{i,t+h}\right] - E^*_{t+j-1}\left[\tilde{x}_{i,t+h}\right]$. While analysts may not update their forecasts every month, it is rare for an analyst to leave a forecast unchanged for an entire year.

Overall, our results align with a long literature on post earnings announcement drift (PEAD). However, our results emphasize that this gradual adjustment over multiple years is relevant for understanding the level of prices and price ratios, not just one-month to one-year returns around quarterly earnings announcements. For example, we can already see evidence of stubborn expectations in the decomposition of price-earnings ratios in Table I. As shown in Column 3 of Table I, one-year earnings growth surprises substantially exceed one-year unexpected returns. Specifically, while 22.8% of the $\tilde{px}_{i,t}$ dispersion is reflected in one-year earnings growth forecast errors, only 11.0% appears in one-year unexpected returns. In other words, disappointment in one-year earnings growth does not translate into an equally large disappointment in one-year returns. This statistically significant 11.8% difference suggests that investors respond stubbornly to earnings disappointment, i.e., stock prices adjust less than 1 for 1 to negative earnings surprises.

This evidence is not limited to cross-sectional differences in valuation ratios. In Appendix G, we sort stocks in portfolios based on 21 different annual anomalies documented in Hou, Xue, and Zhang (2015) and find only moderate revisions after earnings surprises for each of the anomaly portfolios. This evidence, once again, points against models in which unexpected realized returns are highly sensitive to recent earnings surprises and demonstrates the benefit of quantitative decompositions which allow for these types of comparisons.

D. Connection to expectation-formation models

How do these findings compare to common behavioral or learning models of cash flow expectation formation? Agents may overstate the persistence of growth (e.g., Hirshleifer, Li, and Yu 2015) or have diagnostic expectations of growth (e.g., Bordalo et al. 2024). We also consider models in which agents are learning about the mean of an i.i.d. growth process through both Bayesian or non-Bayesian updating (e.g., Lewellen and Shanken 2002 and Nagel and Xu 2022). These mechanisms can all potentially explain the result from Section II.B that investors' cash flow growth expectations overstate the objective relationship between current price-earnings ratios and future earnings growth.[22]

However, as detailed in Appendix F, these models all imply large immediate returns and revisions in expected earnings after an earnings surprise. In the analyst forecast data, a positive earnings surprise lowers expected next period growth, as earnings are expected to largely revert to their previous level. In contrast, these models imply that a positive earnings surprise should raise expected next period earnings growth, meaning that the expected level of future earnings increases more than 1 for 1 with a positive earnings surprise, i.e., $\gamma^*_{h,1} > 1$. Similarly, the data shows that a one unit earnings surprise is only associated with a moderate contemporaneous return, which aligns with the fact that expected earnings are only moderately revised. In these models, a positive surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ raises current returns more than 1 for 1 , i.e., $\theta_1 > 1$, as the earnings surprise raises the expected level of all future cash flows more than 1 for 1.

In Section III, we modify the constant-gain learning model of Nagel and Xu (2022) by

[22]In Appendix F, we also discuss the diagnostic expectations model of Bordalo et al. (2019). Because this model features diagnostic expectations about earnings levels, rather than earnings growth, it predicts that high price-earnings ratio stocks have low one-year earning growth expectations, which is not consistent with our Section II findings.

including temporary shocks to the level of earnings and show that this generates stubborn earnings expectations and gradual returns.[23] This is only a small modification, but it is important for explaining why rational arbitrageurs would not offset these errors in expectations. While the model of Section III focuses on a single representative agent, Appendix C.1 generalizes the model to allow for rational arbitrageurs. In this extended model, non-stubborn errors in expectations have limited impact on asset prices, as they are largely traded away by the rational arbitrageurs. If non-stubborn earnings expectations are too high, arbitrageurs know these expectations will on average be disappointed in the next period, leading to a large revision in expectations and a large one-period return from shorting the asset. In comparison, biases due to stubborn expectations have a large impact on prices. If stubborn earnings expectations are too high, arbitrageurs know that the next period return from shorting the asset will be small, as prices gradually adjust over many periods. If arbitrageurs face shorting or liquidity costs, then they prefer to bet against stocks whose returns are realized quickly.

Importantly, our results do not mean that the biases studied in these previous papers are not relevant for explaining return predictability or other patterns in asset prices. They simply indicate that if one is interested in explaining the level of prices or price ratios, then biases due to stubborn expectations are particularly relevant. In line with the old adage that "markets can remain irrational longer than you can remain solvent," errors in expectations that are stubborn may be more difficult for informed traders to bet against, as they may have to wait years for prices to correct.

III. Model of stubborn expectations and volatile prices

In this section, we introduce a structural model of cash flow expectations and discount rates. The main component of the model is gradual learning about cash flow growth. To match

[23]In this modification, earnings surprises are largely attributed to temporary shocks to the level of earnings. This aligns with evidence of overconfidence (see Daniel and Hirshleifer (2015) for a summary), as overconfident investors will place less weight on new public information and will stick tightly to their prior beliefs.

our results on subjective return expectations, we also include risk premia related to cash flow timing à la Lettau and Wachter (2007). In our baseline specification, the agent updates her beliefs using a simple "rule of thumb" constant gain. Appendix B develops an extended version of the model that microfounds this constant-gain updating through learning with fading memory.

The quantified model fulfills three key purposes. First, quoting Brunnermeier et al. (2021), *"Research focus should be on motivating, building, calibrating, and estimating models with non-RE beliefs rather than on merely rejecting RE models. To make further progress, we need structural models of belief dynamics that can compete with RE models in explaining asset prices and empirically observed beliefs."* This model intends to be a step in this direction. It provides a quantitative model that generates realistic asset pricing moments, outperforms the FIRE models mentioned in Section II in matching the "Realized" decomposition results of Table I, and simultaneously matches the "Expected" decomposition results. Second, the model demonstrates that stubborn errors in expectations increase price volatility, even though these two features seem opposed at first glance. Third, the quantified model allows us to extend the decomposition in equation (3) beyond the four-year horizon to estimate the full role of subjective expected earnings growth and subjective discount rates in accounting for the dispersion in price-earnings ratios.

A. Cash flows and the stochastic discount factor

Throughout this section, we use lowercase letters to denote log values, $z \equiv \log(Z)$. For each firm i, the log cash flow $x_{i,t}$ has an aggregate and a firm-level component,

$$x_{i,t} = x_t^{agg} + \tilde{x}_{i,t} \tag{12}$$

$$x_t^{agg} = \phi x_{t-1}^{agg} + u_t \tag{13}$$

$$\tilde{x}_{i,t} = g_i t + v_{i,t}. \tag{14}$$

The aggregate component is an AR(1) process, which can be thought of as business-cycle fluctuations.[24] The firm-level component is a firm-specific trend $g_i t$ plus noise to capture potential cross-sectional differences in growth rates. The shocks $u_t, v_{i,t}$ are uncorrelated and distributed normally with variances σ_u^2, σ_v^2. Equation (14) implies that the growth of the firm-level component is MA(1),

$$\Delta \tilde{x}_{i,t} = g_i - v_{i,t-1} + v_{i,t}.$$

The agent has a log stochastic discount factor

$$m_{t+1} = -r^f - \frac{1}{2}\gamma^2\sigma_u^2 - \gamma u_{t+1} \tag{15}$$

which depends on the aggregate shock u_{t+1}.

B. *Subjective cash flow expectations*

Objectively, the value of g_i is identical across firms, $g_i = \bar{g}$.[25] However, the agent does not know each firm's g_i. At the end of period $t-1$, the agent has some guess for g_i, which we denote $g^*_{i,t-1}$, and some guess for the previous shock $v_{i,t-1}$, which we denote $v^*_{i,t-1}$. This means her expected growth of the firm-level component is

$$E^*_{t-1}\left[\Delta \tilde{x}_{i,t}\right] = g^*_{i,t-1} - v^*_{i,t-1}.$$

After observing $\Delta \tilde{x}_{i,t}$, we assume that the agent updates her guess for g_i using a simple "rule of thumb" constant-gain update,

$$g^*_{i,t} = g^*_{i,t-1} + \beta\left(\Delta \tilde{x}_{i,t} - E^*_{t-1}\left[\Delta \tilde{x}_{i,t}\right]\right) \tag{16}$$

where β is the constant-gain parameter. For our benchmark model, we assume the agent has no uncertainty about her new guess for g_i and similarly has no uncertainty about her

[24] Appendix C.2 shows an extended version of the model with a unit root in the aggregate component.

[25] Given that our empirical analysis focuses on price-earnings ratios, we normalize $\bar{g}$ to 0 without loss of generality.

guess for the previous shock $v_{i,t-1}$.[26] Given that $\Delta\tilde{x}_{i,t}$ is observable, this means she also has no uncertainty about the inferred current shock,

$$\begin{aligned} v_{i,t}^{*} &= \Delta\tilde{x}_{i,t} - g_{i,t}^{*} + v_{i,t-1}^{*} \\ &= (1-\beta)\left(\Delta\tilde{x}_{i,t} - E_{t-1}^{*}\left[\Delta\tilde{x}_{i,t}\right]\right). \end{aligned} \tag{17}$$

Given these beliefs about g_i and the most recent shock $v_{i,t}$, her expectation for the future growth of the firm-level component is then

$$E_t^{*}\left[\Delta\tilde{x}_{i,t+1}\right] = g_{i,t}^{*} - v_{i,t}^{*}. \tag{18}$$

Her expectation for the future level of the firm-level component is

$$E_t^{*}\left[\tilde{x}_{i,t+n}\right] = \tilde{x}_{i,t} + n g_{i,t}^{*} - v_{i,t}^{*}. \tag{19}$$

C. Prices and risk premia

Appendix A gives the details for all of the equations. To start, let $P_t^{(n)}$ be the price of an n-period aggregate strip, i.e., a claim that pays X_{t+n}^{agg} in n periods. The aggregate strip price is

$$\begin{aligned} P_t^{(n)} &= E_t^{*}\left[\left(\prod_{j=1}^{n} M_{t+j}\right) X_{t+n}^{agg}\right] \\ &= \exp\left\{-nr^f - \gamma\sigma_u^2 \frac{1-\phi^n}{1-\phi} + \frac{1}{2}\sigma_u^2 \frac{1-\phi^{2n}}{1-\phi^2} + \phi^n x_t^{agg}\right\}. \end{aligned} \tag{20}$$

The realized return on the strip is

$$R_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)}}{P_t^{(n)}} = \exp\left\{r^f + \gamma\sigma_u^2\phi^{n-1} - \frac{1}{2}\sigma_u^2\phi^{2(n-1)} + \phi^{n-1}u_{t+1}\right\} \tag{21}$$

and the subjective expected return on the strip is

$$E_t^{*}\left[R_{t+1}^{(n)}\right] = \exp\left\{r^f + \gamma\sigma_u^2\phi^{n-1}\right\}. \tag{22}$$

[26] Our assumptions about uncertainty ensure that beliefs about second moments do not impact asset prices. In Appendix B, we derive equation (16) from a model of learning with fading memory that also includes parameter uncertainty. In the fading memory model, parameter uncertainty adds additional subjective variance to expected cash flows, which impacts asset prices through Jensen's terms, however, we find that the impact on our quantitative results is minor.

The first term (r^f) reflects the risk-free rate and the second term ($\gamma\sigma_u^2\phi^{n-1}$) reflects the risk premium, i.e., the compensation agents require for exposure to risk.

Equation (22) shows the main feature of risk premia in this model: longer horizon strips carry a lower annual risk premium $\gamma\sigma_u^2\phi^{n-1}$. Equation (13) shows that aggregate shocks are persistent but not permanent. This means that short horizon cash flows are disproportionately sensitive to the aggregate shock.[27] Because of this, the annual risk premium is higher for short horizon cash flows. This is similar to the mechanism in Lettau and Wachter (2007).

Each firm i can be viewed as a collection of strips. Specifically, since shocks to the firm-level component $v_{i,t}$ are uncorrelated with the aggregate shock, we can express the firm's price as

$$P_{i,t} = \sum_{n=1}^{\infty} E_t^* \left[\left(\prod_{j=1}^{n} M_{t+j} \right) X_{t+n}^{agg} \tilde{X}_{i,t+n} \right] = \sum_{n=1}^{\infty} P_t^{(n)} E_t^* \left[\tilde{X}_{i,t+n} \right]. \quad (23)$$

Since the agent has no uncertainty about her guess for g_i and the current shock $v_{i,t}$, her expectation for the future $\tilde{X}_{i,t+n}$ is simply

$$E_t^* \left[\tilde{X}_{i,t+n} \right] = E_t^* \left[\exp \left(\tilde{x}_{i,t} + n g_i - v_{i,t} + v_{i,t+n} \right) \right] = \exp \left\{ \frac{1}{2}\sigma_v^2 + \tilde{x}_{i,t} + n g_{i,t}^* - v_{i,t}^* \right\} \quad (24)$$

where the Jensen's term $\frac{1}{2}\sigma_v^2$ comes solely from the volatility of the future shock $v_{i,t+n}$ and the fact that the agent knows the shocks are normal. This means that the firm's price is

$$P_{i,t} = \sum_{n=1}^{\infty} P_t^{(n)} \exp \left\{ \frac{1}{2}\sigma_v^2 + \tilde{x}_{i,t} + n g_{i,t}^* - v_{i,t}^* \right\}. \quad (25)$$

Appendix B shows that we can derive nearly an identical equation for firm prices in a setting of learning with fading memory, with the only distinction being that uncertainty about g_i and the recent shock $v_{i,t}$ increases the Jensen's term.

The subjective expected return on firm i is

$$E_t^* \left[R_{i,t+1} \right] = E_t^* \left[\frac{X_{t+1}^{agg} \tilde{X}_{i,t+1} + P_{i,t+1}}{P_{i,t}} \right] = \sum_{n=1}^{\infty} w_{i,t,n} E_t^* \left[R_{t+1}^{(n)} \right] \quad (26)$$

which is just a weighted average of the subjective expected return on the individual strips

[27] Specifically, the covariance of the log cumulative stochastic discount factor $\log\left(\prod_{j=1}^{n} M_{t+j}\right)$ with the aggregate cash flows x_{t+n}^{agg} is $\frac{1-\phi^n}{1-\phi}\gamma\sigma_u^2$, which increases with horizons but not proportionally.

where the weight $w_{i,t,n} = \frac{\exp\left\{ng^*_{i,t}\right\}P^{(n)}_t}{\sum_{n=1}^{\infty}\exp\left\{ng^*_{i,t}\right\}P^{(n)}_t}$ captures how much of the firm's value in equation (25) comes from its horizon n cash flows. The realized return for firm i is

$$R_{i,t+1} = \frac{X^{agg}_{t+1}\tilde{X}_{i,t+1} + P_{i,t+1}}{P_{i,t}} = \sum_{n=1}^{\infty} w_{i,t,n} R^{(n)}_{t+1} \frac{E^*_{t+1}\left[\tilde{X}_{i,t+n}\right]}{E^*_t\left[\tilde{X}_{i,t+n}\right]}. \tag{27}$$

In addition to depending on a weighted average of realized strip returns $R^{(n)}_{t+1}$, the realized firm return also depends on the change in the expected future firm-level component. From equations (16), (17), and (24), this change in expectations can be expressed entirely in terms of the surprise about one-period growth $\Delta\tilde{x}_{i,t+1} - E^*_t\left[\Delta\tilde{x}_{i,t+1}\right]$, as for $n \geq 2$ we have that

$$\frac{E^*_{t+1}\left[\tilde{X}_{i,t+n}\right]}{E^*_t\left[\tilde{X}_{i,t+n}\right]} = \exp\left\{n\beta\left(\Delta\tilde{x}_{i,t+1} - E^*_t\left[\Delta\tilde{x}_{i,t+1}\right]\right)\right\}. \tag{28}$$

D. Connection to empirical results

Below, we discuss three key results of the model that are relevant to our empirical findings.

First, increases in $g^*_{i,t}$ raise the firm's price in two ways: increasing the expected future cash flows and decreasing the subjective risk premium. From equation (25), a higher $g^*_{i,t}$ naturally increases the value of the firm by increasing the value of future expected cash flows. What is less straightforward is that raising $g^*_{i,t}$ lowers the subjective risk premium. As shown in equation (22), longer horizon cash flow strips carry a lower risk premium in this model, as their annualized return is less sensitive to the aggregate shock u_{t+1}. A higher value for $g^*_{i,t}$ means that more of the firm's value comes from its longer horizon cash flows and therefore the weights $w_{i,t,n}$ in equation (26) are more concentrated on the longer horizon $E^*_t\left[R^{(n)}_{t+1}\right]$. In line with the findings in Table I, this means that both higher expected earnings growth and lower expected returns will help to explain high price-earnings ratios.

Second, if the constant-gain parameter β is small, then expectations are "stubborn." After an earnings surprise $\tilde{x}_{i,t+1} - E^*_t\left[\tilde{x}_{i,t+1}\right]$ (which is equivalent to $\Delta\tilde{x}_{i,t+1} - E^*_t\left[\Delta\tilde{x}_{i,t+1}\right]$), expectations of future earnings are only slightly revised as shown in equation (28). Because expected future earnings are only slightly revised, the earnings surprise does not translate into a large immediate return, as shown in equation (27). Rather than being highly sensitive to recent surprises, agents largely attribute the surprise to the noisy level shock $v_{i,t+1}$. This result is in line with the findings in Table II and highlights how the inclusion of temporary shocks substantially changes the model's predictions relative to the models discussed in Section II.D.

Third, the presence of stubborn errors in expectations increases price volatility. If agents had full-information rational expectations (i.e., agents knew the parameters $g_i = \bar{g}$), then all firms would have the same price and there would be no predictable differences in returns between firms. However, because of errors in earnings expectations, the prices of some firms will rise or fall relative to others due to waves of optimism or pessimism about the firm's underlying growth parameter g_i. Because errors in expectations are stubborn, these deviations in prices away from the FIRE benchmark can be long-lived and a price that is above the FIRE benchmark will predict a gradual sequence of lower returns.

It is important to note that these model implications are the result of two mechanisms: slow updating of future cash flow expectations in response to surprises and a preference for the timing of cash flows. While the model represents these two mechanisms by parameter learning and non-permanent aggregate shocks, there are other mechanisms that could potentially deliver similar implications.

Constant-gain learning about g_i delivers slow updating of future cash flow expectations if the parameter β is small, consistent with the evidence in Table II and prior estimates (Orphanides and Williams, 2005 and Milani, 2007). This approach closely resembles learning from experience, as shown in Malmendier and Nagel (2016). Additionally, slow updating of expectations could be driven by Bayesian learning about a persistent process; however, in the next section, we provide evidence in favor of constant-gain updating over Bayesian updating. The preference for timing is generated from persistent but non-permanent aggregate shocks, which cause short-term cash flows to be disproportionately discounted relative to long horizon cash flows. Importantly, this result depends on perceived, not objective, risk differences. Using survey-based subjective risk measures (Jensen, 2024), we confirm in Appendix J that low price-earnings portfolios are indeed perceived as riskier than high price-earnings portfolios.

Ultimately, we choose to focus on constant-gain parameter learning and differences in perceived risk, as this provides a parsimonious description of these mechanisms and allows us to estimate the model parameters without targeting any of our empirical results from Section II. As shown in the model estimation in Section V, we take the gain β from previous work on constant-gain learning and set risk sensitivity γ to match the aggregate equity premium. As a result, we can fully utilize the empirical decomposition results of Table I to evaluate the quantitative realism of our model.

IV. Constant-gain versus Bayesian learning

One important question is whether the constant-gain learning proposed in our model can be distinguished from more traditional Bayesian learning. As implied by the name, the core feature of constant-gain learning relative to Bayesian learning is that the gain is constant, regardless of the objective dynamics of earnings or how long the investor has been studying the firm. This provides two clear, testable predictions. First, changing the set of analysts should not affect the results of Tables I and II. Specifically, we compare newer analysts to analysts who have been covering each firm for a longer amount of time. Second, changing the set of firms should not affect how much analysts update in Table II, but can affect the gap between analyst updating and the FIRE-predicted updating due to differences in the objective informativeness of earnings surprises. In particular, for firms where earnings surprises are more informative about future earnings (i.e., firms for which analysts are not updating enough), we should find a greater role for “Error” in Table I. We test this prediction by comparing young and mature firms, with the idea that noise may play a larger role in realized earnings for young firms, making their earnings surprises less informative.

Table III

Revisions after earnings surprises: split by analyst coverage time and firm age

Panel A splits analyst–firm forecasts into "New" and "Old" based on whether an analyst's coverage length for firm is below or above the firm-specific median coverage length (coverage length is measured from the first month the analyst issued a forecast for the firm). For each group, we estimate equations (9) and (10). The FIRE columns report the objective informativeness γ_h of the earnings surprise $\tilde{x}_{t+1} - E_t^*[\tilde{x}_{t+1}]$ for earnings at horizons $h = 2, 3, 4$. The $(E_{t+1}^* - E_t^*)$ columns report the coefficient $\gamma_{h,1}^*$ measuring the size of the revision in expected earnings $\tilde{x}_{i,t+h}$ immediately after the surprise $\tilde{x}_{t+1} - E_t^*[\tilde{x}_{t+1}]$. Panels B and C repeat the analysis after splitting firms into "Young" and "Mature" based on firm age according to the first month that the firm enters the CRSP database or the I/B/E/S forecast database. The sample period is 1999 to 2020. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

Panel A: Split by analyst coverage

	New		Old	
	FIRE	$E_{t+1}^* - E_t^*$	FIRE	$E_{t+1}^* - E_t^*$
$\tilde{x}_{i,t+2}$	0.98***	0.16***	0.97***	0.15***
	[0.06]	[0.06]	[0.06]	[0.06]
$\tilde{x}_{i,t+3}$	0.95***	0.12***	0.95***	0.12***
	[0.09]	[0.03]	[0.09]	[0.04]
$\tilde{x}_{i,t+4}$	0.90***	0.08***	0.94***	0.12***
	[0.13]	[0.05]	[0.09]	[0.04]

Panel B: Split by CRSP firm age

	Young		Mature	
	FIRE	$E_{t+1}^* - E_t^*$	FIRE	$E_{t+1}^* - E_t^*$
$\tilde{x}_{i,t+2}$	0.52***	0.09	0.95***	0.08
	[0.15]	[0.06]	[0.03]	[0.05]
$\tilde{x}_{i,t+2}$	0.40***	0.07*	0.94***	0.07**
	[0.14]	[0.04]	[0.05]	[0.03]
$\tilde{x}_{i,t+3}$	0.39***	0.07*	0.95***	0.08**
	[0.13]	[0.04]	[0.05]	[0.03]

Panel C: Split by I/B/E/S firm age

	Young		Mature	
	FIRE	$E_{t+1}^* - E_t^*$	FIRE	$E_{t+1}^* - E_t^*$
$\tilde{x}_{i,t+2}$	0.49***	0.13*	0.96***	0.08
	[0.14]	[0.07]	[0.03]	[0.03]
$\tilde{x}_{i,t+2}$	0.34***	0.10**	0.95***	0.07***
	[0.11]	[0.05]	[0.03]	[0.03]
$\tilde{x}_{i,t+3}$	0.34***	0.10**	0.97***	0.07***
	[0.11]	[0.04]	[0.03]	[0.03]

Table III tests this first prediction. For each firm i and each time t, we calculate the median amount of time that analysts have been covering firm i.[28] We then split analyst forecasts into two groups, "new" and "old" based on whether the analyst is above or below this median amount of coverage time. We drop any firms that only have a single analyst. The first two columns of Panel A show the results when we estimate equations (9) and (10) using only the new analysts' forecasts, and the last two columns of Panel A show the results using only the old analysts' forecasts. The FIRE columns show the objective informativeness of the earnings surprise γ_h. Under a FIRE benchmark, analysts would immediately revise their expectations for $\tilde{x}_{i,t+h}$ by γ_h. The values under the $\left(E^*_{t+1} - E^*_t\right)$ columns show how much analysts actually update their expectations after the surprise is realized. Appendix I provides longer tables showing the full sequence of revisions $\left(E^*_{t+1} - E^*_t\right)$, $\left(E^*_{t+2} - E^*_{t+1}\right)$, and $\left(E^*_{t+3} - E^*_{t+2}\right)$.

Like our main Figure 2, the FIRE columns of Table III Panel A show that one-year earnings surprises are highly informative about forecast errors for longer horizon earnings, with γ_h around 0.9. Despite this, new analysts only slightly update their forecasts after an earnings surprise, with $\gamma^*_{h,1}$ around 0.1. Constant-gain learning would predict that this is a fixed feature of analyst updating whereas Bayesian learning would typically imply that analysts would gradually learn the objective informativeness over time. Looking at the $\gamma^*_{h,1}$ values for older analysts, we find support for constant-gain learning, as older analysts update by nearly the same amount as new analysts.

Similarly, Table IV shows that we find almost identical results for new and old analysts in the decomposition of price-earnings ratio dispersion. In line with our replication of γ_h and $\gamma^*_{h,1}$ from Figure 2, we also replicate the one-year and four-year decomposition from Table I solely using the new analysts' forecasts then again solely using old analysts' forecasts. The expected CF_h, DR_h, and FPE_h closely match for new and old analysts and, as a consequence, the forecast errors shown under the "Error" columns also closely match. Extending the analysis

[28]How long an analyst has been covering a firm is based on the first month in which the analyst provided a forecast for firm i.

Table IV

Decomposition of dispersion in price-earnings ratios by analyst coverage

This table repeats the analysis of Table I after splitting analyst–firm forecasts into "New" and "Old" based on whether an analyst's coverage length for firm is below or above the firm-specific median coverage length (coverage length is measured from the first month the analyst issued a forecast for the firm). For each group, we decompose the variance of price-earnings ratios using equation (3) at multiple horizons. The *Realized* columns report the elements CF_h, DR_h and FPX_h of the decomposition using future earnings growth, future negative returns and future price-earnings ratios. The *Expected* columns report the elements of the decomposition using expected earnings growth, expected returns and expected price-earnings ratios. The *Error* columns report the contribution of the forecast errors. The main sample period is 1999 to 2020. The fourth row shows the element CF_h estimated over the longer sample period of 1982-2020. Driscoll-Kraay and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	New						Old					
	$h=1$			$h=4$			$h=1$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.097***	0.363***	−0.265***	0.092*	0.469***	−0.378***	0.097***	0.363***	−0.265***	0.092*	0.469***	−0.377***
	[0.035]	[0.020]	[0.034]	[0.053]	[0.015]	[0.056]	[0.035]	[0.019]	[0.034]	[0.053]	[0.015]	[0.057]
	[0.047]	[0.023]	[0.047]	[0.070]	[0.018]	[0.079]	[0.050]	[0.021]	[0.050]	[0.073]	[0.018]	[0.076]
DR_h	0.133***	0.030**	0.103*	0.311***	0.127***	0.184**	0.133***	0.030**	0.103*	0.311***	0.127***	0.184**
	[0.050]	[0.012]	[0.054]	[0.078]	[0.040]	[0.082]	[0.050]	[0.012]	[0.054]	[0.078]	[0.040]	[0.082]
	[0.048]	[0.013]	[0.052]	[0.103]	[0.041]	[0.100]	[0.046]	[0.013]	[0.052]	[0.099]	[0.042]	[0.094]
FPX_h	0.761***	0.592***	0.169***	0.565***	0.352***	0.212***	0.761***	0.592***	0.169***	0.565***	0.353***	0.211***
	[0.048]	[0.017]	[0.054]	[0.054]	[0.032]	[0.063]	[0.048]	[0.016]	[0.054]	[0.054]	[0.034]	[0.065]
	[0.042]	[0.021]	[0.048]	[0.065]	[0.035]	[0.071]	[0.042]	[0.021]	[0.048]	[0.063]	[0.039]	[0.075]
	Long Sample: 1982-2020											
CF_h	0.128***	0.339***	−0.212***	0.140***	0.500***	−0.360***	0.128***	0.342***	−0.214***	0.140***	0.509***	−0.369***
	[0.025]	[0.022]	[0.026]	[0.041]	[0.027]	[0.036]	[0.025]	[0.022]	[0.026]	[0.041]	[0.027]	[0.035]
	[0.024]	[0.021]	[0.025]	[0.041]	[0.026]	[0.037]	[0.026]	[0.021]	[0.026]	[0.041]	[0.026]	[0.037]

to the longer sample for CF_h also yields similar conclusions. To summarize, we find no evidence that older analysts have smaller predictable errors for the decomposition of Table I nor smaller gaps between objective informativeness and analyst updating (i.e., $\gamma_h - \gamma^*_{h,1}$).

Panel B of Table III tests the second prediction. Rather than splitting analysts into two groups, we split firms into two groups, Young and Mature, based on their age. The FIRE columns of Panel B show that the two groups of firms differ substantially in the objective informativeness of earnings surprises, with values of γ_h above 0.9 for all horizons for mature firms compared to values of at most 0.52 for young firms. These differences in objective informativeness cannot be explained by random chance, as even after accounting for spatial and serial correlation, we can clearly reject that the two groups have the same γ_h.[29] Despite this difference in objective informativeness, the $\left(E^*_{t+1} - E^*_t\right)$ columns show that analysts update by similar amounts after an earnings surprise for both groups, with $\gamma^*_{h,1}$ around 0.1, consistent with the prediction of constant-gain learning.

To further test the second prediction, Table V Panel A replicates the one-year and four-year decomposition of Table I using only young firms and then repeats the exercise using only mature firms. Under constant-gain learning, analysts should use the same updating rule for all firms. Thus, for the group of firms where earnings surprises are more informative (i.e., the firms with a large $\gamma_h - \gamma^*_{h,1}$ gap), we should find that predictable forecast errors play a larger role in price-earnings ratios. In line with this prediction, we find that price-earnings ratios for mature firms are closely related to expected earnings growth (expected CF_4 of 0.58) but have little connection to realized future earnings growth (realized CF_4 of 0.04). Because of this, over half of all dispersion in price-earnings ratios is attributed to four-year earnings growth forecast errors. Similarly, among mature firms, price-earnings ratios significantly predict future returns as well as return forecast errors. In comparison, while price-earnings ratios for young firms are also connected to subjective earnings growth expectations (expected CF_4 of 0.44), the role of earnings growth forecast errors is much smaller for young firms

[29]For horizons $h = 1, 2, 3$, the differences in γ_h are statistically significant with t-statistics of 2.94, 4.23, and 4.95 respectively based on Driscoll-Kraay standard errors.

Table V

Decomposition of dispersion in price-earnings ratios by firm age

This table repeats the analysis of Table I after splitting firm forecasts into "Young" and "Firms" based on whether a firm's age is below or above the median firm age. Age is measured according to the first month that the firm enters the CRSP database (Panel A) or the I/B/E/S forecast database (Panel B). For each group, we decompose the variance of price-earnings ratios using equation (3) at multiple horizons. The *Realized* columns report the elements CF_h and DR_h of the decomposition using future earnings growth, and future negative returns. The *Expected* columns report the elements of the decomposition using expected earnings growth and expected returns. The *Error* columns report the contribution of the forecast errors. The sample period is 1999 to 2020. Driscoll-Kraay and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	Young Firms						Mature Firms					
	$h=1$			$h=4$			$h=1$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
Panel A: Age by CRSP												
CF_h	0.136**	0.283***	−0.152**	0.309***	0.440***	−0.135*	0.081***	0.530***	−0.449***	0.041	0.581***	−0.540***
	[0.063]	[0.031]	[0.069]	[0.09]	[0.027]	[0.078]	[0.022]	[0.043]	[0.045]	[0.038]	[0.039]	[0.048]
	[0.069]	[0.035]	[0.076]	[0.093]	[0.026]	[0.085]	[0.028]	[0.059]	[0.056]	[0.047]	[0.053]	[0.055]
DR_h	0.103	0.017	0.086	0.072	0.135***	−0.063	0.105***	0.028*	0.077**	0.346***	0.081*	0.265***
	[0.075]	[0.012]	[0.076]	[0.149]	[0.031]	[0.142]	[0.034]	[0.015]	[0.038]	[0.062]	[0.049]	[0.086]
	[0.079]	[0.013]	[0.075]	[0.151]	[0.033]	[0.143]	[0.026]	[0.016]	[0.028]	[0.073]	[0.053]	[0.091]
Panel B: Age by I/B/E/S												
CF_h	0.154**	0.267***	−0.117	0.317***	0.430***	−0.118	0.088***	0.512***	−0.425***	0.052	0.570***	−0.518***
	[0.064]	[0.036]	[0.078]	[0.085]	[0.025]	[0.086]	[0.022]	[0.034]	[0.035]	[0.036]	[0.034]	[0.041]
	[0.074]	[0.039]	[0.096]	[0.097]	[0.025]	[0.103]	[0.027]	[0.047]	[0.044]	[0.050]	[0.048]	[0.049]
DR_h	0.115	0.020*	0.095	0.110	0.151***	−0.041	0.105***	0.033**	0.072*	0.330***	0.090**	0.240***
	[0.083]	[0.012]	[0.088]	[0.155]	[0.029]	[0.157]	[0.033]	[0.015]	[0.037]	[0.060]	[0.045]	[0.078]
	[0.086]	[0.013]	[0.091]	[0.161]	[0.029]	[0.166]	[0.027]	[0.016]	[0.029]	[0.069]	[0.051]	[0.082]

which means that price-earnings ratios are more tightly connected to realized future earnings growth. For young firms, four-year earnings growth errors only explain 13.5% of dispersion in price-earnings ratios. Additionally, the returns and return forecast errors coefficients are smaller and less statistically significant for young firms compared to mature firms.

We repeat our exercise using I/B/E/S age rather than CRSP age and find similar results in Table III Panel C and Table V Panel B.[30] Young and mature firms differ significantly in the objective informativeness of earnings surprises, yet analyst updating $\gamma^*_{h,1}$ is nearly the same for both groups, leading to larger forecast errors in the price-earnings ratio decomposition for the mature firms. We also use another dimension of the data to find differences in the objective informativeness of earnings surprises. In Appendix I, we split firms into different industries and measure the estimates of γ_h for each industry. We then create two groups of firms, one for the industries with high γ_h and one for the industries with low γ_h. We then replicate Tables I and II for each of these two groups. Similar to what we find for young versus mature firms, we find that analyst updating $\gamma^*_{h,1}$ is similar for both groups and that forecast errors play a larger role in the price-earnings ratio decomposition for the high γ_h group. Thus, we find consistent evidence that, in line with constant-gain learning, differences in the objective informativeness of earnings surprises (i) does not noticeably change analyst updating but (ii) does change the importance of errors in expectations by changing the gap between the analysts' constant gain and the optimal gain.

V. Quantitative performance and full decomposition

In this section, we assess the quantitative accuracy of our model and use the model to estimate the full horizon decomposition CF_∞ and DR_∞. Importantly, we do not use any information from the empirical decomposition results to estimate the model parameters. Instead, we set all of the parameters based on time-series moments and previous estimates

[30] I/B/E/S age is based on the first month that the firm enters the I/B/E/S forecast database. CRSP age is based on the first month that the firm enters CRSP.

Table VI

Model estimation

This table shows the value of the six parameters of the model. The parameters for the aggregate cash flow process (ϕ, σ_u) are derived directly from the autocorrelation and standard deviation of the S&P 500 annual earnings growth. The firm-level volatility σ_v is derived directly from the standard deviation over time of the portfolio-level annual earnings growth. The risk-free rate r^f and risk sensitivity γ are set to match the average one-year Treasury yield and average aggregate equity return during the sample period. The constant-gain learning parameter β is taken from Malmendier and Nagel (2016). All moments are estimated over the full sample period of 1982 to 2020.

Parameter	Value	Moments
	Cash flow process	
ϕ	0.83	$AC(\Delta x_{t+1}^{agg})$
σ_u	0.34	$\sigma\left(\Delta x_{t+1}^{agg}\right)$
σ_v	0.10	$\sigma\left(\Delta x_{i,t+1}\right)$
	SDF	
r_f	4.6%	Risk-free rate
γ	1.61	Average aggregate return
	Learning	
β	1.8%	Constant-gain learning (Malmendier and Nagel, 2016)

of the constant gain β. As a result, we can fully utilize the empirical decomposition results of Table I to evaluate the quantitative realism of our model in explaining price-earnings ratio dispersion.

The model only has six parameters, which are all shown in Table VI. The parameters for cash flows $(\phi, \sigma_u, \sigma_v)$ are all estimated directly from realized earnings growth for our full sample of 1982-2020. For the aggregate process, the standard deviation and autocorrelation of S&P 500 earnings growth imply a persistence $\phi = 0.83$ and a volatility $\sigma_u = 0.34$. The volatility of individual shocks $\sigma_v = 0.10$ is obtained from the volatility over time of the portfolio-level earnings growth. Appendix D shows the exact formulas mapping these empirical moments to the model parameters. The constant-gain parameter is obtained from Malmendier and Nagel (2016) as $\beta = 0.018$. Note that this is nearly identical to the gain estimated in Milani (2007) of 0.0183.[31] For the agent's stochastic discount factor, the risk-free rate $r^f = 4.6\%$ and the sensitivity to risk $\gamma = 1.61$ are set to match the average one-year

[31]These papers estimate the constant-gain parameter on a quarterly frequency. We show in Appendix D that this estimation is quantitatively very similar to an estimation using an annual frequency. Conceptually, we are simply imposing that all updating occurs at the end of the year rather than allowing for small amounts of updating within the year.

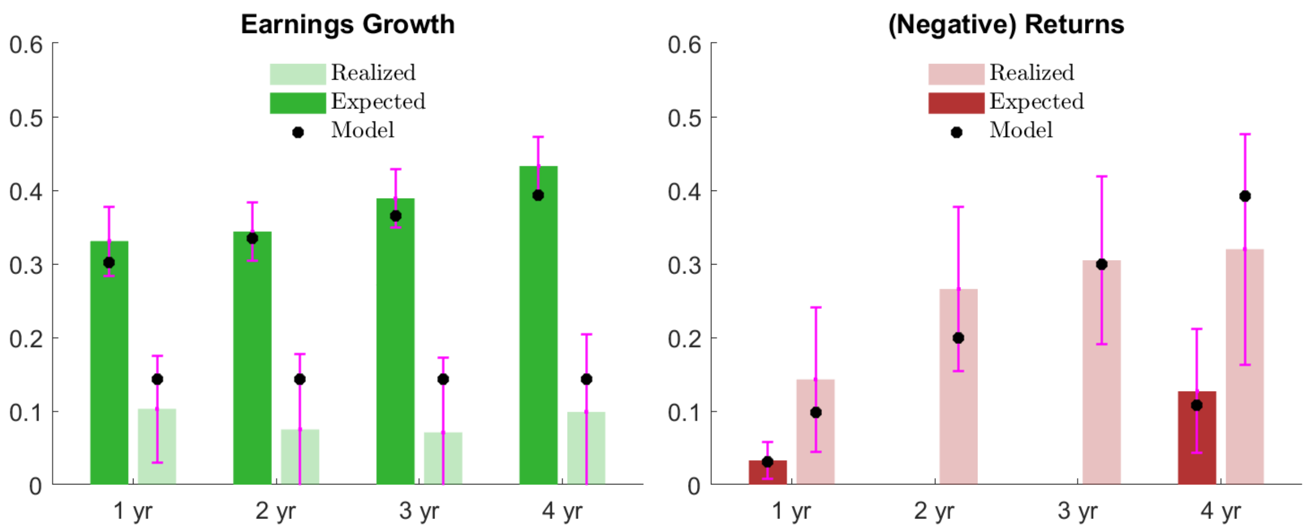

Figure 3. Empirical decomposition and model decomposition. This figure evaluates the decomposition of $\tilde{px}_{i,t}$ dispersion in the model across multiple horizons. The light bars show the contribution of realized earnings growth and realized returns to the dispersion of price-earnings ratios obtained in the *Realized* columns of Table I. The dark bars show the contribution of expected earnings growth and expected returns to the dispersion of price-earnings ratios obtained in the *Expected* columns of Table I. Each bar shows Driscoll-Kraay 95% confidence intervals. The black dots show the values of both the realized and expected decomposition implied by the model.

Treasury yield and average aggregate stock return of 10.5% for 1982-2020.

A. Model performance

A.1. Price ratio decomposition

We first evaluate the joint dynamics of price-earnings ratios, earnings growth, and returns.[32] Figure 3 shows the price-earnings ratio decomposition results from Figure 1, along with their 95% confidence intervals. For comparison, the black dots show the values implied by our model. Overall, the model successfully matches both the objective decomposition of price-earnings ratio dispersion (i.e., comovement of price ratios with future earnings growth and future returns) and the subjective decomposition (i.e., comovement of price ratios with expected earnings growth and expected returns) at every horizon.

[32]The model is simulated yearly over 500 periods for 300 firms. To avoid being impacted by the initial value of the expectations $g^*_{i,0}$, we calculate all moments after $t = 150$.

In the model, high price-earnings ratios are associated with significantly higher expected earnings growth and moderately lower expected returns. Figure 3 shows that a one unit increase in $\tilde{px}_{i,t}$ is associated with a 0.30 (0.39) increase in expected one-year (four-year) earnings growth and a 0.03 (0.11) decrease in expected returns. Because of the temporary shocks to the level of earnings $v_{i,t}$, realized one-year future earnings growth is partly predictable. In line with the data, a one unit increase in the model price-earnings ratio predicts a 0.14 increase in realized one-year earnings growth and this coefficient is unchanged as we increase the horizon to four years. The relationship between price-earnings ratios and expected earnings growth is quantitatively much larger than the relationship between price-earnings ratios and realized future earnings growth, meaning that high price-earnings ratios predict disappointment in future earnings growth. As a result, the relationship between price-earnings ratios and realized negative returns is larger than expected, 0.10 (0.39) at the one-year (four-year) horizon. Overall, this parsimonious model is able to closely match all 14 moments from the decomposition.

The fact that the model matches our decomposition results at multiple horizons highlights its success both in terms of magnitudes and in terms of timing. While the difference in expected and realized one-year earnings growth is large, this does not translate into a large difference between expected and realized one-year negative returns. Instead, agents are slow to adjust their beliefs and the disappointment in earnings growth leads to much lower than expected returns at longer horizons.

Because of this slow adjustment of prices, the model is able to simultaneously match the large one-year earnings growth disappointment shown in Figure 3 and the high empirical persistence of $\tilde{px}_{i,t}$, which is 0.77 in the data and 0.76 in the model. In general, these two facts are difficult to match for models in which growth expectations are sensitive to recent realizations (e.g., overstating the persistence of growth or diagnostic expectations of growth), as disappointing earnings growth for high $\tilde{px}_{i,t}$ firms would cause their price-earnings ratios to quickly fall. Further, our model is still consistent with previous cross-sectional evidence of overreaction. Using the Coibion and Gorodnichenko (2015) regression, we find that revisions in expected long-term growth $E_t^*[g_i] - E_{t-1}^*[g_i]$ negatively predict forecast errors $g_i - E_t^*[g_i]$ with a coefficient of -0.5, which is quantitatively similar to the empirical coefficients estimated in Bordalo et al. (2019) of -0.20 to -0.31. However, the model predicts that these revisions in expectations do not lead to large immediate changes in prices.

A.2. Price volatility and other moments

On top of the 14 untargeted moments shown in Figure 3, Table VII shows that the model performs well in matching a wide array of additional asset pricing moments. Perhaps the most surprising of these results is that the model generates realistic differences in returns $\tilde{r}_{i,t}$ across firms (i.e., the cross-sectional standard deviation) and realistic volatility over time in firms' $\tilde{r}_{i,t}$. This is almost entirely driven by errors in agents' expectations. Under a FIRE benchmark in which agents know the parameters $g_i = \bar{g}$, the cross-sectional standard deviation of $\tilde{r}_{i,t}$ and the idiosyncratic volatility of $\tilde{r}_{i,t}$ are both reduced by a factor of 5: from 5.6% to only 0.9% and from 5.5% to only 1.1%.

Beyond evaluating realized returns, the three panels of Table VII test the model's ability to match moments of price-earnings ratios, realized earnings growth, expected earnings growth, and expected returns. First, despite not using any price information in the estimation other than the average aggregate equity return, Panel A shows that the model generates realistic dynamics for the aggregate price-earnings ratio. The unconditional mean, volatility and autocorrelation of the log price-earnings ratio in the model (2.31, 43.2%, and 0.81) are consistent with the observed values (2.98, 42.5%, and 0.74) and the model generates volatile aggregate returns.

Second, while no information on cross-sectional dispersion was used in the estimation, Panel B shows that the model performs well in matching the empirical dispersion of nearly

Table VII

Model evaluation

This table evaluates the model by comparing the untargeted aggregate and cross-sectional moments in the model simulations with those observed in the data. Panel A shows the mean, standard deviation and autocorrelation of the aggregate price-earnings ratio as well as the standard deviation of aggregate stock returns. Panel B shows the cross-sectional standard deviations of price-earnings ratios, future earnings growth and returns, and expected earnings growth and returns. Panel C shows the idiosyncratic volatility across time of price-earnings ratios, future earnings growth and returns, and expected earnings growth and returns. All moments in the table are untargeted, except for idiosyncratic realized earnings growth volatility. Aggregate moments are estimated over the full sample period of 1982 to 2020. The cross-sectional standard deviation and idiosyncratic volatility moments are estimated over the main sample of 1999 to 2020 due to data availability.

Panel A: Aggregate value				
	Mean px_t	$\sigma(px_t)$	$AC(px_t)$	$\sigma(r_t)$
Model	2.31	43.2%	0.81	11.5%
Data	2.98	42.5%	0.74	15.9%

Panel B: Cross-sectional standard deviation					
	$\tilde{px}_{i,t}$	$\Delta\tilde{x}_{i,t+1}$	$\tilde{r}_{i,t+1}$	$E_t^*[\Delta\tilde{x}_{i,t+1}]$	$E_t^*[\tilde{r}_{i,t+1}]$
Model	20.9%	14.1%	5.6%	11.9%	0.8%
Data	22.6%	12.6%	5.7%	14.0%	2.6%

Panel C: Idiosyncratic volatility					
	$\tilde{px}_{i,t}$	$\Delta\tilde{x}_{i,t+1}$	$\tilde{r}_{i,t+1}$	$E_t^*[\Delta\tilde{x}_{i,t+1}]$	$E_t^*[\tilde{r}_{i,t+1}]$
Model	19.7%	14.0%	5.5%	11.7%	0.7%
Data	19.0%	16.6%	6.3%	12.3%	7.4%

all of our variables.[33] In other words, constant-gain learning with temporary shocks can successfully generate large differences across firms in price-earnings ratios and realized earnings growth, which have model dispersions of 20.9% and 14.1% respectively. Beyond explaining the realized data, we find that the model also accurately captures the large empirical dispersion in expected earnings growth (11.9%).

We do find that the model understates the cross-sectional dispersion in expected returns. For the sake of parsimony, in the model, subjective discount rates are entirely driven by risk premia related to cash flow timing, as shown in equation (26). Expanding the model to incorporate other risks into discount rates could help to better match this moment. However, as shown in Figure 3, the model still succeeds in matching the covariance of price-earnings ratios with expected returns. Thus, while the model does not capture all cross-sectional differences in expected returns, it does capture the portion that is predictable with price-

[33]The empirical dispersion is measured as the median cross-sectional standard deviation for each variable.

earnings ratios, i.e., the portion that is useful for generating large differences in price-earnings ratios.

Third, Panel C shows that the model replicates the measured portfolio-level volatilities. Note that these volatilities, which reflect variation in a single portfolio across time, are distinct from our estimates of dispersion, which capture cross-sectional variation across portfolios. The only information about portfolio-level volatility utilized in the estimation is the volatility of realized earnings growth, which means that the model provides an accurate mapping of how volatility in earnings growth translates into volatility in price-earnings ratios, returns, and expected earnings growth. Similar to our results for dispersion, we find that the model understates the volatility of expected returns.

In summary, we find that the model not only successfully matches the untargeted decomposition moments, but also generates realistic aggregate stock market moments as well as realistic cross-sectional dispersion and portfolio-level volatility. This demonstrates that a relatively parsimonious structural model of belief formation can feasibly improve upon FIRE models in terms of quantitatively matching the realized data while also matching the dynamics of empirically observed beliefs.

B. Full role of objective cash flows, cash flow mistakes, and discount rates

As mentioned in Section II, our empirical decomposition of price-earnings ratio dispersion is limited by the fact that data on subjective expectations is only available out to a four-year horizon. Because of this, roughly 40% of price-earnings ratio dispersion is attributed to subjective expectations of future price-earnings ratios. One of the key benefits of the quantitative model is that it allows us to estimate the full horizon decomposition CF_∞ and DR_∞. As shown in Figure 3, the model successfully matches the empirical decomposition results for the first four years without these moments being targeted, which increases the confidence in using the model to extend the decomposition to longer horizons. Table VIII Panel C shows the decomposition in equation (3) when we extend to the infinite hori-

Table VIII

Infinite-horizon decomposition and counterfactual analysis

Each column shows the decomposition implied by the constant-gain learning model using different key parameter choices. Panel A reports the varying parameters for each specification, with all other parameters fixed at the values from Table VI. The specifications include: (1) no learning or risk sensitivity ($\beta = 0$, $\gamma = 0$), (2) no learning ($\beta = 0$), (3) no risk sensitivity ($\gamma = 0$), and (4) the main specification. For specifications with $\gamma = 0$, we set the risk-free rate r_f to 10.5% to maintain consistent average equity returns across all cases. Panel B reports the mean aggregate price-earnings ratio and mean aggregate return for each specification. Panel C presents the cross-sectional results, showing the variance of $\tilde{r}_{i,t}$, the persistence and variance of $\tilde{px}_{i,t}$, and the decomposition of the cross-sectional variance of $\tilde{px}_{i,t}$ using the infinite-horizon version of equation (3):

$$Var\left(\tilde{px}_{i,t}\right) = Cov\left(\sum_{j=1}^{\infty} \rho^{j-1} E_t^*\left[\Delta \tilde{x}_{i,t+j}\right], \tilde{px}_{i,t}\right) + Cov\left(-\sum_{j=1}^{\infty} \rho^{j-1} E_t^*\left[\tilde{r}_{i,t+j}\right], \tilde{px}_{i,t}\right).$$

The fourth and seventh rows of Panel C show the contributions of expected earnings growth $\sum_{j=1}^{\infty} E_t^*\left[\Delta \tilde{x}_{i,t+j}\right]$ and subjective discount rates $-\sum_{j=1}^{\infty} \rho^{j-1} E_t^*\left[\tilde{r}_{i,t+j}\right]$ in the decomposition. The fifth and eighth rows show the contributions of realized earnings growth and returns, and the sixth and ninth rows show the contribution of earnings growth and return forecast errors. The share of the cross-sectional variance of $\tilde{px}_{i,t}$ is shown in parentheses. For ease of reading, all coefficients in Panel C except for "Persistence $\tilde{px}_{i,t}$" are scaled by 100.

Panel A: Parameter values				
β	0	0	0.018	0.018
γ	0	1.61	0	1.61
r^f	10.5%	4.6%	10.5%	4.6%
Panel B: Levels				
Mean px_t	2.31	2.31	2.31	2.31
Mean r_{t+1}	10.5%	10.5%	10.5%	10.5%
Panel C: Cross section				
Variance $\tilde{r}_{i,t}$	0.01	0.01	0.12	0.31
Persistence $\tilde{px}_{i,t}$	0.00	0.00	0.54	0.76
Variance $\tilde{px}_{i,t}$	0.98	0.98	2.07	4.37
Expected earnings growth	0.98	0.98	2.07	2.87
	(100%)	(100%)	(100%)	(65.7%)
Realized earnings growth	0.98	0.98	0.78	0.65
	(100%)	(100%)	(37.6%)	(14.9%)
Forecast errors	0	0	1.29	2.22
	(0%)	(0%)	(62.4%)	(50.8%)
Subjective discount rates	0	0	0	1.50
	(0%)	(0%)	(0%)	(34.3%)
Negative realized returns	0	0	1.29	3.72
	(0%)	(0%)	(62.4%)	(85.1%)
Negative forecast errors	0	0	-1.29	-2.22
	(0%)	(0%)	(-62.4%)	(-50.8%)

zon. Specifically, it shows the cross-sectional dispersion $Var\left(\tilde{px}_{i,t}\right)$ and the two components $Cov\left(\sum_{j=1}^{\infty}\rho^{j-1}E_t^*\left[\Delta\tilde{x}_{i,t+j}\right],\tilde{px}_{i,t}\right)$, $Cov\left(-\sum_{j=1}^{\infty}\rho^{j-1}E_t^*\left[\tilde{r}_{i,t+j}\right],\tilde{px}_{i,t}\right)$. Additionally, Panel C shows the persistence of $\tilde{px}_{i,t}$, which measures whether cross-sectional differences in price-earnings ratios are transitory or long-lived, and the dispersion in returns $\tilde{r}_{i,t}$.

To start, we focus on the final column, which is our main model parameterization. As shown in the fourth row of Panel C, the model estimates that differences in expected cash flow growth account for two-thirds (65.7%) of all dispersion in price-earnings ratios.[34] Combined with the aggregate time series findings of Delao and Myers (2021), this means that both time series variation in aggregate price ratios and cross-sectional dispersion in price ratios are primarily explained by expected cash flow growth. However, unlike the aggregate time series findings, we also estimate a non-trivial role for subjective discount rates in accounting for price-earnings ratio dispersion. The seventh row of Panel C shows that low subjective discount rates for high price-earnings ratio firms accounts for roughly one-third (34.3%) of all dispersion in price-earnings ratios.

Looking at the breakdown of the 65.7% contribution from expected earnings growth, we see that this largely comes from forecast errors. The comovement of price-earnings ratios with realized future earnings growth only accounts for 14.9% of the dispersion, meaning that the remaining 50.8% comes from price-earnings ratios predicting forecast errors for earnings growth. As a result, high price-earnings ratios are largely associated with low future returns, with negative realized returns accounting for 85.1% of all price-earnings ratio dispersion. Note that at the infinite horizon, forecast errors for earnings growth and forecast errors for returns are equal (i.e., the forecast error row for earnings growth and negative returns are exactly opposite). While gradual learning affects how quickly earnings growth surprises are reflected in unexpected returns, eventually all unexpected earnings growth will appear as unexpected returns.

Conveniently, we can summarize the relative importance of realized future earnings

[34] This is consistent with the empirical results of Table I, where we find that expected earnings growth over just the first four years already accounts for 43.3% of all price-earnings ratio dispersion.

growth, errors in earnings growth expectations, and subjective discount rates. The model estimates that realized earnings growth accounts for roughly 1/6 (14.9%) of price-earnings ratio dispersion, errors in earnings growth expectations account for 1/2 (50.8%), and subjective discount rates account for 1/3 (34.3%). Additionally, besides decomposing differences in price-earnings ratios, the model also decomposes the low realized returns earned by expensive stocks. The estimation implies that 40.3% (34.3/85.1) of the difference in returns between high and low price-earnings ratio stocks reflects subjective discount rates while 59.7% (50.8/85.1) reflects disappointment in earnings growth.

More broadly, by having a structural model, we can investigate the economic role of learning and risk sensitivity in driving the cross-sectional dispersion in price-earnings ratios and returns. The different columns in Table VIII Panel C show the dispersion in returns, the persistence in price-earnings ratios, the dispersion in price-earnings ratios, and the decomposition results when β and/or γ are set to 0, i.e., learning and/or risk sensitivity are turned off. In all cases, the initial expected g_i is set to $\bar{g}$ for all firms, which means that the $\beta = 0$ scenarios are FIRE (i.e., the agent always knows the true parameters).[35] Given that we are interested in cross-sectional dispersion rather than aggregate levels, in the two cases where $\gamma = 0$ we also raise the risk-free rate from 4.6% to 10.5%. As shown in Panel B, this ensures that the aggregate level for price-earnings ratios and equity returns are identical across all four cases and it is only the dispersion that changes. Thus, the two cases where $\gamma = 0$ are equivalent to saying that all firms have the same subjective discount rate of 10.5%.

In the first column, both β and γ are set to 0. In this case, the dispersion in price-earnings ratios is less than 1/4 the value in our main specification (0.98 compared to 4.37) and there is almost no dispersion in returns. The dispersion in price-earnings ratios comes entirely from differences in expected earnings growth, as there are no differences in subjective discount rates. Price-earnings ratios do not predict earnings growth forecast errors. Instead,

[35] If the prior is not $g_i = \bar{g}$, then $\beta = 0$ is not FIRE and actually represents the most stubborn errors in expectations possible. This produces highly persistent dispersion in price-earnings ratios, as agents start with incorrect expectations of g_i and never update.

all differences in expected earnings growth are simply due to the noise shocks $v_{i,t}$. Since these shocks are i.i.d., the autocorrelation in expected earnings growth is zero, which explains why the persistence in $\tilde{px}_{i,t}$ is also zero.

In the second column, the model includes risk sensitivity ($\gamma > 0$) but keeps $\beta = 0$. As shown in Panel C, only including risk sensitivity has no effect on the results relative to the first column. This highlights that, in our model, variation across firms in expected cash flow growth and subjective discount rates are both ultimately related to variation across firms in expected g_i. While agents may be sensitive to risk related to cash flow timing, this only matters if firms are expected to differ in the timing of their cash flows.

In comparison, the third column shows that including learning but keeping $\gamma = 0$ does substantially change the results. The dispersion in price-earnings ratios doubles from 0.98 to 2.07. This largely comes from price-earnings ratios now comoving with future earnings growth forecast errors. However, there is also the interesting result that the comovement of price-earnings ratios with realized earnings growth decreases (0.98 to 0.78). The FIRE expectation for future earnings growth is simply $-v_{i,t}$. With learning, expected earnings growth depends on $g^*_{i,t}$, which comoves positively with $v_{i,t}$, as a positive shock will tend to increase the guess for g_i. Thus, introducing learning means that the price-earnings ratio, which depends on expected earnings growth, will now be less related to future realized earnings growth due to the muted response to shocks $v_{i,t}$. Further, while objective expected earnings growth has 0 persistence over time, subjective expected earnings growth is persistent when agents are learning about g_i. Because of this, cross-sectional differences in price-earnings ratios are moderately long-lived, with a persistence of 0.54. The increased dispersion in price-earnings ratios also leads to a substantial increase in the dispersion of returns, from 0.01 to 0.12.

Finally, the last column shows the interaction from including both risk sensitivity and learning. While risk sensitivity by itself has no effect, once we incorporate learning, increasing γ from 0 to 1.61 more than doubles the dispersion in returns (0.12 to 0.31) and price-earnings ratios (2.07 to 4.37) and makes cross-sectional differences in price-earnings ratios more persistent. Specifically, including risk sensitivity along with learning increases the persistence of $\tilde{px}_{i,t}$ from 0.54 to 0.76, helping the model match the empirical persistence of 0.77. Looking at the contribution of subjective discount rates (the seventh row of Panel C), we clearly see the interaction between risk sensitivity and learning, as dispersion in subjective discount rates now contributes 1.50 (34.3%) to the total dispersion in price-earnings ratios.

More surprisingly, we also find an important interaction between risk sensitivity and learning for the contribution of earnings growth expectations. Given that γ has no impact on equations (16)-(19), changing γ has no effect on expected earnings growth. Thus, the increase in comovement between price-earnings ratios and expected earnings growth (2.07 to 2.87) is entirely due to changes in the price-earnings ratios. Intuitively, incorporating discount rates that depend on expected cash flow timing increases the sensitivity of price-earnings ratios to $g^*_{i,t}$ and decreases their sensitivity to transitory shocks $v_{i,t}$. The increased sensitivity to $g^*_{i,t}$ is reflected in the larger comovement of price-earnings ratios with expected earnings growth, and the reduced sensitivity to $v_{i,t}$ is reflected in an even higher persistence of $\tilde{px}_{i,t}$. This logic extends to any model with preferences for the timing of cash flows and shows that while discount rates may not affect expected earnings growth, they can be quantitatively important for driving the comovement of price ratios with expected earnings growth and earnings growth forecast errors.

Overall, the fact that dispersion in price-earnings ratios and returns for the full model (i.e., $\beta > 0, \gamma > 0$) is more than twice as large as any of the other counterfactuals highlights the natural interaction between preferences for the timing of cash flows and learning about cash flow growth. We find that this interaction is quantitatively important for matching the large empirical dispersion in returns and price-earnings ratios and the persistence of cross-sectional differences in price-earnings ratios.

VI. Robustness checks

A. *Alternative decomposition estimates*

We perform a number of alternative decompositions that serve as robustness checks for the results in Table I.

First, in addition to the Driscoll-Kraay and block-bootstrap standard errors reported in Table I, we also calculate the significance of our results under a worst-case scenario for overlapping observations. Specifically, in Appendix H.2, we perform Bauer and Hamilton (2018) simulations, which account for trends and potential small-sample bias, and assume a worst-case scenario for overlapping observations in which residuals are MA(47). Note that this substantially overstates the measured persistence of our residuals. We find that all of the earnings growth and return coefficients which are significant at the 5% level in Table I remain significant at the 5% level in Table AVII, even under this worst-case scenario.

Second, we estimate an exact decomposition to remove the approximation in equation (3). Table AV shows that this exact decomposition gives nearly identical results to Table I.

Third, in the spirit of Nagel (2024), Table AVIII addresses the concern that dispersion in price-earnings ratios may potentially be driven by fluctuations in one-year earnings rather than cross-sectional differences in prices. We find that the dispersion in price-earnings ratios is nearly identical to the dispersion in price-to-smoothed-three-year-earnings ratios.[36] Further, we show that the decomposition results are not changed in any noticeable way when we repeat the decomposition for price-to-smoothed-three-year-earnings ratios.

Fourth, we address potential survivorship bias. For our main estimation, we require that stocks have observed future prices and future earnings, as this allows us to study forecast errors in subjective expectations. However, in Table AIX, we remove this requirement and calculate future portfolio outcomes by reinvesting delisted stocks based on the delisting

[36]The standard deviation of cross-sectionally demeaned price-earnings ratios is 30.1% and the standard deviation of cross-sectionally demeaned price-to-three-year-earnings ratios is 32.6%. This demonstrates that smoothing the denominator does not reduce the cross-sectional dispersion in price ratios.

return. We find almost no change in our results.

B. Comparison of Value Line to I/B/E/S

This section provides validity checks for the results that rely on long-horizon return expectations constructed from Value Line's three-to-five-year price targets. Long-horizon price targets are not available in I/B/E/S, so Value Line is necessary to obtain return expectations beyond one year. It is important to emphasize that most of the results in the paper do not rely on Value Line expectations. Because each component of the price–earnings ratio decomposition is estimated in separate univariate regressions, Value Line data affect only the estimate of the long-horizon expected discount rate $DR_4$and play no role in the estimation of expected cash flows $CF_1 - CF_4$, short-horizon discount rates DR_1, any of the realized components, nor the stubbornness evidence. Nevertheless, since the Value Line expectations are used for part of the decomposition, it is useful to assess how well the Value Line expectations align with I/B/E/S consensus expectations. Fortunately, Value Line reports earnings expectations in addition to price forecasts starting in 1990, which creates an overlap with I/B/E/S earnings forecasts that we can use to assess whether Value Line shares similar beliefs.

Our first exercise directly compares the overlapping earnings forecasts between I/B/E/S and Value Line. Table AX regresses Value Line expectations on I/B/E/S expectations for one-year and four-year earnings. The estimated coefficients for each horizon are close to one, indicating near one-for-one comovement (0.919^{***},and 0.915^{***}) between Value Line and I/B/E/S consensus expectations. Table AX also performs the same comparison after normalizing expectations by the current price, $E_t^* \left[X_{t+1}\right] / P_t$ and $E_t^* \left[X_{t+4}\right] / P_t$, and finds similarly tight alignment. These results indicate that, for the objects that are jointly observed, Value Line expectations track the I/B/E/S consensus closely in the cross-section.

The relevant question for our decomposition, however, is not only whether the two providers' expectations comove with each other, but whether they comove with price ratios in a similar way. Panel B of Table AX therefore compares the cash flow coefficients obtained when we compute the cash flow components of the cross-sectional decomposition using Value Line rather than I/B/E/S expectations. The resulting CF_1 and CF_4 coefficients are very close to the I/B/E/S-based estimates (e.g., 0.423 versus 0.433 for CF_4), and the differences are not economically meaningful or statistically significant. This evidence implies that replacing I/B/E/S expectations with Value Line expectations does not change the key covariance moments that drive the cash flow component of the decomposition. Taken together with the near one-for-one comovement in Panel A, this provides direct support for using Value Line price targets to extend return expectations to longer horizons.

Finally, we examine time stability in Value Line's long-horizon return expectations. Panel C of Table AX reports the estimated DR_4 coefficient by subperiod. In our main Table I test, the sample ends in 2020, so the last Value Line expectations used are for 2016 as this allows us to compare the expectations to the realized future four-year returns. Thus, for Table AX Panel C, we consider two 13-year subperiods, 1990-2003 and 2004-2016. The estimates are stable across subsamples, providing no evidence that Value Line has systematically altered its long-horizon return expectations over time or that its implied long-horizon discount rate component has trended towards the FIRE benchmark.

Figure A4 also generalizes the subperiod result to show how it changes over time. It plots the expected DR_4 for a moving 13-year window. The very first point in Figure A4 corresponds to the 0.138 value shown in the first column of Panel C for a subperiod of 1990-2003 and the very last point in Figure A4 corresponds to the 0.127 value shown in the second column of Panel C for a subperiod of 2004-2016. The intermediate values show the results for all of the alternative windows. For example, we can consider windows that include both the dot-com bubble and the financial crisis. While the value of DR_4 does change depending on the subsample, the range is fairly minor, moving from a minimum of 0.10 to a maximum of 0.17, meaning that the results stay relatively close to our value from Table I of 0.127.

C. *Analyst updating*

We also assess whether the evidence on the gradual adjustment of earnings expectations can be attributed to infrequent updating by analysts or stale forecasts. Table II Panel B shows that we find a similar gradual adjustment in returns, which already points against the idea that gradual adjustment is purely a mechanical feature of the analyst data. However, to address this more thoroughly, we follow Bouchaud et al. (2019) and construct alternative consensus forecasts using detailed analyst-by-analyst data from the I/B/E/S Detail History File. Specifically, we include only those price forecasts and earnings forecasts that were explicitly updated within the current quarter.[37] This avoids any mechanical staleness in the data and ensures that all included forecasts reflect recent revisions.

Using this alternative consensus forecast, we re-estimate the impact of earnings surprises on the revision of expectations in our main $p\tilde{x}_{i,t}$ portfolios and the other 20 anomalies. Figure A5 reveals that the magnitude of the gradual adjustment in forecast revisions is consistent with the results in Figure 2 and with the extended anomaly results in the Appendix Figure A2. Tables AXI and AXII show the detailed estimates. This indicates that the documented stubbornness cannot be attributed to stale forecasts. We also calculate the subjective and realized decompositions using the alternative consensus forecasts and find no material differences in the main results, as demonstrated by Table AXIII.

The key reason why this change has such a limited impact on our results is because we are studying fairly long horizons. Our shortest horizon test is the one-year forecast revision. While staleness can occur at the monthly level, annual forecast revisions—which underpin our analysis—are infrequently stale, with most analysts providing updates within this timeframe. Combined, these results confirm that the gradual adjustment of expectations and our decomposition results are robust to concerns about staleness in forecasts.

[37] The reason why we do not make this our benchmark specification is that an analyst may choose not to update her forecasts specifically because her beliefs have not changed. In that scenario, dropping the analyst would be removing useful information about expectations.

VII. Conclusion

We find that subjective expectations have substantial potential to explain the cross-section of stock price ratios and shed light on the relative importance of expected future cash flows and discount rates. Using a variance decomposition, we show that cross-sectional dispersion in price-earnings ratios is primarily explained by predictable errors in subjective expectations of earnings growth. Subjective discount rates play a secondary, but non-trivial role. Disappointment in one-year earnings growth does not immediately lead to an equivalent disappointment in one-year returns. Instead, earnings growth surprises are reflected gradually in future returns over time. To understand these findings, we provide a quantitative model which not only outperforms standard FIRE models in matching the dynamics of prices and realized earnings growth and returns, but also outperforms several behavioral models in matching the dynamics of prices and expectations. The model features constant-gain learning about earnings growth and risk premia related to cash flow timing and emphasizes the importance of slow-moving beliefs in order to match the empirical timing of earnings growth expectations and realized returns.

These findings for the cross-section of stock prices are consistent with the aggregate time-series findings of Delao and Myers (2021, 2024) and Bordalo et al. (2023), who emphasize that aggregate stock prices are largely driven by subjective earnings growth expectations and that errors in these expectations play a particularly large role in explaining long-term returns. This harmony between the aggregate time-series and the cross-section indicates that a single mechanism could potentially explain both dimensions of the data and provides a strong motivation for further research understanding how investors form cash flow expectations and discount rates.

References

Amromin, Gene, and Steven A Sharpe, 2014, From the horse's mouth: Economic conditions and investor expectations of risk and return, *Management Science* 60, 845–66.

Bastianello, Federico, 2025, Time-series and cross-section of risk premia expectations: Evidence from financial analysts, *Working Paper,* SSRN Electronic Journal.

Basu, Sanjoy, 1975, The information content of price-earnings ratios, *Financial Management* 4, 53–64.

Bauer, Michael D., and James D. Hamilton, 2018, Robust bond risk premia, *Review of Financial Studies* 31, 399–448.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2025, Finance without exotic risk, *Journal of Financial Economics* 173, 104145.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2019, Diagnostic expectations and stock returns, *Journal of Finance* 74, 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2024, Belief overreaction and stock market puzzles, *Journal of Political Economy* 132, 1450–1484.

Bordalo, Pedro, Nicola Gennaioli Rafael, Rafael La Porta, Matthew Obrien, and Andrei Shleifer, 2023, Long term expectations and aggregate fluctuations, *NBER Macroeconomics Annual* 38.

Bouchaud, Jean Philippe, Philipp Krüger, Augustin Landier, and David Thesmar, 2019, Sticky expectations and the profitability anomaly, *The Journal of Finance* 74, 639–674.

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

Brunnermeier, Markus, Emmanuel Farhi, Ralph S J Koijen, Arvind Krishnamurthy, Sydney C Ludvigson, Hanno Lustig, Stefan Nagel, and Monika Piazzesi, 2021, Review article: Perspectives on the future of asset pricing, *The Review of Financial Studies* 34, 2126–2160.

Büsing, Pascal, and Hannes Mohrschladt, 2023, Stock market risk premia: The perspective of financial analysts, *Working Paper,* .

Campbell, John Y., 1991, A variance decomposition for stock returns, *The Economic Journal* 101, 157–179.

Cassella, Stefano, and Huseyin Gulen, 2018, Extrapolation bias and the predictability of stock returns by price-scaled variables, *Review of Financial Studies* 31, 4345–4397.

Chan, Kin Wai, Charles Chang, and Albert Wang, 2009, Put your money where your mouth is: Do financial firms follow their own recommendations?, *Quarterly Review of Economics and Finance* 49, 1095–1112.

Chan, Louis K C, Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *The Journal of Finance* 51, 1681–1713.

Chen, Huafeng, 2017, Do cash flows of growth stocks really grow faster?, *Journal of Finance* 72, 2279–2330.

Chen, Long, Zhi Da, and Xinlei Zhao, 2013, What drives stock price movements?, *Review of Financial Studies* 26, 841–876.

Cho, Thummim, and Christopher Polk, 2024, Putting the price in asset pricing, *The Journal of Finance* 79, 3943–3982.

Cochrane, John H., 2011, Presidential address: Discount rates, *The Journal of Finance* 66, 1047–1108.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *The Journal of Finance* 58, 609–642.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–2678.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer, 2016, Parameter learning in general equilibrium, *American Economic Review* 106, 664–698.

Cooper, Rick A., Theodore E. Day, and Craig M. Lewis, 2001, Following the leader: A study of individual analysts' earnings forecasts, *Journal of Financial Economics* 61, 383–416.

Couts, Spencer J., Andrei S. Gonçalves, and Johnathan Loudis, 2025, The subjective risk and return expectations of institutional investors, *Working Paper,* Fisher College of Business.

Da, Zhi, and Mitch Warachka, 2011, The disparity between long-term and short-term forecasted earnings growth, *Journal of Financial Economics* 100, 424–442.

Dahlquist, Magnus, and Markus Ibert, 2024, Equity return expectations and portfolios: Evidence from large asset managers, *The Review of Financial Studies* 37, 1887–1928.

Daniel, Kent, and David Hirshleifer, 2015, Overconfident investors, predictable returns, and excessive trading, *Journal of Economic Perspectives* 29, 61–88.

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

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

Delao, Ricardo, Xiao Han, and Sean Myers, 2025, The return of return dominance: Decomposing the cross-section of prices, *Journal of Financial Economics* 169, 104059.

Delao, Ricardo, and Sean Myers, 2021, Subjective cash flow and discount rate expectations, *The Journal of Finance* 76, 1339–1387.

Delao, Ricardo, and Sean Myers, 2024, Which subjective expectations explain asset prices?, *The Review of Financial Studies* 37, 1929–1978.

Décaire, Paul H, and John R Graham, 2024, Valuation fundamentals, *Working Paper,* Arizona State University.

Engelberg, Joseph, R David Mclean, and Jeffrey Pontiff, 2018, Anomalies and news, *The Journal of Finance* 73, 1971–2001.

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff, 2020, Analysts and anomalies, *Journal of Accounting and Economics* 69, 101249.

Fama, Eugene F, and Kenneth R French, 1995, Size and book-to-market factors in earnings and returns, *The Journal of Finance* 50, 131–155.

Fama, Eugene F, and Kenneth R French, 2020, The value premium, *The Review of Asset Pricing Studies* 11, 105–121.

Frankel, Richard, and Charles M.C. Lee, 1998, Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and Economics* 25, 283–319.

Gandhi, Mihir, Niels Joachim Gormsen, and Eben Lazarus, 2025, Forward return expectations, *Working Paper,* University of Chicago.

Graham, John R., 1999, Herding among investment newsletters: Theory and evidence, *The Journal of Finance* 54, 237–268.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

Gómez-Cram, Roberto, and Alastair Lawrence, 2025, The value of software, *American Economic Review* 115, 3514–3558.

Hillenbrand, Sebastian, and Odhrain McCarthy, 2026, Heterogeneous investors and stock market fluctuations, *Working Paper,* Harvard Business School.

Hirshleifer, David, Jun Li, and Jianfeng Yu, 2015, Asset pricing in production economies with extrapolative expectations, *Journal of Monetary Economics* 76, 87–106.

Hommel, Nicolas, Augustin Landier, and David Thesmar, 2025, Corporate valuation: An empirical comparison of discounting methods, *Working Paper,* SSRN Electronic Journal.

Hou, Kewei, G. Andrew Karolyi, and Bong Chan Kho, 2011, What factors drive global stock returns?, *Review of Financial Studies* 24, 2527–2574.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *The Review of Financial Studies* 28, 650–705.

Jensen, Theis Ingerslev, 2024, Subjective risk and return, *Working Paper,* Yale University.

Koijen, Ralph S.J., Hanno Lustig, and Stijn Van Nieuwerburgh, 2017, The cross-section and time series of stock and bond returns, *Journal of Monetary Economics* 88, 50–69.

Kothari, S. P., Eric So, and Rodrigo Verdi, 2016, Analysts' forecasts and asset pricing: A survey, *Annual Review of Financial Economics* 8, 197–219.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *The Journal of Finance* 73, 1183–1223.

La Porta, Rafael, 1996, Expectations and the cross-section of stock returns, *The Journal of Finance* 51, 1715–1742.

Lettau, Martin, and Jessica A. Wachter, 2007, Why is long-horizon equity less risky? a duration-based explanation of the value premium, *The Journal of Finance* 62, 55–92.

Lewellen, Jonathan, and Jay Shanken, 2002, Learning, asset-pricing tests, and market efficiency, *The Journal of Finance* 57, 1113–1146.

Malmendier, Ulrike, and Stefan Nagel, 2016, Learning from inflation experiences, *Quarterly Journal of Economics* 131, 53–87.

Martin, Ian W.R., and Christian Wagner, 2019, What is the expected return on a stock?, *The Journal of Finance* 74, 1887–1929.

Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis, 1999, Does forecast accuracy matter to analysts?, *The Accounting Review* 74, 185–200.

Milani, Fabio, 2007, Expectations, learning and macroeconomic persistence, *Journal of Monetary Economics* 54, 2065–2082.

Nagel, Stefan, 2024, Stock market valuation: Explanations, non-explanations, and some open questions, *Comments given at NBER SI Asset Pricing Meeting* .

Nagel, Stefan, and Zhengyang Xu, 2022, Asset pricing with fading memory, *Review of Financial Studies* 35, 2190–2245.

Orphanides, Athanasios, and John C. Williams, 2005, The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations, *Journal of Economic Dynamics and Control* 29, 1927–1950.

Piazzesi, Monika, Juliana Salomao, and Martin Schneider, 2015, Trend and cycle in bond premia, *Working Paper,* Stanford University.

Schwert, G. William, 2003, Anomalies and market efficiency, *Handbook of the Economics of Finance* 1, 939–974.

Skinner, Douglas J, and Richard G Sloan, 2002, Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio, *Review of Accounting Studies* 7, 289–312.

So, Eric C., 2013, A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts?, *Journal of Financial Economics* 108, 615–640.

Stattman, Dennis, 1980, Book values and stock returns, *The Chicago MBA: A journal of selected papers* 4, 25–45.

van Binsbergen, Jules H., Martijn Boons, Christian C. Opp, and Andrea Tamoni, 2023, Dynamic asset (mis)pricing: Build-up versus resolution anomalies, *Journal of Financial Economics* 147, 406–431.

van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira, 2022, Man versus machine learning: The term structure of earnings expectations and conditional biases, *The Review of Financial Studies* 36, 2361–2396.

Weber, Michael, 2018, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* 128, 486–503.

Zhang, Lu, 2005, The value premium, *The Journal of Finance* 60, 67–103.

Appendix

A. *Model prices and returns*

In the baseline model, after updating, the agent has no uncertainty about her believed value for g_i and the temporary shock $v_{i,t}$, meaning that she does not expect to change her beliefs in the future. Because of this, the agent's beliefs satisfy the law of iterated expectations (LIE); her expectation of her future beliefs is the same as her current beliefs. This means that asset prices are the same regardless of whether the agent considers the payoff of the one-period cash flow plus the future price or considers the payoff of the full stream of future cash flows. In Appendix B, we discuss a model of fading memory in which the agent's beliefs do not satisfy LIE and this distinction between which payoffs the agent considers is relevant.

To derive equation (20), we guess and verify a log-affine form for the strip price, $P_t^{(n)} = \exp\left\{A\left(n\right) + \phi^n x_t^{agg}\right\}$. The strip price is then pinned down by $P_t^{(0)} = \exp\left\{x_t^{agg}\right\}$ (i.e., $A\left(0\right) = 0$) and

$$\begin{aligned} P_t^{(n)} &= E_t^*\left[M_{t+1}P_{t+1}^{(n-1)}\right] \\ &= E_t^*\left[\exp\left\{-r^f - \frac{1}{2}\gamma^2\sigma_u^2 - \gamma u_{t+1} + A\left(n-1\right) + \phi^n x_t^{agg} + \phi^{n-1}u_{t+1}\right\}\right] \\ &= \exp\left\{-r^f - \frac{1}{2}\gamma^2\sigma_u^2 + A\left(n-1\right) + \phi^n x_t^{agg} + \frac{1}{2}\left(\phi^{n-1} - \gamma\right)^2\sigma_u^2\right\}. \end{aligned} \tag{A1}$$

This gives that

$$\begin{aligned} A\left(n\right) &= A\left(n-1\right) - r^f - \gamma\phi^{n-1}\sigma_u^2 + \frac{1}{2}\phi^{2(n-1)}\sigma_u^2 \\ &= -nr^f - \gamma\sigma_u^2\frac{1-\phi^n}{1-\phi} + \frac{1}{2}\sigma_u^2\frac{1-\phi^{2n}}{1-\phi^2}. \end{aligned} \tag{A2}$$

The expected and realized strip returns in equations (21)-(22) then simply utilize the formula for $P_t^{(n)}$. The firm price in equation (23) uses the independence of aggregate and idiosyncratic shocks to simplify $E_t^*\left[\left(\prod_{j=1}^n M_{t+j}\right)X_{t+n}^{agg}\tilde{X}_{i,t+n}\right] = E_t^*\left[\left(\prod_{j=1}^n M_{t+j}\right)X_{t+n}^{agg}\right]E_t^*\left[\tilde{X}_{i,t+n}\right] = P_t^{(n)}E_t^*\left[\tilde{X}_{i,t+n}\right]$.

Given that the firm price is simply a collection of strip prices, the return for a firm is

$$
\begin{aligned}
R_{i,t+1} &= \frac{\tilde{X}_{i,t+1}X_{t+1}^{agg}+P_{i,t+1}}{P_{i,t}} = \frac{\sum_{n=1}^{\infty}P_{t+1}^{(n-1)}E_{t+1}^{*}\left[\tilde{X}_{i,t+n}\right]}{\sum_{n=1}^{\infty}P_{t}^{(n)}E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]} \\
&= \sum_{n=1}^{\infty}\frac{P_{t}^{(n)}E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]}{\sum_{n=1}^{\infty}P_{t}^{(n)}E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]}\frac{P_{t+1}^{(n-1)}}{P_{t}^{(n)}}\frac{E_{t+1}^{*}\left[\tilde{X}_{i,t+n}\right]}{E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]} \\
&= \sum_{n=1}^{\infty}w_{i,t,n}R_{t+1}^{(n)}\frac{E_{t+1}^{*}\left[\tilde{X}_{i,t+n}\right]}{E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]} \qquad \text{(A3)}
\end{aligned}
$$

where the weight is $w_{i,t,n} = \frac{P_t^{(n)}E_t^*\left[\tilde{X}_{i,t+n}\right]}{\sum_{n=1}^{\infty}P_t^{(n)}E_t^*\left[\tilde{X}_{i,t+n}\right]} = \frac{P_t^{(n)}\exp\left\{\frac{1}{2}\sigma_v^2+\tilde{x}_{i,t}+ng_{i,t}^*-v_{i,t}^*\right\}}{\sum_{n=1}^{\infty}P_t^{(n)}\exp\left\{\frac{1}{2}\sigma_v^2+\tilde{x}_{i,t}+ng_{i,t}^*-v_{i,t}^*\right\}}$ from equation (19). Note that the weight reduces to $w_{i,t} = \exp\left\{ng_{i,t}^*\right\}P_t^{(n)}/\left(\sum_{n=1}^{\infty}\exp\left\{ng_{i,t}^*\right\}P_t^{(n)}\right)$ by canceling out the $\exp\left\{\frac{1}{2}\sigma_v^2+\tilde{x}_{i,t}-v_{i,t}^*\right\}$ terms in the numerator and denominator. The agent's subjective expected return is

$$
\begin{aligned}
E_t^*\left[R_{i,t+1}\right] &= E_t^*\left[\sum_{n=1}^{\infty}w_{i,t,n}R_{t+1}^{(n)}\frac{E_{t+1}^{*}\left[\tilde{X}_{i,t+n}\right]}{E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]}\right] \\
&= \sum_{n=1}^{\infty}w_{i,t,n}E_t^*\left[R_{t+1}^{(n)}\right]E_t^*\left[\frac{E_{t+1}^{*}\left[\tilde{X}_{i,t+n}\right]}{E_{t}^{*}\left[\tilde{X}_{i,t+n}\right]}\right] \\
&= \sum_{n=1}^{\infty}w_{i,t,n}E_t^*\left[R_{t+1}^{(n)}\right] \qquad \text{(A4)}
\end{aligned}
$$

where the second line uses the independence of the aggregate shock and idiosyncratic shock and the third line uses LIE.

B. *Fading memory model*

Nagel and Xu (2022) present a model of learning the mean of an i.i.d. process with fading memory. Using $\Delta\tilde{x}_t = g + v_t$ to temporarily denote the hypothetical i.i.d. process, they show that an agent whose posterior is distorted by fading memory in this setting will believe

$$
g \sim N\left(g_t^*, \beta\sigma_v^2\right) \qquad \text{(A5)}
$$

where β is the parameter that controls the degree of fading memory and σ_v^2 is the variance of the normal shocks v_t. The agent's beliefs follow the updating rule

$$g_t^* = g_{t-1}^* + \beta\left(\Delta\tilde{x}_t - E_{t-1}^*\left[\Delta\tilde{x}_t\right]\right). \tag{A6}$$

In the spirit of Nagel and Xu (2022), our goal is to extend this analysis to the scenario of an MA(1) process in a way that preserves equations (A5)-(A6). As discussed in Section II.D and Appendix F, the inclusion of these temporary shocks to the level of the firm-specific component of cash flows (i.e., the MA(1) shocks to the growth of the firm-level component) is only a small modification, but does generate important differences by making prices less sensitive to earnings growth surprises and causing the agent to expect lower next period earnings growth after a positive surprise when β is sufficiently low. For simplicity, we drop the i subscripts as we are focusing on learning g_i for a single given firm.

In our setting, we have that $\Delta\tilde{x}_t$ is MA (1),

$$\Delta\tilde{x}_t = g - v_{t-1} + v_t.$$

Let $\alpha_t \equiv \begin{pmatrix} g \\ g - v_{t-1} \end{pmatrix}$ denote the key variables that the agent needs to learn. These represent the long-term growth g and the average next period growth $g - v_{t-1}$. Let $F = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $G = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Then we have that the law of motion for α_t is

$$\alpha_t = F\alpha_{t-1} + G\Delta\tilde{x}_{t-1}. \tag{A7}$$

The law of motion for $\Delta\tilde{x}_t$ (i.e., the measurement equation) is

$$\Delta\tilde{x}_t = L\alpha_t + v_t \tag{A8}$$

where $L \equiv \begin{pmatrix} 0 & 1 \end{pmatrix}$.

The agent enters the period with prior $\alpha_t|H_{t-1} \sim N\left(\hat{\alpha}_{t|t-1}, P_{t|t-1}\right)$ based on the history H_{t-1} of $\{\Delta\tilde{x}_{t-j}\}_{j=1}^{\infty}$. We assume that the agent forms her posterior $p\left(\alpha_t \mid H_t\right)$ using a distorted version $\tilde{p}\left(\alpha_t \mid H_{t-1}\right)$ of her prior. Specifically, her posterior is formed using

$$p\left(\alpha_t \mid H_t\right) \propto p\left(\Delta\tilde{x}_t \mid \alpha_t\right)\tilde{p}\left(\alpha_t \mid H_{t-1}\right) \tag{A9}$$

where

$$\tilde{p}\left(\alpha_t \mid H_{t-1}\right) \quad \propto \quad \exp\left(-\frac{1}{2}\left(\alpha_t - \hat{\alpha}_{t|t-1}\right)' W' P_{t|t-1}^{-1} W \left(\alpha_t - \hat{\alpha}_{t|t-1}\right)\right). \qquad \text{(A10)}$$

The matrix W reflects how fading memory impacts the agent's updating. In the univariate α_t case of Nagel and Xu (2022), $W = 1 - \beta$ is a constant. In our bivariate α_t case, W is a 2x2 matrix. We choose

$$W \quad = \quad (1-\beta)^{0.5} F$$

because this ensures our setting preserves equations (A5)-(A6) and also maps directly to our baseline model equations (16)-(17), as we will show below. Note that $\det(W) = 1 - \beta$ controls how much the agent's memory fades over time.

Let $\tilde{P}_{t|t-1} \equiv W^{-1} P_{t|t-1} W^{-1\prime}$ denote her distorted prior covariance. Given this distorted prior and equation (A8), the agent's prior belief for $\Delta\tilde{x}_t$ is $\Delta\tilde{x}_t | H_{t-1} \sim N\left(L\hat{\alpha}_{t|t-1}, S_t\right)$ where $S_t = L\tilde{P}_{t|t-1}L' + \sigma_v^2$. The surprise is defined as

$$\begin{aligned} z_t \quad &\equiv \quad \Delta\tilde{x}_t - L\hat{\alpha}_{t|t-1} \\ &= \quad L\left(\alpha_t - \hat{\alpha}_{t|t-1}\right) + v_t. \end{aligned}$$

After observing the surprise, the agent updates her beliefs

$$\hat{\alpha}_{t|t} \quad = \quad \hat{\alpha}_{t|t-1} + K_t z_t$$

choosing K_t to minimize the posterior mean-squared error, $E\left[\left(\alpha_t - \hat{\alpha}_{t|t}\right)' \left(\alpha_t - \hat{\alpha}_{t|t}\right) | H_t\right]$. Given that $\alpha_t - \hat{\alpha}_{t|t} = (I - K_t L)\left(\alpha_t - \hat{\alpha}_{t|t-1}\right) - K_t v_t$, the solution to this minimization gives the classic expression

$$K_t \quad = \quad \tilde{P}_{t|t-1} L' \left(L\tilde{P}_{t|t-1}L' + \sigma_v^2\right)^{-1} \qquad \text{(A11)}$$

which means the posterior covariance matrix is

$$P_{t|t} \quad = \quad (I - K_t L)\,\tilde{P}_{t|t-1}\,(I - K_t L)' + K_t K_t' \sigma_v^2. \qquad \text{(A12)}$$

This also means that the distorted prior for the next period is

$$\tilde{P}_{t+1|t} = W^{-1}FP_{t|t}F'W^{-1\prime} = \frac{1}{1-\beta}P_{t|t}. \tag{A13}$$

and her $\hat{\alpha}_{t+1|t}$ that she carries into the next period is

$$\begin{aligned} \hat{\alpha}_{t+1|t} &= F\hat{\alpha}_{t|t} + G\Delta\tilde{x}_t \\ &= F\hat{\alpha}_{t|t-1} + FK_t z_t + G\left(z_t + L\hat{\alpha}_{t|t-1}\right) \\ &= (F+GL)\,\hat{\alpha}_{t|t-1} + (FK_t + G)\,z_t. \end{aligned} \tag{A14}$$

Equations (A11)-(A13) summarize the agent's learning problem. In steady state, these equations become

$$\begin{aligned} K &= \tilde{P}_{prior}L'\left(L\tilde{P}_{prior}L' + \sigma_v^2\right)^{-1} \\ \tilde{P}_{prior} &= \frac{1}{1-\beta}P_{post} \\ P_{post} &= (I-KL)\,\tilde{P}_{prior}\,(I-KL)' + KK'\sigma_v^2 \end{aligned}$$

which is solved by

$$\begin{aligned} K &= \beta\iota' \\ \tilde{P}_{prior} &= \frac{\beta\sigma_v^2}{1-\beta}\iota'\iota \\ P_{post} &= \beta\sigma_v^2\iota'\iota. \end{aligned}$$

where $\iota \equiv \begin{pmatrix} 1 & 1 \end{pmatrix}$. This means that, in steady state, the agent's posterior belief about g is $g \sim N\left(g_t^*, \beta\sigma_v^2\right)$ where

$$g_t^* = g_{t-1}^* + \beta\left(\Delta\tilde{x}_t - E_{t-1}^*\left[\Delta\tilde{x}_t\right]\right). \tag{A15}$$

So, we can confirm that our extended model preserves equations (A5)-(A6). Further, the updating rule for g_t^* exactly aligns with equation (16) in our baseline model, and we can rearrange the updating rule (A14) for $\hat{\alpha}_{t+1|t}$ to see that the agent's beliefs also align with equation (17) in our baseline model,

$$\begin{aligned} v_t^* &= \begin{pmatrix} 1 & -1 \end{pmatrix} \hat{\alpha}_{t+1|t} \\ &= (1-\beta)\left(\Delta\tilde{x}_t - E_{t-1}^*\left[\Delta\tilde{x}_t\right]\right). \end{aligned} \tag{A16}$$

Additionally, in steady-state the value for $S = L\tilde{P}_{prior}L' + \sigma_v^2$ is simply $\frac{1}{1-\beta}\sigma_v^2$.

B.1. Asset Prices

Throughout, we focus on the steady state pricing results, i.e., the solutions using the steady state K, $\tilde{P}_{prior}$, P_{post}, and S. We also reintroduce the i subscript notation to distinguish firm-level prices from aggregate prices. Our aggregate strip pricing is unchanged from the baseline model, as the aggregate process is unaffected by learning about the firm-level component. Let $P_t^{(n)}$ be the price of an n-period aggregate strip. The aggregate strip price is

$$\begin{aligned} P_t^{(n)} &= E_t^*\left[\left(\prod_{j=1}^{n} M_{t+j}\right) X_{t+n}^{agg}\right] \\ &= \exp\left\{-nr^f - \gamma\sigma_u^2\frac{1-\phi^n}{1-\phi} + \frac{1}{2}\sigma_u^2\frac{1-\phi^{2n}}{1-\phi^2} + \phi^n x_t^{agg}\right\}. \end{aligned} \tag{A17}$$

The realized return on the strip and the subjective expected return on the strip are

$$R_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)}}{P_t^{(n)}} = \exp\left\{r^f + \gamma\sigma_u^2\phi^{n-1} - \frac{1}{2}\sigma_u^2\phi^{2(n-1)} + \phi^{n-1}u_{t+1}\right\} \tag{A18}$$

$$E_t^*\left[R_{t+1}^{(n)}\right] = \exp\left\{r^f + \gamma\sigma_u^2\phi^{n-1}\right\}. \tag{A19}$$

However, for firm-level strip prices, the details of the fading memory model become important. In particular, as discussed in Nagel and Xu (2022), the distortion to the agent's posterior introduced by fading memory means that the agent's beliefs about the firm-specific component of cash flows do not satisfy the law of iterated expectations. Because of this, the price of a firm's strip differs depending whether we take a "resale" approach in which the agent values the asset based on its expected next period payoff or the "buy-and-hold" approach in which the agent values the asset based on its expected cash flow at maturity. As in Nagel and Xu (2022), we focus on the "resale" approach as this ensures firm valuation is time consistent.

For a one-period strip, both approaches give the same result. The price of a one-period strip for firm i is

$$\begin{aligned} P_{i,t}^{(1)} &= E_t^*\left[M_{t+1}X_{t+1}^{agg}\tilde{X}_{i,t+1}\right] \\ &= E_t^*\left[M_{t+1}X_{t+1}^{agg}\right]E_t^*\left[\tilde{X}_{i,t+1}\right] = P_t^{(1)}E_t^*\left[\tilde{X}_{i,t+1}\right] \\ &= P_t^{(1)}\tilde{X}_{i,t}E_t^*\left[\exp\left(\Delta\tilde{x}_{i,t+1}\right)\right] = P_t^{(1)}\tilde{X}_{i,t}\exp\left(L\hat{\alpha}_{t+1|t}+\frac{1}{2}\frac{1}{1-\beta}\sigma_v^2\right) \end{aligned}$$

where the second line uses the independence of the aggregate and firm-level components and the last line uses the fact that the agent believes $\Delta\tilde{x}_{i,t+1}|H_t \sim N\left(L\hat{\alpha}_{t+1|t}, \frac{1}{1-\beta}\sigma_v^2\right)$. For strip prices of horizon 2 and beyond, the distinction between the two approaches does matter. Using the "resale" approach, we guess and confirm the following functional form

$$P_{i,t}^{(n)} = P_t^{(n)}\tilde{X}_{i,t}\exp\left(\tilde{A}(n)+\tilde{B}(n)\hat{\alpha}_{t+1|t}\right). \tag{A20}$$

Let $\tilde{F} \equiv F + GL$, $\tilde{K} \equiv FK + G$. Under the conjectured functional form, strip prices satisfy

$$\begin{aligned} P_{i,t}^{(n)} &= E_t^*\left[M_{t+1}P_{i,t+1}^{(n-1)}\right] \\ &= E_t^*\left[M_{t+1}P_{t+1}^{(n-1)}\right]\tilde{X}_{i,t}E_t^*\left[\exp\left(\Delta\tilde{x}_{i,t+1}+\tilde{A}(n-1)+\tilde{B}(n-1)\hat{\alpha}_{t+2|t+1}\right)\right] \\ &= P_t^{(n)}\tilde{X}_{i,t}E_t^*\left[\exp\left(\Delta\tilde{x}_{i,t+1}+\tilde{A}(n-1)+\tilde{B}(n-1)\left(\tilde{F}\hat{\alpha}_{t+1|t}+\tilde{K}z_{t+1}\right)\right)\right] \\ &= P_t^{(n)}\tilde{X}_{i,t}E_t^*\left[\exp\left(\left(1+\tilde{B}(n-1)\tilde{K}\right)\Delta\tilde{x}_{i,t+1}+\tilde{A}(n-1)+\tilde{B}(n-1)\left(\tilde{F}-\tilde{K}L\right)\hat{\alpha}_{t+1|t}\right)\right] \\ &= P_t^{(n)}\tilde{X}_{i,t}\exp\left(\tilde{A}(n-1)+\left[\tilde{B}(n-1)\tilde{F}+L\right]\hat{\alpha}_{t+1|t}+\frac{1}{2}\left(1+\tilde{B}(n-1)\tilde{K}\right)^2\frac{1}{1-\beta}\sigma_v^2\right). \end{aligned}$$

This implies that

$$\tilde{B}(n) = \tilde{B}(n-1)\tilde{F}+L = \begin{pmatrix} n-1 & 1 \end{pmatrix} \tag{A21}$$

$$\begin{aligned} \tilde{A}(n) &= \tilde{A}(n-1)+\frac{1}{2}\left(1+\tilde{B}(n-1)\tilde{K}\right)^2\frac{1}{1-\beta}\sigma_v^2 \\ &= \tilde{A}(n-1)+\frac{1}{2}(n\beta)^2\frac{1}{1-\beta}\sigma_v^2 \\ &= \frac{1}{2}\frac{1}{1-\beta}\sigma_v^2\left[1+\beta^2\sum_{j=2}^{n}j^2\right]. \end{aligned} \tag{A22}$$

Now, we can combine the prices of the firm-level strips to conclude that the price of the firm is

$$\begin{aligned} P_{i,t} &= \sum_{n=1}^{\infty} P_{i,t}^{(n)} = \sum_{n=1}^{\infty} P_t^{(n)} \tilde{X}_{i,t} \exp\left\{\tilde{A}(n) + \tilde{B}(n)\,\hat{\alpha}_{t+1|t}\right\} \\ &= \sum_{n=1}^{\infty} P_t^{(n)} \exp\left\{\tilde{A}(n) + \tilde{x}_{i,t} + n g_{i,t}^* - v_{i,t}^*\right\} \end{aligned} \tag{A23}$$

where $g_{i,t}^* = \begin{pmatrix} 1 & 0 \end{pmatrix} \hat{\alpha}_{t+1|t}$ and $v_{i,t}^* = \begin{pmatrix} 1 & -1 \end{pmatrix} \hat{\alpha}_{t+1|t}$. As noted earlier, $g_{i,t}^*$ and $v_{i,t}^*$ in the fading memory model are exactly the same as equations (16) and (17) in our baseline model. Thus, the only difference between firm-level prices (25) in the baseline model and (A23) in the fading memory model is the Jensen's term $\frac{1}{2}\sigma_v^2$ versus $\frac{1}{2}\frac{1}{1-\beta}\sigma_v^2\left[1 + \beta^2 \sum_{j=2}^{n} j^2\right]$. The higher value for the Jensen's term in the fading memory model reflects that the agent knows that her future self will have different beliefs about g_i which adds additional variance to the future log strip prices. Note that when $\beta = 0$ (i.e., the agent does not update her beliefs), the Jensen's terms for the fading memory model are identical to the baseline model.

For small values of β, the change to the Jensen's term is relatively small for many values of n. However, there is a relevant technical point that the Jensen's term eventually grows to infinity for large enough n. This is also noted in Nagel and Xu (2022): "Since the price of a dividend claim is convex in dividend growth rates, the subjective growth rate uncertainty in this model could cause the price to be infinite."[38] For our quantitative exercise, we resolve this issue by simply assuming that the stock price is only a claim to firm cash flows for the next 100 years. The quantitative results are not sensitive to the choice of 100 years versus 150 years or 200 years, but simply depend on the choice to use a finite number of periods.

[38] Nagel and Xu (2022) resolve this problem by including a drift in the cash flow that is small enough to not noticeably impact strip prices for most horizons but prevents the strip prices from exploding.

B.2. Returns

Let $\tilde{c}(1) = 1$ and $\tilde{c}(n) = n\beta$ for $n > 1$. The return on a firm strip is

$$\begin{aligned}
R_{i,t+1}^{(n)} &= \frac{P_{i,t+1}^{(n-1)}}{P_{i,t}^{(n)}} = \frac{P_{t+1}^{(n-1)} \tilde{X}_{i,t+1} \exp\left(\tilde{A}(n-1) + \tilde{B}(n-1)\hat{\alpha}_{t+2|t+1}\right)}{P_t^{(n)} \tilde{X}_{i,t} \exp\left(\tilde{A}(n) + \tilde{B}(n)\hat{\alpha}_{t+1|t}\right)} \\
&= \frac{P_{t+1}^{(n-1)} \tilde{X}_{i,t+1} \exp\left(\tilde{A}(n-1) + \tilde{B}(n-1)\left(\tilde{F}\hat{\alpha}_{t+1|t} + \tilde{K} z_{i,t+1}\right)\right)}{P_t^{(n)} \tilde{X}_{i,t} \exp\left(\tilde{A}(n) + \tilde{B}(n)\hat{\alpha}_{t+1|t}\right)} \\
&= R_{t+1}^{(n)} \frac{\tilde{X}_{i,t+1} \exp\left(\tilde{B}(n-1)\tilde{K} z_{i,t+1}\right)}{\tilde{X}_{i,t} \exp\left(\tilde{A}(n) - \tilde{A}(n-1) + L\hat{\alpha}_{t+1|t}\right)} \\
&= R_{t+1}^{(n)} \exp\left(-\frac{1}{2}\tilde{c}(n)^2 \frac{1}{1-\beta}\sigma_v^2 + \tilde{c}(n) z_{i,t+1}\right). \qquad \text{(A24)}
\end{aligned}$$

The return on the firm is a weighted average of the firm-level strip returns,

$$\begin{aligned}
R_{i,t+1} &= \frac{\tilde{X}_{i,t+1} X_{t+1}^{agg} + P_{i,t+1}}{P_{i,t}} = \sum_{n=1}^{\infty} \frac{P_{i,t}^{(n)}}{P_{i,t}} \frac{P_{i,t+1}^{(n-1)}}{P_{i,t}^{(n)}} \\
&= \sum_{n=1}^{\infty} w_{i,t,n} R_{i,t+1}^{(n)} \qquad \text{(A25)}
\end{aligned}$$

where the weight is $w_{i,t,n} = \frac{P_t^{(n)} \exp\left(\tilde{A}(n) + \tilde{B}(n)\hat{\alpha}_{t+1|t}\right)}{\sum_{n=1}^{\infty} P_t^{(n)} \exp\left(\tilde{A}(n) + \tilde{B}(n)\hat{\alpha}_{t+1|t}\right)}$. The subjective expected return on the firm is then

$$\begin{aligned}
E_t^*\left[R_{i,t+1}\right] &= \sum_{n=1}^{\infty} w_{i,t,n} E_t^*\left[R_{i,t+1}^{(n)}\right] \\
&= \sum_{n=1}^{\infty} w_{i,t,n} E_t^*\left[R_{t+1}^{(n)}\right]. \qquad \text{(A26)}
\end{aligned}$$

Thus, the subjective expected return is identical to our baseline model, except for a change in the weights $w_{i,t,n}$.

B.3. Comparison of decomposition

Figure A1 shows a comparison of our empirical decomposition results, the model decomposition results for the baseline model, and the model decomposition results for the fading memory model. Quantitatively, the results for the fading memory model are quite close to

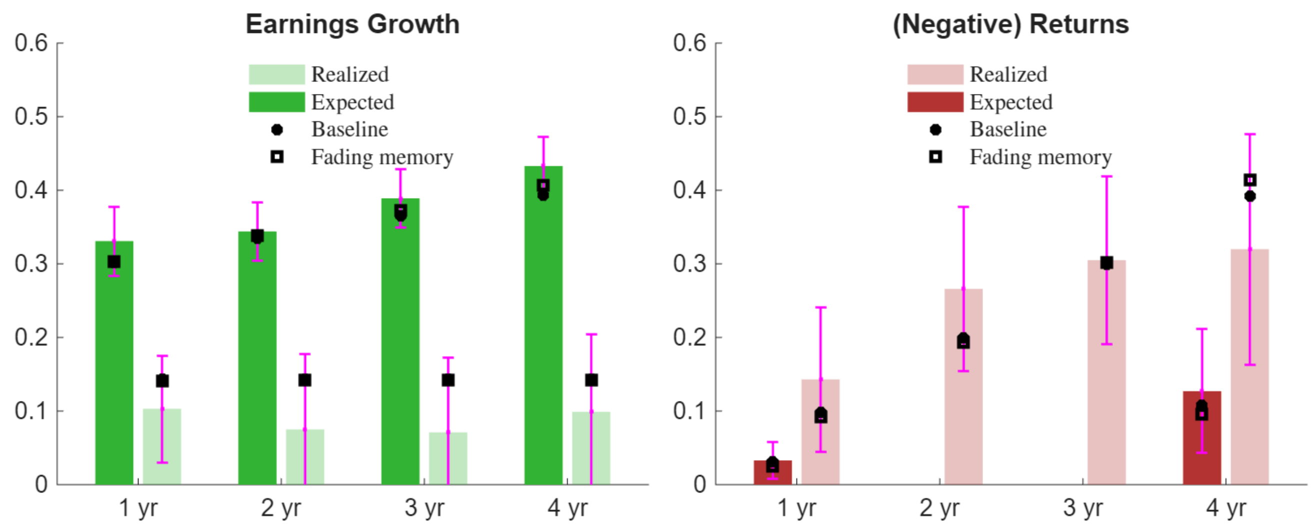

Figure A1. Baseline model and fading memory model decomposition. This figure evaluates the decomposition of $\tilde{px}_{i,t}$ dispersion in the model across multiple horizons. The light bars show the contribution of realized earnings growth and realized returns to the dispersion of price-earnings ratios obtained in the *Realized* columns of Table I. The dark bars show the contribution of expected earnings growth and expected returns to the dispersion of price-earnings ratios obtained in the *Expected* columns of Table I. Each bar shows Driscoll-Kraay 95% confidence intervals. The black dots show the values of both the realized and expected decomposition implied by the baseline model. The black hollow squares show the values of both the realized and expected decomposition implied by the fading memory model.

the baseline model. In fact, for many moments, the hollow square for the fading memory result almost perfectly overlaps with the solid circle for the baseline model result.

To align with the fact that the fading memory model uses the "resale" approach to price assets, we use nested expectations when calculating the Campbell-Shiller decomposition,

$$\begin{aligned}\tilde{px}_{i,t} &\approx E_t^*\left[\Delta\tilde{x}_{i,t+1}\right] - E_t^*\left[\tilde{r}_{i,t+1}\right] + \rho\tilde{px}_{i,t+1} \\ &\approx \sum_{j=1}^{h}\rho^{j-1}E_t^*\left[...E_{t+j-1}^*\left[\Delta\tilde{x}_{i,t+j}\right]\right] - \sum_{j=1}^{h}\rho^{j-1}E_t^*\left[...E_{t+j-1}^*\left[\tilde{r}_{i,t+j}\right]\right] + \rho^h E_t^*\left[...E_{t+j-1}^*\left[\tilde{px}_{i,t+h}\right]\right].\end{aligned}$$

This ensures that the DR_h shown in Figure A1 reflects the discount rates used to price the assets, i.e. a series of one-period required returns. This has no effect on the earnings growth component, as $E_t^*\left[E_{t+1}^*\left[\Delta\tilde{x}_{i,t+2}\right]\right] = E_t^*\left[\Delta\tilde{x}_{i,t+2}\right]$.

C. *Model extensions*

C.1. Stubborn errors and rational arbitrageurs

This extension shows that, in the presence of rational arbitrageurs with shorting costs, a lower value of β limits the scope for arbitrageurs to intervene, allowing errors in expectations to contribute more significantly to price dispersion.

First, consider the model of Section III without any modifications. From equation (20), strip prices are simply functions of the aggregate state variable x_t^{agg}. If agents had FIRE beliefs (i.e., agents knew the true parameters), then all firms would have the same price, $P_t^{FIRE} \equiv \exp\left\{\frac{1}{2}\sigma_v^2\right\} \sum_{n=1}^{\infty} P_t^{(n)}$. Due to errors in expectations $\left(g_{i,t}^*, \tilde{x}_{i,t} - v_{i,t}^*\right)$, prices deviate from this FIRE benchmark,

$$
\begin{aligned}
\frac{P_{i,t}}{P_t^{FIRE}} &= \exp\left\{\tilde{x}_{i,t} - v_{i,t}^*\right\} \left(\sum_{n=1}^{\infty} P_t^{(n)} \exp\left\{n g_{i,t}^*\right\}\right) / \sum_{n=1}^{\infty} P_t^{(n)} \\
&= F\left(x_t^{agg}, g_{i,t}^*, \tilde{x}_{i,t} - v_{i,t}^*\right). \qquad \text{(A27)}
\end{aligned}
$$

Let $b_{i,t} \equiv E_t\left[\Delta\tilde{x}_{i,t+1}\right] - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right] = -g_{i,t}^* - \tilde{x}_{i,t} + v_{i,t}^*$ be the bias in one-period cash flow expectations. The objective expected return is

$$
\begin{aligned}
E_t\left[R_{i,t+1}\right] &= \sum_{n=1}^{\infty} w_{i,t,n} E_t\left[R_{t+1}^{(n)}\right] E_t\left[\frac{E_{t+1}^*\left[\tilde{X}_{i,t+n}\right]}{E_t^*\left[\tilde{X}_{i,t+n}\right]}\right] \\
&= w_{i,t,1} \exp\left\{b_{i,t} + r^f\right\} + \sum_{n=2}^{\infty} w_{i,t,n} \exp\left\{n\beta b_{i,t} + \frac{1}{2}(n\beta)^2 \sigma_v^2 + r^f + \gamma\sigma_u^2 \phi^{n-1}\right\} \\
&= H\left(x_t^{agg}, b_{i,t}\right). \qquad \text{(A28)}
\end{aligned}
$$

Importantly, the function $H\left(x_t^{agg}, b_{i,t}\right)$ depends on the parameter β, whereas the function $F\left(x_t^{agg}, g_{i,t}^*, \tilde{x}_{i,t} - v_{i,t}^*\right)$ does not. In an economy without rational arbitrageurs, changing the stubbornness of errors in expectations (i.e., changing β) does not affect how $g_{i,t}^*$ and $\tilde{x}_{i,t} - v_{i,t}^*$ impact the level of prices but does affect how much these errors in expectations translate into predictable one-period returns.

We now consider risk-neutral rational arbitrageurs. These agents know the true parameters and can construct long-short portfolios to profit on mispricing.[39] We assume these agents face a linear shorting cost with parameter η. This means that, in equilibrium, there cannot be two assets i, j such that $E_t[R_{i,t+1}] - E_t[R_{j,t+1}] > \eta$, otherwise the arbitrageurs would take an infinite long position in i and an infinite short-position in j. It can be shown that this implies in equilibrium that $\left|E_t\left[\tilde{R}_{i,t+1}\right]\right| \leq \eta/2$ for all firms.[40]

Because of the shorting cost, these rational arbitrageurs do not correct all mispricing when they are added to the model. From equation (A28), we see that there are threshold biases such that the arbitrageurs do not trade any firms with $b_{i,t} \in \left[\underline{b}\left(x_t^{agg}\right), \bar{b}\left(x_t^{agg}\right)\right]$. The key takeaway is that these thresholds depend on β. For lower values of β, objective expected one-period returns are less sensitive to $b_{i,t}$ and the range of biases where arbitrageurs do not intervene $\left[\underline{b}\left(x_t^{agg}\right), \bar{b}\left(x_t^{agg}\right)\right]$ is wider. As mentioned above, changing β does not alter the function $F\left(x_t^{agg}, g_{i,t}^*, \tilde{x}_{i,t} - v_{i,t}^*\right)$, so these errors in expectations will still have a large effect on the level of prices even though they do not lead to large enough objective expected one-period returns for arbitrageurs to intervene.

As a final point, we can examine the role of the temporary level shocks. Relative to the belief-formation models of Lewellen and Shanken (2002) and Nagel and Xu (2022), our non-rational agent attributes part of the earnings growth surprise to temporary shocks to the level of earnings. Without this feature, her expectations follow

$$g_{i,t}^* = g_{i,t-1}^* + \beta\left(\Delta\tilde{x}_{i,t} - E_{t-1}^*\left[\Delta\tilde{x}_{i,t}\right]\right) \tag{A29}$$

$$E_t^*\left[\tilde{x}_{i,t+n}\right] = \tilde{x}_{i,t} + n g_{i,t}^* \tag{A30}$$

$$\frac{E_{t+1}^*\left[\tilde{X}_{i,t+n}\right]}{E_t^*\left[\tilde{X}_{i,t+n}\right]} = \exp\left\{(1-\beta+n\beta)\left(\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]\right)\right\}. \tag{A31}$$

Comparing equations (28) and (A31), revisions would be much more sensitive to surprises, as $n\beta$ is changed to $1-\beta+n\beta$. Given equation (27), this means that predictable errors in

[39] In the model, the only tradable assets are shares of the firms. Strip prices are simply mathematical objects to help describe firm prices.

[40] Given that we have a continuum of firms, there will always be some firm with expected demeaned return exactly equal to $\eta/2$ or $-\eta/2$. This means that if any other firm had a demeaned expected return outside $[-\eta/2, \eta/2]$ then it would immediately imply that there is a pair of firms with $E_t[R_{i,t+1}] - E_t[R_{j,t+1}] > \eta$.

expectations would translate much more strongly into predictable one-year returns.

C.2. Incorporating a unit root

This extension shows that the main features of the model continue to hold if there is a unit root in the aggregate process. Specifically, we generalize the aggregate component x_t^{agg} to follow

$$\Delta x_t^{agg} = q_{t-1} + u_t \tag{A32}$$

$$q_t = \phi q_{t-1} + \theta u_t \tag{A33}$$

where $\theta < 0$. This follows the framework of Lettau and Wachter (2007), where the growth of the aggregate component Δx_t^{agg} has a predictable AR(1) element q_{t-1} and the innovation to Δx_t^{agg} is negatively related to the innovation in q_t. This means that positive shocks u_t to the level of x_t^{agg} are, on average, partly reversed over time, as there is an immediate increase in Δx_t^{agg} followed by a decrease in average Δx_{t+1}^{agg}, Δx_{t+2}^{agg}, ... When $\theta = \phi - 1$, equations (A32)-(A33) reduce exactly to our main model equation (13). When $\theta \in (\phi - 1, 0)$, we have a unit root in x_t^{agg}.

Given the stochastic discount factor in equation (15), the price of an aggregate strip is

$$P_t^{(n)} = E_t^* \left[\left(\prod_{j=1}^{n} M_{t+j} \right) X_{t+n}^{agg} \right] = \exp \left\{ D(n) + x_t^{agg} + \frac{1-\phi^n}{1-\phi} q_t \right\}$$

where $D(0) = 0$ and

$$D(n) = D(n-1) - r^f + \frac{1}{2}\sigma_u^2 \left(1 + \theta \frac{1-\phi^{n-1}}{1-\phi} \right)^2 - \sigma_u^2 \gamma \left(1 + \theta \frac{1-\phi^{n-1}}{1-\phi} \right).$$

The realized return and the subjective expected return on the strip are

$$R_{t+1}^{(n)} = \exp \left\{ r^f - \frac{1}{2}\sigma_u^2 \left(1 + \theta \frac{1-\phi^{n-1}}{1-\phi} \right)^2 + \gamma \sigma_u^2 \left(1 + \theta \frac{1-\phi^{n-1}}{1-\phi} \right) + \left(1 + \theta \frac{1-\phi^{n-1}}{1-\phi} \right) u_{t+1} \right\}$$

$$E_t^* \left[R_{t+1}^{(n)} \right] = \exp \left\{ r^f + \gamma \sigma_u^2 \left(1 + \theta \frac{1-\phi^{n-1}}{1-\phi} \right) \right\}.$$

Thus, just as in the main model of Section III, this extended model has the implication that longer horizon strips carry a lower annual risk premium $\gamma\sigma_u^2\left(1+\theta\frac{1-\phi^{n-1}}{1-\phi}\right)$ when $\theta<0$.

Beyond this change to prices, expected returns, and realized returns for aggregate strips, there is no other change to the model of Section III from incorporating a unit root in x_t^{agg}. Because this extended model only changes the aggregate process, the equations for beliefs about the firm-level component of cash flows are unchanged from Section III.B. Equations (23)-(28), which show how to calculate prices, realized returns, and expected returns for individual firms, also continue to hold.

This extended model preserves the three key results of the main model discussed in Section III.D. First, a higher expected g_i raises a firm's price by raising expected future cash flows and also by lowering the risk premium since the required risk premium on longer horizon cash flows is lower. Second, if the constant-gain parameter is small, then expectations are stubborn as the aggregate process has no impact on learning about g_i. Third, these stubborn errors increase price volatility relative to FIRE. Given that our analysis is focused on decomposing cross-sectional differences in price ratios rather than aggregate movements, we use the simpler case of equation (13) for our main analysis to reduce the number of free parameters and make the model as transparent as possible.

D. Model estimation

This section derives the cash flow parameters ϕ, σ_u, and σ_v from the standard deviation and autocorrelation of aggregate earnings growth $\sigma\left(\Delta x_t^{agg}\right)$ and $AC\left(\Delta x_t^{agg}\right)$ and the average across portfolios of the standard deviation over time of earnings growth $\sigma\left(\Delta\tilde{x}_{i,t+1}\right)$. We also relate the constant-gain learning from annual observations used in our model to the evidence on belief updating from quarterly observations.

D.1. Parameter values

According to equation (13), we can express aggregate earnings growth as:

$$\Delta x_t^{agg} = \phi \Delta x_{t-1}^{agg} - u_{t-1} + u_t. \tag{A34}$$

Taking covariance of equation (A34) with current earnings growth on both sides results in:

$$\begin{aligned} Cov(\Delta x_t^{agg}, \Delta x_{t-1}^{agg}) &= \phi Var(\Delta x_{t-1}^{agg}) - \sigma_u^2 \\ AC(\Delta x_t^{agg}) &= \phi - \frac{\sigma_u^2}{Var(\Delta x_t^{agg})}. \end{aligned} \tag{A35}$$

Taking the variance of equation (A34) on both sides gives:

$$\begin{aligned} Var(\Delta x_t^{agg}) &= \phi^2 Var(\Delta x_t^{agg}) + 2\sigma_u^2 - 2\phi\sigma_u^2 \\ Var(\Delta x_t^{agg}) &= \frac{2\sigma_u^2}{1+\phi}. \end{aligned} \tag{A36}$$

From equations (A35) and (A36), we have:

$$\begin{aligned} \phi &= 1 + 2AC\left(\Delta x_t^{agg}\right) \\ \sigma_u &= \left(\frac{1+\phi}{2}\right)^{1/2} \sigma\left(\Delta x_t^{agg}\right). \end{aligned}$$

Finally, to estimate the individual variance, we use equation (14) to obtain the value for σ_v in terms of idiosyncratic earnings growth:

$$\sigma_v = \frac{\sigma\left(\Delta \tilde{x}_{i,t}\right)}{\sqrt{2}}.$$

From the empirical values over the 1982-2020 sample of $\sigma\left(\Delta x_t^{agg}\right) = 0.353$, $AC\left(\Delta x_t^{agg}\right) = -0.086$ and a median portfolio volatility of $\sigma\left(\Delta \tilde{x}_{i,t}\right) = 0.140$ we infer $\phi = 0.828$, $\sigma_u = 0.337$ and $\sigma_v = 0.099$

D.2. Constant-gain parameter

We set our constant-gain parameter based on previous estimates of belief updating from Malmendier and Nagel (2016). In this section, we show that for small gains β, annual updating of beliefs based on annual surprises is quite close to quarterly updating of beliefs based on quarterly surprises.

For intuition, first consider the case of semi-annual updating. In this scenario, the agent is attempting to learn the parameter μ from a semi-annual variable x_t using the updating rule

$$E_t^*[\mu] = E_{t-1/2}^*[\mu] + \beta\left(x_t - E_{t-1/2}^*[\mu]\right). \tag{A37}$$

Iterating this equation, we have

$$\begin{aligned} E_t^*[\mu] &= E_{t-1}^*[\mu] + \beta\left(x_t - E_{t-1/2}^*[\mu]\right) + \beta\left(x_{t-1/2} - E_{t-1}^*[\mu]\right) && \text{(A38)} \\ &= E_{t-1}^*[\mu] + \beta\left(x_{t-1/2} + x_t - 2E_{t-1/2}^*[\mu]\right) - \beta^2\left(x_{t-1/2} - E_{t-1}^*[\mu]\right) && \text{(A39)} \\ &\approx E_{t-1}^*[\mu] + \beta\left(x_{t-1/2} + x_t - 2E_{t-1/2}^*[\mu]\right). && \text{(A40)} \end{aligned}$$

Equation (A40) shows that for small values of β, this semi-annual updating rule is closely approximated by an annual updating rule using the same gain β and the annual surprise $x_{t-1/2} + x_t - 2E_{t-1/2}^*[\mu]$. This is because the effect of within-year updating depends on β^2, which in our case would be quite small at 0.0003. In other words, the adjustment for within-year updating is second order compared to the first order change in beliefs $\beta\left(x_{t-1/2} + x_t - 2E_{t-1/2}^*[\mu]\right)$.

Now, we consider the scenario of quarterly updating. The agent is attempting to learn the parameter μ from a quarterly variable x_t. The agent's updating rule is

$$\begin{aligned} E_t^*[\mu] &= E_{t-1/4}^*[\mu] + \beta\left(x_t - E_{t-1/4}^*[\mu]\right) && \text{(A41)} \\ &= E_{t-1}^*[\mu] + \beta\left(x_{t-3/4} + x_{t-1/2} + x_{t-1/4} + x_t - 4E_{t-1}^*[\mu]\right) \\ &\quad -\beta^2\left[3\left(x_{t-3/4} - E_{t-1}^*[\mu]\right) + 2\left(x_{t-1/2} - E_{t-3/4}^*[\mu]\right) + x_{t-1/4} - E_{t-1/2}^*[\mu]\right]. && \text{(A42)} \end{aligned}$$

Once again, this is approximately equal to an annual updating rule based on the annual surprise $x_{t-3/4} + x_{t-1/2} + x_{t-1/4} + x_t - 4E_{t-1}^*[\mu]$. The adjustments for within-year updating are all scaled by β^2.

E. Connecting returns, earnings growth, and price-earnings ratios

First, we derive the equation for a firm which has zero dividends. For simplicity, we eliminate the index i in this derivation. In this case, the return is equal to the price growth which after log-linearization becomes an exact relationship

$$r_{t+1} \quad = \quad \Delta x_{t+1} - px_t + px_{t+1}. \tag{A43}$$

A high price-earnings ratio px_t must be followed by low future price growth Δp_{t+1} (returns r_{t+1}), high future earnings growth Δx_{t+1}, or a high future price-earnings ratio px_{t+1}.

Now, we consider the case where dividends are non-zero. We start with the one-year return identity of a portfolio

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{\left(\frac{P_{t+1}}{D_{t+1}} + 1\right)\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}},$$

where P_t and D_t are the current price and dividends. Log-linearizing around $\bar{pd}$, we can represent the price-dividend ratio pd_t in terms of future dividend growth, Δd_{t+1}, future returns, r_{t+1}, and the future price-dividend ratio, pd_{t+1}, all in logs:

$$r_{t+1} \quad \approx \quad \kappa^d + \Delta d_{t+1} - pd_t + \rho pd_{t+1}, \tag{A44}$$

where κ^d is a constant, $\rho = e^{\bar{pd}}/\left(1 + e^{\bar{pd}}\right) < 1$. We can then insert the identity $px_t = pd_t + dx_t$, where dx_t is the log payout ratio, into (A44) to obtain

$$r_{t+1} \quad \approx \quad \kappa + \Delta x_{t+1} - px_t + \rho px_{t+1} \tag{A45}$$

where we approximate $(1 - \rho)\, dx_{t+1}$ as a constant given that $(1 - \rho)$ is very close to 0.[41] Here, $\bar{pd}$ does not need to be the mean price-dividend ratio of this specific stock or portfolio. In order to study cross-sectional variation without resorting to portfolio-specific approximation parameters, we use the average price-dividend ratio of the market for $\bar{pd}$ following Cochrane (2011).

While the identity relies on the approximation that $(1 - \rho)\, dx_{t+1}$ is close to a constant,

[41] The zero dividend relationship in equation (A43) is a special case of equation (A45) as $\bar{pd}$ goes to infinity.

empirically equation (A45) holds tightly. For horizons of 1 to 4 years, Table I shows that a one unit increase in px_t is associated with almost exactly a one unit increase in $\sum_{j=1}^{h} \rho^{j-1} \Delta x_{t+j} - \sum_{j=1}^{h} \rho^{j-1} r_{t+j} + \rho^h px_{t+h}$.[42] In other words, the approximation error from ignoring the payout ratio and using a single value for ρ accounts for at most 3.1% of variation in price-earnings ratios in the decomposition of equation (3). For robustness, Appendix H.1 uses an exact relationship instead of equation (3) to ensure the approximation is not driving our results.

An important question when evaluating the decomposition (3) is how one should interpret the results. As discussed in Nagel (2024), one can always consider an altered version of equation (2),

$$\tilde{p}_{i,t} - \tilde{z}_{i,t} \approx \tilde{x}_{i,t} - \tilde{z}_{i,t} + \sum_{j=1}^{\infty} \rho^{j-1} E_t^* \left[\Delta \tilde{x}_{i,t+j}\right] - \sum_{j=1}^{\infty} \rho^{j-1} E_t^* \left[\tilde{r}_{i,t+j}\right] \qquad \text{(A46)}$$

for any arbitrary choice of scaling variable $z_{i,t}$.[43] By changing the choice of $z_{i,t}$, one can alter the size of cross-sectional variation in expected returns relative to the total variation in this new price ratio $\tilde{p}_{i,t} - \tilde{z}_{i,t}$, i.e., $\frac{Var\left(\sum_{j=1}^{\infty} \rho^{j-1} E_t^*[\tilde{r}_{i,t+j}]\right)}{Var(\tilde{p}_{i,t} - \tilde{z}_{i,t})}$. Additionally, if one wants to know how much dispersion in expected returns increases the dispersion in prices above and beyond what is explained by current cash flows and expected future cash flow growth, then one simply needs to measure

$$Var\left(\tilde{p}_{i,t} - \tilde{x}_{i,t} - \sum_{j=1}^{\infty} \rho^{j-1} E_t^* \left[\Delta \tilde{x}_{i,t+j}\right]\right) \approx Var\left(\sum_{j=1}^{\infty} \rho^{j-1} E_t^* \left[\tilde{r}_{i,t+j}\right]\right). \qquad \text{(A47)}$$

Thus, there is a natural question of what one gains from studying the decomposition rather than simply studying $Var\left(\sum_{j=1}^{\infty} \rho^{j-1} E_t^* \left[\tilde{r}_{i,t+j}\right]\right)$.

The key information from the decomposition is that it tells us about the *relative* importance of expected earnings growth compared to expected returns. Changing the choice of denominator $z_{i,t}$ will change $\frac{Var\left(\sum_{j=1}^{\infty} \rho^{j-1} E_t^*[\tilde{r}_{i,t+j}]\right)}{Var(\tilde{p}_{i,t} - \tilde{z}_{i,t})}$ but will not change whether expected

[42] For example, at the one-year horizon, a one unit increase in px_t is associated with a 0.103 increase in Δx_{t+1}, a 0.143 increase in $-r_{t+1}$, and a 0.746 increase in ρpx_{t+1}. At the four-year horizon, a one unit increase in px_t is associated with a with a 0.099 increase in $\sum_{j=1}^{4} \rho^{j-1} \Delta x_{t+j}$, a 0.320 increase in $-\sum_{j=1}^{4} \rho^{j-1} r_{t+j}$, and a 0.550 increase in $\rho^4 px_{t+4}$.

[43] Note that we focus on cross-sectional dispersion but our discussion can equivalently be applied to time-series variation.

returns or expected earnings growth play a larger role, i.e., $\frac{Var\left(\sum_{j=1}^{\infty}\rho^{j-1}E_t^*[\tilde{r}_{i,t+j}]\right)}{Var(\tilde{p}_{i,t}-\tilde{z}_{i,t})}$ versus $\frac{Var\left(\sum_{j=1}^{\infty}\rho^{j-1}E_t^*[\Delta\tilde{x}_{i,t+j}]\right)}{Var(\tilde{p}_{i,t}-\tilde{z}_{i,t})}$.[44] In a similar sense, changing the denominator will not change the relative importance of short-term expectations $\frac{Var(E_t^*[\Delta\tilde{x}_{i,t+j}])}{Var(\tilde{p}_{i,t}-\tilde{z}_{i,t})}$ versus long-term expectations.

If one's only goal is to test how much variation in expected returns impacts stock prices, then it is correct that one only needs to estimate $Var\left(E_t^*\left[\sum_{j=1}^{\infty}\rho^{j-1}\tilde{r}_{i,t+j}\right]\right)$. Many papers have documented that returns are predictable and proposed models that can generate comparable variation in expected returns. However, if one wants to know whether this impact is greater or smaller than the impact of expected earnings growth, then one needs to know $Var\left(E_t^*\left[\sum_{j=1}^{\infty}\rho^{j-1}\Delta\tilde{x}_{i,t+j}\right]\right)$ and how each piece contributes to the total $Var\left(E_t^*\left[\sum_{j=1}^{\infty}\rho^{j-1}\Delta\tilde{x}_{i,t+j}\right]-E_t^*\left[\sum_{j=1}^{\infty}\rho^{j-1}\tilde{r}_{i,t+j}\right]\right)$, which is equivalent to performing a decomposition of price-earnings ratios. Using FIRE expectations and FIRE models, Delao et al. (2025) document that a number of models that are designed to generate realistic dispersion in expected returns also generate counterfactually large $Var\left(E_t\left[\sum_{j=1}^{\infty}\rho^{j-1}\Delta\tilde{x}_{i,t+j}\right]\right)$, leading to a small *relative* importance of FIRE expected returns compared to what we find in the data. In other words, in order to generate realistic dispersion in FIRE expected returns, many papers overstate the dispersion in FIRE expected earnings growth by an order of magnitude. This demonstrates the benefit of focusing not only on the predictability of returns but also on quantitatively matching the dispersion in price-earnings ratios. Matching these relative values is particularly important for any model that wants to tie predictable returns to expectations of earnings growth, either through risk premia related to the timing of earnings or errors in earnings growth expectations.

As a final note, how should one interpret the connection to price dispersion? Equation (A47) shows how dispersion in expected returns increases the dispersion in prices above and beyond what is explained by current earnings and expected future earnings growth. By a similar logic, equation (A48) shows how dispersion in expected earnings growth increases the

[44]For example, using FIRE expectations of future earnings growth and returns, Delao et al. (2025) consider a wide range of potential $z_{i,t}$ and consistently show that FIRE expectations of future returns play a larger role in the decomposition than FIRE expectations of future earnings growth. Changing $z_{i,t}$ simply means that the expected return component and the expected earnings growth component do not sum to 1.

dispersion in prices above and beyond what is explained by current earnings and expected future returns,

$$Var\left(\tilde{p}_{i,t}-\tilde{x}_{i,t}+\sum_{j=1}^{\infty}\rho^{j-1}E_t^*\left[\tilde{r}_{i,t+j}\right]\right) \approx Var\left(\sum_{j=1}^{\infty}\rho^{j-1}E_t^*\left[\Delta\tilde{x}_{i,t+j}\right]\right). \tag{A48}$$

The decomposition of equation (3) then considers the combined impact of dispersion in expected returns and expected earnings growth, $Var\left(\sum_{j=1}^{\infty}\rho^{j-1}E_t^*\left[\Delta\tilde{x}_{i,t+j}\right]-\sum_{j=1}^{\infty}\rho^{j-1}E_t^*\left[\tilde{r}_{i,t+j}\right]\right)$ which is simply $Var\left(\tilde{p}_{i,t}-\tilde{x}_{i,t}\right)$, and shows how we can decompose this to measure the relative contribution of dispersion in expected returns and dispersion in expected earnings growth.

F. *Behavioral and learning model predictions*

In this section, we discuss how the empirical findings relate to several behavioral and learning models in which subjective discount rates are constant and the comovement of current price ratios with future returns is due to non-FIRE beliefs about cash flow growth. When subjective expected returns are constant, $E_t^*\left[\tilde{r}_{i,t+j}\right]=\bar{r}$, equations (1) and (2) imply that realized returns are

$$\tilde{r}_{i,t+1}-\bar{r} \approx \left(\Delta\tilde{x}_{i,t+1}-E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]\right)+\sum_{j=2}^{\infty}\rho^{j-1}\left(E_{t+1}^*\left[\Delta\tilde{x}_{i,t+j}\right]-E_t^*\left[\Delta\tilde{x}_{i,t+j}\right]\right).$$

If one-period cash flow surprises $\tilde{x}_{i,t+1}-E_t^*\left[\tilde{x}_{i,t+1}\right]$ (which are equivalent to $\Delta\tilde{x}_{i,t+1}-E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]$) are positively related to revisions in expected future growth, then one-period cash flow surprises will impact unexpected returns more than 1-1. Empirically, we find that one-period cash flow surprises are negatively related to revisions in expected future growth, as earnings are largely expected to return to their previous levels, and thus translate less than 1-1 into unexpected returns.

Consider the case where agents overstate the persistence of growth. Specifically, the true persistence of growth is ϕ but agents believe the persistence is $\phi^*>\phi$. So long as $\phi^*>0$, then a positive surprise $\Delta\tilde{x}_{i,t+1}-E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]$ will raise expected future growth and translate more than 1-1 into realized returns.

Further, we can consider learning about the mean of an i.i.d. growth process, such as Nagel and Xu (2022). In this setting, agents form expectations of growth based on a weighted average of past realized growth. A higher than expected realization for cash flow growth causes the agent to positively revise her beliefs about mean growth. Because there are no temporary shocks to the level of cash flows, this increase in expected mean growth causes the agent to raise her expectations of all future growth. Specifically, Nagel and Xu (2022) show that the realized unexpected return is

$$r_{t+1} - E_t^*\left[r_{t+1}\right] = \left(1 + \frac{\rho v}{1-\rho}\right)\left(\Delta x_{t+1} - \tilde{\mu}_t\right)$$

where $\tilde{\mu}_t$ is the agent's current expectation of growth and v is the learning gain parameter. Given that $v > 0$, it is immediate that cash flow growth surprises translate more than 1-1 into realized returns.

Finally, we discuss the case of diagnostic growth expectations, as in Bordalo et al. (2024). This model proposes that earnings growth is impacted by tangible news τ_{t+1} and intangible news η_t. Specifically, the process for earnings growth is

$$\Delta x_{t+1} = \mu \Delta x_t + \eta_t + \tau_{t+1}.$$

Subjective expectations of growth are

$$E_t^*\left[\Delta x_{t+j}\right] = \mu^{j-1}\left(\mu \Delta x_t + \eta_t\right) + \mu^{j-1}\epsilon_t$$
$$\epsilon_t = \phi \epsilon_{t-1} + \theta\left(\mu \tau_t + \eta_t\right)$$

where ϵ_t captures biases in expectations and ϕ is assumed to be less than μ. As stated in the paper, realized unexpected returns are

$$r_{t+1} - \bar{r} = \Delta x_{t+1} - E_t^*\left[\Delta x_{t+1}\right] + \sum_{j=2}^{\infty} \rho^{j-1}\left(E_{t+1}^*\left[\Delta x_{t+j}\right] - E_t^*\left[\Delta x_{t+j}\right]\right).$$

The covariance of the price-earnings ratio with future earnings growth surprises and future unexpected returns is then

$$\begin{aligned} Cov\left(px_t, \Delta x_{t+1} - E_t^*\left[\Delta x_{t+1}\right]\right) &= Cov\left(px_t, -\epsilon_t\right) \\ Cov\left(px_t, r_{t+1} - \bar{r}\right) &= Cov\left(px_t, -\left[1 + \frac{\rho}{1-\rho\mu}\left(\mu - \phi\right)\right]\epsilon_t\right) \\ &= \left[1 + \frac{\rho}{1-\rho\mu}\left(\mu - \phi\right)\right] Cov\left(px_t, \Delta x_{t+1} - E_t^*\left[\Delta x_{t+1}\right]\right). \end{aligned}$$

Given the paper's assumption that $\mu > \phi$, this means that the covariance of price-earnings ratios with unexpected returns must be a magnified version of the covariance of price-earnings ratios with earnings growth surprises. In fact, the model implies that for any time t variable, the comovement of that variable with unexpected returns will be $1 + \frac{\rho}{1-\rho\mu}\left(\mu - \phi\right)$ times the comovement of that variable with earnings growth surprises. Thus, the model cannot match our finding that the comovement of price-earnings ratios with one-year unexpected returns is *smaller* than the comovement of price-earnings ratios with earnings growth surprises.

As an extension, we also consider the model of diagnostic expectations of earnings *levels* in Bordalo et al. (2019). In this model, the difference between subjective and objective expectations of the level of log earnings is

$$E_t^*\left[x_{t+j}\right] - E_t\left[x_{t+j}\right] = a^j \frac{1-(b/a)^j}{1-b/a}\left(\hat{f}_t^\theta - \hat{f}_t\right)$$

where $a, b > 0$, $\hat{f}_t$ is an objective inference of an underlying component of earnings, and $\hat{f}_t^\theta$ is the biased inference of this component. Simulating the model using the paper's parameter values, we find that high price-earnings ratio stocks have *lower* subjective expected one-year earnings growth. Further, we find that price-earnings ratios are negatively related to current $\hat{f}_t^\theta - \hat{f}_t$, meaning that high price-earnings ratio stocks have pessimistic expectations of earnings at all horizons. These predictions do not align with our empirical findings that high price-earnings ratio stocks have high subjective expected one-year earnings growth and that high price-earnings ratios are associated with overoptimism in subjective expected earnings growth.

G. *Extending to anomaly returns*

This section shows that the evidence on stubbornness in earnings expectations (i.e., that earnings surprises lead to small revisions) is not limited to cross-sectional differences in valuation ratios. Instead, this pattern is reflected across a wide range of one-year cross-sectional anomalies.

To demonstrate this, we use our high and low price-earnings ratio portfolios from the previous section, as well as 20 other annual anomalies documented in Hou, Xue, and Zhang (2015).[45] We also calculate a representative anomaly that sorts stocks based on the 20 different variables and uses the average ranking across these variables in the sorting and in the regressions. For each anomaly, we sort stocks into five equal-value portfolios based on the anomaly variable and then measure earnings expectations and revisions for each of the portfolios.[46] Using the regression specified in equation (10), we estimate $\gamma^*_{2,1}$, which is the revision in two-year earnings expectations after a $t+1$ earnings surprise.

As illustrated in Figure A2, we find that for most anomalies, positive earnings surprises $\tilde{x}_{i,t+1} - E^*_t\left[\tilde{x}_{i,t+1}\right]$ result in only minor revisions to future earnings expectations $E^*_{t+1}\left[\tilde{x}_{i,t+2}\right] - E^*_t\left[\tilde{x}_{i,t+2}\right]$. Across all anomalies, this coefficient is small and significantly less than one. This means that positive earnings surprises actually *decrease* expected next period growth $E^*_{t+1}\left[\Delta\tilde{x}_{i,t+2}\right]$ as earnings are forecasted to largely return to their previous levels. Since we do not need return expectations to perform this test, we repeat the test over the longer 1982-2020 sample for robustness and find similar results in Table AI.

Section II.C illustrated how this stubbornness in expectations affects market prices through the slow adjustment of returns. Here, we quantify how this sluggish adjustment contributes to the returns of specific anomaly portfolios. Consider an anomaly variable $\tilde{a}_{i,t}$, such as

[45] Beyond the price-earnings ratio, we find 21 anomalies documented in Hou, Xue, and Zhang (2015) which are applicable to annual returns. We then drop the size anomaly because the analyst forecasts are provided primarily for large firms and are thus not suited to cover portfolios over this anomaly.

[46] To perform these tests, stocks are required to have one-year expected and realized earnings growth, returns, and price-earnings ratios. We also require that stocks have a current two-year earnings expectation $E^*_t\left[\tilde{x}_{i,t+2}\right]$ and a future one-year earnings expectation $E^*_{t+1}\left[\tilde{x}_{i,t+2}\right]$ for our test of revisions.

Table AI

Revisions in expectations

This table shows the effect of earnings surprises on revisions. For each anomaly $\tilde{a}_{i,t}$, we sort stocks into five equal-value portfolios based on the anomaly variable. Each row shows the coefficient from regressing the revision in future earnings $E^*_{t+1}\left[\tilde{x}_{i,t+2}\right] - E^*_t\left[\tilde{x}_{i,t+2}\right]$ on the earnings surprise $x_{i,t+1} - E^*_t\left[x_{i,t+1}\right]$. The Representative Anomaly (RA) is the average ranking of each stock across 20 different anomalies. P/X shows the results for the main portfolios used to study cross-sectional variation in price-earnings ratios. Each anomaly variable $\tilde{a}_{i,t}$ is scaled to have unit variance and to positively comove with future returns. The first and third columns show the result of the regressions using the main sample period of 1999 to 2020. The second and fourth columns show the result of the regressions using the long sample period of 1982 to 2020. Driscoll-Kraay standard errors are calculated for each coefficient using a lag equal to the equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

$\tilde{a}_{i,t}$	Main Sample 1999-2020	Long Sample 1982-2020	$\tilde{a}_{i,t}$	Main Sample 1999-2020	Long Sample 1982-2020
rdm	-0.059	-0.010	noa	0.127**	0.122***
	[0.046]	[0.041]		[0.056]	[0.039]
bm	0.194***	0.179***	oaa	0.188***	0.167***
	[0.060]	[0.046]		[0.053]	[0.035]
cfp	0.218***	0.196***	ol	0.121**	0.140***
	[0.046]	[0.039]		[0.058]	[0.037]
adm	0.043	0.083***	pia	0.144***	0.113***
	[0.029]	[0.025]		[0.053]	[0.036]
nop	0.073**	0.070**	poa	0.127*	0.141***
	[0.036]	[0.030]		[0.067]	[0.045]
ia	0.104***	0.137***	pta	-0.002	0.016
	[0.027]	[0.024]		[0.028]	[0.024]
gp	0.257**	0.223***	occ	0.274***	0.251***
	[0.111]	[0.074]		[0.081]	[0.057]
ivc	-0.019	0.053*	dur	0.479***	0.443***
	[0.032]	[0.031]		[0.071]	[0.063]
ivg	0.130***	0.134***	cei	0.082	0.114**
	[0.044]	[0.033]		[0.067]	[0.047]
ig	0.062	0.068***	P/X	0.136*	0.214**
	[0.040]	[0.025]		[0.072]	[0.091]
nsi	0.168***	0.143***	RA	0.189***	0.182***
	[0.030]	[0.031]		[0.049]	[0.060]

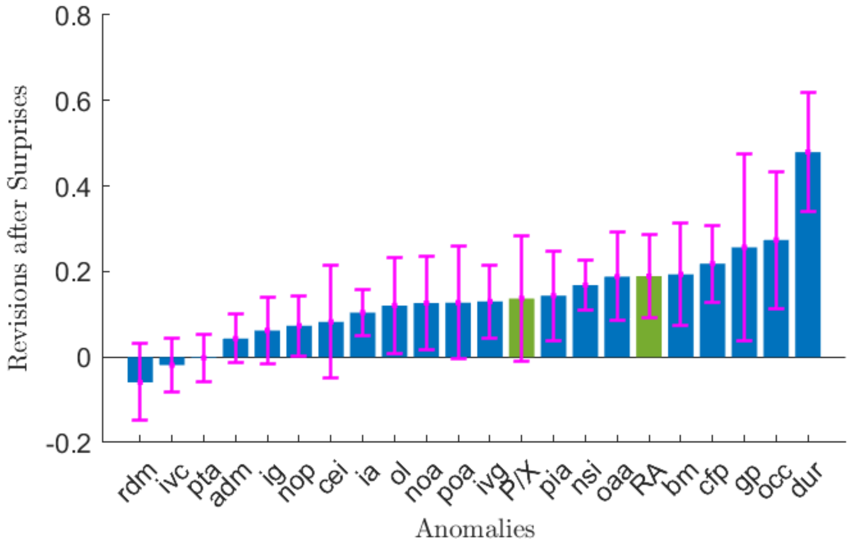

Figure A2. Revisions in anomaly expected future earnings. This figure shows the effect of earnings surprises on revisions for each set of anomaly portfolios. Each bar shows the coefficient from regressing the revision in expected earnings $E_{t+1}^{*}\left[\tilde{x}_{i,t+2}\right]-E_{t}^{*}\left[\tilde{x}_{i,t+2}\right]$ on the earnings surprise $\tilde{x}_{i,t+1}-E_{t}^{*}\left[\tilde{x}_{i,t+1}\right]$. The coefficients are shown in ascending order. Individual anomalies are shown in blue. In green, we show the Representative Anomaly (RA) that sorts stocks based on their average ranking across all of the individual anomalies, as well as the results for the portfolios used to study cross-sectional variation in price-earnings ratios (P/X).

profitability or investment, which predicts next-year returns. To facilitate comparisons, we normalize $\tilde{a}_{i,t}$ to have unit variance and positive covariance with future returns. From equation (1), we have that

$$\underbrace{Cov\left(\tilde{r}_{i,t+1}-E_{t}^{*}\left[\tilde{r}_{i,t+1}\right],\tilde{a}_{i,t}\right)}_{\sigma_{a,r}} \approx \underbrace{Cov\left(\tilde{x}_{i,t+1}-E_{t}^{*}\left[\tilde{x}_{i,t+1}\right],\tilde{a}_{i,t}\right)}_{\sigma_{a,x}} + \underbrace{\rho Cov\left(\tilde{px}_{i,t+1}-E_{t}^{*}\left[\tilde{px}_{i,t+1}\right],\tilde{a}_{i,t}\right)}_{\sigma_{a,px}}. \tag{A49}$$

Note that unexpected earnings growth $\Delta\tilde{x}_{i,t+1}-E_{t}^{*}\left[\Delta\tilde{x}_{i,t+1}\right]$ is identical to unexpected earnings $\tilde{x}_{i,t+1}-E_{t}^{*}\left[\tilde{x}_{i,t+1}\right]$. For robustness, Appendix H.1 shows an exact decomposition based on price growth, which gives very similar results.

Under FIRE, we would have $\sigma_{a,r},\sigma_{a,x},\sigma_{a,px}=0$, i.e., any predictable anomaly returns would be fully anticipated and $\tilde{a}_{i,t}$ would not predict forecast errors. For example, a higher $\tilde{a}_{i,t}$ might be related to higher risk exposure and investors would require higher returns on

Table AII

Unexpected anomaly returns

This table measures and decomposes unexpected anomaly returns. The Representative Anomaly is the average ranking of each stock across 20 different anomalies, and $\tilde{px}_{i,t}$ is the demeaned price-earnings ratio. For each anomaly variable, we sort stocks into five equal-value portfolios based on the anomaly variable. Each column shows the coefficients of regressing each of the dependent variables on a specific anomaly variable. Both anomaly variables are scaled to have unit variance and to positively comove with future returns. The three dependent variables are the unexpected return $\tilde{r}_{i,t+1} - E_t^* [\tilde{r}_{i,t+1}]$, the earnings forecast errors $\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]$, and the price-earnings ratio forecast errors $\rho \left(\tilde{px}_{i,t+1} - E_t^* \left[\tilde{px}_{i,t+1}\right]\right)$. The sample period is 1999 to 2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	Representative Anomaly	$\tilde{px}_{i,t}$
$\tilde{r}_{i,t+1} - E_t^* [\tilde{r}_{i,t+1}]$	0.034^{***}	0.033^{**}
	[0.013]	[0.016]
	[0.013]	[0.016]
$\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]$	0.064^{***}	0.069^{***}
	[0.020]	[0.010]
	[0.020]	[0.014]
$\rho \left(\tilde{px}_{i,t+1} - E_t^* \left[\tilde{px}_{i,t+1}\right]\right)$	-0.032^{***}	-0.038^{**}
	[0.009]	[0.017]
	[0.009]	[0.015]

these stocks as compensation. If expectations deviate from FIRE, then positive values of $\sigma_{a,r}$ indicate that investors understate the relationship between $\tilde{a}_{i,t}$ and future returns. In other words, the high returns on high $\tilde{a}_{i,t}$ stocks are not fully anticipated. In comparison, $\sigma_{a,x}$ and $\sigma_{a,px}$ reflect the extent to which unexpected returns are explained by predictable errors in one-year earnings growth expectations and expectations of the future price-earnings ratio, respectively. A relatively large $\sigma_{a,x}$ indicates that unexpected returns of high $\tilde{a}_{i,t}$ stocks can be largely explained by unexpectedly high earnings.

For each of the 21 anomaly variables, we measure forecast errors for one-year returns, earnings, and price-earnings ratios, and regress each on the anomaly variable to estimate the decomposition in equation (A49). Table AII shows that the results for the representative anomaly are qualitatively and quantitatively similar to those for the $\tilde{px}_{i,t}$ portfolios in Section II.B. The first row shows that a one standard deviation increase in either of the anomaly

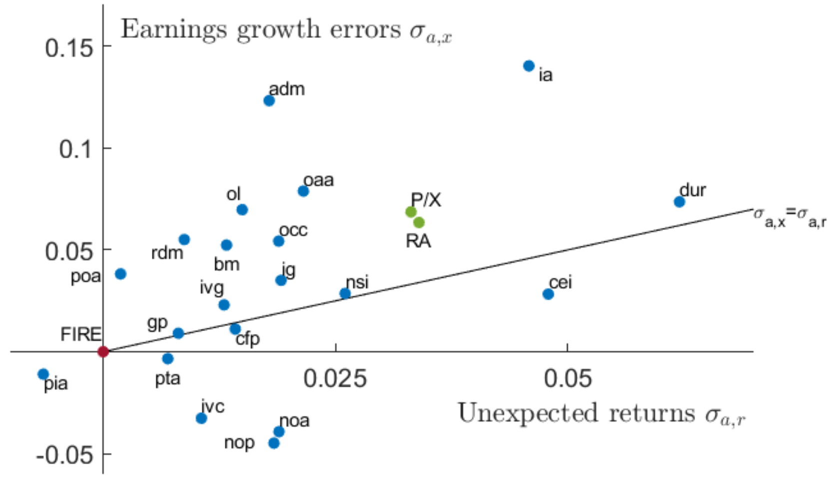

Figure A3. Unexpected anomaly returns and earnings surprises. This figure shows the decomposition results $(\sigma_{a,r}, \sigma_{a,x})$ for each anomaly $\tilde{a}_{i,t}$. The x-axis shows $\sigma_{a,r} = Cov\left(\tilde{r}_{i,t+1} - E_t^*\left[\tilde{r}_{i,t+1}\right], \tilde{a}_{i,t}\right)$, which measures how much the anomaly variable predicts unexpected returns. The y-axis shows $\sigma_{a,x} = Cov\left(\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{a}_{i,t}\right)$, which measures how much the anomaly variable predicts one-year earnings growth forecast errors. The anomalies are shown in blue. In red, we show the FIRE benchmark, which is that $\sigma_{a,r}$ and $\sigma_{a,x}$ should equal 0 for all anomalies. In green, we show a Representative Anomaly (RA) that sorts stocked based on their average ranking across all of the individual anomalies, as well as the results for the portfolios used to study cross-sectional variation in price-earnings ratios (P/X). Each anomaly variable $\tilde{a}_{i,t}$ is scaled to have unit variance and to positively comove with future returns.

variables is associated with a roughly 3pp increase in unexpected returns (0.034 and 0.033, respectively). This increase in unexpected returns is more than accounted for by the roughly 6pp increase in unexpected earnings (0.064 and 0.069, respectively). Since expectations of future earnings are only modestly revised in response to an earnings surprise, as shown in Figure A2, these surprises do not immediately translate 1-1 into unexpected returns. Rephrased, prices do not move 1-1 with the earnings surprise, meaning that the 6pp earnings surprise also leads to the price-earnings ratio being roughly 3pp lower than expected.

Figure A3 shows the results for each of the 22 anomalies (the 20 individual anomalies, our price-earnings ratio portfolios and the representative anomaly). For almost every anomaly, we find positive $\sigma_{a,r}$, meaning that investors do not fully anticipate the high returns on high $\tilde{a}_{i,t}$ stocks. Further, most anomalies (17 out of 22) are associated with large positive one-year earnings forecast errors, as shown by the magnitude of $\sigma_{a,x}$. Table AIII provides for each anomaly the full decomposition of unexpected returns into earnings surprises and unexpected future price-earnings ratios. Table AIV shows an exact decomposition of anomaly price growth rather than anomaly returns.

Comparing $\sigma_{a,r}$ and $\sigma_{a,x}$ across anomalies, we see that anomalies with higher $\sigma_{a,r}$ generally have higher $\sigma_{a,x}$, i.e., larger unanticipated returns are associated with larger one-year earnings surprises, and $\sigma_{a,x}$ is generally larger than $\sigma_{a,r}$, i.e., earnings surprises translate less than 1-1 into unexpected returns. This means that our findings on the dynamics of earnings surprises and returns from Section II.C also extends to most anomaly portfolios. Rephrased, many anomaly returns can be accounted for by prices gradually responding to errors in earnings expectations.

To summarize, consistent with the results from Section II.C, we find only moderate revisions to earnings surprises in anomaly portfolios. Our results highlight that unexpected returns are positively associated with earnings surprises but with a less-than-proportional relationship. This evidence, once again, points against models in which unexpected realized returns are highly sensitive to recent earnings surprises and highlights the benefit of quantitative decompositions which allow for these types of comparisons.

H. Robustness Exercises

H.1. Exact decomposition results

In this section, we derive all the main results using an exact decomposition of price-earnings ratios based on price growth, rather than the approximate decomposition based on returns. For any stock or portfolio of stocks i, the price-earnings ratio $px_{i,t}$ can be expressed in terms of the one-year ahead log price growth $\Delta p_{i,t+1}$, the future earnings growth $\Delta x_{i,t+1}$, and the future price-earnings ratio:

$$px_{i,t} = \Delta x_{i,t+1} + \Delta p_{i,t+1} + px_{i,t+1}. \tag{A50}$$

Table AIII

Unexpected anomaly returns

This table measures and decomposes unexpected anomaly returns using equation (A49). For each anomaly $\tilde{a}_{i,t}$, we sort stocks into five equal-value portfolios based on the anomaly variable. The table shows the coefficients of regressing each of the dependent variables on a specific anomaly variable. The three dependent variables are the unexpected return $\tilde{r}_{i,t+1} - E_t^* [\tilde{r}_{i,t+1}]$, the earnings forecast errors $\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]$, and the price-earnings ratio forecast errors $\rho \left(\tilde{px}_{i,t+1} - E_t^* \left[\tilde{px}_{i,t+1} \right] \right)$. The Representative Anomaly (RA) is the average ranking of each stock across 20 different anomalies. P/X shows the results for the main portfolios used to study cross-sectional variation in price-earnings ratios. Each anomaly variable $\tilde{a}_{i,t}$ is scaled to have unit variance and to positively comove with future returns. The sample period is 1999 to 2020. Driscoll-Kraay standard errors are calculated for each coefficient using a lag equal to the equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	Decomposition				Decomposition		
$\tilde{a}_{i,t}$	$\sigma_{a,r}$	$\sigma_{a,x}$	$\sigma_{a,px}$	$\tilde{a}_{i,t}$	$\sigma_{a,r}$	$\sigma_{a,x}$	$\sigma_{a,px}$
rdm	0.009	0.055***	-0.044***	noa	0.019***	-0.039***	0.057***
	[0.006]	[0.013]	[0.013]		[0.007]	[0.011]	[0.008]
bm	0.013	0.052**	-0.041***	oaa	0.022	0.079***	-0.059**
	[0.014]	[0.023]	[0.016]		[0.018]	[0.029]	[0.024]
cfp	0.014	0.011	-0.002	ol	0.015***	0.070***	-0.051***
	[0.014]	[0.019]	[0.012]		[0.005]	[0.014]	[0.014]
adm	0.018**	0.123***	-0.104***	pia	-0.006	-0.011	0.007
	[0.007]	[0.010]	[0.010]		[0.011]	[0.013]	[0.014]
nop	0.018**	-0.045***	0.057***	poa	0.002	0.038***	-0.035***
	[0.008]	[0.015]	[0.013]		[0.007]	[0.012]	[0.010]
ia	0.046***	0.140***	-0.096***	pta	0.007	-0.003	0.008
	[0.016]	[0.030]	[0.019]		[0.007]	[0.008]	[0.008]
gp	0.008	0.009	0.002	occ	0.019**	0.054***	-0.034***
	[0.010]	[0.016]	[0.015]		[0.009]	[0.016]	[0.012]
ivc	0.011	-0.033***	0.040***	dur	0.062***	0.074**	-0.013
	[0.007]	[0.012]	[0.009]		[0.016]	[0.033]	[0.020]
ivg	0.013	0.023	-0.012	cei	0.048**	0.028	0.016
	[0.010]	[0.015]	[0.012]		[0.021]	[0.031]	[0.015]
ig	0.019	0.035**	-0.018**	P/X	0.033**	0.069***	-0.038**
	[0.012]	[0.014]	[0.008]		[0.016]	[0.010]	[0.017]
nsi	0.026**	0.029	-0.004	RA	0.034**	0.064***	-0.032***
	[0.012]	[0.018]	[0.009]		[0.013]	[0.020]	[0.009]

Table AIV

Individual unexpected anomaly price growth (exact decomposition)

This table measures and decomposes unexpected anomaly price growth for each of the individual anomalies using equation (A53). For each anomaly variable, we sort stocks into five equal-value portfolios based on the anomaly variable. The table shows the coefficients of regressing each of the dependent variables on a specific anomaly variable. The three dependent variables are the unexpected price growth $\Delta\tilde{p}_{i,t+1} - E_t^* [\Delta\tilde{p}_{i,t+1}]$, the earnings forecast errors $\tilde{x}_{i,t+1} - E_t^* [\tilde{x}_{i,t+1}]$, and the price-earnings ratio forecast errors $\tilde{px}_{i,t+1} - E_t^* [\tilde{px}_{i,t+1}]$. The Representative Anomaly (RA) is the average ranking of each stock across 20 different anomalies. P/X shows the results for the main portfolios used to study cross-sectional variation in price-earnings ratios. Each anomaly variable $\tilde{a}_{i,t}$ is scaled to have unit variance and to positively comove with future price growth. The sample period is 1999 to 2020. Driscoll-Kraay standard errors are calculated for each coefficient using a lag equal to the equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	Decomposition				Decomposition		
$\tilde{a}_{i,t}$	$\sigma_{a,r}$	$\sigma_{a,x}$	$\sigma_{a,px}$	$\tilde{a}_{i,t}$	$\sigma_{a,r}$	$\sigma_{a,x}$	$\sigma_{a,px}$
rdm	0.010*	0.055***	-0.045***	noa	0.019***	-0.039***	0.058***
	[0.006]	[0.013]	[0.014]		[0.007]	[0.011]	[0.009]
bm	0.011	0.052**	-0.042***	oaa	0.019	0.079***	-0.060**
	[0.014]	[0.023]	[0.016]		[0.018]	[0.029]	[0.024]
cfp	0.009	0.011	-0.002	ol	0.017***	0.070***	-0.053***
	[0.015]	[0.019]	[0.012]		[0.005]	[0.014]	[0.014]
adm	0.017**	0.123***	-0.106***	pia	-0.004	-0.011	0.007
	[0.007]	[0.010]	[0.011]		[0.011]	[0.013]	[0.014]
nop	0.014*	-0.045***	0.059***	poa	0.002	0.038***	-0.036***
	[0.008]	[0.015]	[0.013]		[0.007]	[0.012]	[0.010]
ia	0.043***	0.140***	-0.098***	pta	0.005	-0.003	0.009
	[0.017]	[0.030]	[0.020]		[0.007]	[0.008]	[0.009]
gp	0.011	0.009	0.002	occ	0.020**	0.054***	-0.035***
	[0.010]	[0.016]	[0.015]		[0.009]	[0.016]	[0.013]
ivc	0.008	-0.033***	0.041***	dur	0.062***	0.074**	-0.013
	[0.007]	[0.012]	[0.010]		[0.016]	[0.033]	[0.021]
ivg	0.011	0.023	-0.012	cei	0.045**	0.028	0.016
	[0.011]	[0.015]	[0.012]		[0.021]	[0.031]	[0.016]
ig	0.017	0.035**	-0.019**	P/X	0.030*	0.069***	-0.039**
	[0.012]	[0.014]	[0.008]		[0.016]	[0.010]	[0.018]
nsi	0.024*	0.029	-0.004	RA	0.031**	0.064***	-0.033***
	[0.012]	[0.018]	[0.009]		[0.013]	[0.020]	[0.009]

This equation is exact and does not contain a log-linearization constant ρ. Applying subjective expectations $E_t^*[\cdot]$, we see that a higher than average price-earnings ratio must be explained by higher than average expected earnings growth, lower than average expected price growth, or a higher than average expected future price-earnings ratio,

$$\tilde{px}_{i,t} = \sum_{j=1}^{h} E_t^*\left[\Delta\tilde{x}_{i,t+j}\right] - \sum_{j=1}^{h} E_t^*\left[\Delta\tilde{p}_{i,t+j}\right] + E_t^*\left[\tilde{px}_{i,t+h}\right]. \tag{A51}$$

Just like the main decomposition, this equation holds under any subjective probability distribution and we can decompose the variance of $\tilde{px}_{i,t}$ into three components:

$$1 = \underbrace{\frac{Cov\left(\sum_{j=1}^{h} E_t^*\left[\Delta\tilde{x}_{i,t+j}\right], \tilde{px}_{i,t}\right)}{Var\left(\tilde{px}_{i,t}\right)}}_{CF_h} + \underbrace{\frac{Cov\left(-\sum_{j=1}^{h} E_t^*\left[\Delta\tilde{p}_{i,t+j}\right], \tilde{px}_{i,t}\right)}{Var\left(\tilde{px}_{i,t}\right)}}_{PG_h} + \underbrace{\frac{Cov\left(E_t^*\left[\tilde{px}_{i,t+h}\right], \tilde{px}_{i,t}\right)}{Var\left(\tilde{px}_{i,t}\right)}}_{FPX_h}. \tag{A52}$$

The coefficients CF_h and PG_h give a quantitative measure of how much dispersion in price-earnings ratios is accounted for by dispersion in expected earnings growth and how much is accounted for by dispersion in expected price growth. We can now estimate this equation using the exact expectations of price growth without an approximation.

Table AV shows that the results of this exact decomposition are very similar to the main decomposition results in Table I. We find that 10.3% of dispersion in price-earnings ratios is accounted for by differences in one-year future earnings growth and 13.2% is accounted for by differences in one-year price growth. Just as in the main decomposition, differences in earnings growth are overestimated, with expected earnings growth accounting for nearly a third (33.1%) of all dispersion in price-earnings ratios. Differences in price growth are underestimated, with expected price growth accounting for only 3.3% of all dispersion in price-earnings ratios. A similar pattern can be observed at longer horizons. Overall, all the coefficients closely align with those reported in Table I.

We can also estimate an exact version of the unexpected anomaly return decomposition

Table AV

Decomposition of dispersion in price-earnings ratios (exact decomposition)

This table decomposes the variance of price-earnings ratios using the exact decomposition (A52) at multiple horizons. The *Realized* columns report the elements CF_h, PG_h and FPX_h of the decomposition using future earnings growth, future price growth and future price-earnings ratios. The *Expected* columns report the elements of the decomposition using expected earnings growth, expected price growth and expected price-earnings ratios. The *Error* columns report the contribution of the forecast errors. For instance, for year $h = 1$, $CF_1 = Cov\left(\Delta\tilde{x}_{i,t+1}, \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$ is shown in the *Realized* column. This component can be split into its expected component $Cov\left(E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$ and its error component $Cov\left(\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$. The main sample period is 1999 to 2020. The fourth row shows the element CF_h of the decomposition estimated over the longer sample period of 1982-2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	$h=1$			$h=2$			$h=3$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.103***	0.331***	−0.228***	0.075	0.344***	−0.269***	0.070	0.391***	−0.321***	0.100*	0.439***	−0.340***
	[0.037]	[0.024]	[0.032]	[0.052]	[0.020]	[0.046]	[0.053]	[0.020]	[0.049]	[0.057]	[0.020]	[0.055]
	[0.052]	[0.027]	[0.044]	[0.069]	[0.022]	[0.065]	[0.073]	[0.022]	[0.071]	[0.080]	[0.022]	[0.077]
DR_h	0.132**	0.033***	0.100*	0.250***			0.284***			0.292***	0.135***	0.157*
	[0.052]	[0.013]	[0.054]	[0.060]			[0.062]			[0.088]	[0.046]	[0.089]
	[0.051]	[0.014]	[0.054]	[0.075]			[0.084]			[0.120]	[0.047]	[0.106]
FPX_h	0.765***	0.636***	0.129**	0.675***			0.646***			0.608***	0.426***	0.182***
	[0.051]	[0.020]	[0.058]	[0.036]			[0.042]			[0.063]	[0.029]	[0.069]
	[0.041]	[0.024]	[0.053]	[0.043]			[0.052]			[0.078]	[0.031]	[0.082]
	Long Sample: 1982-2020											
CF_h	0.137***	0.312***	−0.175***	0.102**	0.335***	−0.234***	0.107**	0.402***	−0.296***	0.149***	0.471***	−0.322***
	[0.026]	[0.021]	[0.027]	[0.049]	[0.023]	[0.042]	[0.044]	[0.025]	[0.035]	[0.041]	[0.028]	[0.035]
	[0.026]	[0.021]	[0.028]	[0.050]	[0.023]	[0.042]	[0.045]	[0.026]	[0.035]	[0.043]	[0.029]	[0.037]

Table AVI

Unexpected anomaly price growth (exact decomposition)

This table measures and decomposes unexpected anomaly price growth. The Representative Anomaly is the average ranking of each stock across 20 different anomalies, and $\tilde{px}_{i,t}$ is the demeaned price-earnings ratio. For each anomaly variable, we sort stocks into five equal-value portfolios based on the anomaly variable. Each column shows the coefficients of regressing each of the dependent variables on a specific anomaly variable. Both anomaly variables are scaled to have unit variance and to positively comove with future price growth. The three dependent variables are the unexpected price growth $\Delta\tilde{p}_{i,t+1} - E_t^*\left[\Delta\tilde{p}_{i,t+1}\right]$, the earnings forecast errors $\tilde{x}_{i,t+1} - E_t^*\left[\tilde{x}_{i,t+1}\right]$, and the price-earnings ratio forecast errors $\tilde{px}_{i,t+1} - E_t^*\left[\tilde{px}_{i,t+1}\right]$. The sample period is 1999 to 2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	Representative Anomaly	$\tilde{px}_{i,t}$
$\Delta\tilde{p}_{i,t+1} - E_t^*\left[\Delta\tilde{p}_{i,t+1}\right]$	0.031^{**}	0.030^{*}
	[0.013]	[0.016]
	[0.013]	[0.015]
$\tilde{x}_{i,t+1} - E_t^*\left[\tilde{x}_{i,t+1}\right]$	0.064^{***}	0.069^{***}
	[0.020]	[0.010]
	[0.020]	[0.014]
$\left(\tilde{px}_{i,t+1} - E_t^*\left[\tilde{px}_{i,t+1}\right]\right)$	-0.033^{***}	-0.039^{**}
	[0.009]	[0.018]
	[0.009]	[0.015]

(A49). Just as in the main identity, we normalize all anomalies $\tilde{a}_{i,t}$ so that they have variance 1 and positively comove with future price growth. From equation (A50), we have the identity

$$\underbrace{Cov\left(\Delta\tilde{p}_{i,t+1} - E_t^*\left[\Delta\tilde{p}_{i,t+1}\right], \tilde{a}_{i,t}\right)}_{\sigma_{a,p}} = \underbrace{Cov\left(\tilde{x}_{i,t+1} - E_t^*\left[\tilde{x}_{i,t+1}\right], \tilde{a}_{i,t}\right)}_{\sigma_{a,x}} + \underbrace{Cov\left(\tilde{px}_{i,t+1} - E_t^*\left[\tilde{px}_{i,t+1}\right], \tilde{a}_{i,t}\right)}_{\sigma_{a,px}}. \quad \text{(A53)}$$

Here, the values for $\sigma_{a,x}$ and $\sigma_{a,px}$ indicate how much the predictable price growth forecast errors are explained by predictable errors in next-year earnings expectations and expectations of the future price-earnings ratio. Table AVI shows the results for the representative anomaly studied in Section G. For nearly every anomaly, we estimate a positive value of $\sigma_{a,p}$, meaning that investors do not fully anticipate the high growth on high $\tilde{a}_{i,t}$ stocks. The predictable errors in one-year earnings growth expectations are more than large enough to account for the unexpected one-year price growth (i.e., $\sigma_{a,x}$ is greater than $\sigma_{a,p}$). In Table AIV we show the exact decomposition results for each of the individual anomalies.

H.2. Overlapping observations and Bauer and Hamilton (2018)

Overlapping forecast horizons can increase the persistence of residuals in Table I. Because of this, we use Driscoll-Kraay and block-bootstrap standard errors to account for any autocorrelation. For additional robustness, in this section we also directly calculate the significance of each coefficient under the worst-case scenario for overlapping observations.

We do this following the methodology proposed by Bauer and Hamilton (2018). Specifically, for expected and realized earnings growth and returns, we run simulations to measure how often we spuriously find a coefficient as large as what we observe in the data. For clarity, we discuss the simulation for the regression of earnings forecast errors $\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right]$ on price-earnings ratios $\tilde{px}_{i,t}$, however, the methodology is identical for the other left hand side variables.

We specify the price-earnings ratio of each portfolio i as an AR(1) process,

$$\tilde{px}_{i,t} = \mu_i + \left(\tilde{px}_{i,t-1} - \mu_i\right) + \sigma_i \varepsilon_{i,t}.$$

The mean, persistence, and variance are set equal to the observed values over our sample. Additionally, the initial value of the simulated price-earnings ratio for portfolio i is set equal to the initial value observed in our data to account for any drift back to the mean which may generate trends in price-earnings ratios over the sample. For example, if the price-earnings ratio for the Growth portfolio is substantially above its mean at the beginning of the sample, then reversion to the mean will create a downward trend in the price-earnings ratio for this portfolio over time. We then simulate one-period forecast errors under the null hypothesis that forecast errors are unpredictable.

If subjective expectations change over time, then there will be little overlap in longer horizon forecast errors. For example, if $E_t^*\left[\Delta\tilde{x}_{i,t+2}\right]$ is very different from $E_{t+1}^*\left[\Delta\tilde{x}_{i,t+2}\right]$, then there is little similarity between the second term of $\sum_{j=1}^{h}\rho^{j-1}\left(\Delta\tilde{x}_{i,t+j} - E_t^*\left[\Delta\tilde{x}_{i,t+j}\right]\right)$ and the

Table AVII

Worst-case decomposition of dispersion in price-earnings ratios

This table decomposes the variance of price-earnings ratios using equation (3) at multiple horizons. The *Realized* columns report the elements CF_h and DR_h of the decomposition using future earnings growth and future negative returns. The *Expected* columns report the elements of the decomposition using expected earnings growth and expected returns. The *Error* columns report the contribution of the forecast errors. The main sample period is 1999 to 2020. The fourth row shows the element CF_h of the decomposition estimated over the longer sample period of 1982-2020. Worst-case p-values using the Bauer and Hamilton (2018) procedure are reported in parentheses. Superscripts indicate significance at the 1% (***), 5% (**), and 10% (*) level.

	$h=1$			$h=2$			$h=3$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.103***	0.331***	−0.228***	0.075	0.344***	−0.268***	0.071	0.389***	−0.317***	0.099	0.433***	−0.335***
	(0.006)	(0.000)	(0.000)	(0.117)	(0.000)	(0.001)	(0.174)	(0.000)	(0.002)	(0.124)	(0.000)	(0.001)
DR_h	0.143***	0.033***	0.110***	0.266***			0.305***			0.320***	0.127***	0.192**
	(0.000)	(0.008)	(0.002)	(0.000)			(0.000)			(0.004)	(0.008)	(0.047)
FPX_h	0.746***	0.620***	0.126**	0.642***			0.599***			0.550***	0.385***	0.165**
	(0.000)	(0.008)	(0.010)	(0.000)			(0.000)			(0.000)	(0.000)	(0.038)
	Long Sample: 1982-2020											
CF_h	0.137***	0.312***	−0.175***	0.102**	0.335***	−0.232***	0.107**	0.399***	−0.291***	0.147**	0.462***	−0.316***
	(0.000)	(0.000)	(0.000)	(0.018)	(0.000)	(0.000)	(0.029)	(0.000)	(0.000)	(0.016)	(0.000)	(0.000)

first term of $\sum_{j=1}^{h} \rho^{j-1} \left(\Delta\tilde{x}_{i,t+1+j} - E^*_{t+1}\left[\Delta\tilde{x}_{i,t+1+j}\right]\right)$. However, in the worst-case scenario in which expectations do not change at all over time, then $\sum_{j=1}^{h} \rho^{j-1} \left(\Delta\tilde{x}_{i,t+j} - E^*_t\left[\Delta\tilde{x}_{i,t+j}\right]\right)$ will be an MA$(h-1)$ process. This will cause the forecast errors at longer horizons to be persistent, which increases the probability of spuriously finding a large coefficient between price-earnings ratios and four-year forecast errors.

For our simulations, we push this worst-case scenario even further by making each period one month instead of one year. This means that $\sum_{j=1}^{h} \rho^{j-1} \left(\Delta\tilde{x}_{i,t+j} - E^*_t\left[\Delta\tilde{x}_{i,t+j}\right]\right)$ will be MA$(12h-1)$ which dramatically increases the persistence. For all of the left-hand side variables in Table I, we find that this worst-case scenario substantially overstates the observed variable persistence. We then set the variance of the monthly forecast errors to match the observed variance of $\sum_{j=1}^{h} \rho^{j-1} \left(\Delta\tilde{x}_{i,t+j} - E^*_t\left[\Delta\tilde{x}_{i,t+j}\right]\right)$. We then run 10,000 simulations and report the probability of spuriously finding a coefficient as large as what we observe in the data.

Table AVII shows the coefficients for realized and expected earnings growth and returns, along with their associated p-values from the simulations. As in Table I, the FIRE CF_2 and CF_3 are not significant at any level. Additionally, under this worst-case scenario, the FIRE CF_4 is not significant at any level. All other coefficients are significant at the 5% level, which again aligns with the results of Table I. Even after accounting for persistence in price-earnings ratios, trends, and a worst-case assumption for overlapping observations, the probability of spuriously generating coefficients as large as what we find in the data is quite small.

H.3. Smoothed earnings

To show that our decomposition results are not influenced by fluctuations in earnings in the denominator of price-earnings ratio, we repeat our analysis normalizing prices with a three-year-smoothed measure of earnings. AVIII shows that the results are very similar to the main decomposition results in Table I. We find that 37.6% of dispersion in price–

Table AVIII

Decomposition of dispersion in price-earnings ratios using smoothed earnings

This table shows the three components of the right hand side of equation (3) using three-year smoothed earnings instead of annual earnings to form the valuation ratios. Let s_t be the three-year smoothed average of earnings. For each period, we form the price-to-smoothed-earnings ratio $\tilde{ps}_{i,t}$. The *Realized* columns report the components CF_h, DR_h and FPE_h of equation (3) using future earnings growth, future returns and future price-earnings ratios. The *Expected* columns report the elements of the equation (3) using expected earnings growth, expected returns and expected price-earnings ratios. The *Error* columns report the contribution of the forecast errors of each element. For instance, $CF_1 = Cov\left(\Delta\tilde{x}_{i,t+1}, \tilde{ps}_{i,t}\right)/Var\left(\tilde{ps}_{i,t}\right)$ is shown in the *Realized* column. This component can be split into its expected component $Cov\left(E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{ps}_{i,t}\right)/Var\left(\tilde{ps}_{i,t}\right)$ and its error component $Cov\left(\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{ps}_{i,t}\right)/Var\left(\tilde{ps}_{i,t}\right)$. The main sample period is 1999 to 2020. The fourth row shows the element CF_h of the decomposition estimated over the longer sample period of 1982-2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	$h=1$			$h=2$			$h=3$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.090**	0.290***	−0.200***	0.053	0.292***	−0.238***	0.045	0.334***	−0.289***	0.066	0.376***	−0.310***
	[0.036]	[0.021]	[0.037]	[0.047]	[0.023]	[0.046]	[0.049]	[0.023]	[0.045]	[0.053]	[0.023]	[0.050]
	[0.047]	[0.024]	[0.053]	[0.067]	[0.031]	[0.062]	[0.065]	[0.029]	[0.046]	[0.069]	[0.028]	[0.066]
DR_h	0.131***	0.032***	0.099**	0.240***			0.276***			0.291***	0.126***	0.165**
	[0.046]	[0.011]	[0.048]	[0.055]			[0.058]			[0.079]	[0.038]	[0.082]
	[0.045]	[0.012]	[0.047]	[0.066]			[0.074]			[0.095]	[0.038]	[0.098]
FPX_h	0.681***	0.573***	0.108**	0.601***			0.565***			0.523***	0.357***	0.166***
	[0.048]	[0.025]	[0.051]	[0.030]			[0.034]			[0.051]	[0.024]	[0.057]
	[0.048]	[0.031]	[0.049]	[0.036]			[0.038]			[0.057]	[0.027]	[0.068]
	Long Sample: 1982-2020											
CF_h	0.120***	0.264***	−0.143***	0.079*	0.275***	−0.196***	0.078**	0.333***	−0.256***	0.109***	0.392***	−0.283***
	[0.025]	[0.018]	[0.029]	[0.042]	[0.022]	[0.040]	[0.038]	[0.023]	[0.033]	[0.038]	[0.024]	[0.032]
	[0.025]	[0.018]	[0.029]	[0.042]	[0.022]	[0.042]	[0.039]	[0.024]	[0.033]	[0.039]	[0.025]	[0.032]

Table AIX

Decomposition of dispersion in price-earnings ratios including exiting firms

This table decomposes the variance of price-earnings ratios including firms that may exit after portfolio formation. To account for these firms, we reinvest the delisting returns of exiting firms in the corresponding portfolio. The *Realized* columns report the elements CF_h, DR_h and FPX_h of the decomposition using future earnings growth, future negative returns and future price-earnings ratios. The *Expected* columns report the elements of the decomposition using expected earnings growth, expected returns and expected price-earnings ratios. The *Errors* columns report the contribution of the forecast errors of each element. For instance, $CF_1 = Cov\left(\Delta\tilde{x}_{i,t+1}, \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$ is shown in the *Realized* column. This component can be split into its expected component $Cov\left(E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$ and its error component $Cov\left(\Delta\tilde{x}_{i,t+1} - E_t^*\left[\Delta\tilde{x}_{i,t+1}\right], \tilde{px}_{i,t}\right)/Var\left(\tilde{px}_{i,t}\right)$. The main sample period is 1999 to 2020. The fourth row shows the element CF_h of the decomposition estimated over the longer sample period of 1982-2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	$h=1$			$h=2$			$h=3$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.100**	0.306***	−0.196***	0.060	0.317***	−0.255***	0.040	0.370***	−0.330***	0.046	0.427***	−0.374***
	[0.044]	[0.026]	[0.046]	[0.071]	[0.020]	[0.070]	[0.075]	[0.023]	[0.069]	[0.073]	[0.031]	[0.070]
	[0.052]	[0.032]	[0.056]	[0.087]	[0.023]	[0.088]	[0.095]	[0.023]	[0.094]	[0.103]	[0.029]	[0.100]
DR_h	0.079	0.046***	0.034	0.200**			0.263***			0.318***	0.167***	0.172**
	[0.071]	[0.009]	[0.070]	[0.088]			[0.083]			[0.084]	[0.028]	[0.078]
	[0.067]	[0.010]	[0.066]	[0.098]			[0.093]			[0.107]	[0.032]	[0.105]
FPX_h	0.808***	0.632***	0.165***	0.704***			0.649***			0.578***	0.352***	0.199***
	[0.064]	[0.023]	[0.060]	[0.061]			[0.059]			[0.061]	[0.032]	[0.065]
	[0.062]	[0.028]	[0.057]	[0.067]			[0.066]			[0.071]	[0.027]	[0.083]
	Long Sample: 1982-2020											
CF_h	0.140***	0.290***	−0.143***	0.089	0.314***	−0.223***	0.078	0.380***	−0.304***	0.089	0.447***	−0.359***
	[0.033]	[0.019]	[0.035]	[0.066]	[0.019]	[0.062]	[0.062]	[0.022]	[0.055]	[0.057]	[0.027]	[0.050]
	[0.032]	[0.018]	[0.035]	[0.067]	[0.019]	[0.062]	[0.061]	[0.022]	[0.055]	[0.059]	[0.026]	[0.052]

to-smoothed-earnings ratios is accounted for by differences in expected four-year earnings growth and 12.6% is accounted for by differences in four-year returns. Just as our main results, differences in cash flow growth are overestimated, with errors in expected cash flows accounting for nearly a third (31.0%) of all dispersion in price–to-smoothed-earnings ratios, and differences in return are underestimated.

H.4. Delisting firms

As explained in Section VI, we require in the main analysis that firms have an observed future price and future earnings. This allows us to calculate direct forecast errors for the subjective expectations. To test whether survivorship bias is impacting the results, we repeat our analysis without the requirement that firms must have a future observed price and observed earnings. Instead, when firms exit the sample, we measure the delisting return and reinvest those funds into the remaining firms in the portfolio. We then calculate earnings growth and returns under this reinvestment strategy.

As shown in Table AIX, the results are quite close to our main estimation in Table I. We find that 42.7% (4.6%) of dispersion in price-earnings ratios is explained expected (realized) four-year earnings growth and 16.7% (31.8%) is explained by expected (realized) four-year returns.

H.5. Comparison of Value Line and I/B/E/S forecasts

Table AX and Figure A4 show the exercises performed in Section VI.B to validate Value Line forecasts.

Table AX Panel A shows the coefficients of regressing Value Line earnings forecasts on the I/B/E/S consensus forecasts for both the one-year and four-year horizons. The estimated slopes are close to one across specifications, indicating near one-for-one comovement between the two expectation sources in the cross-section. Panel B repeats our cross-sectional cash flow decomposition using Value Line instead of I/B/E/S and reports the estimated cash flow

Table AX

Value Line expectations

Panel A regresses Value Line earnings forecasts on the I/B/E/S consensus for overlapping horizons. We report results for one-year and four-year expected earnings, both in levels ($E_t^*[X_{t+h}]$) and scaled by the current price ($E_t^*[X_{t+h}]/P_t$). Panel B repeats our cross-sectional price-earnings ratio decomposition using Value Line expectations instead of I/B/E/S expectations over the overlapping horizons, and reports the estimated cash flow coefficients CF_1 and CF_4. The sample period for both panels is 1999 to 2020. Panel C examines time stability in Value Line long-horizon expectations by estimating the four-year discount rate coefficient DR_4 over the long sample (1990-2016) split over two subperiods: 1990–2003 and 2004–2016. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

Panel A: Expectations				
	Value Line			
	$h=1$		$h=4$	
	$E^*[X_{t+h}]$	$\frac{E^*[X_{t+h}]}{P_t}$	$E^*[X_{t+h}]$	$\frac{E^*[X_{t+h}]}{P_t}$
I/B/E/S	0.919***	0.900***	0.915***	0.856***
	[0.009]	[0.021]	[0.010]	[0.022]
	[0.008]	[0.018]	[0.012]	[0.024]
R^2	0.902	0.706	0.902	0.652
Panel B: CF coefficients				
	Value Line		I/B/E/S	
	CF_1	CF_4	CF_1	CF_4
	0.336***	0.423***	0.331***	0.433***
	[0.023]	[0.031]	[0.024]	[0.020]
	[0.023]	[0.034]	[0.027]	[0.022]
Panel C: Value Line subsample				
	$1990-2003$		$2004-2016$	
DR_4	0.138***		0.139***	
	[0.032]		[0.077]	
	[0.033]		[0.080]	

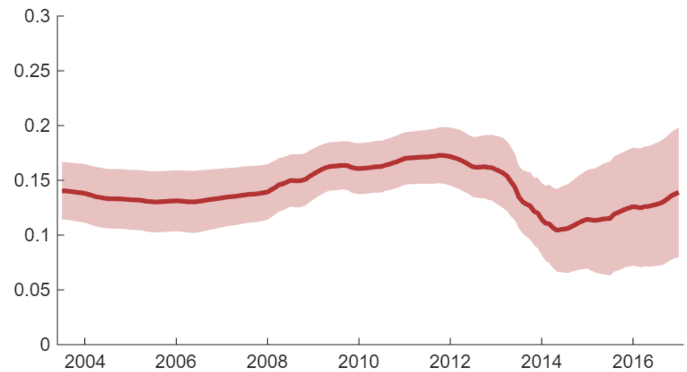

Figure A4. Time variation in the long-horizon discount-rate component implied by Value Line forecasts. The figure plots the estimated four-year discount rate coefficient DR_4 computed over rolling 13-year windows. The x-axis shows the time t that the 13-year window ends. The y-axis shows the estimated DR_4 using Value Line expectations over a 13-year subsample of $t-12$ to t.

coefficients CF_1 and CF_4. The resulting estimates are very similar to the I/B/E/S-based benchmarks, implying that the covariance moments that identify the cash flow component are essentially unchanged when using Value Line. Table AX Panel C examines time stability in Value Line long-horizon expectations by estimating the four-year discount-rate coefficient DR_4 separately over the subperiods 1990–2003 and 2004–2016, and finds stable estimates across subsamples.

Figure A4 plots the estimated four-year discount-rate coefficient DR_4 using Value Line expectations computed over a rolling 13-year windows. The first point uses the 1990–2003 window and the last point uses the 2004–2016 window, closely matching the two subperiod estimates in Panel C of Table AX. Intermediate points show all overlapping 13-year windows, illustrating that DR varies modestly across subsamples and exhibits no systematic trend.

H.6. Removing stale forecasts

The possibility that the observed stubbornness in expectations is mechanically driven by stale analyst forecasts is not supported by the evidence in Table II Panel B which shows a similar gradual adjustment in market returns. Moreover, Figure A5 replicates the main stubbornness results of Figures 2 and A2 using only actively updated forecasts within the

Table AXI

Gradual adjustment of expectations using only actively updated forecasts

Panel A shows the gradual adjustment of expectations about future earnings $\tilde{x}_{i,t+h}$ after an earnings surprise at $t+1$, i.e., the coefficients $\gamma^*_{h,j}$ estimated using equation (10). For example, the first row shows $\gamma^*_{2,1}$, the effect of an earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ on the revisions to two-year earnings $E^*_{t+1}[\tilde{x}_{i,t+2}] - E^*_t[\tilde{x}_{i,t+2}]$. The second row shows $\gamma^*_{3,1}$ and $\gamma^*_{3,2}$, the effect of an earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ on revisions about $\tilde{x}_{i,t+3}$ occurring in years $t+1$ and $t+2$. Panel B shows the estimated coefficient θ_j from equation (11) which estimates the reaction of returns $\tilde{r}_{i,t+1}, \tilde{r}_{i,t+2}$, and $\tilde{r}_{i,t+3}$ after an earnings surprise at $t+1$. To eliminate the possibility that the decomposition results are driven by stale forecasts, consensus expectations are constructed exclusively from those individual earnings forecasts that were actively updated each quarter. The sample period is 1999 to 2020. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

Panel A: Earnings revisions			
	$j=1$	$j=2$	$j=3$
$\left(E^*_{t+j} - E^*_{t+j-1}\right)[\tilde{x}_{i,t+2}]$	0.13**		
	[0.06]		
$\left(E^*_{t+j} - E^*_{t+j-1}\right)[\tilde{x}_{i,t+3}]$	0.11**	0.11***	
	[0.04]	[0.04]	
$\left(E^*_{t+j} - E^*_{t+j-1}\right)[\tilde{x}_{i,t+4}]$	0.11***	0.08***	0.11**
	[0.04]	[0.03]	[0.04]
Panel B: Returns			
$\tilde{r}_{i,t+j}$	0.10***	0.14***	0.09***
	[0.04]	[0.04]	[0.03]

current quarter to form the consensus expectations. Tables AXI and AXII show the detailed results for Figure A5. Furthermore, in Table AXIII we also replicate the main decomposition results from Table I using this subset of actively updated forecasts.

Our findings remain consistent: as in Table I, a substantial share (46.2%) of the dispersion in price–earnings ratios is explained by differences in expected four-year earnings growth, while 16.7% is attributed to differences in four-year returns. Similar to our main results, errors in expected cash flows account for a significant portion of the dispersion, with one-third (36.2%) of the total dispersion reflecting overestimation of differences in cash flow growth. Conversely, differences in returns are underestimated.

Table AXII

Revisions in expectations using only actively updated forecasts

This table shows the effect of earnings surprises on revisions using only actively updated forecasts. To eliminate the possibility that the decomposition results are driven by stale forecasts, consensus expectations are constructed exclusively from those individual earnings forecasts that were actively updated each quarter. For each anomaly $\tilde{a}_{i,t}$, we sort stocks into five equal-value portfolios based on the anomaly variable. Each row shows the coefficient from regressing the revision in future earnings $E^*_{t+1}[\tilde{x}_{i,t+2}] - E^*_t[\tilde{x}_{i,t+2}]$ on the earnings surprise $x_{i,t+1} - E^*_t[x_{i,t+1}]$. The Representative Anomaly (RA) is the average ranking of each stock across 20 different anomalies. P/X shows the results for the main portfolios used to study cross-sectional variation in price-earnings ratios. Each anomaly variable $\tilde{a}_{i,t}$ is scaled to have unit variance and to positively comove with future returns. The sample period is 1999 to 2020. Driscoll-Kraay standard errors are calculated for each coefficient using a lag equal to the equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

$\tilde{a}_{i,t}$	Coefficient	$\tilde{a}_{i,t}$	Coefficient
rdm	-0.058	noa	0.159***
	[0.060]		[0.060]
bm	0.202***	oaa	0.177***
	[0.060]		[0.057]
cfp	0.243***	ol	0.156**
	[0.052]		[0.061]
adm	0.049	pia	0.212***
	[0.040]		[0.075]
nop	0.079	poa	0.134**
	[0.049]		[0.067]
ia	0.101***	pta	0.043
	[0.026]		[0.052]
gp	0.258**	occ	0.308***
	[0.104]		[0.07]
ivc	0.013	dur	0.178***
	[0.035]		[0.063]
ivg	0.144***	cei	0.104
	[0.051]		[0.068]
ig	0.113***	P/X	0.131**
	[0.044]		[0.063]
nsi	0.202***	RA	0.221***
	[0.043]		[0.053]

Table AXIII

Decomposition of dispersion in price-earnings ratios using only actively updated forecasts

This table decomposes the variance of price-earnings ratios when the expectations are constructed using only actively updated forecasts. The *Realized* columns report the elements CF_h, DR_h and FPX_h of the decomposition using future earnings growth, future negative returns and future price-earnings ratios. The *Expected* columns report the elements of the decomposition using expected earnings growth, expected returns and expected price-earnings ratios. To eliminate the possibility that the decomposition results are driven by stale forecasts, consensus expectations are constructed exclusively from those individual forecasts that were actively updated each quarter. The *Errors* columns report the contribution of the forecast errors of each element. The main sample period is 1999 to 2020. The fourth row shows the element CF_h of the decomposition estimated over the longer sample period of 1982-2020. Driscoll-Kraay standard errors and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	$h=1$			$h=2$			$h=3$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.109**	0.356***	−0.247***	0.083	0.374***	−0.291***	0.063	0.420***	−0.357***	0.100**	0.462***	−0.362***
	[0.040]	[0.027]	[0.034]	[0.055]	[0.019]	[0.051]	[0.054]	[0.017]	[0.053]	[0.050]	[0.016]	[0.054]
	[0.033]	[0.027]	[0.029]	[0.045]	[0.021]	[0.043]	[0.043]	[0.020]	[0.042]	[0.038]	[0.018]	[0.040]
DR_h	0.157**	−0.002	0.159**	0.293***			0.327***			0.319***	0.167***	0.152
	[0.063]	[0.010]	[0.066]	[0.064]			[0.065]			[0.088]	[0.036]	[0.099]
	[0.060]	[0.011]	[0.067]	[0.058]			[0.051]			[0.060]	[0.031]	[0.077]
FPX_h	0.726***	0.630***	0.096	0.609***			0.588***			0.552***	0.317***	0.234***
	[0.058]	[0.026]	[0.068]	[0.037]			[0.038]			[0.059]	[0.030]	[0.074]
	[0.057]	[0.026]	[0.073]	[0.037]			[0.032]			[0.043]	[0.027]	[0.059]
	Long Sample: 1982-2020											
CF_h	0.143***	0.355***	−0.211***	0.112**	0.378***	−0.266***	0.108**	0.442***	−0.334***	0.152***	0.502***	−0.350***
	[0.027]	[0.028]	[0.025]	[0.052]	[0.029]	[0.039]	[0.048]	[0.031]	[0.034]	[0.042]	[0.032]	[0.031]
	[0.020]	[0.021]	[0.018]	[0.036]	[0.022]	[0.028]	[0.031]	[0.021]	[0.024]	[0.026]	[0.023]	[0.022]

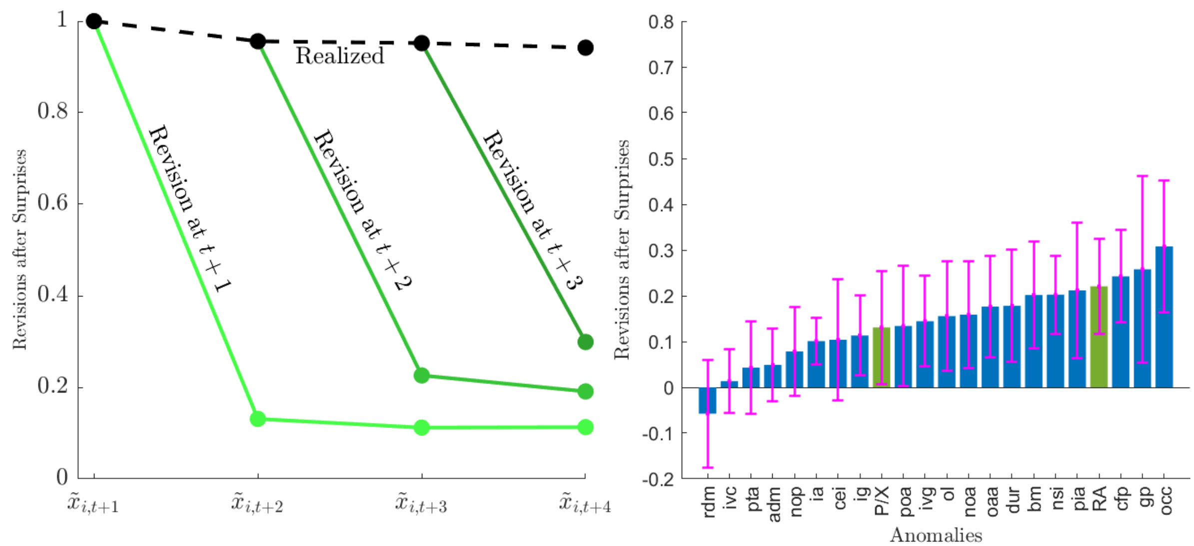

Figure A5. Evidence of stubborn expectations using only actively updated forecasts. These figures replicate Figures 2 and A2. To eliminate the possibility that revisions to expected earnings are driven by stale forecasts, the consensus expectations are constructed exclusively from those individual earnings forecasts that were actively updated each quarter.

I. Constant-gain versus Bayesian learning

This appendix expands the evidence on constant-gain learning by reporting the full path of forecast revisions following an earnings surprise, split across analyst and firm groups. Table AXIV extends Table III by estimating the response of analyst forecasts not only in the immediate revisions $E^*_{t+1}[\tilde{x}_{i,t+h}] - E^*_t[\tilde{x}_{i,t+h}]$ to the earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$, but also in the subsequent revisions $E^*_{t+2}[\tilde{x}_{i,t+h}] - E^*_{t+1}[\tilde{x}_{i,t+h}]$ and $E^*_{t+3}[\tilde{x}_{i,t+h}] - E^*_{t+2}[\tilde{x}_{i,t+h}]$. The dynamics reinforce the main conclusions of Section IV.

In Panel A, older analysts do not converge faster or slower than newer analysts through larger second- and third-step revisions. In Panels B and C, forecast revisions for $j = 1$ are similar for young and mature firms even though young firms exhibit substantially lower objective informativeness of earnings surprises. Similar to our original Table II, the large gap between γ_h and $\gamma^*_{h,1}$ for mature firms means that there is on average a large $(t+2)$ forecast error, which causes analysts to again revise their forecast errors at time $t+2$, and then revise yet again at time $t+3$. In comparison, the smaller gap between γ_h and $\gamma^*_{h,1}$ for young firms

Table AXIV

Full path of revisions: split by analyst coverage time and firm age

Panel A splits analyst–firm forecasts into "New" and "Old" based on whether an analyst's coverage length for firm is below or above the firm-specific median coverage length. For each group, we estimate equations (9) and (10). The FIRE columns report the objective informativeness γ_h of the earnings surprise $\tilde{x}_{t+1} - E_t^*[\tilde{x}_{t+1}]$ for earnings at horizons $h = 2, 3, 4$. The $(E_{t+j}^* - E_{t+j-1}^*)$ columns report the coefficient $\gamma_{h,j}^*$ measuring the size of the immediate ($j = 1$) and subsequent ($j = 2, 3$) revisions in expected earnings $\tilde{x}_{i,t+h}$ following the surprise $\tilde{x}_{t+1} - E_t^*[\tilde{x}_{t+1}]$. Panels B and C repeat the analysis after splitting firms into "Young" and "Mature" based on firm age according to the first month that the firm enters the CRSP database or the I/B/E/S forecast database. The sample period is 1999 to 2020. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

Panel A: Split by Analysts

	New				Old			
	FIRE	$E_{t+j}^* - E_{t+j-1}^*$			FIRE	$E_{t+j}^* - E_{t+j-1}^*$		
		$j=1$	$j=2$	$j=3$		$j=1$	$j=2$	$j=3$
$\tilde{x}_{i,t+2}$	0.98***	0.16***			0.97***	0.15***		
	[0.06]	[0.06]			[0.06]	[0.06]		
$\tilde{x}_{i,t+3}$	0.95***	0.12***	0.13***		0.95***	0.12***	0.13***	
	[0.09]	[0.03]	[0.05]		[0.09]	[0.04]	[0.04]	
$\tilde{x}_{i,t+4}$	0.90***	0.08***	0.08***	0.12***	0.94***	0.12***	0.08***	0.12***
	[0.13]	[0.05]	[0.03]	[0.04]	[0.09]	[0.04]	[0.03]	[0.04]

Panel B: Split by CRSP firm age

	Young				Mature			
	FIRE	$E_{t+j}^* - E_{t+j-1}^*$			FIRE	$E_{t+j}^* - E_{t+j-1}^*$		
		$j=1$	$j=2$	$j=3$		$j=1$	$j=2$	$j=3$
$\tilde{x}_{i,t+2}$	0.52***	0.09			0.95***	0.08		
	[0.15]	[0.06]			[0.03]	[0.05]		
$\tilde{x}_{i,t+3}$	0.40***	0.07*	0.01		0.94***	0.07**	0.07*	
	[0.14]	[0.04]	[0.02]		[0.05]	[0.03]	[0.04]	
$\tilde{x}_{i,t+4}$	0.39***	0.07*	0.03	0.01	0.95***	0.08**	0.05*	0.10**
	[0.13]	[0.04]	[0.03]	[0.02]	[0.05]	[0.03]	[0.03]	[0.04]

Panel C: Split by I/B/E/S firm age

	Young				Mature			
	FIRE	$E_{t+j}^* - E_{t+j-1}^*$			FIRE	$E_{t+j}^* - E_{t+j-1}^*$		
		$j=1$	$j=2$	$j=3$		$j=1$	$j=2$	$j=3$
$\tilde{x}_{i,t+2}$	0.49***	0.13*			0.96***	0.08		
	[0.14]	[0.07]			[0.03]	[0.03]		
$\tilde{x}_{i,t+3}$	0.34***	0.10**	0.01		0.95***	0.07***	0.09**	
	[0.11]	[0.05]	[0.02]		[0.03]	[0.03]	[0.04]	
$\tilde{x}_{i,t+4}$	0.34***	0.10**	0.02	0.02**	0.97***	0.07***	0.07**	0.09**
	[0.11]	[0.04]	[0.03]	[0.01]	[0.03]	[0.03]	[0.03]	[0.04]

Table AXV

Industry informativeness

The table reports estimates of the objective informativeness of earnings surprises for future earnings by firm groups defined using the Fama-French five-industry classification. The estimates report the coefficient γ_h obtained after estimating equation (9) for $h = 2, 3, 4$. The sample period is 1999 to 2020. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	Industry informativeness				
	Consumer	Manufacturing	High-Tech	Healthcare	Other
$\tilde{x}_{i,t+2}$	1.010***	0.945***	0.945***	0.915***	0.682***
	[0.058]	[0.087]	[0.075]	[0.039]	[0.205]
$\tilde{x}_{i,t+3}$	0.909***	0.931***	1.015***	0.831***	0.512***
	[0.072]	[0.114]	[0.097]	[0.049]	[0.181]
$\tilde{x}_{i,t+4}$	0.762***	0.827***	0.971***	0.782***	0.465***
	[0.055]	[0.093]	[0.117]	[0.050]	[0.156]

Table AXVI

Gradual adjustment of expectations split by industry informativeness

This table repeats the analysis of Table II after splitting firm into "High informativeness" and "Low Informativeness" based on the value of γ_h for the firm industry group in Table AXV. The table shows the gradual adjustment of expectations about future earnings $\tilde{x}_{i,t+h}$ after an earnings surprise at $t+1$, i.e., the coefficients $\gamma^*_{h,j}$ estimated using equation (10). For example, the first row shows $\gamma^*_{2,1}$, the effect of an earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ on the revisions to two-year earnings $E^*_{t+1}[\tilde{x}_{i,t+2}] - E^*_t[\tilde{x}_{i,t+2}]$. The second row shows $\gamma^*_{3,1}$ and $\gamma^*_{3,2}$, the effect of an earnings surprise $\tilde{x}_{i,t+1} - E^*_t[\tilde{x}_{i,t+1}]$ on revisions about $\tilde{x}_{i,t+3}$ occurring in years $t+1$ and $t+2$. The sample period is 1999 to 2020. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	High informativeness			Low informativeness		
	$E^*_{t+j} - E^*_{t+j-1}$			$E^*_{t+j} - E^*_{t+j-1}$		
	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$
$\tilde{x}_{i,t+2}$	0.078*			0.207***		
	[0.041]			[0.029]		
$\tilde{x}_{i,t+3}$	0.074**	0.038		0.073***	0.070*	
	[0.037]	[0.024]		[0.019]	[0.037]	
$\tilde{x}_{i,t+4}$	0.084**	0.030	−0.007	0.067***	−0.011	0.082***
	[0.040]	[0.028]	[0.025]	[0.020]	[0.016]	[0.023]

Table AXVII

Decomposition of dispersion in price-earnings ratios by industry informativeness

This table repeats the analysis of Table I after splitting firm into "High informativeness" and "Low Informativeness" based on the value of γ_h for the firm industry group in Table AXV. For each group, we decompose the variance of price-earnings ratios using equation (3) at multiple horizons. The *Realized* columns report the elements CF_h, DR_h and FPX_h of the decomposition using future earnings growth, future negative returns and future price-earnings ratios. The *Expected* columns report the elements of the decomposition using expected earnings growth, expected returns and expected price-earnings ratios. The *Error* columns report the contribution of the forecast errors. The main sample period is 1999 to 2020. The fourth row shows the element CF_h estimated over the longer sample period of 1982-2020. Driscoll-Kraay and block-bootstrap standard errors are calculated for each coefficient. In both cases, we use a lag equal to the maximum lag with any significant residual autocorrelation. Superscripts indicate Driscoll-Kraay significance at the 1% (***), 5% (**), and 10% (*) level.

	High informativeness						Low informativeness					
	$h=1$			$h=4$			$h=1$			$h=4$		
	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error	Realized	Expected	Error
	Main Sample: 1999-2020											
CF_h	0.084***	0.582***	−0.498***	0.182***	0.697***	−0.516***	0.508***	0.496**	0.012	0.596***	0.667**	−0.070
	[0.029]	[0.085]	[0.067]	[0.032]	[0.075]	[0.062]	[0.196]	[0.238]	[0.101]	[0.228]	[0.262]	[0.135]
	[0.029]	[0.093]	[0.074]	[0.031]	[0.079]	[0.068]	[0.202]	[0.266]	[0.161]	[0.248]	[0.275]	[0.216]
DR_h	0.106*	−0.016	0.122**	0.190*	−0.0190	0.208**	0.120*	0.025	0.096**	0.236*	0.154	0.082
	[0.057]	[0.013]	[0.057]	[0.110]	[0.049]	[0.096]	[0.07]	[0.044]	[0.044]	[0.136]	[0.117]	[0.075]
	[0.057]	[0.012]	[0.057]	[0.118]	[0.050]	[0.097]	[0.075]	[0.042]	[0.047]	[0.150]	[0.111]	[0.104]
FPX_h	0.800***	0.423***	0.376***	0.590***	0.285***	0.305**	0.372***	0.468**	−0.095	0.172	0.147	0.026
	[0.045]	[0.079]	[0.110]	[0.086]	[0.045]	[0.123]	[0.144]	[0.195]	[0.098]	[0.110]	[0.137]	[0.076]
	[0.039]	[0.088]	[0.123]	[0.094]	[0.049]	[0.132]	[0.145]	[0.231]	[0.158]	[0.123]	[0.167]	[0.108]
	Long Sample: 1982-2020											
CF_h	0.083**	0.556***	−0.473***	0.164***	0.691***	−0.529***	0.368**	0.509***	−0.139**	0.412**	0.629***	−0.215***
	[0.034]	[0.05]	[0.042]	[0.039]	[0.047]	[0.051]	[0.167]	[0.139]	[0.071]	[0.205]	[0.156]	[0.081]
	[0.033]	[0.047]	[0.042]	[0.038]	[0.048]	[0.051]	[0.159]	[0.136]	[0.076]	[0.200]	[0.151]	[0.084]

tells us that on average there is not as large of a forecast error realized at time $t+2$, which means we see a smaller set of subsequent revisions for $j = 2, 3$. In short, comparing young and mature firms, we find that the initial updating in response to an earnings surprise is similar for both groups, but the objective informativeness of the earnings growth surprise differs between the two. For the group where there is a bigger gap between the analyst updating and the objective informativeness, we see a larger series of subsequent revisions as this group on average experiences larger future surprises.

In Tables AXV-AXVII, we use a different dimension, namely the Fama-French 5 industry classifications, to generate differences in objective informativeness and test the implications of constant-gain learning. Table AXV reports estimates of γ_h across firm groups formed from their Fama-French industry classification: Consumer, Manufacturing, High-Tech, Healthcare, and other. Four industry groups exhibit high objective informativeness, with γ_1 between 0.92 and 1.0, whereas the residual "Other" group exhibits substantially lower informativeness ($\gamma_1 = 0.68$). Tables AXVI and AXVII replicate Tables III and IV after splitting firms into high- versus low-informativeness industries. Consistent with Section IV, higher objective informativeness does not increase analyst updating $\gamma^*_{h,j}$, but it does affect the importance of predictable forecast errors by changing the gap between the analysts' constant gain and the optimal gain.

J. Subjective risk for our portfolios

One of the main components of our model is that investors perceive lower risk for the high $\tilde{px}_{i,t}$ firms. While the relation is supported by our main evidence in Table I that high price-earnings ratios are associated with lower subjective expected returns, we can also look at more direct measures of subjective risk. In this section we explore the relation between our portfolios and two subjective risk measures: the absolute risk index assigned to firms by Value Line, and a cross-sectionally standardized risk index created by Jensen (2024).

The first measure of risk is the "Safety Rank," directly taken from Value Line. This

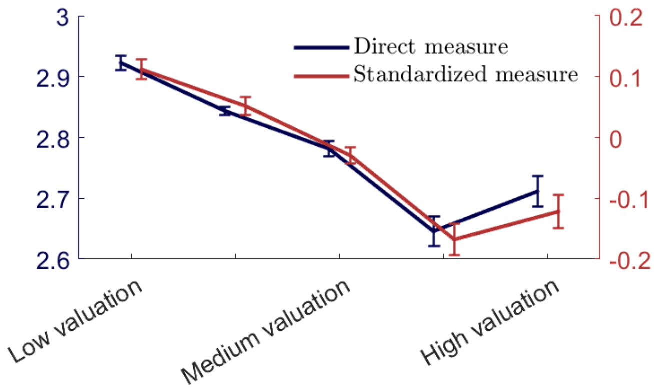

Figure A6. Subjective risk across portfolios. This figure plots the average subjective risk for each of the five main portfolios. The direct measure in blue is the "Safety Rank" measure from Value Line. This measure ranges from 1 to 5, where 5 is the highest perceived risk. The standardized measure in red takes the average of the 'Price Stability' and 'Financial Strength' measures and it is cross-sectionally rescaled to have mean zero and unit standard deviation. This measure increases with perceived risk. Each portfolio shows the 95% confidence intervals. The 5 portfolios are shown in ascending order of price-earnings ratios.

measure ranges from 1 to 5, where 5 denotes a high perceived risk, and it equals the average of the analyst score for price stability and financial strength, two perceived characteristics for each firm. We take a value–weighted average of this measure across all firms in each portfolio to obtain our first subjective risk measure. To account for time effects, we also construct a second standardized measure of risk following Jensen (2024). For each firm, we define the subjective risk as the average cross-sectional rank of price stability and financial strength, and then we standardize this measure every period (i.e., we rescale the measure so that the cross-sectional mean and cross-sectional standard deviation are 0 and 1 in every period).

Figure A6 shows that our high $\tilde{px}_{i,t}$ portfolios indeed have lower subjective expected risk using both the direct and the standardized measures of expected risk. The direct measure of risk is on average 2.90 for the lowest $\tilde{px}_{i,t}$ portfolio and 2.69 for the highest $\tilde{px}_{i,t}$ portfolio, while the standardized measure is 0.11 for the lowest $\tilde{px}_{i,t}$ portfolio and -0.12 for the highest $\tilde{px}_{i,t}$ portfolio. As shown by the 95% confidence interval bars, these differences are highly significant for both measures.