Finance without exotic risk

Pedro Bordalo, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer[1]

September 18, 2024

Abstract

We address the joint hypothesis problem in cross-sectional asset pricing by using measured analyst expectations of earnings growth. We construct a firm-level measure of Expectations Based Returns (EBRs) that uses analyst forecast errors and revisions and shuts down any cross-sectional differences in required returns. We obtain three results. First, variation in EBRs accounts for a large chunk of cross-sectional return spreads in value, investment, size, and momentum factors. Second, time variation in these spreads is predictable, and proxied by predictable time variation in EBRs. This result holds even controlling for scaled price variables, which may capture time varying required return differentials. Third, firm characteristics typically viewed as capturing risk predict disappointment of expectations (and of EBRs). Overall, return spreads typically attributed to exotic risk factors are explained by predictable movements in non-rational expectations of firms' earnings growth.

[1] The authors are from Oxford Said Business School, Università Bocconi, Brown University, and Harvard University, respectively. Gennaioli thanks the European Commission for Financial support through the ERC Advanced Grant (GA 101097578).

1. Introduction.

The textbook version of the efficient market hypothesis (EMH) holds that the realized return on a generic security such as a stock i can be written as:

$$r_{it} = r_i + \Delta E_{it},$$

where r_i is the required return, which increases in the stock's riskiness, and ΔE_{it} embeds news and revisions of rational expectations of dividends. Because under the EMH news and revisions are unpredictable ($\Delta E_{it} = 0$ on average over time), return predictability is tied to r_i. A stock earns a higher average return than another if it is riskier. But what determines risk?

In the capital asset pricing model (CAPM, Sharpe 1964), risk increases with a stock's exposure to market movements. Starting in the 1970s, evidence on return predictability not tied to such exposure challenged the CAPM, casting doubt on the underlying EMH (Basu 1977, 1983; Rosenberg et al. 1985; Banz 1981). However, as pointed out by Fama (1970) and Fama and French (1993), this evidence is not necessarily a rejection of the EMH, but possibly of the CAPM model of risk. This came to be known as the joint hypothesis problem: without observing expectations or risk, any test of market efficiency is also a test of a model of risk.

One solution to this puzzle, which has powered research in cross-sectional asset pricing for the past 30 years, is to keep rational expectations and view r_i as capturing exposure to additional "risk factors" (Fama and French 1993, 2015). It has proved challenging, however, to link these factors to tangible risks such as bankruptcy and distress (La Porta et al 1997). A second solution, pursued in behavioural finance, is to relax rational expectations, generating systematic return differentials via belief extrapolation, over- and underreaction, or other biases in the expectations term ΔE_{it} (e.g., Lakonishok et al 1994, Barberis et al. 1998, Hong and Stein 1999, Jegadeesh and Titman 2011, Daniel and Hirshleifer 2015, Barberis et al 2015, Kozak et al. 2018, van Binsbergen et al 2023). Even in this approach, returns are not matched to measured expectations. The joint hypothesis problem remains.

To make progress, we use measured expectations of the future earnings growth of US listed firms, an empirical proxy for ΔE_{it}. This allows us to assess the rationality of expectations and their ability to account for returns, testing the departures from the EMH directly. In recent years, the use of surveys of analyst and investor expectations has become common (La Porta 1996, Greenwood and Shleifer 2014, Giglio and Kelly 2018, Bordalo, Gennaioli, La Porta, and Shleifer BGLS 2019, 2024, de la O and Myers 2021, 2024, Nagel Xu 2022, Jiang et al 2022, Bianchi et al. 2024). Here we ask: can cross-sectional return spreads typically attributed to risk factors come from expectations? We show that, to a large extent, the answer is yes.

In Section 2, we use the Campbell-Shiller decomposition and measured expectations to construct firm level *expectations-based returns* (EBRs), the proxy for ΔE_{it}. EBRs attribute all cross-sectional return variation to observed belief errors and revisions, while shutting down any cross sectional and time series variation in required returns. We show that variation in EBRs quantitatively explains most of the contemporaneous returns of the long-short book-to-market and size portfolios (HML and SMB, Fama and French 1993). The residual spread after accounting for expectations is close to zero and insignificant. HML and SMB are puzzles of expectations, with little room left for risk.

We then extend the analysis to the other standard factors, investment and profitability (making up the Fama French 2015 five factor model), as well as momentum (Jegadeesh and Titman 1993). Here as well, after accounting for variation in EBRs there is little left to explain, except for profitability, where we face the challenge of attrition of low profitability firms with negative earnings. These results hold for short horizons typical of cross-sectional analyses (Fama and French 1993, 2015) but also longer horizons common to reversal anomalies (De Bondt and Thaler 1985). They also hold controlling for aggregate market optimism, showing that EBRs capture systematic cross-sectional variation that affects cross sectional spreads.

The explanatory power of EBRs for cross-sectional spreads suggests that market inefficiency is at play: average spreads materialize because the realized earnings growth of stocks in the portfolio's short arm systematically disappoints compared to that of stocks in its long arm. This is in line with growing evidence of systematic overreaction of long-term earnings growth forecasts, LTGs (BGLS 2024). Consistent with this mechanism, we show that realized spreads have large and significant loadings on forecast errors, again to an extent that leaves little room for risk to explain observed average return differentials.

One challenge to this interpretation is the measurement of expectations: if analysts mechanically extract expectations from prices, risk-driven price movements may be erroneously interpreted as informative about future earnings growth. BGLS (2024) already found strong evidence against this possibility at the aggregate and firm levels. Here we perform new tests in our cross-sectional setting. In particular, we first ask whether future cross-sectional spreads and EBRs are predictable using current portfolio-level expectations *and* scaled price variables. Because the latter capture both market expectations and required returns, in an efficient market they should be the key source of predictability, even if expectations are mechanically inferred from prices and capture shocks to required returns.

We find that expectations predict future return spreads in the direction consistent with the non-rational mechanism: spreads are predictably high after periods when analysts are very optimistic about the portfolio's short arm and they are predictably low (including negative) otherwise. Crucially, expectations completely dwarf price-ratio-based predictability. This finding is contrary to the inference from prices hypothesis and suggests that cross-sectional spreads arise from to non-rationality of expectations, with little room left for risk.

To further characterize the non-rational expectations mechanism, we next show that future firm level EBRs are predictable from firm characteristics such as the book to market ratio, size, investment, profitability and momentum. Rather than proxies for risk factors, these characteristics proxy for non-rational beliefs, predicting belief disappointment and downward forecast revisions. Market inefficiency links characteristics to predictable returns.

Our results have significant implications for the risk-based analysis of cross-sectional return spreads. They suggest, most importantly, that if documented return patterns reflect expectations and not risk, then the corrections for "risk factors" in the computation of abnormal returns are inappropriate. Many empirical findings might have been erroneously discarded because they were attributed to compensation for risk and thus deemed unsurprising. They could instead be driven by expectations or other forces.[2]

We contribute to a fast-growing program showing that measured expectations allow an empirically and theoretically disciplined approach to asset prices. Following La Porta (1996), research showed that high LTG predicts low returns due to overreaction to good news (BGLS 2019). BGLS (2024) show that forecast errors of aggregate LTG are predictable and account for most of the predictability of market level returns; see also Nagel and Xu (2022) and Adam and Nagel (2023). The volatility of aggregate LTG can also quantitatively explain Shiller's (1981) excess volatility puzzle (Bordalo et al. 2024). De la O and Myers (2021) show that volatility of valuation ratios reflects short term earnings expectations. Overall, non-rational measured beliefs unify cross sectional and aggregate return predictability without the need for unobserved variation in aggregate and cross sectional required returns.

BGLS (2024) show that fluctuations in aggregate LTG account for some cross-sectional predictability. Here we use EBRs to develop a systematic analysis of standard cross-sectional return spreads. We show that these can be explained by cross sectional co-movement of expectations and forecast errors that goes beyond the market level waves of optimism.[3]

---

[2] Conversely, other work systematically documents patterns of price distortions, or alphas, relative to simple risk models (e.g., van Binsbergen and Opp 2019, van Binsbergen et al 2023). Our results suggest that such findings may reflect systematic expectation errors. Linking returns to measured expectations is a productive way forward.

[3] Frey (2023) finds that many suggested risk factors predict convergence of earnings growth forecasts between the long and short arms of the factor, consistent with our results.

Our findings also connect to recent attempts to add new risk factors such as duration (Lettau and Wachter 2007, van Binsbergen and Koijen 2017, Gormsen and Lazarus 2023), or intertemporal versions of CAPM (Campbell and Vuolteenaho 2004, Campbell et al. 2023). These papers do not offer direct measures of risk, nor use data on expectations, and therefore cannot rule out these patterns being generated by incorrect beliefs, which often yield pricing biases that are horizon-dependent and time varying (Giglio and Kelly 2018, BGLS 2019, 2024). Instead, and in line with our results, Engelberg et al (2018) show that cross-sectional returns associated with a wide range of risk factors accrue mainly during cash flow news events when forecast errors materialize, consistent with growth expectations being predictably surprised.

Section 2 describes our framework and data. Section 3 revisits the Fama French three factor model and shows that both value and size premia can be explained in terms of expectations, with no need for risk premia. Section 4 extends this analysis to additional factors. Section 5 ties the evidence on factor returns directly to predictable corrections of expectations errors and shows that in fact characteristics capture biased expectations. Section 6 concludes.

2. Concepts and Methods

2.1 Risk in Efficient Markets Finance

Following Campbell and Shiller (1987, 1988), the log return $r_{i,t+1}$ obtained from holding the stock of a generic firm i between t and $t+1$ can be approximated as:

$$r_{i,t+1} = \alpha\left(p_{i,t+1} - d_{i,t+1}\right) + g_{i,t+1} - \left(p_{i,t} - d_{i,t}\right) + k, \tag{1}$$

where $p_{i,t}$ is the firm's log price at t, $d_{i,t}$ is its log dividend, $g_{i,t+1} = d_{i,t+1} - d_{i,t}$ is its dividend growth between t and $t+1$, while constants k and α depend on the mean log price dividend ratio.[4] By iterating (1) forward and ruling out "bubbles" we obtain the ex-post identity:

[4] As in Campbell and Mei (1993), we equalize k and α across firms. Specifically, we set $\alpha = \frac{e^{pd}}{1+e^{pd}} = 0.9981$ where $pd = 6.2634$ is the average price dividend ratio in the sample.

$$p_{i,t} - d_{i,t} = \frac{k}{1-\alpha} + \sum_{s\geq 0} \alpha^s g_{i,t+1+s} - \sum_{s\geq 0} \alpha^s r_{i,t+1+s}. \quad (2)$$

Ex-ante, the equilibrium price $p^e_{i,t}$ for the stock is obtained by taking the expectation of Equation (2) using market beliefs $\tilde{\mathbb{E}}_t(.)$. If the market requires a constant compensation r_i for the firm's risk, and thus expects r_i as a future return, we obtain:

$$p^e_{i,t} - d_{i,t} = \frac{k - r_i}{1-\alpha} + \sum_{s\geq 0} \alpha^s \tilde{\mathbb{E}}_t\left(g_{i,t+1+s}\right). \quad (3)$$

Plugging the equilibrium price into Equation (1), the realized stock return satisfies:

$$r_{i,t+1} = r_i + \left[g_{i,t+1} - \tilde{\mathbb{E}}_t\left(g_{i,t+1}\right)\right] + \sum_{s\geq 1} \alpha^s \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\left(g_{i,t+1+s}\right), \quad (4)$$

which is higher if the firm is riskier, r_i is higher, or if good news are received on the firm, either as a positive dividend growth surprise (in square brackets), or as an upward revision of expectations of future growth (the sum). The sum of the last two terms is what we called ΔE_{it}.

Traditional asset pricing builds on efficient markets, the hypothesis that market expectations are rational. This implies that news in Equation (4) cannot be systematically positive or negative, so that the average one period return $\overline{r}_{i,1}$ is simply the required return:

$$\overline{r}_{i,1} = r_i. \quad (5)$$

With rational expectations, if stock or portfolio i earns a higher average return than j, $\overline{r}_{i,1} > \overline{r}_{j,1}$, then it must have a higher risk exposure and thus a higher required return, $r_i > r_j$.

Research maintaining the rational expectations assumption starts from Equation (5), and searches for the model of risk r_i matching the observed average return $\overline{r}_{i,1}$. This has led to the creation of "risk factors" often based on firm-level characteristics such as book to market, size, investment, and profitability (Fama and French 1993, 2015), but also on their recent returns (momentum, Jegadeesh and Titman 1993). The interpretation of these factors as capturing genuine sources of investor risk, however, remains problematic.

By using measured expectations of future firm level fundamentals, we work with Equation (4) without having to assume (5). This allows us to assess how much of the average return spreads can be accounted for by measured expectations, and hence what is left for risk to explain. We first lay out our strategy, and then describe how we implement it empirically.

2.2 Expectations Based Returns (EBRs)

The pillar of our approach is the theoretical concept of "Expectations Based Returns" or EBR. We define EBR as the part of a firm's realized stock return that is *exclusively* due to the forecast error and expectation revisions in Equation (4), while deliberately shutting down any cross-firm variation in risk, as captured by r_i. The realized EBR is obtained from Equation (4) by replacing the firm-specific required return r_i by the average market return r:

$$\text{EBR}_{i,t+1} = r + [g_{i,t+1} - \tilde{\mathbb{E}}_t(g_{i,t+1})] + \sum_{s\geq 1} \alpha^s (\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t)(g_{i,t+1+s}), \tag{6}$$

Substituting (6) into (4) we obtain the key equation that decomposes realized return into the true excess returns $(r_i - r)$ and the expectations-based return:

$$r_{i,t+1} = (r_i - r) + \text{EBR}_{i,t+1}. \tag{7}$$

Suppose we could perfectly measure $\text{EBR}_{i,t+1}$. A regression of a firm's realized return $r_{i,t+1}$ on its contemporaneous $\text{EBR}_{i,t+1}$ should then give a unit slope and, crucially, a regression constant that offers an unbiased estimate of $(r_i - r)$. Constructing EBRs for the long (L) and short (S) factor portfolios would then yield:

$$(r_{L,t+1} - r_{S,t+1}) = (r_L - r_S) + (\text{EBR}_{L,t+1} - \text{EBR}_{S,t+1}), \tag{8}$$

so the constant term now identifies the pure risk-based return spread. A finding that $r_L - r_S = 0$ says that non-rational expectations are enough to account for observed return differences. If instead $r_L - r_S > 0$, the intercept gives us a magnitude of the "needed" cross sectional risk premium. On the other hand, $r_L - r_S < 0$ means that the standard factor return is entirely due to non-rationality: the long arm is safer than the short arm, a theoretical possibility.

We use measured expectations of future earnings growth to construct a proxy for $\text{EBR}_{i,t+1}$. We next show how such a proxy is constructed. Of course, analyst expectations are likely to imperfectly proxy for market expectations and hence for $\text{EBR}_{i,t+1}$. Our strategy takes this into account, and proposes a way to adjust for specific forms of measurement error.

2.3 Data and Construction of EBRs

Expectations data. We obtain monthly firm level data on analyst forecasts of future earnings growth of listed firms from the IBES Unadjusted US Summary Statistics file. We focus on the median forecasts of a firm's earnings per share (EPS_{it}) and of its long-term earnings growth (LTG_{it}), defined as the "...expected annual increase in operating earnings over the company's next full business cycle. These forecasts refer to a period of between three to five years." This data is available starting on 3/1976 for EPS_{it} and 12/1981 for LTG_{it}. EPS_{it} forecasts are for fixed horizons. To work with monthly data, and to fill in any missing forecasts, we interpolate EPS_{it} at horizons of 1 to 5 years (in one-month increments).

We collect median forecast data on dividends for the upcoming fiscal year from IBES, and use them to compute the stock's expected payout ratio. Although IBES began tracking dividend forecasts in 1994, the data do not become broadly available until 2002. Our dataset includes expected payout data for approximately 56% of the observations from 2002 to 2023, and for 25% of observations across the entire sample.

Other data. We obtain monthly data on shares outstanding and returns from CRSP, from 1981 to 12/2022. We obtain quarterly and annual accounting data from COMPUSTAT (also through 12/2022) and data on the risk-free rate (the return of the 90-day t-bill) from CRSP.

We define book to market (BM) and investment following Fama and French (2015) and use NYSE breakpoints to assign stocks to quintile portfolios of BM and investment.

To match our data to the method used for computing portfolio returns in Ken French's website, we define the raw monthly expectation-based return $\mathrm{EBR}^{\mathrm{r}}_{i,t,t+1}$ of firm i between months t and $t+1$ as:

$$\mathrm{EBR}^{\mathrm{r}}_{i,t,t+1} = \frac{D_{i,t+1} + \tilde{P}^{a}_{i,t+1}}{\tilde{P}^{a}_{i,t}}, \tag{9}$$

where, critically, $\tilde{P}^{a}_{i,t}$ is a price index that depends on analyst earnings growth expectations but shuts down any variation in required returns, across firms and over time. We construct this index, which we dub "analyst' price", based on the Campbell Shiller decomposition and the assumption of constant (across firms and over time) returns in Equation (3).

We set the required return r for all firms at the average in-sample realized annual market return, $r = 10.72\%$. We then write the analyst price $\tilde{P}^{a}_{i,t}$ as the present value of the firm's expected cash flows. Specifically, for each firm i at each time t (in months) we set:

$$\tilde{P}^{a}_{i,t} = \sum_{s=1,\dots,5} \frac{\tilde{\mathbb{E}}^{a}_{t} DPS_{i,t+12s}}{(1+r)^{s}} + \frac{1+g}{(1+r)^{5}} \frac{\tilde{\mathbb{E}}^{a}_{t} DPS_{i,t+60}}{r-g}. \tag{10}$$

We derive expected dividends per share from expected earnings per share, as follows. We proxy expected earnings per share with analyst short term earnings expectations $\tilde{\mathbb{E}}^{a}_{t} EPS_{i,t+12s}$ up to the second fiscal year; starting with the last non-missing positive EPS forecast and up to five years out, analysts expect EPS_{it} to grow at the rate LTG_{it}. To translate expected earnings into expected dividends, we use the expected payout ratio inferred from analysts' expectations of dividends and earnings. Specifically, we assume a constant ratio equal to the average expected payout ratio $\frac{\tilde{\mathbb{E}}_t^a DPS_{i,t+12}}{\tilde{\mathbb{E}}_t^a EPS_{i,t+12}}$ in our sample for those firms which paid dividends that year, $DPS_{i,t+12} > 0$, which equals 0.41.[5]

We do not observe analyst forecasts for very long horizons. For the terminal value which captures cash flows beyond year five, we again assume a terminal payout ratio of 0.41, and we set the continuation value of expected cash flow growth g to match the average stock price across all firms and months in 1981-2022. Since the required return r and growth in the very long term g are constant and common to all firms, differences in the price index $\tilde{P}_{i,t}^a$ across firms arise exclusively from differences in expectations.

The firm level raw EBR is extended to the monthly raw return of a portfolio π using an equal weighted average, $\mathrm{EBR}_{\pi,t,t+1}^{\mathrm{r}} = \frac{1}{|\pi|}\sum_{i\in\pi} \mathrm{EBR}_{i,t,t+1}^{\mathrm{r}}$.[6] Our results also hold when using value weighted portfolios (see Appendix B1). To compute (log) returns over longer horizons t to $t+h$, we compute firm level monthly raw returns $\mathrm{EBR}_{i,t+j-1,t+j}^{\mathrm{r}}$ for each of the next h months and aggregate up to portfolio monthly raw returns. We then rebalance portfolios at the end of each month *j* and compound monthly EBRs to obtain the *log* return (where for simplicity we use the same notation as for EBRs based on market beliefs, Equation (6):

$$\mathrm{EBR}_{\pi,t,t+h} = \sum_{j=1}^{h} \alpha^{j-1} \ln\left(\mathrm{EBR}_{\pi,t+j-1,t+j}^{\mathrm{r}}\right)$$

For example, the one-month log EBR is $\ln\left(\mathrm{EBR}_{\pi,t,t+1}^{\mathrm{r}}\right)$ under the Campbell Shiller approximation in (1) takes the form of Equation (6). We obtain EBRs for factor portfolios as the difference between the returns of the portfolio's long and short arms:

[5] Our results are robust to different specifications of the payout ratio. An alternative specification sets the expected payout ratio to zero if the firm did not pay a dividend the previous year. This has a correlation with our main specification of over 97%. The Appendix shows our results are unchanged with this measure.

[6] IBES surveys analysts in the middle of each month (i.e. the Thursday before the third Friday of every month, see IBES Unadjusted US Summary Statistics file). We use CRSP daily file to compute actual returns over the same periods as EBRs. Results are similar if we compute actual returns using calendar months, but the correlation between one- and three-months EBRs and returns is slightly stronger when using IBES.

1. High-Minus-Low book-to-market (HML): EBR of a portfolio that is long value stocks ($\pi = V$, top quintile book-to-market firms) and short growth stocks ($\pi = G$, bottom quintile). Thus, $\text{EBR}_{HML,t,t+h} = \text{EBR}_{V,t,t+h} - \text{EBR}_{G,t,t+h}$.
2. Small Minus Big Size (SMB): EBR of a portfolio that is long small stocks ($\pi = S$, bottom quintile market equity) and short big stocks ($\pi = B$, top quintile). Thus, $\text{EBR}_{SMB,t,t+h} = \text{EBR}_{S,t,t+h} - \text{EBR}_{B,t,t+h}$.
3. Conservative Minus Aggressive Investment (CMA): EBR of a portfolio that is long conservative stocks ($\pi = C,$ bottom quintile investment-to-asset ratio) and short aggressive ones ($\pi = A$, top quintile). Thus, $\text{EBR}_{CMA,t,t+h} = \text{EBR}_{C,t,t+h} - \text{EBR}_{A,t,t+h}$.
4. Robust Minus Weak Profitability (RMW): EBR of a portfolio that is long robust profitability ($\pi = R$, top quintile operating profitability) and short weak profitability stocks ($\pi = W$,, bottom quintile).Thus, $\text{EBR}_{RMW,t,t+h} = \text{EBR}_{R,t,t+h} - \text{EBR}_{W,t,t+h}$.
5. Winners Minus Losers momentum (WML): EBRs of a portfolio that is long winning stocks ($\pi = W$, top quintile returns between periods $t - 11$ and $t - 1$) and short losing stocks ($\pi = L$, bottom quintile). Thus, $\text{EBR}_{WML,t,t+h} = \text{EBR}_{W,t,t+h} - \text{EBR}_{L,t,t+h}$.

Following this notation, we refer to the generic long-short portfolio as LMS. Our sample consists of monthly firm level observations from 1981 to 2023 for which LTG_t and LTG_{t+h} exist.[7] This requirement restricts our sample from the 2 million observations in the CRSP/Compustat database to about 1.3 million observations for $h = 1$ and 1.1 million for $h = 12$. The sample drops firms that tend to be smaller in market cap, but the samples are comparable in characteristics such as book to market (0.63 in the full sample, 0.61 in our sample), size ($7.2bn vs $7.8bn), investment (0.19 for both), and others, see Appendix B1. As

[7] We also restrict the sample to firms with data on size and positive book-to-market in June of year *t* plus standard CRSP requirements (i.e. common stock listed on a major US exchange).

a robustness check, we dropped the requirement that firms have data on LTG_{t+h} and computed actual returns for the sample of firms for which LTG_t exists. This sample is very similar to our sample in all characteristics.

2.4 Raw Portfolio EBRs and Correlations with Actual Returns

Table 1 reports the average return of factor portfolios in our sample, the target of our exercise, and the average EBRs of the same portfolios.

Table 1. Average returns and EBRs of portfolios

Note: Panel A presents sample means of log portfolio returns over holding horizons h ranging from one month to five years, following the methodology outlined on the website of Ken French. Portfolios are formed independently based on quintiles. Results are displayed for the following five quintile portfolios: (1) book-to-market, with *Growth* stocks in bottom quintile and *Value* stocks in the top quintile, (2) investment, *Aggressive* stocks in the bottom quintile and *Conservative* ones in the top quintile, (3) size, *Big* stocks in the top quintile and *Small* ones in the bottom quintile, (4) Profitability, *Weak* profitability in the bottom quintile and *Robust* profitability in the top quintile), and (5) Momentum, *Losers* stocks in the bottom quintile and *Winners* stocks in the top quintile. Panel B presents sample means of log expectation-based returns (EBRs) returns computed following Equation 10 in the text for the same groupings of stocks. Portfolio returns and EBRs are equally weighted with monthly rebalancing. The sample period extends from December 1981 to December 2023.

Panel A. Average portfolio returns

Holding Horizon	Growth	Value	Aggr.	Cons.	Big	Small	Weak	Robust	Losers	Winners
1 Month	10.3%	15.7%	9.0%	14.9%	11.8%	14.4%	11.3%	13.5%	9.4%	15.4%
3 Months	10.1%	15.0%	8.7%	14.5%	11.4%	14.3%	11.0%	13.1%	9.2%	14.3%
1 Year	11.2%	15.3%	9.7%	14.9%	11.9%	15.3%	11.7%	13.4%	12.6%	12.9%
3 Years	11.7%	15.0%	10.8%	14.1%	12.1%	14.5%	12.5%	13.1%	13.2%	12.4%
5 Years	11.6%	14.1%	11.0%	13.4%	11.6%	13.8%	12.6%	12.4%	12.8%	12.0%

Panel B. Average portfolio expectation-based returns

Holding Horizon	Growth	Value	Aggr.	Cons.	Big	Small	Weak	Robust	Losers	Winners
1 Month	10.8%	13.7%	7.3%	15.5%	11.4%	10.4%	16.0%	9.4%	-16.0%	33.1%
3 Months	9.9%	13.2%	6.5%	15.0%	11.0%	10.0%	15.0%	9.1%	-13.4%	30.0%
1 Year	9.2%	13.7%	6.7%	14.5%	10.2%	11.6%	14.2%	9.3%	-0.6%	20.0%
3 Years	9.4%	13.0%	8.1%	13.1%	10.2%	11.2%	13.2%	9.8%	8.1%	12.7%
5 Years	9.6%	12.3%	8.8%	12.6%	10.1%	10.8%	13.1%	9.7%	9.2%	11.6%

In line with existing work, Panel A shows that portfolios in the long arm exhibit higher average returns than those in the short arm, at both long and short horizons. An assessment of the statistical significance of the long-short return spreads (Tables 4 and 8 below) reveals that the value and investment spreads are large and significant at all horizons in our sample, with annualized spreads between 3 and 5%. Momentum spreads are large and significant at horizons of under a year, with annualized spreads of around 5%. These are key targets of our analysis. The size spreads are instead not significant in our sample, which is also in line with the literature. Fama and French (2015) note that the size anomaly has weakened in recent decades relative to the earlier period in Fama French (1993). In our sample and period, average profitability spreads are not significant either.[8]

Panel B shows that EBRs display patterns that are directionally similar to those in Panel A for HML, SMB, CMA, and WML. For profitability the average EBR spread is the opposite of that in Panel A, a level mismatch which is however compatible with a strong correlation over time as we show below. The magnitudes of EBR spreads are also aligned with actual return spreads, with the exception of WML where spreads in EBRs are higher than in returns. While analyst expectations are an imperfect proxy for market expectations, a point we return to below, the broad agreement between Panels A and B suggests that EBRs may help account for both cross sectional and time series variation in spreads.

Indeed, while average return spreads have been the focus of the literature, for EBRs to be a good proxy they should also be positively correlated with returns *over time*. This is important in light of recent work exploring time variation in spreads (Gormsen 2021, Campbell, Giglio, and Polk 2023). The "perfect proxy" benchmark of Equation (8) implies that with a constant required return the correlation coefficient should be one. We compute the correlation between $EBR_{\pi,t+h}$ and $r_{\pi,t+h}$ for the long and short portfolios of HML, SMB, CMA, RMW and WML. We consider horizons h of $\{1, 3, 12, 36, 60\}$ months, covering the short

[8] There is also no systematic profitability spread when forming quintile portfolios on the full CRSP / COMPUSTAT sample in our sample period of 1981 – 2023. Using double sorts on size and profitability (as in Fama French 2015), a profitability spread emerges within big firms.

horizons typical of the cross-sectional analysis (Fama and French 1993, 2015) as well as longer horizons typical of reversal anomalies (De Bondt and Thaler 1985) and aggregate stock market variation. Table 2 reports the results.

Table 2. Portfolio level correlations for actual and expectation based returns.
Note: The table presents pairwise correlations between log returns and expectation-based returns (EBRs) for portfolios of stocks formed on book-to-market, investment, size, profitability, and momentum sorts over holding horizons ranging from one month to five years. The sample period is from December 1981 to December 2023.

Holding Horizon	Growth	Value	Aggr.	Cons.	Big	Small	Weak	Robust	Losers	Winners
1 Month	8%	19%	10%	16%	7%	16%	12%	12%	11%	6%
3 Months	22%	35%	24%	31%	23%	30%	29%	22%	28%	25%
1 Year	36%	52%	41%	37%	35%	46%	43%	34%	48%	42%
3 Years	43%	52%	52%	36%	34%	61%	46%	37%	54%	51%
5 Years	36%	41%	43%	28%	24%	48%	39%	28%	36%	37%

The correlation coefficients between contemporaneous returns and EBRs are positive and large at long horizons. This is also the case for profitability portfolios, whose time variation in returns is well captured by EBRs. Correlations are however always less than one, the theoretical "perfect proxy" benchmark. There are three potential reasons for this. First, analyst beliefs may depart from market beliefs due to (unobserved) disagreement between the marginal investor and the analyst consensus. Second, analyst beliefs only cover horizons up to 5 years out, so EBRs miss longer term variation in market expectations. Third, the required return in Equation (8) may be time varying, yielding a distinct source of return variation not captured by earnings growth expectations. This seems unlikely given that both fundamental risk and investor preferences toward it arguably vary more at lower frequencies, but correlations between returns and EBRs are uniformly and counterfactually higher at these frequencies.[9] We examine the link between market and analyst expectations in detail in Section 2.5.

[9] The fact that the correlation is lower than one may also capture other drivers of stock returns such as market liquidity or investor demand, especially at high frequencies.

Figure 1 offers a visual illustration of the time series correlation between EBRs and returns for the five Fama-French long minus short portfolios.

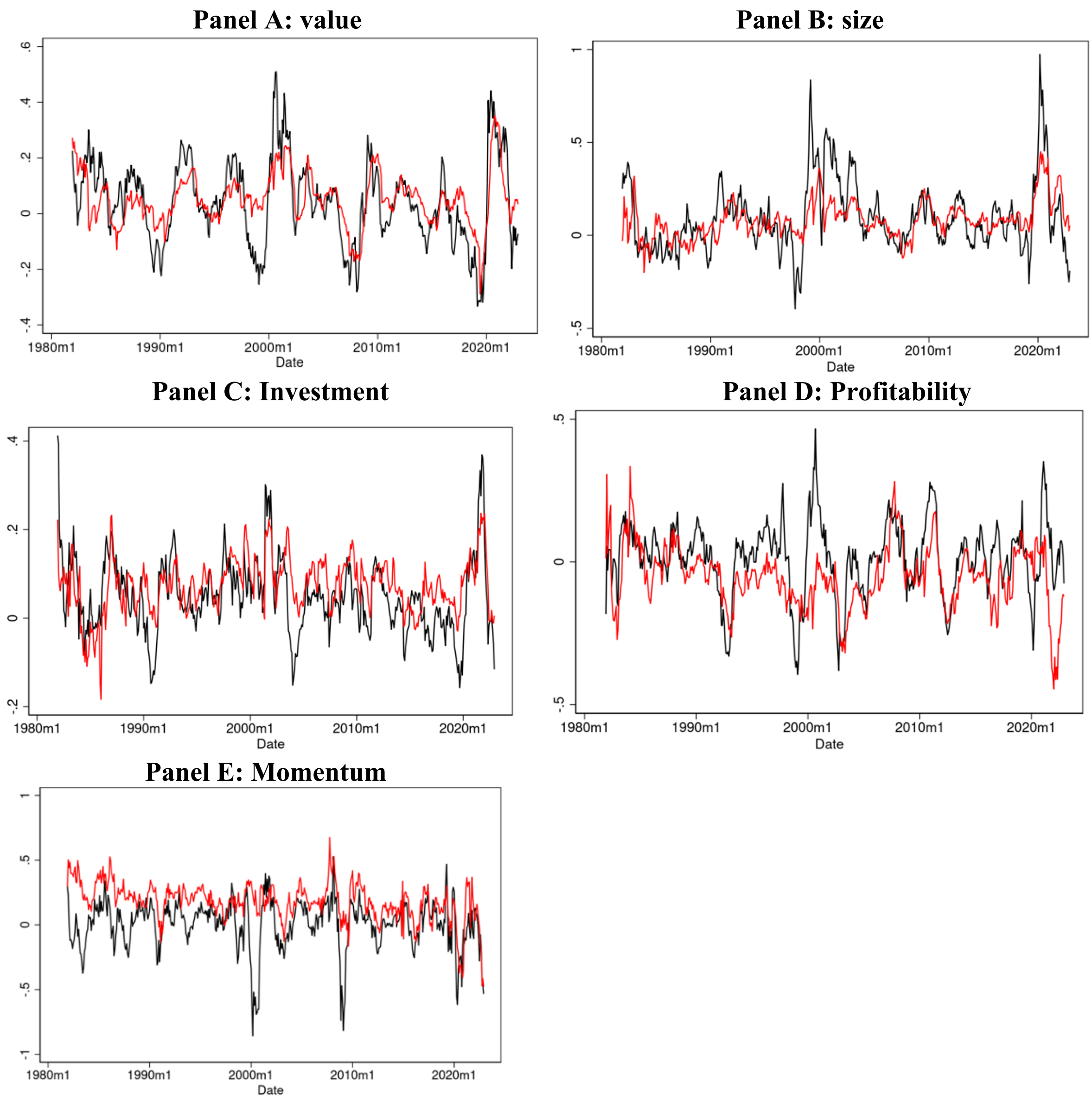


Figure 1. Actual and expectation-based return spreads for FF Portfolios

Note: Panel A plots one-year log returns (black line) and expectations-based returns (EBRs) following Equation (10) in the text EBRs (red line) for the portfolio that is long value stocks and short growth stocks (HML). Panel B plots one-year log returns (black line) and EBRs (red line) for the portfolio that is long small stocks and short big stocks (SMB). Panel C plots one-year log returns (black line) and expectations-based returns (EBRs) for the portfolio that is long conservative stocks and short aggressive stocks (CMA). Panel D plots one-year log returns (black line) and EBRs (red line) for the portfolio that is long robust stocks and short weak stocks (RMW). Panel E plots one-year log returns (black line) and EBRs (red line) for the portfolio that is long winners and short losers (WML). The sample period 1981:12–2020:12.

Return spreads are more volatile than EBR spreads, perhaps due to our expectations proxy being imperfect (e.g., EBRs do not account for expected growth beyond 5 years) or to the presence of other factors driving short term returns. Importantly, in line with Table 2, EBR and return spreads are strongly positively correlated. EBR spreads appear to track return spreads both at times when return spreads are positive – consistent with the average return differences in Table 1 Panel A – and when returns spreads are negative, contrary to the conventional risk explanation.

We next systematically assess the extent to which the variation in EBRs can account both for average return spreads and their predictable time variation. Studying the two phenomena through the prism of expectations data offers a unique opportunity to disentangle risk-based accounts, focused on r_i, from expectations-based accounts, focused on ΔE_{it}.

2.5 Measured Expectations versus Market Expectations

One concern with our analysis is the use of analyst forecasts as a proxy for market expectations of earnings growth. One issue is that analyst forecasts may measure market expectations with noise. To deal with this issue, we extract the contemporaneous information contained in expectations by regressing return spreads on contemporaneous EBR spreads. Still, the measurement error contaminates the regression constant in (8), which is our measure of the risk-based spread. In the next section we present one method to adjust the estimated constant under specific assumptions about noise and other possible measurement distortions in EBRs. These adjustments do not affect our key findings. More broadly, measurement noise would work against finding any explanatory power of EBRs for return spreads, making our estimates more conservative than they would be with a better proxy for market beliefs.

The second concern is "price-based inference": analysts might infer earnings growth expectations from stock prices by inverting a valuation formula such as the Campbell-Shiller approximation, which involves using a model for required returns. It is a priori implausible that analysts behave in this way. Producing cash flow forecasts for individual firms is their main task, and available evidence suggests that analysts perform it using a "bottom up" approach based on accounting and management information including earnings calls (Ben-David and Chinco 2024) and forecasts of the aggregate economy (Decaire and Graham 2024).

Nevertheless, "price-based inference" has implications for the relationship between analyst forecasts and stock returns, which we can test for in the data. These implications depend on which model of required returns would be used by analysts. If analysts use the same required return as investors, then inference from prices would reveal investors' cash flow expectations. Our measured expectations would then be a perfect measure of investor cash flow expectations. In this case, even if analysts engaged in "price-based inference" our analysis would be valid.

A more pressing case is if analysts mechanically infer expectations of earnings growth from prices using a mis-specified model of required returns, so analysts' expectations would erroneously incorporate priced information about required returns. For instance, analysts may assume that r_i is constant across stocks or follows the CAPM when instead the true r_i incorporates the Fama-French risk factors. This is a leading case because it would imply that, consistent with the data: i) EBRs would track actual portfolio returns and ii) forecast errors would be systematically lower for the portfolio's long arm. These features however would not reflect non-rational beliefs about cash flows but rather contamination of analyst beliefs with discount rates under mechanical inference from prices, invalidating our market efficiency tests.

We assess this second case by exploiting its two key testable implications. First, if analysts mechanically infer cash flow expectations from prices, then forecast revisions for a firm's cash flows should be driven by its contemporaneous stock return. In particular, returns should: i) have strong explanatory power for EBRs and, ii) controlling for cash flow news should add no explanatory power. In Appendix B we show that both predictions are rejected at the portfolio and firm levels. In both cases, the contemporaneous correlation between EBRs and returns is well below one (for portfolios we already saw this in Table 2). Moreover, future

EBRs are strongly predicted by contemporaneous cash flow growth and/or forecast errors (both measures of news), controlling for returns. Importantly, the increase in explanatory power by adding news to univariate regressions of EBRs on returns is substantial, with e.g. 1-year R2s increasing from 26% to 38% at the firm level, and 32% to 57% on average across portfolios. This indicates that, independent of current returns, current tangible cash flow news predicts forecast errors and revisions, consistent with non-rationality. In a similar vein, BGLS (2024) show that cash flow news is essential to account for LTG revisions at the aggregate and firm levels, and that stock returns in the past 1 or 3 years play a much smaller role. The first key prediction of mechanic inference from prices does not hold in the data.[10],[11]

The second key prediction of price-based inference is that in an efficient market the price dividend (or earnings) ratio is the leading predictor of returns, because it contains priced information about the required return r_i. If expectations are mechanically inferred from prices by inverting a valuation formula under a mis-specified model of required returns, then such expectations should contain no predictive power for returns beyond that contained in the price itself. Thus, when predicting future return spreads using current measured expectations and current scaled price variables, the latter should exhibit strong predictive power, dwarfing that of expectations. In Section 5, we show that the opposite is true in the data: current expectations have strong predictive power for future return spreads, while scaled prices have virtually no predictive power.[12] Consistent with this result, BGLS (2024) found that measured expectations

[10] Nagel (2024) argues that the explanatory power of cash flow news could surreptitiously capture time varying required returns (that vary countercyclically with the firm, portfolio, or aggregate conditions). Again, since time varying required returns would directly affect actual returns, this view implies that there should be no added explanatory power for analyst forecasts coming from cash flow news, in contrast to our findings.

[11] Chaudhry (2024) uses indexation events (Pavlova and Sikorskaya 2023) to construct proxies for non-cash flow driven price changes. He finds that such price changes impact analyst forecasts (more for the short term, less so for LTG). This identification strategy is based on the null that "analyst revisions only react to current cash flow news". This null however rules out non-rational analyst updating, which is the heart of the matter. Absent current good cash flow news, for instance, analysts may revise expectations up because they overreacted to past bad cash flow news, as in models of diagnostic expectations. Chaudhry also uses a mutual fund flow instrument (Lou 2012), which is subject to similar endogeneity concerns.

[12] The same holds when replacing scaled prices with current returns. In fact, under price-based inference EBRs should be collinear with contemporaneous returns, which is strongly rejected in the data. Nagel (2024) argues that

predict *future* returns both in the aggregate and at the firm level even controlling for current aggregate or firm level scaled price variables.

Other evidence on analysts also contradicts the price-based inference hypothesis. First, analysts make "buy" recommendations for stocks for which they have high LTG expectations, consistent with them viewing these firms as undervalued by the market (Bradshaw 2004).[13] Second, analysts do not expect these same firms to exhibit lower stock return in the future, which would be the case if analysts attributed at least some of the price increase to lower required returns by investors (BGLS 2019, De la O and Myers 2021, Decaire and Graham 2024). In sum, the analysis of expectations data and their role in return predictability rejects the hypothesis of "price-based inference." Measured expectations contain genuine information about non-rational market beliefs that affects prices and helps predict future returns.

3. EBRs and the Value and Size Premia

To study whether EBRs account for cross sectional spreads, we proceed as follows. In Section 3.1, we assess the roles of expectations versus risk by constructing EBRs for the value minus growth and small minus big long-short portfolios (from Fama and French's three factor model) and using them to estimate Equation (8). We then allow analyst expectations to be an imperfect proxy for market beliefs about earnings growth, and correct for measurement error the coefficients obtained from estimating Equation (8). Our empirical analysis shows that such discrepancies indeed exist but that accounting for them usually makes little difference for the entailed required return differential $r_L - r_S$.

---

expectations may exhibit independent predictive power because prices may be influenced by short term factors, such as short-term cash flow variation. It seems implausible that cash flow expectations inferred under a wrong required return model would be a better proxy for market required returns than the prices from which the same expectations are inferred. This seems all the more implausible for EBRs, which embed variation in short term expectations that heavily depend on short term cash flows.

[13] Jung, Shane, and Yang (2012) further show that prices respond more strongly to analyst recommendations when these are accompanied by LTG forecasts, suggesting a similar view by market participants.

In Section 3.2 we assess the contribution of different components of EBRs, in particular the forecast errors and revisions, by separately introducing them into equation (8). After accounting for EBRs, there is little systematic variation in the value and size spreads left for risk premia to explain. Forecast errors play a key role in this result. In Section 4, we extend the analysis to the other long-short portfolios.

3.1 EBRs Explain the Value and Size Premia

Table 3, Panel A reports regressions of the actual value and size long-short portfolio spreads $r_{LMS,t+h}$ on $\text{EBR}_{LMS,t+h}$, as specified in Equation (8), for various horizons h. In this regression, the constant term represents a first estimate of the required return spread, under the assumption that analyst beliefs are a perfect proxy for market beliefs.[14] To deal with overlapping observations for horizons $h > 1$ month, we correct standard errors using the Newey-West (1987) procedure.

Table 3
Expectation based returns and the HML and SMB spreads

Note: Panel A presents univariate regression results of log returns for the portfolio that is long value and short growth (HML) on expectation-based returns (EBRs) for that portfolio (columns 1 to 5) and similarly for the portfolio long small firms and short big firms (columns 6 to 10). Separate regressions are estimated for horizons h of one-month, three-month, one-year, three-year, and five-year horizon. Panel B extends the analysis by adding the expectation-based returns for the market portfolio ($\text{EBR}_{Mkt,t+h}$), which includes all the stocks in the sample. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period is December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A

			$r_{HML,t+h}$					$r_{SMB,t+h}$		
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\text{EBR}_{LMS,t,t+h}$	0.5067[a]	0.8313[a]	1.0274[a]	1.1719[a]	1.1723[a]	0.5698[a]	0.7813[a]	1.1470[a]	1.1858[a]	1.0915[a]
	(0.1527)	(0.1638)	(0.1156)	(0.2274)	(0.1842)	(0.1304)	(0.1559)	(0.1779)	(0.1787)	(0.2321)
Constant	0.0032[c]	0.0052	-0.0055	-0.0269	-0.0308	0.0027	0.0092[c]	0.0177	0.0365	0.0698
	(0.0018)	(0.0046)	(0.0140)	(0.0462)	(0.0424)	(0.0017)	(0.0049)	(0.0134)	(0.0291)	(0.0485)
Obs	504	502	493	469	445	504	502	493	469	445
Adj R^2	4%	16%	46%	45%	50%	5%	12%	34%	55%	49%

[14] The right hand side is reported in log returns over the appropriate horizon, so in the first column the constant should be read as a return of 0.32% over 1 month.

Panel B

	$r_{HML,t+h}$					$r_{SMB,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\text{EBR}_{LMS,t+h}$	0.3965[b]	0.7954[a]	1.0884[a]	1.2214[a]	1.1425[a]	0.4514[a]	0.6816[a]	1.1747[a]	1.2365[a]	1.2018[a]
	(0.1561)	(0.1732)	(0.1452)	(0.2369)	(0.2260)	(0.1390)	(0.1640)	(0.1902)	(0.2469)	(0.2735)
$\text{EBR}_{Mkt,t+h}$	0.3438[c]	0.0824	-0.1408	-0.2122	-0.3558[b]	0.3564[c]	0.2144	-0.0546	-0.0923	-0.1987
	(0.1813)	(0.1511)	(0.1374)	(0.1629)	(0.1676)	(0.1839)	(0.1532)	(0.1387)	(0.2143)	(0.2551)
Constant	0.0002	0.0032	0.0065	0.0344	0.1591[c]	-0.0005	0.0037	0.0230	0.0644	0.1705
	(0.0026)	(0.0059)	(0.0167)	(0.0661)	(0.0893)	(0.0025)	(0.0066)	(0.0225)	(0.0857)	(0.1586)
Obs	504	502	493	469	445	504	502	493	469	445
Adj R^2	6%	16%	47%	47%	57%	6%	12%	34%	55%	50%

EBRs and actual returns are significantly positively correlated for the HML and SMB portfolios, especially at longer horizons. The R^2 also sharply rises with the horizon: it is 6% at one month and 40% or more at horizons of one to five years. Expectations thus contain substantial information about the news perceived by the market. Consistent with Equation (8), the estimated coefficient on EBRs is close to one, and is statistically indistinguishable from one, at most horizons. It is smaller than one at the monthly horizon, suggesting that analyst forecasts are a noisier proxy for market beliefs at higher frequencies.

The estimate for the risk premium for HML, the regression constant, is small in magnitude and statistically indistinguishable from zero at all horizons over one month. The estimate for the risk premium for SMB is also statistically indistinguishable from zero at all horizons, except three months (recall that in our sample there is no SMB spread to begin with). After accounting for the average difference in "perceived cash flow news" with EBRs, there is no systematic value or size spreads left for risk to explain. This is our first joint hypothesis assessment: after accounting for expectations, no risk premium difference is needed to explain the average HML and SMB return spreads.

BGLS (2024) show that lagged aggregate optimism, as measured by high expectations of long-term aggregate earnings growth, high LTG_t, predicts both aggregate disappointment and a larger subsequent HML spread, suggesting that at least part of the value spread is driven by a predictable, aggregate expectations based ‘factor’. Is the explanatory power of portfolio level EBR in Table 3, Panel A due to this aggregate factor, or does it point to further sources of expectation based cross sectional spreads?

To answer this question, we compute the market-level expectations-based return $\text{EBR}_{Mkt,t,t+h}$,[15] and run a horse race between the aggregate EBR, $\text{EBR}_{Mkt,t,t+h}$, and the portfolio ones $\text{EBR}_{HML,t,t+h}$ and $\text{EBR}_{SMB,t,t+h}$, in accounting for the contemporaneous observed return spread, Table 3 Panel B. The question is not only whether the portfolio EBR survives in the regression, but also how much adding $\text{EBR}_{Mkt,t,t+h}$ affects the regression R^2 and the estimated constant compared to Panel A.

The results of Panel B indicate that the cross-sectional value spread is mostly accounted for by cross sectional movement in expectations. The estimated coefficients on $\text{EBR}_{HML,t+h}$ and $\text{EBR}_{SMB,t+h}$ remain similar to those in Panel A and are always statistically significant. The proxy for market-wide growth in optimism, $\text{EBR}_{Mkt,t+h}$ plays some role at the one month and 5 years horizons: it adds a bit of explanatory power in terms of R^2, and it causes the regression constant to shrink, further reducing the need for required return differentials. Broadly speaking, the value and size premia are largely due to cross sectional cycles in expectations, as proxied by $\text{EBR}_{HML,t,t+h}$ and $\text{EBR}_{SMB,t,t+h}$. This finding is consistent with the results in BGLS (2024), who find similar expectational boom and bust patterns in the aggregate stock market.

As seen in Table 3, EBRs are a good proxy for market expectations, consistent with research showing their explanatory power for prices and returns (BGLS 2019, 2024, De la O and Myers 2021, 2024, Nagel and Xu 2022), but not a perfect measure, especially at short horizons. We next allow analyst expectations to constitute an imperfect proxy of market beliefs

[15] Analogously to portfolio EBRs, the market EBR is $\text{EBR}_{Mkt,t,t+h} = \ln\left(\sum_{j=1}^{h} \alpha^{j-1} \frac{1}{|M|} \sum_{i \in Mkt} \text{EBR}_{i,t+j-1,t+j}\right)$, see Section 2.2.

and develop a correction for the estimated required return differential that takes measurement error into account. Our correction assumes the following affine stochastic relation between analyst forecasts $\widetilde{\mathbb{E}}_{it}^{a}$ and market forecasts $\widetilde{\mathbb{E}}_{it}$ at time t about stock i:

$$\widetilde{\mathbb{E}}_{it}^{a} = \beta + \tau \cdot \widetilde{\mathbb{E}}_{it} + \sigma \cdot \varepsilon_{it}. \quad (11)$$

where ε_{it} is an iid possibly stock specific white noise shock. This specification allows for three distortions: $\beta > 0$ may capture analysts' systematic over-optimism relative to the market, which may be due to agency problems, τ captures analysts' distorted reaction to news compared to the market, where analyst reaction is excessive relative to the market for $\tau > 1$ and insufficient for $\tau < 1$, while $\sigma > 0$ is the volatility of the iid white noise term.

As we show in Appendix A, given the measurement error structure in (11), we can use the estimated constant κ and slope γ in Table 3 and other moments estimated in the data to adjust the original estimate of the required return spread κ for distortion parameters (β, τ, σ). The analysts' systematic bias β is irrelevant for the adjustment because it cancels out when comparing different portfolios. Parameters σ and τ can be recovered using two moments of the data: the deviation of the estimated slope γ from one and the gap between $cov(g_{it}, EBR_{i,t})$ and $cov(g_{it}, r_{i,t})$.[16] Under the maintained assumption that the required return is constant, there are two equations in two unknowns, τ and σ. This estimation only uses the restriction that the required return is constant but allows for any possible risk model satisfying that restriction, including the Fama and French model.[17]

---

[16] The adjusted estimate is equal to (where variables capture differences between long and short portfolios):

$$r = \kappa + \left[\gamma - \frac{cov(r_t, g_t) - var(g_t)}{cov(\mathrm{EBR}_t, g_t) - var(g_t)}\right]\overline{\mathrm{EBR}} + \frac{cov(r_t, g_t) - cov(\mathrm{EBR}_t, g_t)}{cov(\mathrm{EBR}_t, g_t) - var(g_t)}\overline{g}$$

See Appendix A for a proof. If there is no measurement error, $\gamma = 1$ and the covariance of actual return and cash flow growth is identical to the covariance of EBRs and cash flow growth, $cov(r_t, g_t) = cov(\mathrm{EBR}_t, g_t)$, so the estimate for the required risk premium is equal to the estimated regression constant, $r = \kappa$, as in Equation (7).

[17] This exercise also allows to estimate the extent to which time variation in required returns unrelated to time variation in non-rational analyst expectations may be needed to account for the return spread observed on average. A direct horse race between possibly time varying required returns and time varying expectations is performed in our return-spread predictability tests in Section 5.

Applying this method, we find that the average τ is 1.1 so that analysts tend to respond to news on average slightly more strongly than the market. We can use these estimated parameters to adjust the estimated constant κ of the required return premium of the HML and SMB portfolios that account for measurement error in analyst expectations. Table 4 compares the actual return spread between the value and growth portfolios, and the small and big firm portfolios, to the adjusted required return measure (both measures are annualized).

Table 4
Expectation based estimates of the HML and SMB required return spreads

Note: the table presents estimates of the required return premia for the portfolio that is long value and short growth (HML, Columns 1 to 5) and for the portfolio that is long small firms and short big firms (SMB, Columns 6 to 10). The adjustment allows for three distortions in expectation-based returns (EBRs) as described in Equation (11). As benchmarks, we report (in the first row) the sample long-short spreads for the relevant portfolios for horizons of one-month, three-month, one-year, three-year, and five-year horizons. The second row reports the intercept from a univariate regression of annualized log returns of relevant long-short portfolio and horizon h on their EBRs. The last row reports annualized estimates of the required risk premia. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period extends from December 1981 to December 2018. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

			$r_{HML,t+h}$					$r_{SMB,t+h}$		
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Average	0.0535^{a}	0.0484^{b}	0.0411^{b}	0.0338^{b}	0.0259^{b}	0.0263	0.0290	0.0340	0.0241	0.0216
spread	(0.0206)	(0.0196)	(0.0200)	(0.0153)	(0.0127)	(0.0213)	(0.0202)	(0.0210)	(0.0182)	(0.0154)
Constant κ	0.0387^{c}	0.0207	-0.0055	-0.0090	-0.0062	0.0319	0.0368^{c}	0.0177	0.0122	0.0140
	(0.0213)	(0.0183)	(0.0140)	(0.0154)	(0.0085)	(0.0209)	(0.0188)	(0.0134)	(0.0097)	(0.0097)
Adjusted κ	0.0257	0.0150	-0.0091	-0.0029	-0.0007	0.0362	0.0379	0.0130	0.0207	0.0298

Correcting for measurement error confirms our previous results, and yields estimated true spreads for HML that are even closer to zero than those in Table 3. The value-growth puzzle appears to be entirely about expectations of future earnings growth being bullish for growth stocks and bearish for value stocks, compared to reality. In our sample, the average SMB spread is not significant (yet as shown in Table 3 its variation over time is well captured by expectations).

In sum, the evidence shows that the observed cross-sectional differences in expectations fully account for HML and SMB spreads. Growth stocks do worse when optimism for them drops relative to that for value stocks, and similarly for big stocks relative to small ones. Growth stocks do worse on average because such optimism happens more than its reverse, not because growth stocks are less risky. In principle, it is entirely possible that the value and size spreads are predictably negative when the prospects of high growth (resp. big) firms are underestimated relative to those of value (resp. small) firms. We later assess this possibility.

3.3 Decomposing EBRs into Forecast Errors and Revisions

Instead of computing EBRs, in the spirit of Equation (4) we can directly regress realized returns on contemporaneous forecast errors and revisions at different horizons. This strategy is informative for two reasons. First, it allows us to separately assess the explanatory power of forecast errors, which are directly linked to belief biases, and of forecast revisions at different horizons. Second, it relaxes the parametric restrictions embedded in our computation of EBRs, in particular by allowing analyst forecasts (available up to 5 years ahead) to capture potentially correlated but unmeasured variation in longer term beliefs.

We perform the decomposition at the yearly horizon or above (earnings are published quarterly so we refrain from computing forecast errors at a 1- or 3-month frequency). For one and three years horizons, $h = 12, 36$, we compute the firm level forecast error as the difference between realized one or three year earnings growth and the growth expected one or three years prior $FE_{i,t+h} = \ln\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right) - \ln\left(\mathbb{E}_t\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right)\right)$. At five year horizons, we compute the forecast error using LTG as $FE_{i,5} = \ln\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right)/5 - LTG_{i,t}$. We compute revisions of short term growth forecasts, $h = 12{,}24$, as $\Delta_h STG_{i,t+h} = (\mathbb{E}_{t+h} - \mathbb{E}_t) ln\left(\frac{EPS_{i,t+h+1}}{EPS_{i,t+h}}\right)$ and revisions of long term forecasts as $\Delta_h LTR_{i,t+h} = LTG_{i,t+h} - LTG_{i,t}$.

We aggregate each measure of forecast error and revision at the portfolio level, e.g. for the forecast error we compute $FE_{\pi,t+h} = \frac{1}{|\pi|}\sum_{i\in\pi} FE_{i,t+h}$ and we analogously aggregate forecast revisions. We use the differences in these aggregated forecast errors and revisions between the long and short portfolios as explanatory variables for contemporaneous long minus short return spreads.[18] Table 5 shows the results.

Table 5
Portfolio level forecast errors and revisions predict spreads

Note: This table presents multivariate regressions of log returns for the long-short value minus growth (HML) and small minus big (SMB) portfolios for horizons (h) of one-year, three-years, and five-years. The independent variables include: (a) spreads in forecast errors between t and $t+h$, (b) spreads in forecast revisions between t and $t+h$ of one-year earnings growth in year $t+h+1$, and (c) spreads in changes in long-term growth forecasts between t and $t+h$. Panel B presents analogous results for the portfolio that small stocks and short big stocks (SMB). Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{HML,t,t+12}$	$r_{HML,t,t+36}$	$r_{HML,t,t+60}$	$r_{SMB,t,t+12}$	$r_{SMB,t,t+36}$	$r_{SMB,t,t+60}$
$(1-E_t)\ \Delta_h e_{LMS,t+h}$	0.1318[a]	0.1743[a]	0.1413[a]	0.1108[a]	0.1386[a]	0.0616
	(0.0134)	(0.0261)	(0.0376)	(0.0202)	(0.0347)	(0.0475)
$(E_{t+h}-E_t)\Delta_h e_{LMS,t+h+12}$	0.0808[a]	0.0575[b]	-0.0240	0.0313[c]	0.0144	-0.0255
	(0.0127)	(0.0248)	(0.0289)	(0.0188)	(0.0261)	(0.0365)
$\Delta_h LTG_{LMS,t+h}$	0.0276[a]	0.0182	0.0665[c]	-0.0217	0.0213	0.0783[b]
	(0.0094)	(0.0207)	(0.0346)	(0.0137)	(0.0288)	(0.0364)
Constant	-0.1111[a]	-0.2847[a]	-0.3308[a]	-0.0728[a]	-0.1219[b]	-0.1031
	(0.0213)	(0.0634)	(0.0872)	(0.0179)	(0.0473)	(0.0749)
Obs	493	469	445	493	469	445
Adj R^2	53%	48%	45%	32%	40%	21%

Two results stand out. First, the expectation components are strongly predictive of HML and SMB spreads, with forecast errors playing a dominant role in the multivariate regression (note that the regressors are standardized). The pre-eminence of forecast errors for explaining both HML and SMB suggests that returns reflect a disappointment in the short arm compared to the long arm, e.g. of growth stocks compared to value stocks for HML. Portfolio expectations revisions also play a role, so part of the HML spread is accounted for by

[18] Following the logic of the Campbell-Shiller firm-level decomposition, we are averaging logs, which implicitly drops firms with negative $EPS_{i,t}$ and/or $EPS_{i,t+h}$.

systematically lower upward revisions or larger downward revisions of future prospects for growth firms relative to value firms, particularly over the first two years.

Consistent with already reported findings, the constant terms in Table 5 suggest that the average returns spreads are entirely explained by expectations. The regression constants in Table 5 are all negative, and often statistically significant.

In sum, estimating Equation (7) shows that the systematic return spreads on the Fama French HML and SMB factors are explained by EBRs, and in particular by systematic differences in forecast errors and revisions (Tables 3 and 5). Taking these into account, there is little evidence of systematic differences in required returns (Table 4).

4. Other FF factors

We repeat the analysis of Section 3 for investment, profitability and momentum. Table 6 presents our baseline estimates of equation (8) for these factors.

Table 6
Actual and expectations based long short portfolio return spreads

Note: This table presents univariate regression results for log returns against expectation-based returns (EBRs) for three distinct long-short (LMS) portfolios. The portfolios examined are: (1) CMA, which is long stocks in lowest quintile of one-year asset growth and short stocks in highest quintile, (2) RMW, which is long stocks in the highest quintile of operating profitability and short stocks in the lowest quintile, and (3) WML, which is long stocks in the top quintile of returns during period $t-11$ through $t-1$ and short stocks in the bottom quintile of returns during the same period. We estimate separate regressions for one-month, three-months, one-year, three-years, and five-years horizons. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period extends from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{LMS,t+1}$	$r_{LMS,t+3}$	$r_{LMS,t+12}$	$r_{LMS,t+36}$	$r_{LMS,t+60}$
Investment (CMA)	(1)	(2)	(3)	(4)	(5)
$EBR_{LMS,t,t+h}$	0.2743[a]	0.4859[a]	0.8019[a]	0.7562[a]	0.8602[a]
	(0.0945)	(0.0931)	(0.1291)	(0.1857)	(0.1404)
Constant	0.0030[b]	0.0041	-0.0116	-0.0142	-0.0424
	(0.0012)	(0.0032)	(0.0130)	(0.0411)	(0.0282)
Adj R^2	3%	10%	36%	27%	42%
Profitability (RMW)					
$EBR_{LMS,t,t+h}$	0.2975[a]	0.4276[a]	0.4877[a]	0.5502[a]	0.6400[a]
	(0.1087)	(0.1101)	(0.1464)	(0.0756)	(0.0830)
Constant	0.0035[b]	0.0118[a]	0.0406[a]	0.0755[a]	0.0965[a]

	(0.0014)	(0.0035)	(0.0107)	(0.0161)	(0.0202)
Adj R^2	2%	7%	16%	33%	42%
Momentum (WML)					
$EBR_{LMS,t,t+h}$	0.1214	0.5786[a]	0.7418[a]	0.7299[a]	0.5738[a]
	(0.0991)	(0.1268)	(0.1692)	(0.1427)	(0.1343)
Constant	0.0000	-0.0499[a]	-0.1499[a]	-0.1251[a]	-0.1055[a]
	(0.0044)	(0.0166)	(0.0412)	(0.0313)	(0.0323)
Adj R^2	0%	10%	29%	50%	33%
Obs	504	502	493	469	445

As with HML and SMB, expectation-based returns have strong explanatory power for actual returns. The slope coefficients are large, statistically significant, and increase with the holding horizon. For CMA and RMW, their magnitudes are comparable to those obtained for HML and SMB returns: for investment the coefficients are close to, or statistically indistinguishable from, the benchmark value of 1 at longer horizons. For momentum, and especially for profitability, they are lower than 1 although still substantial throughout.

Turning to our main test, the intercepts are either small and statistically indistinguishable from zero, or negative, except for investment at the one-month horizon and for profitability. This result for profitability is in line with Table 1, where the average EBR is higher for low profitability firms, while actual returns go in the opposite direction. Momentum has a negative spread, which may be consistent with winners being deemed safer than losers.

We next adjust the estimated required return for measurement error across portfolios and present the results in Table 7.

Table 7
Expectation based estimates of long-short portfolios required return spread

Note: the table estimates of the required return premia (adjusted κ) for the portfolio that is long conservative and short aggressive investment stocks (CMA), long robust and short weak profitability stocks (RMW), and long winners and short momentum stocks (RMW). The adjustment allows for three distortions in expectation-based returns (EBRs) as described in Equation (11). As benchmarks, we report (in the first row) the sample long-short spreads for the relevant portfolios for horizons of one-month, three-month, one-year, three-year, and five-year horizons. The second row reports the intercept from a univariate regression of annualized log returns of relevant long-short portfolio and horizon h on their EBRs. The last row reports annualized estimates of the required risk premia. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period extends from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{LMS,t+1}$	$r_{LMS,t+3}$	$r_{LMS,t+12}$	$r_{LMS,t+36}$	$r_{LMS,t+60}$
Investment (CMA)	(1)	(2)	(3)	(4)	(5)
Average	0.0588[a]	0.0571[a]	0.0515[a]	0.0329[a]	0.0249[a]
spread	(0.0124)	(0.0113)	(0.0106)	(0.0081)	(0.0068)
Constant κ	0.0361[b]	0.0162	-0.0116	-0.0047	-0.0085
	(0.0140)	(0.0123)	(0.0130)	(0.0137)	(0.0056)
Adjusted κ	-0.0212	-0.0240	-0.0264	-0.0076	-0.0068
Profitability (RMW)					
Average	0.0221	0.0219	0.0168	0.0061	-0.0029
spread	(0.0169)	(0.0143)	(0.0140)	(0.0063)	(0.0063)
Constant κ	0.0419[b]	0.0474[a]	0.0406[a]	0.0252[a]	0.0193[a]
	(0.0167)	(0.0135)	(0.0107)	(0.0054)	(0.0040)
Adjusted κ	0.0510	0.0289	-0.0127	0.0237	0.0075
Momentum					
Average	0.0601[b]	0.0512[b]	0.0025	-0.0082	-0.0077
spread	(0.0271)	(0.0232)	(0.0216)	(0.0119)	(0.0084)
Constant κ	0.0006	-0.1996[a]	-0.1499[a]	-0.0417[a]	-0.0211[a]
	(0.0532)	(0.0618)	(0.0412)	(0.0104)	(0.0065)
Adjusted κ	-0.2795	-0.2904	-0.1709	-0.0541	-0.0316

The earlier findings are broadly confirmed. For investment the correction proves important for spreads at short rather than long horizons. For profitability and momentum, EBRs may have more noise and the corrections are accordingly larger. For profitability the estimated required return spreads decline, especially at longer horizons. The corrections for momentum are in line with the earlier interpretation that firms in the long portfolio (winners) are if anything viewed as safer than those in the short portfolio.

These patterns are confirmed in the EBR decomposition exercise, which is reported in the Appendix B2: spreads in forecast errors and revisions positively and significantly predict return spreads and the intercepts are either small and insignificant -- for investment and profitability -- or negative for size, momentum, as well as for HML (Table 4).

In sum, the investment, size and momentum puzzles are solved with expectations. As with HML, the market does not see conservative firms as riskier than aggressive ones, nor winners as riskier than losers. Instead, analysts and the market appear to hold systematically bullish expectations about firms in the short portfolios, compared to firms in the long portfolios, and the former do worse on average because that relative optimism systematically decreases.

5. Predictable returns and market efficiency.

The central puzzle motivating our paper is the predictability of cross-sectional return spreads from firm characteristics, which are sometimes interpreted are proxies for risk. While the literature has focused on the average predictability, cross sectional spreads exhibit dramatic variation over time, as shown in Figure 1, and recent work suggests that this time variation is in part predictable (Lochstoer and Tetlock 2020, Campbell, Giglio, and Polk 2023). Because expectations of earnings growth are not rational, and display revisions and errors that are predictable from lagged expectations (BGLS 2019, 2024), market inefficiency is a possible explanation for such predictable time variation.

We thus ask two questions. First, can *future* cross-sectional returns be predicted – conditional on price scaled variables – using current measured expectations, and in a way that is consistent with non-rational expectations? An affirmative answer would reject the joint hypothesis of efficient markets plus "price-based inference". Second, what is the link between expectations and firm characteristics that have been shown to predict returns? Do standard firm characteristics predict future EBRs (and hence future forecast errors and revisions) and not just returns? An affirmative answer would help explain why characteristics matter: they capture non-rational beliefs, rather than an elusive connection to tangible risks.

5.1 Predicting Portfolio Return Spreads from Expectations

We examine return predictability by first running a horse race between *current* measured expectations and price dividend ratio in predicting *future* return spreads. As discussed in Section 2.5, if the market is efficient and analysts mechanically infer expectations from prices under a wrong required return model, then the price dividend ratio should exhibit overwhelming predictive power, with higher spreads in portfolio $p - d$ forecasting lower return spreads. If instead current optimism as captured by high current expectations negatively predicts future returns but the $p - d$ ratio does not, then analysts do not infer expectations from $p - d$, indicating that time variation in return spreads mostly reflect market inefficiency, not variation in required returns incorporated in prices. This test is very demanding, both because current prices also contain information about longer term unmeasured investor expectations and because analyst expectations are only a proxy for market expectations embedded in prices.

We regress the future return spread $r_{LMS,t,t+h}$ of the long short portfolio LMS on the difference in current dividend to price ratio between the long and short portfolios, $dp_{LMS,t} = \ln DP_{L,t} - \ln DP_{S,t}$, and on the expectation components of current EBR spreads used in Table 5, namely: revisions of forecasts of long term growth, $\Delta_h LTG_{LMS,t}$, and of short term growth, $(\mathbb{E}_t - \mathbb{E}_{t-h})\Delta_h e_{LMS,t+12}$, and forecast errors $FE_{LMS,t-l,t}$. Following BGLS (2024), who show that expectation *levels*, as well as expectations of aggregate earnings, predict forecast errors, we further control for lagged expectations for long term growth, $LTG_{LMS,t-h}$, and short term growth, $\mathbb{E}_{t-h}\Delta_h e_{LMS,t+12}$, as well as the current revision $\Delta_h LTG_{Mkt,t}$ and lagged value $LTG_{Mkt,t-h}$ of expectations for the aggregate market. All variables are standardized to allow for a quantitative comparison of explanatory power.

Table 8 presents the results for the 1-month and 1-year horizon for each factor, with other horizons reported in the Appendix C1. The last row of Table 8 presents the R^2 of a univariate regression of $r_{LMS,t,t+h}$ on $dp_{LMS,t}$.

Table 8

Predicting future return spreads from expectations data

Note: the table presents regressions of log returns for portfolios that are long value and short growth stocks (HML), long small and short big stocks (SMB), long conservative and short aggressive investment stocks (CMA), long robust and short weak profitability stocks (RMW), and long winner and short momentum stocks (WML). Separate regressions are estimated for horizons (h) one-month and one year. The set of independent variables includes: (a) the ranked forecast error in portfolio earnings between $t-h$ and t, (b) the change in the portfolio forecast for short-term growth in earnings between $t-h$ and t, (c) the lagged portfolio forecast for one-year growth in earnings at $t-h$, (d) the change in the portfolio forecast for long-term growth in earnings between $t-h$ and t, (e) the lagged portfolio forecast for long-term growth in earnings at $t-h$, (f) the change in the aggregate forecast for long-term growth in earnings between $t-h$ and t, (g) the forecast for long-term growth in aggregate earnings at $t-h$, and (g) the portfolio log dividend-to-price ratio ($dp_{LMS,t}$) at time t. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). All independent variables are standardized. The last row reports the R^2 from a univariate regression of $r_{LMS,t+h}$ on $dp_{LMS,t}$. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] 0% level.

	$r_{HML,t+h}$		$r_{SMB,t+h}$		$r_{CMA,t+h}$		r_{RMWt+h}		$r_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$(1-E_t)\ \Delta_h e_{LMS,t}$		0.0821[c]		0.0142		-0.0246		-0.0309[c]		-0.0031
		(0.0492)		(0.0576)		(0.0232)		(0.0184)		(0.0320)
$(\mathbb{E}_{t+h}-\mathbb{E}_t)$ $\Delta_h e_{LMS,t+h+12}$	-0.0022	0.0396[b]	0.0020	0.0456[a]	-0.0007	0.0113	0.0039	0.0035	0.0075[b]	0.0706[a]
	(0.0023)	(0.0167)	(0.0026)	(0.0157)	(0.0017)	(0.0118)	(0.0025)	(0.0120)	(0.0032)	(0.0236)
$\mathbb{E}_{t-h}\Delta_h e_{LMS,t+12}$	0.0158	0.2155[a]	0.0304[b]	0.1286[c]	0.0081	0.1187[b]	0.0234[c]	0.0130	0.0558[a]	0.2527[a]
	(0.0105)	(0.0651)	(0.0132)	(0.0737)	(0.0092)	(0.0479)	(0.0139)	(0.0683)	(0.0176)	(0.0704)
$\Delta_h LTG_{LMS,t+h}$	0.0045[c]	-0.0762[b]	0.0060[b]	0.0057	0.0038[b]	0.0088	0.0014	-0.0029	0.0031	-0.0279
	(0.0026)	(0.0341)	(0.0025)	(0.0277)	(0.0016)	(0.0161)	(0.0023)	(0.0159)	(0.0030)	(0.0249)
$LTG_{LMS,t-h}$	-0.0190	-0.2657	0.0052	0.0348	-0.0098[c]	-0.1410[a]	0.0034	0.0794	-0.0134[c]	-0.1470[b]
	(0.0143)	(0.1730)	(0.0068)	(0.0708)	(0.0057)	(0.0516)	(0.0080)	(0.0724)	(0.0075)	(0.0572)
$\Delta_h LTG_{Mkt,t}$	-0.0046[c]	0.0257	0.0085[a]	-0.0033	-0.0023[c]	0.0176	-0.0082[a]	0.0373	0.0022	0.0073
	(0.0025)	(0.0307)	(0.0028)	(0.0248)	(0.0012)	(0.0112)	(0.0023)	(0.0257)	(0.0025)	(0.0184)
$LTG_{Mkt,t-h}$	0.0041[c]	0.0346	0.0030	0.0483[b]	0.0038[a]	0.0246[a]	0.0026	0.0356	0.0028	-0.0205
	(0.0022)	(0.0294)	(0.0032)	(0.0203)	(0.0014)	(0.0092)	(0.0028)	(0.0272)	(0.0027)	(0.0297)
$dp_{LMS,t}$	-0.0075	0.0305	0.0098	0.1491	-0.0087	0.0828	0.0232	0.1470	0.0118	0.0880
	(0.0243)	(0.2334)	(0.0153)	(0.1159)	(0.0152)	(0.0972)	(0.0215)	(0.2000)	(0.0163)	(0.0922)
Constant	-0.0401[b]	-0.4177[b]	-0.0307	-0.2504[c]	-0.0255[a]	-0.2406[a]	-0.0102	-0.1953	0.0027	0.2259
	(0.0164)	(0.1890)	(0.0271)	(0.1497)	(0.0086)	(0.0675)	(0.0170)	(0.1537)	(0.0160)	(0.2092)
Obs	444	433	444	433	444	433	444	433	444	433
Adjusted R^2	5%	16%	4%	23%	6%	27%	5%	9%	7%	26%
Univariate R^2	0%	3%	0%	7%	0%	8%	0%	0%	0%	4%

Lagged expectations have strong predictive power for the future HML return spread, even at the short 1-month horizon, a challenging test, and increasing at the 1-year horizon. This is driven in particular by the role of lagged optimism about both short- and long-term earnings growth. Contrary to the efficient markets plus “price-based inference” hypothesis, the predictive power of expectations dwarfs that of prices. In the horse race of Table 8, the dividend price ratio is never significant, whereas one standard deviation of individual expectations variables can explain up to 0.25 standard deviations of spreads, depending on the factor (as before, predictive power is lower for profitability).

Critically, the dividend price ratio’s predictive power in univariate regressions is negligible, as seen by the univariate R^2 shown in the last row. Adding lagged expectations leads to a dramatically higher adjusted R^2 in all specifications, and the dividend price ratio is insignificant. This suggests that priced portfolio risk is not time varying enough to explain time variation in the spread, which is instead explained by time varying beliefs measured by analyst forecasts. At longer horizons, expectations account for up to 60% of variation in return spreads across factors, and these results are robust to using other valuation ratios, such as price earnings or book to market (Appendix C1).[19]

Taken together, our results show that returns are predictable by measured expectations of earnings growth at the portfolio level (as well as at the aggregate and firm level, BGLS 2024). There is no evidence that these expectations erroneously capture discount rates inferred from prices. The Appendix offers further evidence against this hypothesis.[20]

How then do expectations predict future returns? Linking to our earlier evidence that EBRs explain contemporaneous spreads, future return spreads may be explained, at least in

---

[19] Price earnings ratio spreads are not significant in the horse race with expectations, while book to market is significant at longer horizons for some factors (size and momentum). The predictive power of expectations is very similar to that in Table 8. As noted above, this is consistent with market growth expectations driving cross sectional spreads, because prices incorporate those expectations. Note that aggregate returns *are* strongly predicted by aggregate pd, yet also in that case current measured expectations explain future aggregate returns controlling for prices (BGLS 2024). To further control for inference from prices, we repeat the analysis replacing the valuation ratio by the return spread, which is in closer correspondence with the forecast revisions and errors over the corresponding horizon. The results are very similar, see Appendix C.

[20] Appendix C extends the predictability analysis in Tables 8 and 9 to the firm level. Specifically, we predict firm level EBRs using the expectation variables in Table 8, and then run a horse race between predicted future firm level EBRs and current firm level valuation ratios to explain future firm level returns. Predicted EBRs again explain future returns at all horizons and firm valuation ratios have little predictive power except at long horizons, confirming that expectations are not spuriously capturing information about required returns.

part, by the unfolding of predictable market expectations errors as proxied by predictable EBRs. To assess this mechanism, we perform a two-step exercise. In the first stage, we use the variables in Table 8 to predict future EBR spreads of long-short portfolios at time $t+1$. This assesses the non-rationality of expectations, because EBRs, as a combination of forecast errors and revisions, should not be predictable if expectations are rational. We also control for the price dividend ratio, which would absorb predictability of future EBRs under the price-based inference hypothesis. In the second stage, we test the ability of the EBRs predicted from expectations, which we denote by $\widehat{\mathrm{EBR}}_{LMS,t\to t+1}$, to explain contemporaneous returns, again controlling for the price dividend ratio:

$$r_{L,t+1} - r_{S,t+1} = \beta_0 + \beta_1 \cdot \widehat{\mathrm{EBR}}_{LMS,t\to t+1} + dp_{LMS,t} + v_{t+1} \tag{9}$$

Compared to Equation (8), this test ties return differentials to error predictability, the hallmark of non-rationality, and allows for a quantitative assessment of the role of predictable EBRs in explaining actual spreads. The results are reported in Table 9.

Table 9
Non-rational expectations and predictable returns

Note: the table presents instrumental variable regressions of log returns for the HML, SMB, CMA, RMW, and WML portfolios on expectations-based returns (EBRs) for the relevant portfolio. The set of instrumental variables includes all independent variables in Table 8. Separate regressions are estimated for one- month and one-year horizons. The last row of the table presents the R2 from the first stage regressions. Superscripts: [a] significant at the 1% level, [b] 5% level, [c] 10% level.

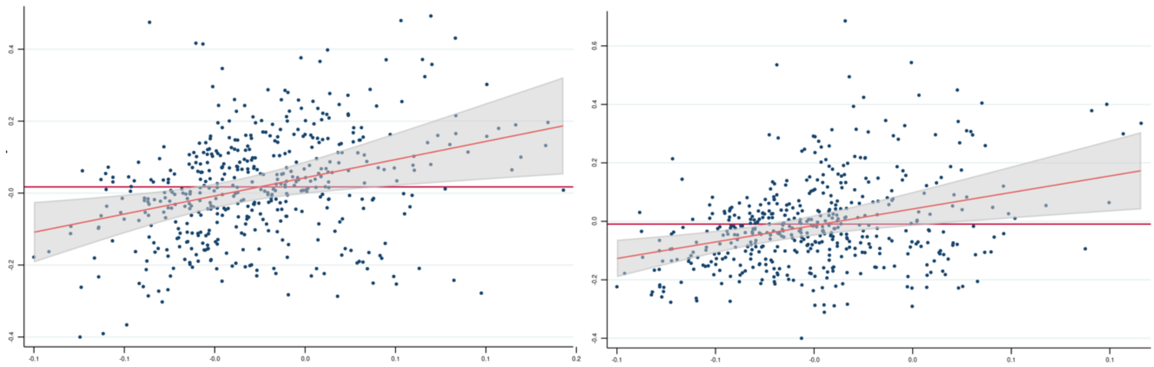
	$r_{HML,t+h}$		$r_{SMB,t+h}$		$r_{CMA,t+h}$		$r_{RMW,t+h}$		$r_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\widehat{EBR}_{LMS,t,t+h}$	1.3267[a]	1.0300[a]	1.4470[a]	0.8276[b]	0.9758[a]	1.4032[a]	0.9479	-0.2098	1.1135[b]	0.8457[a]
	(0.4452)	(0.3119)	(0.4084)	(0.3781)	(0.2722)	(0.3526)	(0.7012)	(0.2690)	(0.5053)	(0.2920)
$dp_{LMS,t}$	0.0171	0.1747[b]	0.0003	0.0763	0.0160	0.1420[b]	0.0178	-0.0176	0.0078	0.0998
	(0.0126)	(0.0853)	(0.0112)	(0.1082)	(0.0104)	(0.0580)	(0.0157)	(0.0840)	(0.0103)	(0.0778)
Constant	-0.0039	-0.0562[b]	0.0035	0.0774	-0.0065[c]	-0.1044[a]	0.0015	0.0015	-0.0397[c]	-0.1604[a]
	(0.0040)	(0.0267)	(0.0070)	(0.0714)	(0.0037)	(0.0335)	(0.0033)	(0.0284)	(0.0205)	(0.0600)
Obs	444	433	444	433	444	433	444	433	444	433
Adjusted R2	4%	16%	2%	12%	4%	25%	0%	0%	2%	19%
1[st] stage R2	14%	30%	9%	47%	14%	18%	5%	30%	7%	50%

The first stage R2 is presented in the last row of the Table (see Appendix C for the full first stage results). EBRs are strongly predicted by lagged expectations, reflecting systematic predictability of forecast revisions and errors. In particular, analyst differential optimism about long term growth of the short arm negatively predicts subsequent forecast errors and revisions, and hence high EBRs, in line with previously documented analyst overreaction to news (BGLS 2019, 2024). Turning to the second stage, Table 9 shows that actual spreads load strongly on predicted EBRs, with all coefficients strongly significant and indistinguishable from 1, except for profitability where we knew from Table 7 that EBRs do not predict returns.

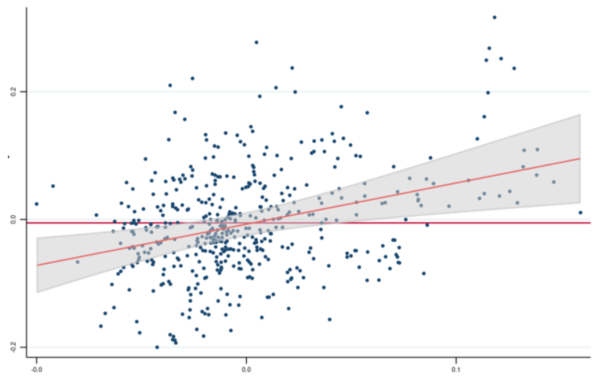
The loadings on the price dividend ratio are instead small and insignificant. Intercepts for actual return spreads are small and insignificant, and in some cases negative, confirming that EBRs, and in particular their predictable component, can account for the observed spreads in actual returns. Overall, these results show that predictable expectational errors and revisions explain not only average spreads, but also their time variation. This is consistent with the fact that future return spreads capture, at least in part, the unfolding of predictable market expectations errors, as proxied by predictable EBRs.

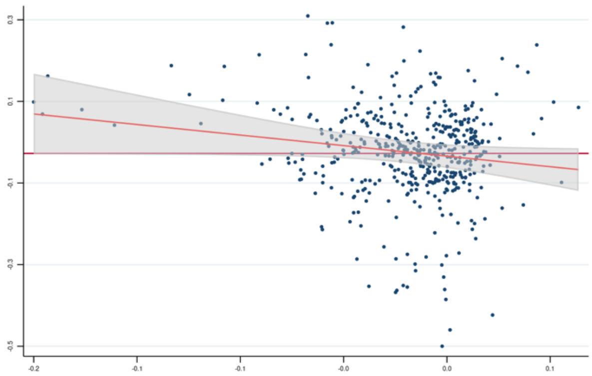
To visualize this point, Figure 3 plots the predicted 1-year return spreads $\widehat{EBR}_{LMS,t,t+h}$ (residualized of the dividend to price ratio spread dp_t) in the x-axis, against the realized spreads (also residualized of the dividend to price ratio spread dp_t) in the y-axis,

Panel A: HML **Panel B: SMB**



Panel C: CMA **Panel D: RMW**





Panel E: WML

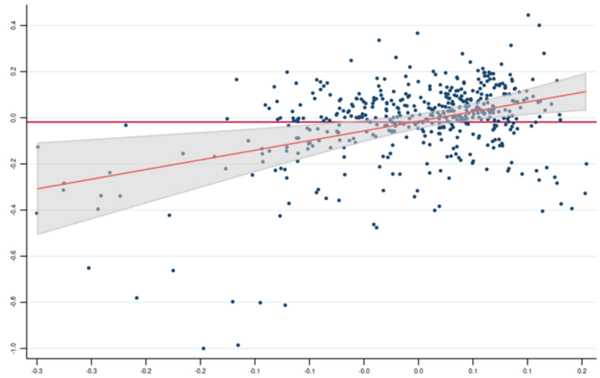


Figure 3. Predicted and realized return spreads

Note. The Figure plots one-year returns for portfolios that are long value and short growth stocks (HML), long small and short big stocks (SMB), long conservative and short aggressive investment stocks (CMA), long robust and short weak profitability stocks (RMW), and long winner and short momentum stocks (WML) against the predicted expectation-based return (EBR) for the relevant portfolio. See Table 8 for the list of variables used to generate predicted EBRs. Both portfolio returns and predicted EBRs are residualized of the log dividend to price ratio. We show 95% confidence bands around predicted EBRs. Standard errors are based on Newey-West (1987) standard errors.

Figure 3 shows that the predicted long minus short return spread covers a large range, and helps account not only for substantial variation in positive spreads but also for *negative* spreads. Large spreads systematically arise following periods when beliefs about firms in the short arm are particularly bullish, in line with a correction to overly optimistic beliefs about such firms, as in the overreaction mechanism documented by BGLS (2024) for the aggregate market. Negative spreads instead arise following periods when beliefs about aggregate growth are optimistic but not disproportionately so for firms in the short arm. These periods may be followed by increases in optimism about the short arm, as illustrated by the underperformance of value portfolios during the late 1990s and mid 2010s stock market boom (see Figure 1).

Predictable time variation in return spreads is a challenge to the view that characteristics predict returns by capturing discount rates (Fama French 1993). A growing literature seeks to explain such time variation under a decomposition of returns into shocks to discount rates or shocks to cash flow expectations, similar to Equation (4). This approach typically assumes that changes in prices or characteristics such as book to market capture shocks to discount rates, and to assign residual movement in returns to expectations (Vuolteenaho 2002, Campbell and Vuolteenaho 2004). This work has found a small role for discount rate shocks in accounting for time variation in cross-sectional spreads (Lochstoer and Tetlock 2020, Campbell et al. 2023), consistent with our finding that valuation ratios fail to predict return spreads.

Our approach additionally shows that significant predictive power comes from directly measured non-rational cash flow expectations. These expectations account for time variation due to predictable reversals of forecast errors: excess optimism about a portfolio is systematically disappointed in the future, leading to low predictable returns. This effect is at times so strong that the spread becomes predictably negative: a finding inconsistent with any model of a risk averse marginal investor. Consistent with this view, Engelberg et al (2018) shows that such cross-sectional returns accrue mainly during cash flow news events when forecast errors materialize, consistent with measured expectations capturing market expectations and being systematically surprised.

5.2 Characteristics and EBRs.

We next assess the predictability of EBRs, and hence of expectations errors, from characteristics. This test closes the circle in our investigation of the extent to which standard characteristics proxy for market inefficiency as opposed to risk.

Table 10 regresses, at the firm level, future EBRs on current characteristics. In Columns 1 through 5 we examine how firm level book to market, size, investment, profitability and momentum predict future EBRs. We do not include firm fixed effects here, because they would potentially absorb the role of the measured characteristics themselves.

Table 10
Characteristics predict firm level expectation based returns

Note: This table presents regressions of firm level log expectations-based returns (EBRs) at horizons (h) of one-month, three-months, one-year, three-years, and five-years. The independent variables include: (a) log book-to-market ($\ln bm_{i,t}$) at time t, (b) one-year growth in assets between $t-1$ and t ($Inv_{i,t}$), (c) log market value of equity at time t, (d) operating profitability at time t ($op_{i,t}$), and (e) returns between periods $t-11$ and $t-1$ ($r_{i,t-11\to t-1}$). All specifications have firm fixed effects. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$EBR_{i,t,t+1}$	$EBR_{i,t,t+3}$	$EBR_{i,t,t+12}$	$EBR_{i,t,t+36}$	$EBR_{i,t,t+60}$
	(1)	(2)	(3)	(4)	(5)
$bm_{i,t}$	0.0019[a]	0.0074[a]	0.0450[a]	0.1110[a]	0.1543[a]
	(0.0005)	(0.0024)	(0.0101)	(0.0194)	(0.0191)
$size_{i,t}$	-0.0021[a]	-0.0081[a]	-0.0364[a]	-0.0660[a]	-0.0704[a]
	(0.0002)	(0.0011)	(0.0031)	(0.0056)	(0.0093)
$Inv_{i,t}$	-0.0018[a]	-0.0047[a]	-0.0050	0.0055	0.0130
	(0.0003)	(0.0010)	(0.0045)	(0.0101)	(0.0116)
$op_{i,t}$	0.0090[a]	0.0135[a]	-0.0708[a]	-0.2604[a]	-0.3516[a]
	(0.0008)	(0.0029)	(0.0149)	(0.0551)	(0.0795)
$r_{i,t-12\to t-1}$	0.0136[a]	0.0387[a]	0.0652[a]	0.0242[b]	0.0206[b]
	(0.0007)	(0.0030)	(0.0098)	(0.0110)	(0.0102)
Obs	875,404	815,293	772,240	594,850	474,100
Adj R^2	0%	2%	2%	7%	13%

On their own, characteristics have strong and highly significant predictive power for EBRs (columns 1 to 5). Low book to market, high investment and low past returns predict subsequent disappointment and low EBRs at all horizons, consistent with the average spread of the corresponding factors.[21] Interestingly, large firms have higher short term EBRs but lower EBRs at horizons of one year and longer. Of all characteristics, only profitability does not reliably predict EBRs once other characteristics are controlled for.

[21] These results are consistent with recent work linking characteristics and expectations data. Frey (2023) examines a large number of factors and finds that short term growth expectations between the long and short arm to converge. Gormsen and Lazarus (2023) find that characteristics associated with the short arm of factors, such as low book to market, high investment, low profitability, high beta and low payout, predict high LTG.

In Appendix D, we further examine the link between characteristics and market inefficiency. In a mediation exercise (MacKinnon 2012) we estimate the share of return predictability from characteristics that works through their ability to predict analyst expectations (versus the share that works through their direct predictive ability after controlling for EBRs). The exercise shows that, to a large extent, characteristics predict returns precisely because they capture distorted expectations (Table D1). Finally, following BGLS (2024), we assess the extent to which firm-level return predictability from the book to market ratio is accounted for by future expectation errors and revisions, and find that, particularly at longer horizons, the strong predictability of $bm_{i,t}$ is entirely captured by expectations (Table D2).

6.Taking stock

We started the paper with a simple question: does understanding the cross-section of stock returns need exotic risk factors, first introduced by Fama and French (1993)? The evidence we presented says no. Rather, the risk premia identified by Fama and French appear to reflect corrections of measurable expectations errors about earnings growth. Relaxing the assumption of rational expectations allows us to use the classical dividend discount model and observed expectations of future cash flows to account for the cross-sectional evidence on stock returns. We view this result as a victory for financial economics, because it shows that we do not need exotic risk factors to explain the data.

Our evidence shows that spreads are generated because expectations about future growth of firms in the short arm of the portfolios are systematically too optimistic, and those about firms in the long arm too pessimistic, so that the long portfolio outperforms the short one as expectation errors are corrected in the future. Characteristics such as book to market or investment predict returns at least in part because they predict differential optimism and forecast errors. Notably, the same mechanism helps account for momentum. Predictability from other characteristics may also work through expectations. This, of course, has significant implications not just for cross sectional (and time series) asset pricing, but also for firm investment policies, financial policies, and other decisions. We conclude by highlighting two follow-up questions.

The first question concerns the structure of expectations. Analyst beliefs can reproduce return co-movements across firms sharing similar characteristics because expectations themselves comove within groups of firms identified by those characteristics. Where does such co-movement come from, and why does it lead firms with certain characteristics to be over-priced? One possibility is that co-movement reflects the non-rational reaction of beliefs to common shocks hitting groups of firms or sectors. Another possibility is that co-movement in beliefs reflects spurious similarity of firms to their peers (Sarkar 2024). Understanding the structure of expectations may also shed light on the evidence that idiosyncratic risk is priced (Campbell et al 2001), because such firm-specific return differentials may also reflect time varying optimism about firm growth rather than compensation for firm specific risk.

The second question concerns the required rate of return that the dividend discount model relies on. What are its properties and determinants? In standard theory one component is the risk premium, which depends on the curvature of the utility of wealth and the quantity of risk, another component is interest rates, which are determined by time preference and technology. Yet, a large body of work using experimental and field data, including applications to the stock market (Benartzi and Thaler 1995, Barberis 2018), shows that risk attitudes depend on factors other than the marginal utility of wealth. It is also well known that interest rates themselves are highly volatile (Shiller 1980, Singleton 1980, Giglio and Kelly 2018). Psychology may also help us understand where the required return comes from.

REFERENCES

Adam, Klaus, and Stefan Nagel, 2023, Expectations data in asset pricing. In *Handbook of Economic Expectations*, Academic Press, 477-506.

Banz, Rolf, 1981, The relationship between return and market value of common stocks. *Journal of Financial Economics* 9 (1): 3-18.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *Journal of Finance* 32 (3): 663-682.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of Financial Economics* 12 (1): 129-156.

Barberis, Nicholas, 2018, Psychology-based models of asset prices and trading volume. *Handbook of Behavioral Economics: Applications and Foundations,* North-Holland 1 (1): 79-175.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics* 115 (1): 1-24.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment. *Journal of Financial Economics* 49 (3): 307-343.

Benartzi, Shlomo, and Richard Thaler, 1995, Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics* 110 (1): 73-92.

Ben-David, Itzhak and Alexandre Chinco, 2024, Expected EPS x Trailing P/E, working paper.

Bianchi, Francesco, Sydney Ludvigson, and Sai Ma, 2024, What Hundreds of Economic News Events Say About Belief Overreaction in the Stock Market. NBER w32301.

van Binsbergen, Jules and Ralph Koijen, 2017, The term structure of returns: facts and theory. *Journal of Financial Economics* 124 (1): 1-21.

van Binsbergen, Jules, Martijn Boons, Christian Opp, and Andrea Tamoni, 2023, Dynamic asset (mis) pricing: Build-up versus resolution anomalies. *Journal of Financial Economics* 147 (2): 406-431.

De Bondt, Werner, and Richard Thaler, 1985, Does the stock market overreact? *Journal of finance* 40 (3): 793-805.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2019, Diagnostic expectations and stock returns. *Journal of Finance* 74 (6): 2839-2874.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2024, Belief overreaction and stock market puzzles. *Journal of Political Economy* 132 (5): 1450-1484.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, Matthew O'Brien, and Andrei Shleifer, 2024, Long term expectations and aggregate fluctuations. *NBER Macroeconomics Annual*, volume 38, University of Chicago Press.

Bradshaw, Mark, 2004, How do Analysts Use Their Earnings Forecasts in Generating Stock Recommendations? *Accounting Review* 79 (1): 25–50.

Campbell, John, Stefano Giglio, and Christopher Polk, 2023, What drives booms and busts in value? NBER w31859.

Campbell, John, Martin Lettau, Burton Malkiel, and Yexiao Xu, 2001, Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56 (1): 1-43.

Campbell, John and Jianping Mei, 1993, Where do betas come from? Asset price dynamics and the sources of systematic risk. *Review of Financial Studies* 6 (3): 567-592.

Campbell, John, and Robert Shiller, 1987, Cointegration and tests of present value models, *Journal of Political Economy* 95 (5): 1062-1088.

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

Campbell, John, and Tuomo Vuolteenaho, 2004, Bad beta, good beta. *American Economic Review* 94 (5): 1249-1275.

Chaudhry, Aditya, 2023, The impact of prices on analyst cash flow expectations: Reconciling subjective beliefs data with rational discount rate variation." *Available at SSRN* 4443349.

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

De la O, Ricardo, and Sean Myers, 2021, Subjective cash flow and discount rate expectations. *Journal of Finance* 76 (3): 1339-1387.

De la O, Ricardo, and Sean Myers, 2024, Which Subjective Expectations Explain Asset Prices? *Review of Financial Studies*, forthcoming.

De la O, Ricardo, Xiao Han, and Sean Myers, 2024, The return of return dominance: Decomposing the cross section of prices. *Available at SSRN* 4206091.

Décaire, Paul, and John Graham, 2024, Valuation Fundamentals. Working paper

Driscoll, John and Aart Kraay, 1998, Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80 (4): 549-560.

Engelberg, Joseph, David McLean, and Jeffrey Pontiff, 2018, Anomalies and news. *Journal of Finance* 73 (5): 1971-2001.

Fama, Eugene, 1970, Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25 (2): 383–417.

Fama, Eugene, and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

Fama, Eugene, and Kenneth French, 2015, A five-factor asset pricing model. *Journal of Financial Economics,* 116 (1): 1-22.

Frey, Jonas, 2023, Which stock return predictors reflect mispricing? *working paper,* University of Oxford.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer, 2016, Expectations and investment. *NBER Macroeconomics Annual* 30 (1): 379-431.

Giglio, Stefano, and Bryan Kelly, 2018, Excess volatility: Beyond discount rates. *Quarterly Journal of Economics* 133 (1): 71-127.

Gormsen, Niels, 2021, Time variation of the equity term structure. *Journal of Finance* 76 (4): 1959-1999.

Gormsen, Niels, and Eben Lazarus, 2023, Duration-driven returns. *Journal of Finance,* 78 (3): 1393-1447.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns. *Review of Financial Studies* 27 (3): 714-746.

Harvey, Campbell, Yan Liu, and Heqing Zhu, 2016, … And the cross-section of expected returns. *Review of Financial Studies* 29 (1): 5-68.

Hong, Harrison, and Jeremy Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54 (6): 2143-2184.

Jiang, Zhengyang, Hongqi Liu, Cameron Peng, and Hongjun Yan, 2022, Investor memory and biased beliefs: Evidence from the field. *Working paper*.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of finance* 48 (1): 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2011, Momentum. *Annual Review of Financial Economics* 3 (1): 493 – 509.

Jung, Boochun, Philip Shane, and Yanhua Sunny Yang, 2012, Do financial analysts' long-term growth forecasts matter? Evidence from stock recommendations and career outcomes. *Journal of Accounting and Economics* 53 (1,2): 55-76.

Kleibergen, Frank, and Richard Paap, 2006, Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* 133 (1): 97-126.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models. *Journal of Finance* 73 (3): 1183-1223.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation, and risk. *Journal of Finance* 49 (5): 1541-1578.

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

La Porta, Rafael La, Josef Lakonishok, Andrei Shleifer, and Robert Vishny, 1997, Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52 (2): 859-874.

Lettau, Martin, and Jessica Wachter, 2007, Why is long-horizon equity less risky? A duration-based explanation of the value premium. *Journal of Finance* 62 (1): 55-92.

Lochstoer, Lars, and Paul Tetlock, 2020, What drives anomaly returns? *Journal of Finance* 75 (3): 1417-1455.

Ma, Yueran, 2019, Nonfinancial firms as cross-market arbitrageurs. *Journal of Finance* 74 (6): 3041-3087.

MacKinnon, David, 2012, Introduction to statistical mediation analysis. *Routledge*.

Nagel, Stefan, 2024, Stock Market Valuation: Explanations, Non-Explanations, and Some Open Questions. *Comments given at NBER SI Asset Pricing Meeting.*

Nagel, Stefan, and Xu, Zhengyang, 2022, Asset pricing with fading memory. *Review of Financial Studies* 35 (5): 2190-2245.

Newey, Whitney and Kenneth West, 1987, Hypothesis testing with efficient method of moments estimation. *International Economic Review* 28 (3): 777-787.

Olea, José Luis Montiel, and Carolin Pflueger, 2013, A robust test for weak instruments. *Journal of Business & Economic Statistics* 31 (3): 358-369.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11 (3): 9-16.

Sarkar, Suproteem, 2024, Semantic representations and economic beliefs. *Working paper* Harvard University.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of finance* 19 (3): 425-442.

Shiller, Robert, 1981, Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review, Papers & Proceedings*, 71 (3): 421-436.

Singleton, Kenneth, 1980, Expectations models of the term structure and implied variance bounds. *Journal of Political Economy* 88 (6): 1159-1176.

Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns? *Journal of Finance* 57 (1): 233-264.

Appendix

Appendix A: Adjusting for measurement noise in analysts' expectations

Here we present a method for recovering the expected return differential $r_L - r_S$ from Equation (8), in particular from the corresponding regression constant κ, the slope γ, and other known moments, under two assumptions: first, that the true return for firm i is equal to:

$$r_{i,t+1} = r_i + \left[g_{i,t+1} - \tilde{\mathbb{E}}_t\left(g_{i,t+1}\right)\right] + \sum_{s\geq 1} \alpha^s \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\left(g_{i,t+1+s}\right).$$

where $\tilde{\mathbb{E}}_{it}$ are market expectations about firm i, and second, that measured (analysts') expectations $\tilde{\mathbb{E}}^m_{ti}$ relate to $\tilde{\mathbb{E}}_{it}$ as:

$$\tilde{\mathbb{E}}^m_{ti} = \beta + \tau \cdot \tilde{\mathbb{E}}_{it} + \sigma \cdot \varepsilon_{it}.$$

The revision in measured expectations between t and $t+1$ is given by:

$$\tilde{\mathbb{E}}^m_{t+1,i} - \tilde{\mathbb{E}}^m_{t,i} = \tau \cdot \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right) + \sigma \cdot \left(\varepsilon_{it+1} - \varepsilon_{it}\right).$$

The measured forecast error at $t+1$ is equal to:

$$g_{i,t+1} - \tilde{\mathbb{E}}^m_{i,t} = g_{i,t+1} - \beta - \tau \cdot \tilde{\mathbb{E}}_{it} - \sigma \cdot \varepsilon_{i,t}.$$

So the measured EBR for firm i is equal to:

$$\begin{aligned}\mathrm{EBR}_{i,t+1} = r &+ \left[g_{i,t+1} - \beta - \tau \cdot \tilde{\mathbb{E}}_{i,t}\left(g_{i,t+1}\right) - \sigma \cdot \varepsilon_{it}\right] \\ &+ \sum_{s=1,\ldots,5} \alpha^s \left[\tau \cdot \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\left(g_{i,t+1+s}\right) + \sigma \cdot \left(\varepsilon_{it+1} - \varepsilon_{it}\right)\right]\end{aligned}$$

We can then aggregate at the level of portfolio p:

$$\begin{aligned}r_{p,t+1} &= r_p + \left[g_{p,t+1} - \tilde{\mathbb{E}}_t\left(g_{p,t+1}\right)\right] + \sum_{s\geq 1} \alpha^s \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\left(g_{p,t+1+s}\right) \\ &= \left(r_p - r\right) + \mathrm{EBR}_{p,t+1}.\end{aligned}$$

Where $\mathrm{EBR}_{p,t+1}$ are true EBRs. The measured EBR for firm i is equal to:

$$\begin{aligned}\mathrm{EBR}^m_{p,t+1} &= r + g_{p,t+1} - \beta - \tau \cdot \tilde{\mathbb{E}}_{pt} - \sigma \cdot \varepsilon_{p,t} \\ &+ \sum_{s\geq 1} \alpha^s \left(\tau \cdot \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\right)\left(g_{p,t+1+s}\right) + \alpha^s \sigma \cdot \left(\varepsilon_{pt+1} - \varepsilon_{pt}\right) \\ &= r - \beta + (1-\tau) g_{p,t+1} + \tau\left[g_{p,t+1} - \tilde{\mathbb{E}}_{pt}\left(g_{p,t+1}\right)\right] - \sigma \cdot \varepsilon_{p,t} \\ &+ \tau \sum_{s\geq 1} \alpha^s \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\left(g_{p,t+1+s}\right) + \alpha^s \sigma \cdot \left(\varepsilon_{it+1} - \varepsilon_{it}\right) \\ &= r - \beta + (1-\tau) g_{p,t+1} + \tau \mathrm{EBR}_{p,t+1} - \sigma \cdot \varepsilon_{p,t} + + \alpha^s \sigma \cdot \left(\varepsilon_{it+1} - \varepsilon_{it}\right)\end{aligned}$$

Drop for notational simplicity the portfolio label except for r_p that denotes the portfolio required return. We can then redefine these equations as

$$r_{t+1} = r_p + v_{t+1}$$

where $v_{t+1} = \text{EBR}_{p,t+1} - r$, and

$$\text{EBR}_{t+1}^{m} = r - \beta + (1-\tau)g_{p,t+1} + \tau \cdot v_{t+1} + k_1 \cdot \varepsilon_t + k_2 \cdot \varepsilon_{t+1},$$

Note that v is not EBR as defined in the paper, it is just the expectations component (without the aggregate required return r).

Next, the estimated portfolio regression gives (denoting sample averages by upper bars)

$$\overline{r} = \kappa + \gamma \cdot \overline{\text{EBR}^{m}},$$

And we also know $\overline{r} = r_p + \overline{v}$, which implies:

$$r_p = \kappa + \gamma \cdot \overline{\text{EBR}^{m}} - \overline{v}$$

If we know $\overline{v}$, we get to know the required return r_p, which is the first ingredient for computing the spread we are looking for. To find $\overline{v}$, first note that:

$$\overline{\text{EBR}^{m}} = r - \beta + (1-\tau)\overline{g} + \tau \cdot \overline{v}$$

Thus, because r, $\overline{g}$ and $\overline{\text{EBR}^{m}}$ are known, if we find out τ and β then we can backup $\overline{v}$. Assume that market expectations are on average unbiased. This implies that *measured* forecast errors, which satisfy:

$$\left[g_{t+1} - \widetilde{\mathbb{E}}_t^{m}(g_{t+1})\right] = g_{t+1} - \beta - \tau \cdot \widetilde{\mathbb{E}}_t - \sigma \cdot \varepsilon_t,$$

on average obey:

$$\overline{\left[g_{t+1} - \widetilde{\mathbb{E}}_t^{m}(g_{t+1})\right]} = \overline{g} - \beta - \tau \cdot \overline{g},$$

So that:

$$\beta = (1-\tau) \cdot \overline{g} - \overline{e},$$

where $\overline{e}$ is the average forecast error. We can then plug this expression for beta in the previous equation and obtain:

$$\overline{\text{EBR}^{m}} = r + \overline{e} + \tau \cdot \overline{v}$$

To recover τ note that:

$$cov(\text{EBR}^{m}, g) = (1-\tau) \cdot var(g) + \tau \cdot cov(v, g)$$

where we know $cov(\text{EBR}^{\text{m}}, g)$ and $var(g)$ but not $cov(v, g)$. Under our assumptions on the true required returns, the latter can be obtained from:

$$cov(v, g) = cov(r_{t+1}, g).$$

This yields:

$$\tau = \frac{cov(\text{EBR}^m, g) - var(g)}{cov(r_{t+1}, g) - var(g)}$$

This allows us to back up $\overline{v}$ from $\overline{\text{EBR}^m}$, and thus the required return spread:

$$r_L - r_S = \kappa_L - \kappa_S + \overline{\text{EBR}_L^m - \text{EBR}_S^m}\left[\gamma - \frac{cov(r_{t+1}, g) - var(g)}{cov(\text{EBR}^m, g) - var(g)}\right]$$
$$+ \frac{cov(r_{t+1}, g) - cov(\text{EBR}^m, g)}{cov(\text{EBR}^m, g) - var(g)}\left(\overline{g}_L - \overline{g}_S\right).$$

Appendix B: Expectation Based Returns (EBRs), assumptions and robustness

B.1 Samples and robustness.

In this Section, we present descriptive statistics, assess the robustness of our construction of EBRs to alternative choices regarding sample selection, assumptions about payout ratios, and value weighting portfolio EBRs. Table B.1 compares our main sample, defined in Section 2.2, to the full CRSP / COMPUSTAT sample.

Table B1.
Sample descriptive statistics

Note: The table presents descriptive statistics for the sample of firms that meet the CRSP/COMPUSTAT data requirements of Fama and French (1992) and the subsample with an expectation-based return (EBR) for July of year t (*our sample*). We report sample means for: (1) the book-to-market ratio, (2) the one-year change in total assets in fiscal year *t-1* divided by t-2 total assets (*investment*), (3) annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses divided by book equity (*operating profitability*), (4) the return between *t-11* and *t*, (5) the ratio of the 3-year moving average of earnings per share to price, (6) , (7) the dividend to price ratio, and (8) the average number of observations. For each variable, we report sample mans for the two samples, their difference, the standard error for the difference, the z-stat, and the significance of the difference of the two means.

	Sample mean		Difference	Standard error	z-stat	p
	All CRSP/ COMPUTAT	Our Sample				
book-to-market	0.6365	0.6109	0.0256	0.0038	6.72	0.0%
investment	0.1905	0.1918	-0.0013	0.0020	-0.67	50.3%
operating profitability	0.3320	0.3256	0.0064	0.0378	0.17	86.6%
market capitalization (mill)	7,256	7,842	-586	65	-8.99	0.0%
Lagged one-year return	0.1527	0.1687	-0.0160	0.0020	-8.12	0.0%
earnings-to-price ratio	0.0202	0.0452	-0.0250	0.0035	-7.06	0.0%
cape ratio	0.0300	0.0461	-0.0161	0.0024	-6.63	0.0%
dividend-to-price ratio	0.0159	0.0165	-0.0006	0.0001	-6.87	0.0%
observations	2,162	1,786	375.8	34.8	10.80	0.0%

Next, we show that the results of Tables 1 and 2 on how portfolio level EBRs compare to actual returns are qualitatively similar under an alternative payout definition. We use the expected payout ratio implied from analysts' expectations of dividends and earnings, namely $\widetilde{\mathbb{E}}_t DPS_{i,t+12}/\widetilde{\mathbb{E}}_t EPS_{i,t+12}$ for observations with available expectations of dividends. When this data is not available, we set the expected payout ratio to zero if the firm did not pay a dividend the previous year. If the firm did pay a dividend the previous year, we set the payout ratio to the average expected payout ratio $\frac{\widetilde{\mathbb{E}}_t DPS_{i,t+12}}{\widetilde{\mathbb{E}}_t EPS_{i,t+12}}$ in our sample for those firms which paid dividends that year, $DPS_{i,t+12} > 0$, which equals 0.41. In the text we assumed a constant payout ratio, which has a correlation with the above specification of over 97%.

Table B2.
Expectation Based Returns with Alternative Measures of Payout Ratio

Note: Panel A presents sample means of log expectation-based returns (EBR) returns for portfolios of stocks formed on book-to-market, investment, size, profitability and momentum over holding horizons ranging from one month to five years. EBRs are computed following Equation (10) in the text with the payout ratio set to: (1) the ratio of expected dividends-to-earnings in period $t+12$ (i.e. $\widetilde{\mathbb{E}}_t DPS_{i,t+12}/\widetilde{\mathbb{E}}_t EPS_{i,t+12}$) for those observations where expectations of dividends are available, or (2) zero if the firm did not pay a dividend the previous year and $\widetilde{\mathbb{E}}_t DPS_{i,t+12}$ is unavailable, or (3) the average expected payout ratio $\frac{\widetilde{\mathbb{E}}_t DPS_{i,t+12}}{\widetilde{\mathbb{E}}_t EPS_{i,t+12}}$ in our sample for those firms which paid dividends in the previous year but do not have $\widetilde{\mathbb{E}}_t DPS_{i,t+12}$. Portfolio EBRs are equally weighted with monthly rebalancing. Panel B shows pairwise correlations between log returns and EBR for portfolios of stocks formed on book-to-market, investment, size, profitability and momentum sorts over holding horizons ranging from one month to five years. The sample period extends from December 1981 to December 2023.

Panel A: EBRs

Holding Horizon	Growth	Value	Aggressive	Conservative	Small	Big	Weak	Robust	Losers	Winners
1 Month	10.7%	13.5%	15.0%	7.4%	10.2%	11.5%	15.2%	9.7%	-16.0%	33.0%
3 Months	9.9%	13.0%	14.6%	6.7%	9.9%	11.1%	14.4%	9.4%	-13.4%	30.0%
1 Year	9.2%	13.7%	14.4%	6.9%	11.6%	10.3%	13.9%	9.6%	-0.8%	20.2%
3 Years	9.4%	13.1%	13.0%	8.2%	11.3%	10.3%	13.1%	10.0%	7.9%	12.9%
5 Years	9.6%	12.5%	12.6%	8.9%	10.9%	10.2%	13.0%	9.9%	9.2%	11.8%

Panel B: Correlation between returns and EBRs

Holding Horizon	Growth	Value	Aggressive	Conservative	Small	Big	Weak	Robust	Losers	Winners
1 Month	7%	18%	14%	9%	14%	9%	11%	12%	11%	5%
3 Months	22%	34%	30%	24%	29%	23%	28%	21%	27%	25%
1 Year	35%	50%	36%	40%	45%	33%	43%	32%	46%	42%
3 Years	43%	49%	35%	51%	60%	30%	46%	33%	52%	49%

5 Years	34%	37%	25%	40%	47%	16%	38%	23%	35%	32%

Next, we show our baseline analysis in Tables 3 and 6 is robust to value weighting portfolio EBRs.

Table B3.
Value weighted EBRs and Returns

Note: This table presents univariate regression results of log value-weighted returns for portfolios of stocks formed on book-to-market and size (Panel A), investment and profitability (Panel B), and momentum (Panel C) over holding horizons ranging from one month to five years. The independent variable is the value-weighted expectation-based returns (EBR) for the relevant portfolio. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period is December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A: book to market and size

	Book-to-market (HML)					Size (Small minus Big)				
	1 Mo	3 Mo	1 Yr	3 Yr	5 Yr	1 Mo	3 Mo	1 Yr	3 Yr	5 Yr
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
EBR_{LMS}	0.2398[b]	0.5377[a]	0.8245[a]	1.0648[a]	1.0778[a]	0.6007[a]	0.9965[a]	1.1643[a]	1.2857[a]	1.1787[a]
	(0.0991)	(0.1401)	(0.1314)	(0.1598)	(0.1757)	(0.1438)	(0.1868)	(0.1879)	(0.2032)	(0.2162)
Constant	0.0012	0.0028	0.0025	-0.0131	-0.0106	0.0049[b]	0.0098[c]	-0.0020	-0.0615	-0.0677
	(0.0020)	(0.0051)	(0.0157)	(0.0368)	(0.0469)	(0.0019)	(0.0051)	(0.0193)	(0.0423)	(0.0516)
Obs	504	502	493	469	445	504	502	493	469	445
Adj R^2	2%	11%	37%	56%	59%	6%	19%	38%	54%	49%

Panel B: profitability and investment

	Profitability (Robust minus Weak)					Investment (CMA)				
	1 Mo	3 Mo	1 Yr	3 Yr	5 Yr	1 Mo	3 Mo	1 Yr	3 Yr	5 Yr
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
EBR_{LMS}	0.0645	0.3204[a]	0.5553[a]	0.5242[a]	0.5076[a]	0.3024[a]	0.4724[a]	0.8755[a]	0.9631[a]	0.9690[a]
	(0.0806)	(0.0898)	(0.1197)	(0.1159)	(0.0944)	(0.0711)	(0.0852)	(0.1293)	(0.1317)	(0.0974)
Constant	0.0016	0.0080[b]	0.0417[a]	0.0654[a]	0.0603	0.0018	0.0049	0.0071	-0.0022	-0.0061
	(0.0016)	(0.0038)	(0.0120)	(0.0235)	(0.0367)	(0.0014)	(0.0038)	(0.0126)	(0.0269)	(0.0336)
Obs	504	502	493	469	445	504	502	493	469	445
Adj R^2	0%	6%	23%	33%	40%	4%	11%	33%	52%	56%

Panel C: momentum

	1 Mo	3 Mo	1 Yr	3 Yr	5 Yr
	(1)	(2)	(3)	(4)	(5)
EBR_{LMS}	0.1561	0.5484[a]	0.5909[a]	0.7103[a]	0.5882[a]
	(0.1035)	(0.0979)	(0.0914)	(0.1351)	(0.1518)
Constant	-0.0034	-0.0440[a]	-0.1040[a]	-0.1347[a]	-0.1151[b]
	(0.0043)	(0.0111)	(0.0230)	(0.0446)	(0.0535)
Obs	504	502	493	469	445
Adj R^2	1%	11%	21%	37%	25%

B2. EBR spreads and realized return spreads

We reproduce the EBR decomposition analysis of Table 5 for all factors. We further control for the spread in portfolios' price dividend ratios as a means to control for spreads in discount rates incorporated in lagged valuation ratios.

Table B4.
Portfolio level forecast errors and revisions predict spreads

Note: This table presents univariate regression results of log value-weighted returns for portfolios of stocks formed on book-to-market and size (Panel A), investment, profitability and momentum (Panel B) over holding horizons ranging from one year to five years. The independent variables include the long-short: (a) forecast error between t and $t+h$, (b) forecast revisions between t and $t+h$ of one-year earnings growth in year $t+h+1$, and (c) changes in long-term growth forecasts between t and $t+h$, and (d) log dividend-to-price ratio. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A: Book to market and Size

	Book-to-Market (HML)			Size (SMB)		
	$r_{HML,t+12}$	$r_{HML,t+36}$	$r_{HML,t+60}$	$r_{HML,t+12}$	$r_{HML,t+36}$	$r_{HML,t+60}$
$(1-E_t)\ \Delta e_{LMS,t+h}$	0.1302[a]	0.1745[a]	0.1351[a]	0.1035[a]	0.1295[a]	0.0381
	(0.0136)	(0.0267)	(0.0376)	(0.0204)	(0.0367)	(0.0486)
$(E_{t+h}-E_t)\ \Delta e_{LMS,t+h+1}$	0.0800[a]	0.0572[b]	-0.0249	0.0334[c]	0.0109	-0.0340
	(0.0128)	(0.0247)	(0.0284)	(0.0185)	(0.0262)	(0.0351)
$\Delta_h LTG_{LMS,t+h}$	0.0278[a]	0.0197	0.0486	-0.0280[b]	0.0175	0.0799[b]
	(0.0094)	(0.0212)	(0.0345)	(0.0139)	(0.0291)	(0.0351)
$\ln(dp_{LMS,t})$	0.0030	-0.0040	0.0500	0.0252[c]	0.0239	0.0662
	(0.0106)	(0.0263)	(0.0351)	(0.0148)	(0.0342)	(0.0485)
Constant	-0.1149[a]	-0.2827[a]	-0.3671[a]	-0.1410[a]	-0.1896[c]	-0.2954[c]
	(0.0259)	(0.0661)	(0.0916)	(0.0442)	(0.1069)	(0.1545)
Obs	493	469	445	493	469	445
Adj R^2	53%	48%	48%	34%	40%	26%

Panel B: Investment, Profitability, Momentum

	Investment (CMA)			Profitability (RMW)			Momentum (WML)		
	$r_{HML,t+12}$	$r_{HML,t+36}$	$r_{HML,t+60}$	$r_{HML,t+12}$	$r_{HML,t+36}$	$r_{HML,t+60}$	$r_{HML,t+12}$	$r_{HML,t+36}$	$r_{HML,t+60}$
$(1-E_t)\ \Delta e_{HML,t+h}$	0.0649[a]	0.0737[a]	0.1023[a]	0.0489[a]	0.0652[a]	0.0143	0.1325[a]	0.1796[a]	0.0706[b]
	(0.0089)	(0.0126)	(0.0199)	(0.0155)	(0.0173)	(0.0167)	(0.0269)	(0.0292)	(0.0342)
$(E_{t+h}-E_t)\ \Delta e_{HML,t+h+1}$	0.0497[a]	0.0400[a]	0.0333[b]	0.0225	0.0075	0.0215	0.0833[a]	0.0668[a]	0.0239
	(0.0081)	(0.0123)	(0.0134)	(0.0159)	(0.0173)	(0.0164)	(0.0228)	(0.0236)	(0.0257)

$\Delta_h LTG_{HML,t+h}$	0.0149[b]	0.0428[a]	-0.0207	-0.0201	-0.0135	-0.0095	0.0187	0.0428[c]	0.0012
	(0.0065)	(0.0137)	(0.0201)	(0.0127)	(0.0147)	(0.0217)	(0.0176)	(0.0249)	(0.0393)
$\ln(dp_{HML,t})$	0.0128[c]	-0.0131	0.0198	-0.0127	-0.0344[b]	-0.0099	0.0150	-0.0325	0.0602
	(0.0067)	(0.0135)	(0.0185)	(0.0132)	(0.0164)	(0.0255)	(0.0164)	(0.0237)	(0.0371)
Constant	0.0225	0.0451	0.0524	0.0326	0.0439	0.0594	-0.1582[a]	-0.0900[b]	0.0145
	(0.0157)	(0.0319)	(0.0515)	(0.0245)	(0.0274)	(0.0460)	(0.0377)	(0.0352)	(0.0478)
Obs	493	469	445	493	469	445	493	469	445
Adj R^2	46%	47%	39%	11%	20%	4%	30%	46%	16%

Section 2.5 discusses the hypothesis that analyst expectations may surreptitiously capture discount rates that may be incorporated in prices. Here we show, following BGLS (2024), that changes in measured beliefs respond to realized fundamentals, and do not appear to mechanically respond to prices. Table B5 regresses portfolio level EBRs on *contemporaneous* portfolio level returns (which would drive the results if analysts mechanically infer forecasts from prices), as well as on contemporaneous cash flow news.

Table B5.

Expectation based portfolio returns and contemporaneous news

Note: The table presents portfolio-level regressions of expectation-based returns (EBR) at horizons (h) of one- and 3-months, and one-, three- and five-years. In each panel, the independent variables in the first regression are: (a) log returns between t and $t + h$. (b) earnings growth between t and $t + h$, and (c) the forecast error for earnings growth between t and $t + h$. The independent variable in the second regression is (a) log returns between t and $t + h$. Panel A presents results for book to market and size, Panel B for investment and profitability, and Panel C for momentum. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: a significant at the 1% level, b significant at the 5% level, c significant at the 10% level.

Panel A: book to market and size

	$EBR_{HML,t+h}$					$EBR_{SMB,t+h}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1 Mo	3 Mo	1 yr	3 yrs	5 yrs	1 Mo	3 Mo	1 yr	3 yrs	5 yrs
$r_{LMS,t+h}$	0.6313[a]	0.4407[a]	0.2819[a]	0.2631[a]	0.2765[a]	0.5551[a]	0.4168[a]	0.1457[a]	0.2315[a]	0.3035[a]
	(0.1111)	(0.0753)	(0.0492)	(0.0577)	(0.0637)	(0.0901)	(0.0646)	(0.0349)	(0.0534)	(0.0690)
$\Delta e_{HML,t+h}$	1.7713[a]	0.8254[a]	0.2801[a]	0.2021[a]	0.2096[b]	1.2244[a]	0.5266[a]	0.0126	-0.0411	-0.2125
	(0.3345)	(0.1355)	(0.0686)	(0.0767)	(0.0912)	(0.3545)	(0.1452)	(0.0685)	(0.1020)	(0.1933)
$(1 - E_t)$ $\Delta_h e_{LMS,t+h}$			0.1342[b]	0.5288[c]	0.8856[c]			0.3297[a]	1.6494[a]	3.4592[a]
			(0.0639)	(0.2761)	(0.5174)			(0.0535)	(0.2744)	(0.8469)
Constant	0.0338[a]	0.0266[a]	-0.0097	0.0085	0.0126	0.0135	0.0083	0.0355[a]	0.0914[a]	0.1208[b]
	(0.0106)	(0.0094)	(0.0131)	(0.0250)	(0.0326)	(0.0098)	(0.0087)	(0.0098)	(0.0296)	(0.0613)

Obs	493	493	493	469	445	493	493	493	469	445
Adj R^2	24%	36%	57%	56%	63%	17%	27%	58%	76%	70%

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1 Mo	3 Mo	1 yr	3 yrs	5 yrs	1 Mo	3 Mo	1 yr	3 yrs	5 yrs
$r_{LMS,t+h}$	0.7452[a]	0.5792[a]	0.4476[a]	0.1523[a]	0.4260[a]	0.6280[a]	0.4788[a]	0.2982[a]	0.1954[a]	0.4516[a]
	(0.1312)	(0.0905)	(0.0498)	(0.0530)	(0.0665)	(0.0982)	(0.0694)	(0.0426)	(0.0354)	(0.1002)
Constant	0.0418[a]	0.0379[a]	0.0270[a]	0.0279[b]	0.0817[a]	0.0127	0.0103	0.0041	-0.0036	-0.0138
	(0.0121)	(0.0108)	(0.0086)	(0.0120)	(0.0195)	(0.0104)	(0.0093)	(0.0080)	(0.0097)	(0.0377)
Obs	493	493	493	469	445	493	493	493	469	445
Adj R^2	8%	18%	46%	12%	50%	9%	20%	34%	34%	49%

Panel B: investment and profitability

	$EBR_{CMA,t+h}$					$EBR_{RMW,t+h}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1 Mo	3 Mo	1 yr	3 yrs	5 yrs	1 Mo	3 Mo	1 yr	3 yrs	5 yrs
$r_{LMS,t+h}$	0.6677[a]	0.5822[a]	0.3262[a]	0.1717[b]	0.3588[a]	0.6522[a]	0.5484[a]	0.2675[a]	0.4881[a]	0.5255[a]
	(0.1323)	(0.0955)	(0.0528)	(0.0707)	(0.0763)	(0.1411)	(0.1137)	(0.0650)	(0.1101)	(0.0865)
$\Delta e_{HML,t+h}$	0.4480	0.2204[c]	0.0280	0.2330[a]	0.1873[c]	0.5070[c]	0.3237[b]	0.2613[a]	0.2475[a]	0.4244[a]
	(0.2841)	(0.1154)	(0.0521)	(0.0777)	(0.1088)	(0.2972)	(0.1271)	(0.0544)	(0.0913)	(0.0884)
$(1-E_t)$ $\Delta_h e_{LMS,t+h}$			0.2402[a]	0.5612[b]	0.2571			0.1635[a]	0.1859	-0.3512
			(0.0470)	(0.2245)	(0.4179)			(0.0621)	(0.3016)	(0.3792)
Constant	0.0675[a]	0.0612[a]	0.0318[a]	0.0951[a]	0.1250[a]	-0.0322[b]	-0.0230	-0.0048	-0.0430	-0.0153
	(0.0087)	(0.0080)	(0.0069)	(0.0140)	(0.0269)	(0.0154)	(0.0153)	(0.0171)	(0.0385)	(0.0287)
Obs	493	493	493	469	445	493	493	493	469	445
Adj R^2	8%	19%	51%	42%	48%	7%	16%	44%	42%	60%

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1 Mo	3 Mo	1 yr	3 yrs	5 yrs	1 Mo	3 Mo	1 yr	3 yrs	5 yrs
$r_{LMS,t+h}$	0.7165[a]	0.6256[a]	0.4534[a]	0.1313[b]	0.4898[a]	0.6498[a]	0.5682[a]	0.3335[a]	0.2644[a]	0.6561[a]
	(0.1372)	(0.0978)	(0.0581)	(0.0582)	(0.0699)	(0.1430)	(0.1173)	(0.0821)	(0.0820)	(0.1186)
Constant	0.0750[a]	0.0694[a]	0.0553[a]	0.0626[a]	0.1330[a]	-0.0500[a]	-0.0517[a]	-0.0545[a]	-0.0481[a]	-0.1639[a]
	(0.0074)	(0.0068)	(0.0068)	(0.0111)	(0.0127)	(0.0116)	(0.0110)	(0.0108)	(0.0107)	(0.0228)
Obs	493	493	493	469	445	493	493	493	469	445
Adj R^2	6%	17%	36%	8%	42%	5%	11%	16%	14%	42%

Panel C: momentum (WML)

	$EBR_{WML,t+h}$			
(1)	(2)	(3)	(4)	(5)
1 Mo	3 Mo	1 yr	3 yrs	5 yrs

$r_{LMS,t+h}$	0.7864[a]	0.7271[a]	0.2403[a]	0.4357[a]	0.5007[a]
	(0.1248)	(0.0931)	(0.0456)	(0.0851)	(0.1161)
$\Delta e_{HML,t+h}$	0.9235[b]	0.4931[a]	0.1464[c]	-0.0318	0.1481
	(0.4523)	(0.1677)	(0.0749)	(0.1073)	(0.2198)
$(1-E_t)$ $\Delta_h e_{LMS,t+h}$			0.4035[a]	1.3463[a]	0.9940
			(0.0553)	(0.2896)	(0.8039)
Constant	0.1360[a]	0.1146[a]	0.0634[a]	0.0818[a]	0.1166[b]
	(0.0357)	(0.0309)	(0.0190)	(0.0258)	(0.0513)
Obs	493	493	493	469	445
Adj R^2	12%	29%	62%	63%	40%

	(1) 1 Mo	(2) 3 Mo	(3) 1 yr	(4) 3 yrs	(5) 5 yrs
$r_{LMS,t+h}$	0.7611[a]	0.7223[a]	0.3904[a]	0.2582[a]	0.5731[a]
	(0.1269)	(0.0974)	(0.0627)	(0.0679)	(0.1135)
Constant	0.2016[a]	0.1964[a]	0.2045[a]	0.2151[a]	0.1389[a]
	(0.0161)	(0.0141)	(0.0142)	(0.0219)	(0.0442)
Obs	493	493	493	469	445
Adj R^2	8%	24%	29%	20%	33%

Consistent with our portfolio level results (Tables 3 and 7), portfolio level EBRs have significant loadings on contemporaneous portfolio level returns (Panel A). In turn, Panel B shows that, controlling for actual return spreads $r_{LMS,t+h}$, EBRs strongly responds to news, in terms of both contemporaneous realized growth $\Delta_h e_{LMS,t+h}$ and realized forecast errors, which are a broader proxy for news including forward looking news. In particular, accounting for forecast errors leads to a substantial increase in the adjusted R^2, as well as a drop in the correlation of EBR and actual returns in most cases. Thus, stock returns are not mechanically incorporated into expectations. In the next section we present an analogous firm level result.

B3. Firm level results

In this section we complement our portfolio level analysis with three sets of exercises at the firm level. We first repeat the baseline analysis of Table 3 regressing returns on EBRs, because

Equations (8) and (4) for realized returns hold at that level, too. This exercise confirms our findings that EBRs explain contemporaneous returns, and addresses a possible critique of our approach, namely that there it relies on a relatively small number of observations. Next, we show (following Table 5) that the components into of EBRs – forecast errors and revisions – explain returns.

Finally, we repeat the analysis in Table B5, assessing the extent to which firm level EBRs respond to firm level news, controlling for firm level returns, which confirms that expectations respond strongly to news controlling for prices and is evidence for the validity of analyst expectations as proxies for market expectations.

Starting with the baseline analysis, there is substantial firm level variation in beliefs that can be exploited to detect the link between expectations and returns. In BGLS (2024) we already showed robust predictability of returns from lagged expectations at the firm level. Here, in line with Equations (4) and (8), we similarly report the explanatory power of contemporaneous firm level EBR for actual firm level returns.

Table B6.
Expectation based returns and actual returns

Note: This table presents firm level univariate regression results for log returns against expectation-based returns (EBRs) We estimate separate regressions for one-month, three-months, one-year, three-years, and five-years horizons. Each regression is run with (odd columns) and without (even columns) time and firm fixed effects. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period extends from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	1 Mo		3 Mo		1 Yr		3 Yrs		5 Yrs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$EBR_{i,t+h}$	0.1552[a]	0.1734[a]	0.2064[a]	0.2438[a]	0.4141[a]	0.4715[a]	0.5951[a]	0.6612[a]	0.6752[a]	0.7486[a]
	(0.0045)	(0.0099)	(0.0069)	(0.0148)	(0.0152)	(0.0220)	(0.0222)	(0.0238)	(0.0225)	(0.0203)
Constant		0.0027		0.0088		0.0455[b]		0.1207[a]		0.1691[a]
		(0.0028)		(0.0074)		(0.0206)		(0.0380)		(0.0493)
Obs	899,458	899,977	837,316	837,792	790,868	791,014	604738	604,871	479,198	479,282
Adj R^2	3%	3%	7%	7%	25%	26%	45%	47%	51%	54%
F-stat	0.0	0.0	0.1	0.1	0.2	0.3	0.4	0.5	0.5	0.5
Time FE	Y	N	Y	N	Y	N	Y	N	Y	N
Firm FE	Y	N	Y	N	Y	N	Y	N	Y	N

To focus on firm level variation in beliefs, in some specifications we include time and firm fixed effects in firm level regressions. In those specifications, R^2s capture explanatory power relative to the demeaned variables. Firm level EBRs have strong explanatory power for realized returns, and more so at longer horizons. The explanatory power is similar with and without fixed effects. This is an important finding: first, it suggests that firm level differences in required returns are negligeable (since firm fixed effects do not matter). Second, and related, it suggests that aggregate changes in required returns are also negligeable (since time fixed effects do not matter).

Next, we present a firm-level counterpart to Table 5, which decomposes the link between EBRs and returns into EBR's forecast error and forecast revision components.

Table B7.
Forecast errors and revisions predict returns at the firm level

Note: This table present firm-level multivariate regressions of expectation-based returns (EBRs) over holding horizons (h) of one-year, three-years, and five-years. The independent variables include: (a) the forecast error for earnings growth between t and $t+h$, (b) forecast revisions between t and $t+h$ of one-year earnings growth in year $t+h+1$, and (c) changes in long-term growth forecasts between t and $t+h$,. Regressions include time and firm fixed effects in columns (1), (4), and (7), only time fixed effects in columns (2), (5), and (8), and exclude fixed effects in columns (3), (6), and (9). All independent variables are standardized. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

		1 Yr			3 Yrs			5 Yrs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$(1-E_t)\ \Delta e_{i,t+h}$	0.2946[a]	0.2989[a]	0.3070[a]	0.4556[a]	0.4643[a]	0.4725[a]	0.3996[a]	0.4269[a]	0.4445[a]
	(0.0103)	(0.0107)	(0.0134)	(0.0122)	(0.0138)	(0.0121)	(0.0113)	(0.0132)	(0.0160)
$(E_{t+h}-E_t)\ \Delta e_{i,t+h+1}$	0.2073[a]	0.2158[a]	0.2262[a]	0.2378[a]	0.2552[a]	0.2636[a]	0.1713[a]	0.1553[a]	0.1691[a]
	(0.0076)	(0.0082)	(0.0114)	(0.0074)	(0.0081)	(0.0092)	(0.0058)	(0.0086)	(0.0127)
$\Delta_h LTG_{i,t+h}$	0.0340[a]	0.0351[a]	0.0394[a]	0.0532[a]	0.0490[a]	0.0498[a]	0.0291	0.0196	0.0234
	(0.0044)	(0.0044)	(0.0043)	(0.0122)	(0.0101)	(0.0105)	(0.0250)	(0.0181)	(0.0181)
Constant			0.1432[a]			0.4276[a]			0.6330[a]
			(0.0171)			(0.0325)			(0.0451)
Obs	695,893	696,107	696,107	570,256	570,431	570,431	447,310	447,421	447,421
Adj R^2	28%	30%	28%	44%	46%	43%	35%	37%	36%
Time FE	Y	Y	N	Y	Y	N	Y	Y	N
Firm FE	Y	N	N	Y	N	N	Y	N	N

Consistent with Equation (4), all three expectations measures have strong explanatory power for the variation of firm level returns over time, with a large aggregate adjusted R^2. Each specification is run with and without time and fixed effects; remarkably, and similar to Table B6, the explanatory power of expectations – both the coefficients and R2 – remains unchanged. In sum, Tables B6 and B8 place sharp limits on the role of required returns in the Campbell Shiller decomposition. In fact, the results are thus consistent with the decomposition (4) with r_{ih} being a constant that scales with horizon h, as captured by the constant term in columns 3, 6, and 9.[22]

Finally, we replicate Table B5 above at the firm level. Specifically, we examine whether firm level EBRs are explained entirely by *contemporaneous* firm level returns (which would hold if analysts mechanically infer forecasts from prices), or whether they reflect contemporaneous cash flow growth.

Table B8
Expectation based returns and contemporaneous news

Note: Panel A presents firm-level univariate regressions of expectation-based returns (EBR) at horizons (h) of one-month, three-months, one-year, three-years, and five-years. The independent variable is: (a) log returns between t and $t+h$. Panel B presents firm-level multivariate regressions of expectation-based returns (EBR) at horizons (h) of one-month, three-months, one-year, three-years, and five-years. The independent variables include: (a) log returns between t and $t+h$, (b) earnings growth between t and $t+h$, and (c) the forecast error for earnings growth between t and $t+h$. All specifications include firm fixed effects. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A

	$EBR_{i,t+1}$	$EBR_{i,t+3}$	$EBR_{i,t+12}$	$EBR_{i,t+36}$	$EBR_{i,t+60}$
	(1)	(2)	(3)	(4)	(5)
$r_{i,t+h}$	0.1221[a]	0.2265[a]	0.4741[a]	0.6325[a]	0.6424[a]
	(0.0015)	(0.0093)	(0.0151)	(0.0168)	(0.0110)
Obs	721,554	683,366	731,977	540,744	406,309
Adj R^2	2%	7%	26%	47%	55%

Panel B

	$EBR_{i,t+1}$	$EBR_{i,t+3}$	$EBR_{i,t+12}$	$EBR_{i,t+36}$	$EBR_{i,t+60}$
	(1)	(2)	(3)	(4)	(5)
$r_{i,t+h}$	0.1185[a]	0.2113[a]	0.3449[a]	0.3952[a]	0.4648[a]
	(0.0014)	(0.0091)	(0.0140)	(0.0117)	(0.0119)

[22] In line with the prediction, the coefficients for 1 year horizon are larger for forecast errors and smaller for long term forecasts (note that for long horizons, information about long term forecasts is already included in the other two regressors).

$\Delta e_{i,t+h}$	0.0098[a]	0.0319[a]	0.0219[a]	-0.0170[b]	0.0059
	(0.0002)	(0.0014)	(0.0026)	(0.0071)	(0.0183)
$(1-E_t)\Delta_h e_{i,t+h}$			0.1451[a]	0.3095[a]	0.2334[a]
			(0.0056)	(0.0103)	(0.0173)
Obs	721,347	683,145	731,768	540,564	406,189
Adj R^2	2%	7%	36%	59%	60%

Consistent with our portfolio level results (Tables 3 and 7 in the text, and Table B7 above), firm level EBRs have significant loading on firm level contemporaneous returns, which increase at longer horizons (Panel A). In turn, Panel B shows that, controlling for actual returns $r_{i,t+h}$, EBRs strongly responds to news, in terms of both contemporaneous realized growth $\Delta_h e_{t+h}$ and realized forecast errors, which are a broader proxy for news including forward looking news. In fact, accounting for news leads to a significant drop in the correlation of EBR and actual returns, as well as a substantial increase in the adjusted R^2. Table B8 shows that stock returns are not mechanically incorporated into expectations.

Appendix C: Return predictability

C1. Portfolio level results.

Table C1 presents the full results of the return predictability exercise described in Table 8, in which return spreads are predicted from lagged expectations variables and from the dividend-to-price ratio (as a proxy for required returns).

Table C1

Predicting return spreads from expectations data and the price dividend ratio spreads

Note: The table presents portfolio-level regressions of log returns ($r_{LMS,t+h}$) at horizons (h) of one-and 3-months, and one-, three- and five-years. In each panel, the independent variables in the first regression are: (a) log dividend-to-price ratio at time t ($dp_{LMS,t}$), (b) the change in the portfolio forecast for long-term growth in earnings between $t-$h and t, (c) the lagged portfolio forecast for long-term growth in earnings at $t-h$, (d) the change in the portfolio forecast for one-year growth in earnings between t-h and t, (e) the ranked forecast error in portfolio earnings between $t-h$ and t, (f) the change in the aggregate forecast for long-term growth in earnings between $t-h$ and t, and (g) the aggregate forecast for long-term growth in earnings at t. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). In the second regression, the independent variable is (a) log dividend-to-price ratio at time t ($dp_{LMS,t}$). Panel A presents results for portfolios that are long value and short growth stocks (HML) and long small and short big stocks (SMB), Panel B for portfolios that are long conservative and short aggressive investment stocks (CMA) and long robust and short weak profitability stocks (RMW), and Panel C for portfolios that are long winner and short momentum stocks (WML). All independent variables are standardized. The last row reports the R^2 from an univariate regression of $ret_{LMS,t+h}$ on $dp_{LMS,t}$. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] 0% level.

Panel A: value and size

			$r_{HML,t+h}$					$r_{SMB,t+h}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
$\ln(dp_{LS,t})$	-0.0075	0.0392	0.0305	0.0753	0.4096^{b}	0.0098	0.0815^{c}	0.1491	0.1727	0.7869^{a}
	(0.0243)	(0.0591)	(0.2334)	(0.3420)	(0.1880)	(0.0153)	(0.0465)	(0.1159)	(0.1093)	(0.1819)
$\Delta_h LTG_{LS,t}$	0.0045^{c}	0.0093	-0.0762^{b}	-0.0571^{c}	0.0470	0.0060^{b}	-0.0020	0.0057	-0.0192	-0.0227
	(0.0026)	(0.0074)	(0.0341)	(0.0337)	(0.0364)	(0.0025)	(0.0104)	(0.0277)	(0.0542)	(0.0613)
$LTG_{LS,t-h}$	-0.0190	-0.0395	-0.2657	-0.4045	-0.1988	0.0052	0.0280	0.0348	-0.1523	0.1743
	(0.0143)	(0.0364)	(0.1730)	(0.2788)	(0.2063)	(0.0068)	(0.0172)	(0.0708)	(0.1073)	(0.1513)
$\Delta_h STG_{LS,t}$	-0.0022	0.0017	0.0396^{b}	0.0829^{a}	0.1156^{a}	0.0020	0.0028	0.0456^{a}	0.0494^{a}	0.0192
	(0.0023)	(0.0075)	(0.0167)	(0.0263)	(0.0343)	(0.0026)	(0.0079)	(0.0157)	(0.0153)	(0.0183)
$STG_{LS,t-h}$	0.0158	0.0676^{b}	0.2155^{a}	0.4322^{a}	0.2980^{b}	0.0304^{b}	0.0407	0.1286^{c}	-0.0153	-0.1179
	(0.0105)	(0.0341)	(0.0651)	(0.0902)	(0.1306)	(0.0132)	(0.0261)	(0.0737)	(0.1290)	(0.1442)
$FE_{LS,t-h}$ rank			0.0821^{c}	0.1261	-0.2389^{b}			0.0142	0.0451	0.0478
			(0.0492)	(0.1359)	(0.0978)			(0.0576)	(0.0853)	(0.0704)
$\Delta_h LTG_{Agg,t}$	-0.0046^{c}	0.0047	0.0257	0.0586	0.2336^{a}	0.0085^{a}	0.0008	-0.0033	0.0843^{a}	0.1584^{a}
	(0.0025)	(0.0072)	(0.0307)	(0.0377)	(0.0507)	(0.0028)	(0.0106)	(0.0248)	(0.0233)	(0.0362)
$LTG_{Agg,t-h}$	0.0041^{c}	0.0115	0.0346	0.1279^{a}	0.1237^{a}	0.0030	0.0064	0.0483^{b}	0.0992^{a}	0.0667^{b}
	(0.0022)	(0.0071)	(0.0294)	(0.0351)	(0.0428)	(0.0032)	(0.0069)	(0.0203)	(0.0249)	(0.0259)
Constant	-0.0401^{b}	-0.1147^{b}	-0.4177^{b}	-1.0655^{a}	-1.2508^{a}	-0.0307	-0.0314	-0.2504^{c}	-0.2601^{c}	0.0421
	(0.0164)	(0.0533)	(0.1890)	(0.3428)	(0.3593)	(0.0271)	(0.0542)	(0.1497)	(0.1421)	(0.2209)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	5%	7%	16%	27%	67%	4%	3%	23%	43%	58%

$\ln(dp_{p,t})$	0.0039	0.0365	0.1756	0.1500	0.3681	0.0119	0.0519^{c}	0.2031^{b}	0.4859^{a}	0.6227^{a}
	(0.0118)	(0.0372)	(0.1171)	(0.2480)	(0.3079)	(0.0102)	(0.0293)	(0.0913)	(0.1871)	(0.1859)
Constant	0.0024	0.0002	-0.0103	0.0338	-0.0107	0.0101	0.0413^{b}	0.1707^{a}	0.4201^{a}	0.5668^{a}
	(0.0033)	(0.0106)	(0.0301)	(0.0644)	(0.0725)	(0.0064)	(0.0187)	(0.0634)	(0.1413)	(0.1514)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	0%	0%	3%	1%	5%	0%	2%	7%	19%	26%

Panel B: investment and profitability

			$r_{CMA,t+h}$					$r_{RMW,t+h}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
$\ln(dp_{p,t})$	-0.0087	-0.0525	0.0828	-0.3400^{c}	-0.0501	0.0232	0.0903	0.1470	-0.3627	-0.1907
	(0.0152)	(0.0393)	(0.0972)	(0.1787)	(0.2076)	(0.0215)	(0.0552)	(0.2000)	(0.3296)	(0.2050)
$LTG_{p,t-1}$	0.0038^{b}	0.0051	0.0088	-0.0781^{a}	-0.0588	0.0014	-0.0026	-0.0029	0.0006	0.0214
	(0.0016)	(0.0045)	(0.0161)	(0.0174)	(0.0364)	(0.0023)	(0.0043)	(0.0159)	(0.0213)	(0.0207)

ΔLTG$_{p,t}$	-0.0098[c]	-0.0386[a]	-0.1410[a]	-0.2884[b]	-0.1174	0.0034	0.0210	0.0794	-0.1677	0.3496[a]
	(0.0057)	(0.0138)	(0.0516)	(0.1162)	(0.1259)	(0.0080)	(0.0172)	(0.0724)	(0.1077)	(0.0972)
STG$_{p,t}$	-0.0007	0.0044	0.0113	0.0030	0.0201	0.0039	0.0052	0.0035	0.0065	-0.0261[c]
	(0.0017)	(0.0043)	(0.0118)	(0.0227)	(0.0223)	(0.0025)	(0.0047)	(0.0120)	(0.0208)	(0.0135)
FE$_{p,t-1}$ rank	0.0081	0.0197	0.1187[b]	-0.0077	0.1047	0.0234[c]	0.0410[c]	0.0130	0.0292	-0.1556
	(0.0092)	(0.0243)	(0.0479)	(0.1204)	(0.1458)	(0.0139)	(0.0238)	(0.0683)	(0.1351)	(0.1241)
FE$_{p,t-3}$ rank			-0.0246	0.0138	0.1186[b]			-0.0309[c]	0.0234	-0.0327
			(0.0232)	(0.0580)	(0.0484)			(0.0184)	(0.0267)	(0.0222)
FE$_{p,t-5}$ rank	-0.0023[c]	0.0026	0.0176	0.0044	0.0475[b]	-0.0082[a]	-0.0072	0.0373	0.0004	0.0502[c]
	(0.0012)	(0.0030)	(0.0112)	(0.0168)	(0.0205)	(0.0023)	(0.0066)	(0.0257)	(0.0248)	(0.0269)
ΔLTG$_{agg,t}$	0.0038[a]	0.0120[a]	0.0246[a]	0.0177	0.0399[b]	0.0026	0.0108[c]	0.0356	-0.0294	0.1315[a]
	(0.0014)	(0.0035)	(0.0092)	(0.0180)	(0.0162)	(0.0028)	(0.0059)	(0.0272)	(0.0361)	(0.0366)
LTG$_{agg,t-1}$	-0.0255[a]	-0.0722[a]	-0.2406[a]	-0.0913	-0.1604	-0.0102	-0.0579[c]	-0.1953	0.1942	-0.7219[a]
	(0.0086)	(0.0223)	(0.0675)	(0.1801)	(0.1361)	(0.0170)	(0.0311)	(0.1537)	(0.2336)	(0.2458)
Constant	444	442	433	409	385	444	442	433	409	385
	6%	11%	27%	18%	16%	5%	4%	9%	8%	28%
Obs	-0.0087	-0.0525	0.0828	-0.3400[c]	-0.0501	0.0232	0.0903	0.1470	-0.3627	-0.1907
Adjusted R^2	(0.0152)	(0.0393)	(0.0972)	(0.1787)	(0.2076)	(0.0215)	(0.0552)	(0.2000)	(0.3296)	(0.2050)

ln(dp$_{p,t}$)	-0.0012	0.0238	0.1889[a]	0.0746	0.1146	0.0001	0.0172	-0.0051	-0.1520	-0.4238[b]
	(0.0093)	(0.0265)	(0.0561)	(0.1014)	(0.0744)	(0.0106)	(0.0260)	(0.0813)	(0.1044)	(0.1687)
Constant	0.0052[b]	0.0089	0.0056	0.0617[c]	0.0614[b]	0.0010	-0.0023	0.0097	0.0511	0.0928[b]
	(0.0024)	(0.0070)	(0.0161)	(0.0336)	(0.0277)	(0.0033)	(0.0086)	(0.0255)	(0.0317)	(0.0434)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	0%	0%	8%	0%	1%	0%	0%	0%	3%	13%

Panel C: momentum

	$r_{WML,t+h}$				
	(1)	(2)	(3)	(4)	(5)
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
ln(dp$_{p,t}$)	0.0118	0.0125	0.0880	0.1077	0.4152[a]
	(0.0163)	(0.0318)	(0.0922)	(0.0876)	(0.0991)
LTG$_{p,t-1}$	0.0031	0.0018	-0.0279	-0.0166	0.0161
	(0.0030)	(0.0082)	(0.0249)	(0.0338)	(0.0634)
ΔLTG$_{p,t}$	-0.0134[c]	-0.0372[b]	-0.1470[b]	-0.1787[a]	-0.0013
	(0.0075)	(0.0149)	(0.0572)	(0.0502)	(0.1105)
STG$_{p,t}$	0.0075[b]	0.0199[b]	0.0706[a]	0.0971[a]	0.0877[b]
	(0.0032)	(0.0093)	(0.0236)	(0.0238)	(0.0392)
FE$_{p,t-1}$ rank	0.0558[a]	0.1102[a]	0.2527[a]	0.4252[a]	0.4535
	(0.0176)	(0.0387)	(0.0704)	(0.1387)	(0.3245)
FE$_{p,t-3}$ rank			-0.0031	-0.0026	-0.0313
			(0.0320)	(0.0371)	(0.0462)

$FE_{p,t-5}$ rank	0.0022	0.0002	0.0073	-0.0787[a]	-0.1070[a]
	(0.0025)	(0.0069)	(0.0184)	(0.0298)	(0.0350)
$\Delta LTG_{agg,t}$	0.0028	0.0077	-0.0205	-0.0593[a]	-0.0609[c]
	(0.0027)	(0.0059)	(0.0297)	(0.0223)	(0.0353)
$LTG_{agg,t-1}$	0.0027	0.0056	0.2259	0.4656[a]	0.4755[c]
	(0.0160)	(0.0335)	(0.2092)	(0.1683)	(0.2567)
Constant	444	442	433	409	385
	7%	16%	26%	36%	23%
Obs	0.0118	0.0125	0.0880	0.1077	0.4152[a]
Adjusted R^2	(0.0163)	(0.0318)	(0.0922)	(0.0876)	(0.0991)

$\ln(dp_{p,t})$	0.0074	0.0219	0.1333[c]	0.1800	0.1798[b]
	(0.0104)	(0.0219)	(0.0704)	(0.1100)	(0.0910)
Constant	0.0040	0.0093	-0.0068	-0.0487	-0.0726[b]
	(0.0025)	(0.0067)	(0.0216)	(0.0336)	(0.0363)
Obs	444	442	433	409	385
Adjusted R^2	0%	0%	4%	4%	3%

The results in Table C1 generalize those of Table 8 to other horizons: the dividend price ratio is nearly never significant and often of the wrong sign, while the coefficients on expectations variables are large, particularly at long horizons. The dividend price ratio's predictive power in univariate regressions is very low, while adding lagged expectations dramatically increases it in all specifications, as measured by higher adjusted R^2. The same results obtain when replacing the dividend-to-price ratio with the book-to-market ratio or the lagged return spread, as shown in Table C2.

Table C2
Robustness of return predictability from expectations data

Note: the table presents regressions of log returns for portfolios that are long value and short growth stocks (HML), long small and short big stocks (SMB), long conservative and short aggressive investment stocks (CMA), long robust and short weak profitability stocks (RMW), and long winner and short momentum stocks (WML). Separate regressions are estimated for horizons (h) one-month and one year. The set of independent variables includes: (a) the ranked forecast error in portfolio earnings between $t-h$ and t, (b) the change in the portfolio forecast for short-term growth in earnings between $t-h$ and t, (c) the lagged portfolio forecast for one-year growth in earnings at $t-h$, (d) the change in the portfolio forecast for long-term growth in earnings between $t-h$ and t, (e) the lagged portfolio forecast for long-term growth in earnings at $t-h$, (f) the change in the aggregate forecast for long-term growth in earnings between $t-h$ and t, (g) the forecast for long-term growth in aggregate earnings at $t-h$, and (g) the portfolio log book-to-market ratio ($bm_{LMS,t}$) at time t in Panel A, and the portfolio log return spread ($r_{LMS,t}$) at time t in Panel B. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). All independent variables are standardized. The last row reports the R^2 from a univariate regression of $r_{LMS,t+h}$ on $bm_{LMS,t}$ (panel A) or on $r_{LMS,t}$ (panel B). Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: ^{a} significant at the 1% level, ^{b} significant at the 5% level, ^{c} 0% level.

Panel A: controlling for lagged book-to-market ratio spreads

	$r_{HML,t+h}$		$r_{SMB,t+h}$		$r_{CMA,t+h}$		r_{RMWt+h}		$r_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$bm_{LMS,t}$	-0.0144^{c}	0.0850	-0.0123^{b}	0.1425^{a}	-0.0080^{b}	0.0171	-0.0096^{b}	0.0058	-0.0076^{b}	0.0292
	(0.0079)	(0.0824)	(0.0059)	(0.0438)	(0.0037)	(0.0276)	(0.0039)	(0.0268)	(0.0037)	(0.0259)
$\Delta_h LTG_{LMS,t+h}$	0.0046^{c}	-0.0789^{b}	0.0063^{b}	-0.0151	0.0043^{a}	0.0050	0.0017	-0.0088	0.0036	-0.0232
	(0.0027)	(0.0308)	(0.0026)	(0.0260)	(0.0016)	(0.0164)	(0.0024)	(0.0179)	(0.0032)	(0.0244)
$LTG_{LMS,t-h}$	-0.0058	-0.3733^{a}	0.0016	0.0099	-0.0110^{b}	-0.1537^{a}	-0.0002	0.0300	-0.0227^{a}	-0.1346^{b}
	(0.0083)	(0.1274)	(0.0056)	(0.0510)	(0.0049)	(0.0420)	(0.0057)	(0.0351)	(0.0074)	(0.0560)
$(\mathbb{E}_{t+h}-\mathbb{E}_t)$ $\Delta_h e_{LMS,t+h+12}$	-0.0025	0.0413^{b}	0.0021	0.0362^{b}	-0.0010	0.0125	0.0042^{c}	0.0063	0.0084^{a}	0.0698^{a}
	(0.0024)	(0.0170)	(0.0026)	(0.0168)	(0.0017)	(0.0110)	(0.0024)	(0.0132)	(0.0032)	(0.0217)
$\mathbb{E}_{t-h}\Delta_h e_{LMS,t+12}$	0.0119	0.2466^{a}	0.0338^{a}	0.1325^{c}	0.0078	0.1052^{b}	0.0221	0.0115	0.0620^{a}	0.1990^{a}
	(0.0091)	(0.0690)	(0.0125)	(0.0709)	(0.0088)	(0.0459)	(0.0141)	(0.0698)	(0.0172)	(0.0749)
$(1-E_t)\ \Delta_h e_{LMS,t}$		0.0727^{c}		0.0404		-0.0148		-0.0285		0.0051
		(0.0372)		(0.0492)		(0.0247)		(0.0198)		(0.0322)
$\Delta_h LTG_{Mkt,t}$	-0.0044^{c}	0.0156	0.0085^{a}	-0.0393	-0.0024^{b}	0.0206^{c}	-0.0081^{a}	0.0325	0.0018	0.0084
	(0.0024)	(0.0292)	(0.0029)	(0.0248)	(0.0012)	(0.0118)	(0.0023)	(0.0223)	(0.0025)	(0.0189)
$LTG_{Mkt,t-h}$	0.0050^{b}	0.0248	0.0059	0.0050	0.0028^{c}	0.0276^{a}	0.0029	0.0260	0.0010	-0.0091
	(0.0022)	(0.0301)	(0.0036)	(0.0184)	(0.0015)	(0.0094)	(0.0026)	(0.0248)	(0.0028)	(0.0315)
Constant	-0.0132	-0.5732^{b}	-0.0398	-0.1901	-0.0184^{c}	-0.2523^{a}	-0.0272	-0.1132	0.0029	0.1818
	(0.0176)	(0.2240)	(0.0268)	(0.1259)	(0.0102)	(0.0655)	(0.0175)	(0.1658)	(0.0162)	(0.2108)
Obs	442	433	442	433	442	433	442	433	442	433
Adjusted R^2	6%	18%	5%	27%	8%	27%	6%	8%	9%	27%
Univariate R^2	2%	0%	0%	16%	2%	1%	1%	0%	0%	9%

Panel B: controlling for lagged return spreads

	$r_{HML,t+h}$		$r_{SMB,t+h}$		$r_{CMA,t+h}$		r_{RMWt+h}		$r_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t}$	0.2180^{a}	0.0466	0.2591^{a}	-0.0676^{b}	0.2133^{a}	0.0054	0.2392^{a}	-0.0835^{b}	0.1708^{a}	0.0348
	(0.0329)	(0.0379)	(0.0256)	(0.0336)	(0.0262)	(0.0353)	(0.0338)	(0.0355)	(0.0294)	(0.0377)
$\Delta_h LTG_{LMS,t+h}$	0.0051^{b}	-0.0837^{b}	0.0023	-0.0078	0.0024^{c}	0.0060	-0.0020	-0.0072	0.0017	-0.0175
	(0.0026)	(0.0327)	(0.0021)	(0.0258)	(0.0013)	(0.0163)	(0.0020)	(0.0153)	(0.0027)	(0.0296)
$LTG_{LMS,t-h}$	-0.0118	-0.2688^{a}	0.0023	-0.0067	-0.0064	-0.1607^{a}	-0.0002	0.0453	-0.0129^{c}	-0.1866^{a}
	(0.0082)	(0.0886)	(0.0045)	(0.0542)	(0.0044)	(0.0386)	(0.0046)	(0.0343)	(0.0068)	(0.0527)
$(\mathbb{E}_{t+h}-\mathbb{E}_t)$ $\Delta_h e_{LMS,t+h+12}$	0.0012	0.0374^{b}	0.0039^{b}	0.0503^{a}	0.0005	0.0112	0.0025	0.0116	0.0050	0.0770^{a}
	(0.0022)	(0.0179)	(0.0019)	(0.0157)	(0.0014)	(0.0118)	(0.0020)	(0.0144)	(0.0040)	(0.0239)
$\mathbb{E}_{t-h}\Delta_h e_{LMS,t+12}$	0.0163^{c}	0.1675^{a}	0.0228^{b}	0.1569^{b}	0.0090	0.1027^{b}	0.0088	0.0273	0.0315^{c}	0.2083^{a}

	(0.0089)	(0.0641)	(0.0102)	(0.0720)	(0.0080)	(0.0442)	(0.0119)	(0.0723)	(0.0165)	(0.0696)
$(1-E_t)\ \Delta_h e_{LMS,t}$		0.0645		0.0840		-0.0266		-0.0165		0.0060
		(0.0533)		(0.0692)		(0.0247)		(0.0211)		(0.0323)
$\Delta_h LTG_{Mkt,t}$	0.0005	0.0285	0.0030	-0.0027	-0.0002	0.0186	-0.0006	0.0182	0.0019	-0.0040
	(0.0022)	(0.0284)	(0.0026)	(0.0247)	(0.0012)	(0.0116)	(0.0022)	(0.0230)	(0.0024)	(0.0205)
$LTG_{Mkt,t-h}$	0.0029	0.0319	0.0021	0.0511^{b}	0.0023^{c}	0.0244^{b}	0.0015	0.0257	0.0021	-0.0354
	(0.0019)	(0.0290)	(0.0025)	(0.0201)	(0.0013)	(0.0095)	(0.0022)	(0.0234)	(0.0026)	(0.0332)
Constant	-0.0289^{b}	-0.3911^{b}	-0.0241	-0.3043^{b}	-0.0174^{b}	-0.2284^{a}	-0.0057	-0.1179	0.0012	0.2084
	(0.0137)	(0.1746)	(0.0208)	(0.1345)	(0.0078)	(0.0664)	(0.0131)	(0.1265)	(0.0151)	(0.2091)
Obs	444	433	444	433	444	433	444	433	444	433
Adjusted R^2	23%	17%	29%	24%	24%	26%	25%	13%	24%	26%
Univariate R^2	21%	2%	29%	2%	23%	1%	26%	9%	21%	0%

Table C3 presents the first stage of the exercise of Table 9, in which EBRs spreads are predicted from lagged expectations variables and from the dividend-to-price ratio.

Table C3

Note: The table presents portfolio-level regressions of log expectations-based returns (EBRs) at horizons (h) of one- and 3-months, and one-, three- and five-years. In each panel, the independent variables in the first regression are: (a) log dividend-to-price ratio at time t ($dp_{LMS,t}$), (b) the change in the portfolio forecast for long-term growth in earnings between $t-$h and t, (c) the lagged portfolio forecast for long-term growth in earnings at $t-h$, (d) the change in the portfolio forecast for one-year growth in earnings between t-h and t, (e) the ranked forecast error in portfolio earnings between $t-h$ and t, (f) the change in the aggregate forecast for long-term growth in earnings between $t-h$ and t, and (g) the aggregate forecast for long-term growth in earnings at t. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). In the second regression, the independent variable is (a) log dividend-to-price ratio at time t ($dp_{LMS,t}$). Panel A presents results for portfolios that are long value and short growth stocks (HML) and long small and short big stocks (SMB), Panel B for portfolios that are long conservative and short aggressive investment stocks (CMA) and long robust and short weak profitability stocks (RMW), and Panel C for portfolios that are long winner and short momentum stocks (WML). All independent variables are standardized. The last row reports the R^2 from an univariate regression of log $EBR_{LMS,t+h}$ on $dp_{LMS,t}$. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] 0% level.

	$EBR_{HML,t+h}$					$EBR_{SMB,t+h}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
$\ln(dp_{LS,t})$	-0.0102	-0.0400	-0.2259^{c}	-0.1714	0.0089	-0.0042	0.0065	-0.0353	0.0673	0.4009^{a}
	(0.0086)	(0.0246)	(0.1166)	(0.1779)	(0.0947)	(0.0053)	(0.0166)	(0.0525)	(0.0568)	(0.0727)
$\Delta_h LTG_{LS,t}$	0.0033^{a}	0.0125^{a}	-0.0623^{a}	-0.0491^{b}	-0.0177	0.0011	0.0029	0.0051	-0.0375	-0.0547^{b}
	(0.0010)	(0.0037)	(0.0177)	(0.0249)	(0.0222)	(0.0008)	(0.0031)	(0.0100)	(0.0242)	(0.0267)
$LTG_{LS,t-h}$	-0.0052	-0.0267	-0.2889^{a}	-0.3699^{b}	-0.1250	-0.0040^{c}	-0.0098	-0.0576^{a}	-0.1481^{a}	-0.0514
	(0.0049)	(0.0166)	(0.0973)	(0.1587)	(0.0963)	(0.0022)	(0.0066)	(0.0188)	(0.0429)	(0.0516)
$\Delta_h STG_{LS,t}$	-0.0039^{a}	-0.0094^{a}	0.0386^{a}	0.0590^{a}	0.0654^{a}	-0.0004	-0.0020	0.0322^{a}	0.0677^{a}	0.0579^{a}
	(0.0009)	(0.0025)	(0.0099)	(0.0131)	(0.0125)	(0.0008)	(0.0029)	(0.0075)	(0.0114)	(0.0072)

$STG_{LS,t-h}$	-0.0000	0.0121	0.1617^{a}	0.2786^{a}	0.2212^{b}	0.0170^{a}	0.0363^{a}	0.1110^{a}	0.1738^{a}	0.3351^{a}
	(0.0041)	(0.0119)	(0.0434)	(0.0682)	(0.0909)	(0.0045)	(0.0133)	(0.0269)	(0.0537)	(0.0691)
$FE_{LS,t-h}$ rank			0.0839^{b}	0.0427	-0.0692			0.0913^{a}	0.1016^{b}	0.0515
			(0.0378)	(0.0461)	(0.0435)			(0.0261)	(0.0470)	(0.0860)
$\Delta_h LTG_{Agg,t}$	-0.0001	0.0028	0.0309^{b}	0.0337^{b}	0.0946^{a}	0.0014	0.0007	0.0215^{b}	0.0311^{b}	0.0286^{c}
	(0.0008)	(0.0032)	(0.0132)	(0.0140)	(0.0208)	(0.0010)	(0.0032)	(0.0088)	(0.0121)	(0.0170)
$LTG_{Agg,t-h}$	0.0024^{a}	0.0072^{a}	0.0319^{b}	0.0554^{a}	0.0190	0.0000	0.0007	0.0212^{a}	-0.0024	-0.0335^{a}
	(0.0008)	(0.0026)	(0.0134)	(0.0138)	(0.0178)	(0.0007)	(0.0023)	(0.0060)	(0.0127)	(0.0115)
Constant	-0.0174^{a}	-0.0588^{a}	-0.3480^{a}	-0.4974^{a}	-0.1977	-0.0087	-0.0122	-0.0902^{c}	0.2324^{b}	0.4529^{a}
	(0.0055)	(0.0210)	(0.0954)	(0.1520)	(0.1524)	(0.0067)	(0.0192)	(0.0501)	(0.1113)	(0.1596)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	14%	28%	30%	34%	55%	9%	16%	47%	57%	68%
R^2 univariate	1%	2%	0%	0%	0%	1%	6%	16%	30%	46%

	$EBR_{CMA,t+h}$					$EBR_{RMW,t+h}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
$\ln(dp_{LS,t})$	-0.0151^{c}	-0.0576^{a}	0.0039	-0.0848	0.2493^{b}	0.0046	0.0237	0.0587	-0.0079	0.0546
	(0.0080)	(0.0222)	(0.0618)	(0.1104)	(0.1139)	(0.0100)	(0.0257)	(0.0842)	(0.1284)	(0.1970)
$\Delta_h LTG_{LS,t}$	0.0010	0.0014	0.0154	-0.0412^{a}	-0.0113	0.0012	0.0078^{a}	0.0211^{b}	0.0021	0.0442^{b}
	(0.0007)	(0.0036)	(0.0102)	(0.0150)	(0.0164)	(0.0009)	(0.0022)	(0.0087)	(0.0136)	(0.0219)
$LTG_{LS,t-h}$	-0.0039	-0.0156^{b}	-0.0773^{b}	-0.1276^{c}	-0.0366	0.0062^{c}	0.0178^{b}	-0.0356	-0.0859	0.0979
	(0.0035)	(0.0071)	(0.0362)	(0.0736)	(0.0576)	(0.0032)	(0.0081)	(0.0385)	(0.0532)	(0.0845)
$\Delta_h STG_{LS,t}$	-0.0045^{a}	-0.0029	0.0187^{c}	0.0157^{c}	0.0504^{a}	-0.0009	0.0006	0.0191^{b}	0.0319^{b}	0.0275^{c}
	(0.0009)	(0.0023)	(0.0100)	(0.0092)	(0.0091)	(0.0009)	(0.0023)	(0.0090)	(0.0131)	(0.0155)
$STG_{LS,t-h}$	-0.0058	0.0065	0.1111^{a}	0.1619^{a}	0.3005^{a}	0.0055	0.0229^{c}	0.1493^{a}	0.2473^{a}	0.1441
	(0.0048)	(0.0124)	(0.0343)	(0.0471)	(0.0473)	(0.0047)	(0.0123)	(0.0404)	(0.0714)	(0.1012)
$FE_{LS,t-h}$ rank			-0.0149	-0.0363	0.0265			-0.0294^{c}	-0.0263	-0.0325
			(0.0201)	(0.0402)	(0.0461)			(0.0165)	(0.0233)	(0.0280)
$\Delta_h LTG_{Agg,t}$	-0.0003	0.0014	0.0114	0.0021	0.0196	-0.0012	-0.0025	-0.0137	0.0017	-0.0089
	(0.0007)	(0.0026)	(0.0084)	(0.0105)	(0.0132)	(0.0008)	(0.0025)	(0.0094)	(0.0149)	(0.0268)
$LTG_{Agg,t-h}$	0.0018^{a}	0.0071^{a}	0.0160^{a}	-0.0059	0.0403^{a}	0.0007	0.0024	-0.0152	-0.0047	0.0187
	(0.0007)	(0.0021)	(0.0061)	(0.0095)	(0.0087)	(0.0009)	(0.0026)	(0.0140)	(0.0215)	(0.0397)
Constant	-0.0049	-0.0202	-0.0893^{c}	0.1119	-0.1918^{b}	-0.0035	-0.0131	0.0922	0.0181	-0.1791
	(0.0046)	(0.0138)	(0.0468)	(0.0902)	(0.0747)	(0.0059)	(0.0149)	(0.0787)	(0.1526)	(0.2510)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	14%	15%	18%	16%	30%	5%	13%	30%	18%	23%
R^2 univariate	3%	4%	1%	0%	0%	3%	5%	1%	2%	7%

$EBR_{HML,t+h}$				
(1)	(2)	(3)	(4)	(5)
$h=1$	$h=3$	$h=12$	$h=36$	$h=60$

$\ln(dp_{LS,t})$	-0.0102	-0.0400	-0.2259^{c}	-0.1714	0.0089
	(0.0086)	(0.0246)	(0.1166)	(0.1779)	(0.0947)
$\Delta_h LTG_{LS,t}$	0.0033^{a}	0.0125^{a}	-0.0623^{a}	-0.0491^{b}	-0.0177
	(0.0010)	(0.0037)	(0.0177)	(0.0249)	(0.0222)
$LTG_{LS,t-h}$	-0.0052	-0.0267	-0.2889^{a}	-0.3699^{b}	-0.1250
	(0.0049)	(0.0166)	(0.0973)	(0.1587)	(0.0963)
$\Delta_h STG_{LS,t}$	-0.0039^{a}	-0.0094^{a}	0.0386^{a}	0.0590^{a}	0.0654^{a}
	(0.0009)	(0.0025)	(0.0099)	(0.0131)	(0.0125)
$STG_{LS,t-h}$	-0.0000	0.0121	0.1617^{a}	0.2786^{a}	0.2212^{b}
	(0.0041)	(0.0119)	(0.0434)	(0.0682)	(0.0909)
$FE_{LS,t-h}$ rank			0.0839^{b}	0.0427	-0.0692
			(0.0378)	(0.0461)	(0.0435)
$\Delta_h LTG_{Agg,t}$	-0.0001	0.0028	0.0309^{b}	0.0337^{b}	0.0946^{a}
	(0.0008)	(0.0032)	(0.0132)	(0.0140)	(0.0208)
$LTG_{Agg,t-h}$	0.0024^{a}	0.0072^{a}	0.0319^{b}	0.0554^{a}	0.0190
	(0.0008)	(0.0026)	(0.0134)	(0.0138)	(0.0178)
Constant	-0.0174^{a}	-0.0588^{a}	-0.3480^{a}	-0.4974^{a}	-0.1977
	(0.0055)	(0.0210)	(0.0954)	(0.1520)	(0.1524)
Obs	444	442	433	409	385
Adjusted R^2	14%	28%	30%	34%	55%
R^2 univariate	1%	2%	0%	0%	0%

C2. Firm level results

Here we perform, at the firm level, the analysis of Table 10: we first show that firm level future EBRs are predictable from current expectations (Panel A), and then run a horse race between predicted EBRs and current dividend price ratio (Panels B), the latter being a proxy for required returns.

Table C4.

Predicted EBRs predict returns at the firm level

Note: Panel A presents firm-level regressions of log expectations-based returns (EBRs) over holding horizons (h) of one-month, three-months, one-year, three-years, and five-years. The set of independent variables includes: (a) log dividend-to-price ratio at time t ($dp_{LMS,t}$), (b) the change in the portfolio forecast for long-term growth in earnings between $t-$h and t, (c) the lagged portfolio forecast for long-term growth in earnings at $t-h$, (d) the change in the portfolio forecast for one-year growth in earnings between t-h and t, (e) the ranked forecast error in portfolio earnings between $t-h$ and t, (f) the change in the aggregate forecast for long-term growth in earnings between $t-h$ and t, and (g) the aggregate forecast for long-term growth in earnings at t. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). All independent variables are standardized. The dependent variable in Panel B is firm-level returns (EBRs) over holding horizons (h) of one month, three months, one year, three years, and five years The independent variables in Panel B include the log dividend-price ratio. In columns 6 to 10, the EBRs predicted by the Panel A regressions are also included. All regressions include firm fixed effects. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans from December 1981 to December 2023. Superscripts: ^{a} significant at the 1% level, ^{b} significant at the 5% level, ^{c} 0% level.

Panel A: Predicting firm level expectation based returns

	(1) 1 Mo	(2) 3 Mo	(3) 1 Yr	(4) 3 Yrs	(5) 5 Yrs
$\ln(dp_{i,t})$	-0.0043^{a}	-0.0132^{a}	-0.0206^{a}	-0.0224^{a}	-0.0268^{c}
	(0.0005)	(0.0016)	(0.0054)	(0.0081)	(0.0143)
$\Delta_h LTG_{i,t}$	0.0019^{a}	0.0010	-0.0547^{a}	-0.1408^{a}	-0.1594^{a}
	(0.0003)	(0.0012)	(0.0058)	(0.0115)	(0.0235)
$LTG_{i,t-h}$	-0.0050^{a}	-0.0221^{a}	-0.1084^{a}	-0.1494^{a}	-0.1605^{a}
	(0.0007)	(0.0027)	(0.0097)	(0.0143)	(0.0263)
$\Delta_h STG_{i,t}$	-0.0107^{a}	-0.0343^{a}	0.0019	0.0133	0.0137
	(0.0007)	(0.0045)	(0.0137)	(0.0170)	(0.0222)
$STG_{i,t-h}$	-0.0117^{a}	-0.0197^{a}	0.0278^{b}	0.0312^{b}	0.0566^{b}
	(0.0010)	(0.0049)	(0.0136)	(0.0158)	(0.0246)
$FE_{i,t-h}$ rank			0.0329^{a}	-0.0473^{a}	-0.0835^{a}
			(0.0044)	(0.0083)	(0.0092)
$\Delta_h LTG_{Agg,t}$	0.0040^{a}	0.0136^{a}	-0.0039	-0.0632^{a}	-0.0288
	(0.0008)	(0.0032)	(0.0077)	(0.0173)	(0.0294)
$LTG_{Agg,t-h}$	0.0002	-0.0025	-0.0296^{a}	0.0013	-0.0691^{c}
	(0.0006)	(0.0022)	(0.0091)	(0.0245)	(0.0376)
Observations	501,342	467,941	443,350	355,688	273,440
Adj R^2	0%	2%	3%	4%	5%
F-Stat	69	41	55	24	27

Panel B: Predicted expectation based returns and actual returns

Dependent Variable: $r_{i,t+h}$

	(1) 1 Mo	(2) 3 Mo	(3) 1 Yr	(4) 3 Yr	(5) 5 Yr	(6) 1 Mo	(7) 3 Mo	(8) 1 Yr	(9) 3 Yr	(10) 5 Yr
$\ln(dp_{i,t})$	0.0014	0.0057	0.0301^{a}	0.0625^{a}	0.1195^{a}	0.0015	0.0070^{c}	0.0316^{a}	0.0544^{a}	0.0837^{a}
	(0.0014)	(0.0037)	(0.0101)	(0.0211)	(0.0273)	(0.0016)	(0.0037)	(0.0099)	(0.0184)	(0.0183)
Hat $EBR_{i,t}$						0.0363	0.1370	0.4184^{a}	0.5527^{a}	0.8322^{a}
						(0.1650)	(0.1171)	(0.0911)	(0.1364)	(0.2091)
Obs	501,342	467,941	443,334	355,657	273,400	501,342	467,941	443,334	355,657	273,400
Adj R^2	-1.0%	-1.0%	-0.8%	-0.6%	0.2%	-1.0%	-1.0%	-0.1%	1.2%	3.7%
F-stat	1.0	2.3	8.8	8.8	19.1	0.5	2.3	11.5	12.4	13.4

APPENDIX D: Characteristics, expectations and market efficiency

The next step of our analysis is a mediation exercise (MacKinnon 2012). To obtain an estimate of the share of return predictability from characteristics that works through their ability to predict analyst expectations (versus the share that works through their direct predictive ability after controlling for EBRs), we regress firm level realized returns on contemporaneous EBRs and lagged firm characteristics. The exercise shows that, to a large extent, characteristics predict returns precisely because they capture distorted expectations.

Specifically, we run the regression:

$$r_{i,t+h} = a_r + b \cdot \mathrm{EBR}_{i,t,t+h} + c_{bm} \cdot bm_{i,t} + c_{inv} \cdot inv_{i,t} + c_{size} \cdot size_{i,t} + c_{prof} \cdot op_{i,t} + c_{mom} \cdot r_{i,t-12\to t-1} + \epsilon_{t+h}, \qquad (D1)$$

where coefficients c_χ capture the predictive power of characteristic χ for returns that is independent of the firm level EBR. The predictive power of a characteristic such as book to market working through EBRs can then be quantified as $b \cdot d_{bm}$, where d_{bm} is the coefficient on the regression that predicts EBRs from characteristics (Table 11 in the text):

$$\mathrm{EBR}_{i,t,t+h} = a_{ebr} + d_{bm} \cdot bm_{i,t} + d_{inv} \cdot inv_{i,t} + d_{size} \cdot size_{i,t} + d_{prof} \cdot op_{i,t} + d_{mom} \cdot r_{i,t-12\to t-1} + \varepsilon_{t+h}. \qquad (D2)$$

Finally, the predictive power of book to market working through EBRs, $b \cdot d_{bm}$, can be compared to the independent predictive power c_{bm} of book to market alone, and similarly for other characteristics. This exercise offers a lower bound for the role of growth expectations (and their predictable errors and reversals) on the documented return predictability from characteristics: our measured analyst beliefs in fact contain only partial information about market beliefs, not only due to measurement noise, but also because we observe expectations only for specific forecast horizons.

Table D1 shows the empirical results, reporting Equation (D2) in Panel A.

Table D1

Return predictability from characteristics is mediated by expectations

Note: Panel A presents regressions of log firm-level returns at horizons (h) of one month, three months, one year, three years, and five years. The independent firm-level variables include: (a) log book-to-market ($\ln bm_{i,t}$) at time t, (b) log market value of equity at time t, (c) one-year growth in assets between $t-1$ and t ($Inv_{i,t}$), (d) operating profitability at time t ($op_{i,t}$), and (e) returns between periods $t-11$ and $t-1$ ($r_{i,t-11\to t-1}$). All specifications have firm fixed effects. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans from December 1981 to December 2023. Panel B shows the share of predictability of log firm-level returns at each horizon h accounted for by $Inv_{i,t}$ and $\ln bm_{i,t}$ as detailed in Equations (12,13) and in the text. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A: Explaining Returns

	$r_{i,t+1}$ (1)	$r_{i,t+3}$ (2)	$r_{i,t+12}$ (3)	$r_{i,t+36}$ (4)	$r_{i,t+60}$ (5)
$\text{EBR}_{i,t,t+h}$	0.1701[a]	0.2440[a]	0.4456[a]	0.5609[a]	0.6016[a]
	(0.0012)	(0.0138)	(0.0221)	(0.0227)	(0.0223)
$\ln bm_{i,t}$	0.0126[a]	0.0213[a]	0.0460[a]	0.0758[a]	0.1107[a]
	(0.0003)	(0.0054)	(0.0097)	(0.0149)	(0.0217)
$\ln size_{i,t}$	-0.0029[a]	-0.0081[a]	-0.0189[a]	-0.0191[a]	-0.0124[b]
	(0.0002)	(0.0018)	(0.0049)	(0.0043)	(0.0050)
$Inv_{i,t}$	0.0034[a]	0.0088[a]	0.0284[a]	0.0378[a]	0.0513[a]
	(0.0002)	(0.0017)	(0.0051)	(0.0084)	(0.0079)
$op_{i,t}$	0.0038[a]	-0.0364[a]	-0.1610[a]	-0.2681[a]	-0.3424[a]
	(0.0003)	(0.0068)	(0.0215)	(0.0414)	(0.0610)
$r_{i,t-12\to t-1}$	-0.0022[a]	-0.0115[a]	-0.0560[a]	-0.0646[a]	-0.0821[a]
	(0.0001)	(0.0033)	(0.0108)	(0.0120)	(0.0094)
Obs	875,404	815,293	772,217	594,792	474,032
Adj R^2	2%	7%	31%	51%	59%

Panel B: Share of predictability from characteristics via expectations

bm	113%	70%	46%	38%	43%
$size$	17%	24%	22%	16%	10%
inv	-58%	-54%	-253%	385%	237%
op	7%	-65%	102%	58%	59%
$r_{i,t-12\to t-1}$	-3%	-7%	-38%	-148%	-239%

Panel A shows that EBR has substantial explanatory power: conditioning on characteristics, b is large and significant, consistent with Table 9. The converse is also true (c_χ are large and significant) which, on its own, is consistent both with a characteristic based required return and with the earlier remark that our measures of market beliefs are partial. Momentum has the wrong sign in Table D1 panel A but the correct sign in Table 11, suggesting all of the predictability is captured by EBR.

We next compute a lower bound for the expectation-channel share of the predictability of characteristic $\chi = bm, size, inv, prof, mom$ as $\frac{b \cdot d_\chi}{b \cdot d_\chi + c_\chi}$, which are reported in Table D1 panel B. At horizons of 1 year or longer, most predictability from bm, $size$ and mom, and a substantial share of predictability from inv, works through the expectations channel.[23] There is no explanatory power for profitability, in line with the result from Table C2 that, controlling for other characteristics, dp does not predict EBRs. These results offer direct evidence that analyst expectations help explain the documented predictive power of firm characteristics for future returns.

We next complement the analysis in Table D1 concerning the link between characteristics and expectations. We first predict returns $r_{i,t+h}$ at the firm level using a saturated specification of *contemporaneous* expectations measures: forecast errors $FE_{i,t+j}$, revisions of long term forecasts $\Delta_j LTG_{i,t+j}$ and of short term forecasts $\Delta_j E_{t+j} \Delta e_{i,t+j+1}$ for $j = 1, \dots, h$. We also include measures of *lagged* aggregate optimism LTG_t, ΔLTG_{t-1} which BGLS (2024) show predict the future value premium. These regressions achieve R^2s ranging from 27% to 48%. We next regress the residual firm level returns on firm level book to market $bm_{i,t}$. Since $bm_{i,t}$ is scaled by market price, and thus includes a measure of market expectations, this sequential procedure helps better identify the component of returns that is orthogonal to the (noisier) measured expectations component. The regressions do not include time or firm fixed effects, since they seek to capture cross-sectional variation in returns arising from firm level characteristics.

Table D2.

[23] As a further test, we offer another lower bound on the role of expectations by following the residualization strategy in BGLS (2024). Specifically, we first regress returns $r_{i,t+h}$ at the firm level using a saturated specification of *contemporaneous* expectations measures. These regressions achieve R^2s ranging from 27% to 48%. We next regress the residuals of this regression on firm level book to market $bm_{i,t}$. Appendix C2 shows that the predictive power of characteristics drops dramatically in magnitude and significance and ceases to be significant for horizons of 1 year or above, once expectations are controlled for in this way. Under efficient markets, the predictive power of characteristics for returns should be unaffected.

Residualization exercise

Note: Panel A presents univariate regressions of firm level log returns at horizons (h) of one-month, three-months, one-year, three-years, and five-years on log book-to-market at time t. The dependent variable in Panel B is the residualized return from separate regressions –which are not shown-- of returns at horizon $t+h$ on forecast errors and revisions in both long-term and short-term growth between t and $t+h$. The independent variables in both panels is log book-to-market at time t. All variables are standardized. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A. Book to market and firm level returns

	Dep. Variable: $r_{i,t+h}$				
	(1)	(2)	(3)	(4)	(5)
	h =1 Mo	h = 3 Mo	h = 1 Yr	h = 3 Yrs	h = 5 Yrs
ln($bm_{i,t}$)	0.0049[a]	0.0118[a]	0.0402[a]	0.0746[a]	0.1006[a]
	(0.0012)	(0.0029)	(0.0084)	(0.0183)	(0.0235)
Constant	0.6959[a]	2.0899[a]	8.3633[a]	24.5599[a]	40.0203[a]
	(0.0015)	(0.0037)	(0.0124)	(0.0244)	(0.0302)
Obs	477,256	445,657	506,844	478,353	416,251
Adj R^2	0%	0%	1%	1%	1%
F-stat	18.1	16.8	22.9	16.6	18.3

Panel B. Book to market and firm level residual returns

	Dep. Variable: residual $r_{i,t+h}$				
	(1)	(2)	(3)	(4)	(5)
	h =1 Mo	h = 3 Mo	h = 1 Yr	h = 3 Yrs	h = 5 Yrs
ln($bm_{i,t}$)	0.0036[a]	0.0058[b]	0.0120[c]	0.0049	0.0024
	(0.0012)	(0.0028)	(0.0064)	(0.0133)	(0.0192)
Constant	-0.0020	-0.0031	-0.0063	-0.0025	-0.0012
	(0.0015)	(0.0032)	(0.0083)	(0.0178)	(0.0231)
Obs	477,256	445,657	401,537	352,841	302,816
Adj R^2	0%	0%	0%	0%	0%
F-stat	9.8	4.3	3.5	0.1	0.0

The Table shows $bm_{i,t}$ has strong predictive power for firm level returns in our sample (panel A). The predictive power of $bm_{i,t}$ drops substantially, both in magnitude and in significance, when contemporaneous measured expectations are controlled for (panel B). At the 1 year horizon, the slope coefficient drops by more than 40% which, given the limits in the measures of expectations, this is a lower bound for the role of expectations. For longer horizons, the strong predictability of $bm_{i,t}$ is entirely captured by expectations.

These results also complement the finding in BGLS (2024) that the return predictability of the aggregate price dividend ratio for returns on the aggregate market disappears once contemporaneous expectations are taken into account. These two conceptually similar exercises show that return predictability from valuation ratios is to a large extent a predictability of (differential) expectation revisions. This fact implies that characteristics encode information about beliefs, and specifically about departures from rationality.