Finance without exotic risk

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

May 26, 2025

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 from that in EBRs, holding constant scaled price variables (as proxies for time varying required returns). Third, firm characteristics often seen as capturing risk premia predict disappointment of expectations and low EBRs. Overall, return spreads typically attributed to exotic risk factors are explained by predictable movements in non-rational expectations of firms' earnings growth.

[1] Pedro Bordalo: Oxford Saïd Business School, pedro.bordalo@sbs.ox.ac.uk; Nicola Gennaioli: Università Bocconi, nicola.gennaioli@unibocconi.it; Rafael La Porta: Brown University, rafael_la_porta@brown.edu; and Andrei Shleifer: Harvard University, shleifer@fas.harvard.edu. We are grateful to Stefan Nagel, Richard Sloan, and two anonymous referees for extremely helpful comments. 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 on average zero and unpredictable, return predictability is tied to r_i. One 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 does not necessarily reject the EMH, but possibly 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.

To solve this problem, most research on cross-sectional asset pricing maintains rational expectations and views 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 distress (La Porta et al 1997). A second solution, pursued in behavioural finance, is to relax rational expectations and to generate systematic return differentials with 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, Xiaohong et al 2020, van Binsbergen et al 2023, Cho and Polk 2024). Even in this approach, however, returns are not matched to measured expectations. The joint hypothesis problem remains.

To make progress, we address the joint hypothesis problem by using analyst forecasts of future firm-level earnings growth as an empirical proxy for ΔE_{it}, while also allowing for rich models of required returns.[2] In one class of such models required returns vary only in the cross section, which allows for many characteristics to matter (Fama and French 1993, 2015, Jegadeesh and Titman 1993, Harvey, Liu, and Zhu 2015). In another class, cross-sectional required returns also vary over time (Merton 1973, Lettau and Ludvigson 2001), consistent with the evidence on predictable time variation in both the aggregate market and the cross section (Cochrane 2011, Campbell et al. 2023). We show that, for each class of models, market efficiency imposes joint restrictions on the explanatory power of required returns and of expectations. The main goal of this paper is to test these restrictions.

A key question for this approach is whether analyst forecasts are valid proxies for market expectations. In particular, if analysts mechanically extract forecasts from prices, price differences driven by required returns might be erroneously interpreted as informative about future earnings growth. We show that our tests of market efficiency are valid even if analysts engage in such inference to some extent, provided they do not rely on it exclusively.

Our tests rely on firm level *expectations-based returns* (EBRs), our proxy for ΔE_{it}. EBRs attribute all cross-sectional return variation to observed belief errors and revisions, while shutting down *any* variation in required returns. We first ask: can EBRs account for the long-short book-to-market portfolio return spread (HML, Fama and French 1993)? In an efficient market, EBRs can covary with contemporaneous portfolio returns, but they should not explain average portfolio spreads. We show instead that variation in EBRs quantitatively accounts both for much of contemporaneous return variation, but also for the average HML spread: EBRs are on average higher for value than for growth stocks, and fully explain the average spread between those portfolios at horizons ranging from 1 month to 5 years.

[2] The use of survey-based measures 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 and Xu 2022, Jiang et al 2022, Bianchi et al. 2024).

Expectations data leaves little room for risk to explain observed average return differentials. This suggests market inefficiency is at play in driving HML returns: 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 prior evidence of overreaction of long-term earnings growth forecasts or LTGs (Bordalo, Gennaioli, La Porta, and Shleifer, BGLS 2024).

We next turn to models of required returns that vary both in the cross section and over time. In these models, market efficiency imposes that future returns should be negatively predictable from current prices and positively predictable from current cash flow expectations. Crucially, this remains the case even if analyst expectations are to some extent inferred from prices and thus capture shocks to required returns – in fact, in this case predictability from prices should be even stronger. Running this test with the HML spread rejects market efficiency: scaled price variables have no predictive power while expectations strongly predict returns, and in particular expectations of long-term growth predict returns *negatively*. The evidence clearly rejects the sign restrictions entailed by the EMH. While fully characterizing the mispricing mechanism is beyond the scope of this paper, the evidence is consistent with overreaction in expectations of the long term (as in BGLS 2024) and under-reaction in expectations of the short term (as in Bouchaud et al 2019).[3]

In sum, HML is a puzzle of expectations: these explain predictable time variation in spreads, with little room left for risk. In Section 5, we show these results extend to other major factors, including size, investment, momentum, and to some extent profitability. This suggests that these characteristics capture not risk but expectations. Consistent with this view, in Section 6 we show that firm characteristics strongly predict future firm-level EBRs, capturing systematic disappointment and revisions in expectations about individual firms. W

[3] In Appendix A.2, we extend our analysis of the joint distribution of realized returns and growth expectations to account for over and underreaction in the latter. We show that, under plausible assumptions, these mechanisms can produce our empirical findings.

While it has been challenging to tie characteristics to tangible risk, they are clearly associated with errors in beliefs. In line with this non-rational belief mechanism, we further show that firm level tangible news negatively predicts future returns, as in (BGLS 2019, 2024).

The fact that systematic return differences based on characteristics reflect expectations and not risk implies that "correcting" for such factors in the computation of abnormal returns is misleading. This does not mean that asset prices only move due to expectations. They can also move due to liquidity, market frictions (e.g., Gabaix and Koijen 2023), or risk perceptions about firms or sectors. But elusive systematic risks do not appear to be needed to explain the expected return variation associated with standard characteristics.

We contribute to rapidly-growing research showing that measured expectations allow an empirically disciplined and theoretically structured approach to asset pricing. Following La Porta (1996), research showed that high LTG predicts low cross-sectional 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). Aggregate LTG is also excessively volatile, leading to volatile valuations that explain Shiller's (1981) excess volatility puzzle (Bordalo et al. 2024). In turn, short-term earnings expectations help explain the volatility of valuation ratios (De la O and Myers 2021) and the profitability anomaly (Bouchaud et al 2019).

Relative to this work, we show that non-rational measured beliefs provide a unified account of cross sectional and aggregate return puzzles without the need for unobserved time and cross-sectional variation in required returns. The primacy of expectations in these puzzles is consistent with the evidence that investors fail to attend to the co-movement of stock returns that is due to consumption shocks (Chinco, Hartzmark, and Sussman 2022).

A large body of evidence links return spreads between long and short portfolios, some of which are related to standard characteristics to systematic forecast errors or other deviations from rational forecasts (Engelberg, McClean and Pontiff 2018, Dechow and Sloan 1997, Da and Warachka 2011, Sloan and Wang 2022, De la O, Han, and Myers 2023, van Binsbergen, Han, and Lopez-Lira 2023, Frey 2023). These papers use different methods and samples to support and corroborate a common idea: return predictability reflects market inefficiency. Our paper complements this work. We develop a formal framework to address the joint hypothesis problem using expectations data, and implement it for leading cross-sectional spreads. Compared to BGLS (2024), who use fluctuations in the aggregate LTG to test for cross-sectional predictability, we show that portfolio-level expectations and forecast errors exhibit strong cross-sectional co-movement beyond market-level waves of optimism.[4]

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). Because these papers do not offer direct measures of risk or use data on expectations, they cannot reject these patterns being generated by incorrect beliefs that yield pricing biases that are horizon-dependent and time varying (Giglio and Kelly 2018, BGLS 2019, 2024).

Section 2 describes our framework and data. Section 3 shows that, within the class of models with constant required returns over time, the value premium can be fully explained by expectations, with no need for cross sectional risk premia. Section 4 shows that the value premium is predictable but also that market efficiency with time-varying required returns is rejected. Instead, predictable returns reflect systematic forecast errors in growth expectations. Section 5 extends the analysis to additional factors. Section 6 ties the evidence on factor returns to predictable corrections of expectations errors at the firm level. Section 7 concludes.

[4] Other work shows 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 these findings may reflect systematic expectation errors.

2. Concepts and Methods

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

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

where $p_{i,t}$ is the log price at t, $d_{i,t}$ its log dividend, $g_{i,t+1} = d_{i,t+1} - d_{i,t}$ its dividend growth between t and $t+1$. Constants k and ρ depend on the mean log price dividend ratio.[5] By iterating (1) forward under no bubbles and by taking expectations we obtain:

$$p_{i,t}^{e} - d_{i,t} = \frac{k}{1-\rho} + \sum_{s\geq 0} \rho^{s} \widetilde{\mathbb{E}}_{t}(g_{i,t+1+s}) - \sum_{s\geq 0} \rho^{s} \widetilde{\mathbb{E}}_{t}\left(r_{i,t+1+s}\right), \tag{2}$$

which describes the equilibrium price-dividend ratio $p_{i,t}^{e} - d_{it}$ in terms of current expectations about future dividend growth and returns. Market beliefs $\widetilde{\mathbb{E}}_{t}(.)$ are allowed to depart from rational beliefs $\mathbb{E}_{t}(.)$. Plugging (2) into (1) we obtain the realized return at time $t+1$:

$$r_{i,t+1} = \widetilde{\mathbb{E}}_{t}\left(r_{i,t+1}\right) - \sum_{s\geq 1} \rho^{s}\left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_{t}\right)\left(r_{i,t+1+s}\right) +$$

$$+\left[g_{i,t+1} - \widetilde{\mathbb{E}}_{t}\left(g_{i,t+1}\right)\right] + \sum_{s\geq 1} \rho^{s}\left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_{t}\right)\left(g_{i,t+1+s}\right). \tag{3}$$

On the top line, the realized return is higher when the time t required and expected return $\widetilde{\mathbb{E}}_{t}\left(r_{i,t+1}\right)$ is higher, or when news at $t+1$ signal higher future returns. On the bottom line, the realized return is also higher with better cash flow news at $t+1$, such as a better dividend surprise (in square brackets), or a better revision of growth expectations (the sum).

With efficient markets, expectations are rational, so news are on average zero and unpredictable based on past information, i.e. they are a martingale difference sequence. Cross section or time series predictability comes entirely from the structure of required returns

---

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

which must match expected returns $\mathbb{E}_t(r_{i,t+1})$. The precise predictability pattern hinges on the model of expected returns. Assume the following AR(1) structure (the results are more general):

$$\mathbb{E}_t(r_{i,t+1}) = (1-\eta)r_i + \eta \cdot \mathbb{E}_{t-1}(r_{i,t}) + \omega_{i,t}, \tag{4}$$

which allows for cross sectional and time series variation with persistence $\eta \in [0,1]$. To capture characteristics-based spreads, researchers have developed models of required returns in which r_i varies between portfolios sorted on firms' book to market, size, investment, profitability, and so on (so i is a portfolio of stocks with a particular characteristic). We test two main classes of these models in our analysis.

In the first class, required returns vary across characteristics but are constant over time, so $\mathbb{E}_t(r_{i,t+1}) = r_i$ with $= \omega_{i,t} = 0$. The original risk factor models (Fama and French 1993, 2015) belong to this class and allow for many characteristics to matter, including momentum (Jegadeesh and Titman 1993) and others (Cochrane 2011, Harvey, Liu, and Zhu 2015). A slight variation with $\eta = 0$ and $\omega_{i,t}$ iid white noise allows for non-persistent shocks such as portfolio-level liquidity or demand, while still allowing for stable differences in r_i.

A second class of models allows required returns to also vary over time in a persistent way, depending on market conditions or investor preferences (Merton 1973, Lettau and Ludvigson 2001, Lewellen and Nagel 2006). These models are motivated in part by the failure of the unconditional CAPM to capture average cross-sectional spreads. They also speak to the fact that average returns appear to exhibit predictable time variation in both the aggregate stock market and the cross section (Cochrane 2011, Campbell et al. 2023). Here required return shocks $\omega_{i,t}$ to characteristics-based portfolios are persistent, $\eta > 0$, creating predictable time variation. Average characteristics-based differences in r_i are still possible.

As this discussion illustrates, model (4) is very flexible. This flexibility allows to fit the data but is problematic when the connection between realized spreads and tangible risks is elusive, as is the case for characteristics-based portfolios. In line with the joint hypothesis problem, to assess market efficiency as an explanation for characteristics-based return spreads it seems necessary to measure expectations and test: i) their rationality, and ii) their ability to account for predictable returns.

As we already suggested, required returns and the prices of particular stocks, or even groups of stocks sorted on characteristics, may vary for many reasons other than expectations or characteristics-based risk factors. This is especially true in the short run, when trading can lead to return movements from market frictions (e.g., Gabaix and Koijen 2023).[6] Our analysis does not require that all price variation comes from expectations. Frictions, liquidity, or risk perceptions surely impact realized returns of sectors or stocks, including factor portfolios. Our focus here is very specific: to test for market efficiency, and in particular the connection between characteristics-based predictable return spreads and growth expectations.

To test the joint hypothesis in this context, we use measured expectations about firm cash flows. Section 2.1 outlines the logic of our tests under perfect measurement of expectations. Section 2.2 accounts for potential concerns regarding our empirical proxies.

2.1 Market efficiency tests using measured expectations

Consider first models of time-invariant required returns. With rational expectations, news in Equation (3) cannot be systematically positive or negative, so the time average of the cash flow news term satisfies $\overline{\Delta E_{it+1}} = 0$. As a result, the average (over time) of one period return $\overline{r}_{i,1}$ is simply the constant required return, $\overline{r}_{i,1} = r_i$. To test the joint hypothesis in this class, we define "Expectations Based Returns" or EBR, the part of a firm's realized stock

[6] It is straightforward to generalize (4) to model required returns featuring both a persistent component (which accommodates characteristics-based factors) and a transient component, i.e. $\mathbb{E}_t(r_{i,t+1}) = \tilde{r}_{i,t} + e_{i,t}$, where $\tilde{r}_{i,t} = \eta \cdot \tilde{r}_{i,t} + \omega_{i,t}$, where $\omega_{i,t}$ capture persistent macro risk shocks while $e_{i,t}$ capture transient liquidity shocks.

return that is due to the forecast error and expectation revisions in Equation (4), while deliberately shutting down any cross-firm variation in risk:

$$\mathrm{EBR}_{i,t+1} = r + \left[g_{i,t+1} - \tilde{\mathbb{E}}_t\left(g_{i,t+1}\right)\right] + \sum_{s\geq 1} \rho^s \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)\left(g_{i,t+1+s}\right), \quad (5)$$

Substituting (5) into (3) we decompose the realized return into the true excess returns $(r_i - r)$ and the expectations-based return, $r_{i,t+1} = (r_i - r) + \mathrm{EBR}_{i,t+1}$. Note that under the realistic assumption that even non-rational investors hold a correct expectation about the constant return r_i, this equation is also valid with inefficient markets. By differencing this equation for any long minus short Fama and French portfolio, we obtain our first test.

Time-invariant Required Returns Test. *In the regression,*

$$r_{LMS,t+1} = \alpha_{LMS} + \gamma \cdot \mathrm{EBR}_{LMS,t+1} + \nu_{LMS,t+1} \quad (6)$$

the constant captures the constant required return spread $\alpha_{LMS} = r_L - r_S$ *and* $\gamma = 1$.

The time-invariant required return is estimated by the regression constant. Any return predictability due to non-rational expectations is captured by the average EBR spread. In time invariant models of required returns, this test is appropriate for two reasons. First, under the null of efficient markets, $\mathrm{EBR}_{LMS,t+1}$ is on average zero,[7] and hence cannot account for average return spreads. Second, EBRs are orthogonal to the contemporaneous shock $\nu_{LMS,t+1}$ because, in this class of models, the latter is a zero-mean iid shock to portfolio returns. The estimated α_{LMS} thus quantifies the risk premium needed to account for observed return differentials. We expect $\alpha_{LMS} > 0$ if standard factor portfolios indeed capture risk. If instead $\alpha_{LMS} \leq 0$, the standard factor return is entirely due to non-rationality. In particular, if the inequality is strict the long arm is safer than the short arm, which is a theoretical possibility.

In the second class of models, required returns vary across firms and over time. In this case, the joint distribution of returns and expectations depends on the exact way in which

[7] Recall that under the null of rational expectations $\mathrm{EBR}_{LMS,t+1}$ is the difference between two martingales (i.e. it is a martingale difference sequence), each with mean r and unpredictable based on past information.

belief updating departs from rationality. Our focus here is to test for market efficiency, so we derive our test for the null in which the required return is allowed to vary but return and cash flow expectations are rational. Introducing expected returns (4) into the equilibrium price-dividend ratio (2), we obtain:

$$p_{i,t}^{e} - d_{i,t} = \frac{k_i'}{1-\rho} + \sum_{s\geq 0} \rho^s \mathbb{E}_t\left(g_{i,t+1+s}\right) - \frac{\mathbb{E}_t\left(r_{i,t+1}\right)}{1-\rho\eta}, \tag{7}$$

where $k_i' = k + \frac{(1-\eta)\rho}{1-\rho\eta} r_i$. We can use (7) to write $\mathbb{E}_t\left(r_{i,t+1}\right)$ as a function of $p_{i,t}^{e} - d_{i,t}$ and of current expectations of future dividend growth, which gives our second test.

Time Varying Required Return Test. *Under the null of efficient markets, the regression:*

$$r_{LMS,t+1} = \alpha + \beta \cdot \left(p_{LMS,t} - d_{LMS,t}\right) + \gamma \cdot \sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t\left(g_{LMS,t+1+s}\right) + \nu_{LSM,t+1}, \tag{8}$$

entails a joint coefficient restriction $\beta = -(1-\rho\eta) < 0$ *and* $\gamma = (1-\rho\eta) > 0$.

In efficient markets, future portfolio returns $r_{LMS,t+1}$ should be negatively predictable from the portfolio price divided ratio $p_{LMS,t} - d_{LMS,t}$, and positively predictable from cash flow growth expectations. A well-specified predictive model should control for both prices and expectations because the price dividend ratio can be high when cash flow expectations are optimistic or when required returns are low. By controlling for expectations, one can isolate required return movements, which are the only source of predictability under efficient markets. Equation (8) can only be estimated if the price dividend ratio $p_{LMS,t} - d_{LMS,t}$ and market expectations are not collinear. Given Equations (4) and (7), this is indeed the case if portfolio required returns are subject to shocks, either transient or persistent. If markets are not efficient, non-rational cash flow forecasts at t may predict future errors and revisions $\Delta E_{i,t+1}$, possibly violating the joint restriction on estimated coefficients.

Our chief goal is to use expectations data to address the joint hypothesis problem when testing for market efficiency. If the estimated coefficients reject market efficiency, two further questions arise. First, to what extent do expectations explain characteristics-based average return spreads, and their time variation? Second, what is the mispricing mechanism? While it is beyond the scope of this paper to fully characterize the latter, in Appendix A.2 we consider the impact of belief overreaction or under-reaction. Under plausible conditions, belief overreaction reduces γ in our return predictability tests of Equation (8), and may even yield $\gamma < 0$ if strong enough. Belief under-reaction instead reinforces $\gamma > 0$. The intuition is that, when beliefs overreact, high cash flow expectations are associated with excess optimism and future disappointment, predicting low returns (underreaction produces the reverse pattern). Overreaction of measured expectations of long-term cash flows and their negative predictive power for returns has been documented at both firm and aggregate levels (La Porta 1996, BGLS 2019, 2024), as has under-reaction of expectations of short-term cash flows and their positive predictive power for returns (Bouchaud et al. 2019).

2.2 Using measured expectations to test for market efficiency

We discuss how to implement these tests using measured expectations, while allowing such proxies to depart from market expectations. Data on expectations of firms' earnings growth comes from reports by equity analysts (IBES), as described in detail in the next Section. One problem with this data is that analyst forecasts may be noisier or exhibit excess sluggishness or sensitivity compared to market beliefs. This would attenuate or magnify estimated coefficients, without changing them directionally. In Appendix A.3 we desscribe a method for correcting our estimates for some forms of mis-measurement.

A more important concern is that analysts may at least in part infer earnings forecasts from stock prices, so that forecasts may surreptitiously incorporate information on required returns. The analyst may start from a model of required returns, and then retrieve a firm's cash flow forecast from the current price dividend ratio by inverting a valuation formula such as Equation (2). Based on what we know about analysts, it is implausible that they behave in this way. Producing cash flow forecasts is their main task, for which they rely heavily on accounting and management information including earnings calls (Ben-David and Chinco 2024) and aggregate economic forecasts (Decaire and Graham 2024). A large accounting literature takes analysts forecasts as genuine estimates of expected fundamentals and uses them to back out firms' cost of capital from stock prices (Lee et al 2020).[8]

That said, we show that analyst price inference can be incorporated within the framework of the previous section and, even if present, does not invalidate our tests of market efficiency. To see this, consider first the extreme view in which analysts *exclusively* rely on price inference. If they use the true model of returns, they retrieve market expectations, so our tests would still be valid. Problems arise if analysts use a wrong model for returns, in particular if they assume a constant required return r when true required returns vary across firms or over time. Analyst expectations, denoted $\widetilde{\mathbb{E}}_t^a(.)$, would then be inferred as:

$$\sum_{s\geq 0}\rho^s\widetilde{\mathbb{E}}_t^a\left(g_{i,t+1+s}\right)=p_{i,t}^e-d_{i,t}-\frac{k-r}{1-\rho}$$

$$=\frac{r}{1-\rho}+\sum_{s\geq 0}\rho^s\widetilde{\mathbb{E}}_t\left(g_{i,t+1+s}\right)-\sum_{s\geq 0}\rho^s\widetilde{\mathbb{E}}_t\left(r_{i,t+1+s}\right). \quad (9)$$

According to (9), lower required returns would cause higher analyst expectations, so that part or even all of the latter's explanatory power could be due to required returns. Analyst forecasts would also exhibit predictable forecast errors that, as shown in previous work (BGLS 2024), take the form of systematic disappointment after high optimism.[9]

---

[8] Other evidence also contradicts the price-based inference hypothesis. 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).[8] Moreover, 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).

[9] Chaudhry (2024) construct proxies for non-cash flow driven price changes using indexation events (Pavlova and Sikorskaya 2023). He finds that analyst forecasts respond to such price changes (more for the short term, less so for LTG) and interprets the result as consistent with price inference. This conclusion 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. Even absent current good cash flow news analysts may revise

The case in which analysts exclusively rely on price inference is however extreme, and inconsistent with the data. In this case: i) analyst forecasts are collinear with prices, which is strongly rejected (see Table 2 in Section 2); and ii) price changes entirely explain forecast revisions. In contrast to ii), observed earnings news greatly improve explanatory power for forecast changes (see Appendix B, Table B7). This finding is consistent with the evidence in Ben-David and Chinco (2024), Decaire and Graham (2024) and many others that analysts pay great attention to firms' earnings news when making their forecasts.

In a more realistic case, analysts partially rely on price inference, which we can formalize as follows:

$$\sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t^a\left(g_{i,t+1+s}\right) = (1-\varphi)\lambda \sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t\left(g_{i,t+1+s}\right) + \varphi\left(pd_t - \frac{k-r}{1-\alpha}\right). \tag{10}$$

Based on earnings data the analyst makes an estimate $\lambda\widetilde{\mathbb{E}}_t\left(g_{i,t+1+s}\right)$ similar to market beliefs but possibly more sluggish ($0 < \lambda < 1$) or more sensitive ($\lambda > 1$). She then mixes this estimate with price inference, with weight $0 \leq \varphi < 1$ on the latter.

Equation (10) yields testable implications for the observed relationship between analyst forecasts and stock returns. Equations (6) and (8) become (all derivations are in Appendix A.1):

$$r_{LMS,t+1} = (1-\varphi)\cdot r_{LMS} + \mathrm{EBR}_{i,t+1}^a, \tag{11}$$

$$\mathbb{E}_t\left(r_{LMS,t+1}\right) = \alpha' + \beta'\cdot\left(p_{LMS,t} - d_{LMS,t}\right) + \gamma'\cdot\sum_{s\geq 0}\rho^s\widetilde{\mathbb{E}}_t^a\left(g_{LMS,t+1+s}\right), \tag{12}$$

with $\beta' = -(1-\rho\eta)\left(1+\frac{\varphi}{(1-\varphi)\lambda}\right)$ and $\gamma' = \frac{1-\rho\eta}{(1-\varphi)\lambda}$.

Equation (11) shows that, in the class of time-invariant expected return models, price inference $\varphi > 0$ reduces the regression constant below the full required return but maintains

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.

the directional implication of a positive and significant regression constant. Only the very extreme case of full price inference $\varphi = 1$ is consistent with a constant of zero.

Equation (12) shows that, under market efficiency, the sign of coefficients in Equation (8) are robust to price inference, since $\beta' = \beta\left(1 + \frac{\varphi}{(1-\varphi)\lambda}\right)$ and $\gamma' = \gamma/(1-\varphi)\lambda$. In particular, the future return continues to be predicted negatively by the current price dividend ratio and positively by earnings expectations, with a strength that *increases* in the extent of price inference (higher φ). Again, estimating Equation (12) requires the price dividend ratio and analyst expectations to not be collinear. Equation (10) implies that this is the case for any persistent or transient return shocks uncorrelated with analyst expectations.

In arbitrating between rational models, time series predictability in portfolio spreads favours the class of models with time varying returns. But then, a violation of the sign restrictions in (12) would constitute evidence against the hypothesis that such spreads are exclusively driven by required returns. The extent and sign of expectations' predictive power is informative about the actual drivers of returns, as discussed in Sections 5 and 6.

2.3 Data and Construction of EBRs

Expectations data. We obtain firm level data on median analyst forecasts of future earnings per share and their long-term growth from the IBES Unadjusted US Summary Statistics file and use CRSP's daily share adjustment factor to adjust forecasts of earnings per share for stock splits. We use forecasts of a firm's earnings per share (EPS_{it}) and 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 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 did 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.

We follow two approaches to measuring expectations-based returns (EBRs) and testing market efficiency (particularly for the time-constant required return class). The first is to construct analyst prices that incorporate measured expectations into a dividend discount model with a constant discount rate, and to derive EBRs from them. This method disciplines the use of the term structure of expectations, and matches the method used for computing returns. The second approach is to follow Equation (4) and directly use the components of EBRs, the forecast errors and forecast revisions, as regressors. We pursue the first approach in this section and the second in Section 3.2. The two approaches yield very similar results.

The raw monthly expectations-based return $\mathrm{EBR}^{\mathrm{r}}_{i,t,t+1}$ of firm i between months t and $t+1$ is defined as:

$$\mathrm{EBR}^{\mathrm{r}}_{i,t,t+1} = \frac{D_{i,t+1} + \tilde{P}^{a}_{i,t+1}}{\tilde{P}^{a}_{i,t}}, \tag{13}$$

where $\tilde{P}^{a}_{i,t}$ is the "analyst price", constructed by discounting analyst earnings growth expectations using a firm and time-invariant return. We set this return r at the average in-sample realized annual market return, $r = 10.72\%$. We then compute (monthly) $\tilde{P}^{a}_{i,t}$ as:

$$\tilde{P}^a_{i,t} = \sum_{s=1,\ldots,5} \frac{\tilde{\mathbb{E}}^a_t DPS_{i,t+12s}}{(1+r)^s} + \frac{1+g}{(1+r)^5}\frac{\tilde{\mathbb{E}}^a_t DPS_{i,t+60}}{r-g}. \quad (14)$$

We derive expected dividends per share by multiplying expected earnings per share by the average expected payout ratio $\frac{\tilde{\mathbb{E}}^a_t DPS_{i,t+12}}{\tilde{\mathbb{E}}^a_t EPS_{i,t+12}}$ among firms in our sample that paid dividends that year, $DPS_{i,t+12} > 0$, which equals 0.41.[10] 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}.

We do not observe analyst forecasts at 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. This assumption captures the intuitive idea that at long enough horizons the analysts (and the market) expect mean reversion of a firm's earnings growth rate to a constant g calibrated to match a realistic price level. Of course, firms exhibit differences in the long run *level* of earnings. 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 $\tilde{P}^a_{i,t}$ across firms arise exclusively from differences in expectations.[11]

The firm level raw EBR is extended to the monthly raw return of a portfolio π using an equal weighted average, $\mathrm{EBR}^{\mathrm{r}}_{\pi,t,t+1} = \frac{1}{|\pi|}\sum_{i\in\pi} \mathrm{EBR}^{\mathrm{r}}_{i,t,t+1}$.[12] Our results also hold when

[10] 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%. Appendix B, Table B2 shows our results are unchanged with this measure.

[11] Given that mean reversion of growth rates occurs five years out, which is discounted by about $(1.1)^{-5} = 0.62$, our results are not very sensitive to deviations of expected medium term earnings from g, provided of course that expectations eventually revert to a value close to it. Our results are also robust to different specifications of the terminal period, for instance extrapolating LTG forecasts to an 8 or 10 year horizon, or allowing for a decay of forecasted growth rates between years 5 and 10 from LTG to g (Appendix B, Table B3).

[12] 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

using value weighted portfolios (see Appendix B, Table B4). To compute (log) returns over longer horizons t to $t+h$, we compute firm level monthly raw returns $\mathrm{EBR}^{\mathrm{r}}_{i,t+j-1,t+j}$ 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 (with some abuse of notation we use the same notation as for market EBRs, Equation (5)):[13]

$$\mathrm{EBR}_{\pi,t,t+h} = \sum_{j=1}^{h} \rho^{j-1} \ln\left(\mathrm{EBR}^{\mathrm{r}}_{\pi,t+j-1,t+j}\right) \tag{15}$$

For example, the one-month log EBR is $\ln(\mathrm{EBR}^{\mathrm{r}}_{\pi,t,t+1})$. We use log, as opposed to raw, returns for several reasons. First, when compounding over multiple periods, the use of log returns is more appropriate empirically and in line with standard practice. Second, as we show in Appendix B.4, the log transformation does not alter the martingale-difference property of EBRs under the null of rational expectations, and thus allows for a valid test of market efficiency. This holds to first approximation in the variance of the distribution of EBRs, so we also confirm empirically that our results are robust to using raw returns, particularly when computed over longer horizons (Tables B11 to B13).

We obtain EBRs for factor portfolios as the difference between the returns of the portfolio's long and short arms:

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, $\mathrm{EBR}_{HML,t,t+h} = \mathrm{EBR}_{V,t,t+h} - \mathrm{EBR}_{G,t,t+h}$.

---

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.

[13] The primary role of monthly rebalancing of portfolio EBRs is to address changes in analyst coverage over time. To make an apples-to-apples comparison we set to missing stock returns ($r_{i,t+1}$) if EBRs ($\mathrm{EBR}^{\mathrm{r}}_{i,t,t+1}$) are missing. As a robustness check, we calculated returns and EBRs for firms with uninterrupted sequences of one-, three-, and five-year returns, and the results are robust.

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, $\text{EBR}_{WML,t,t+h} = \text{EBR}_{W,t,t+h} - \text{EBR}_{L,t,t+h}$.

We refer to the generic long-short portfolio as LMS. Our sample consists of monthly firm level data from 1981 to 2023 for which LTG_t and LTG_{t+h} exist.[14] This requirement restricts the 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 B Table B1. As 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 similar to ours in all characteristics.

[14] 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).

If the approximation in Equation (14) to analyst expectations of earnings beyond 5 years out is valid, then under efficient markets (i.e., rationality of expectations) the average EBR of any portfolio is equal to $(1 + r)$. As a result, the long minus short difference in EBR has mean zero, and is again a martingale difference sequence.[15] Our second method for computing EBRs by directly regressing returns on measures of forecast errors and revisions does not rely on any assumption about long run growth expectations and hence provides an important robustness check, which confirms our key results.

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 portfolio 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. Portfolios are formed independently based on quintiles. Results are displayed for the following five quintile portfolios: (1) book-to-market, *Growth* stocks in the bottom quintile and *Value* stocks in the top quintile, (2) investment, *Aggressive* stocks in the top quintile and *Conservative* ones in the bottom 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 expectations-based returns (EBRs) computed following Equation (13) 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 expectations-based returns

[15] Formally, $\tilde{P}^a_{i,t+12n} = \sum_{s=1,\dots,5} \frac{\widetilde{\mathbb{E}}^a_{t+12n} DPS_{i,t+12n+12s}}{(1+r)^s} + \frac{1+g}{(1+r)^5} \frac{\widetilde{\mathbb{E}}^a_{t+12n} DPS_{i,t+12n+60}}{r-g}$, and so if expectations are rational the law of iterated expectations ensures that $\widetilde{\mathbb{E}}^a_t\left(\tilde{P}^a_{i,t+12n}\right) = \sum_{s=1,\dots,5} \frac{\widetilde{\mathbb{E}}^a_t DPS_{i,t+12n+12s}}{(1+r)^s} + \frac{1+g}{(1+r)^5} \frac{\widetilde{\mathbb{E}}^a_t DPS_{i,t+12n+60}}{r-g}$. Then, if the approximation $\widetilde{\mathbb{E}}^a_t DPS_{i,t+72} = (1+g)\widetilde{\mathbb{E}}^a_t DPS_{i,t+60}$ is valid, we obtain $\widetilde{\mathbb{E}}^a_t\left(DPS_{i,t+12} + \tilde{P}^a_{i,t+12}\right) = \tilde{P}^a_{i,t}(1+r)$. Using Equation (13), this implies that any portfolio has an average EBR equal to $(1+r)$.

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. Tables 4 and 8 below reveal 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) and others note that the size anomaly has weakened in recent decades relative to the earlier sample in Fama French (1993). In our sample period, average profitability spreads are not significant either.[16]

Panel B shows that EBR spreads 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, as we discuss below, the broad agreement between Panels A and B suggests that EBRs i) capture non-rational beliefs in that differences in forecast errors and revisions are predictable across portfolios, and ii) therefore may help account for cross sectional and time series variation in spreads.

[16] 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.

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 approaches exploring time variation in spreads (Lettau and Ludvigson 2001, Lewellen and Nagel 2006, Gormsen 2021, Campbell, Giglio, and Polk 2023). The "perfect proxy" benchmark of Equation (8) implies that, if the required return is truly constant, 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 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 expectations-based returns.

Note: The table presents pairwise correlations between log returns and expectations-based returns (EBRs) for quintile portfolios of stocks formed based 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. See Table 1 for a description of the portfolios.

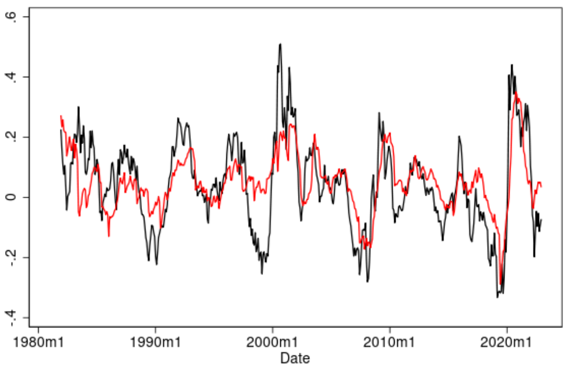
Holding Horizon	Growth	Value	Aggr.	Cons.	Big	Small	Weak	Robust	Losers	Winners
1 Month	8%	19%	10%	16%	7%	16%	12%	12%	11%	6%
3 Months	22%	35%	24%	31%	23%	30%	29%	22%	28%	25%
1 Year	36%	52%	41%	37%	35%	46%	43%	34%	48%	42%
3 Years	43%	52%	52%	36%	34%	61%	46%	37%	54%	51%
5 Years	36%	41%	43%	28%	24%	48%	39%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

The correlations between contemporaneous returns and EBRs are positive and large at long horizons, 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). 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 of Equation (3). One possible reason is that there may be significant required return variation, perhaps also capturing transient drivers of stock returns unrelated to the Fama and French factors, such as market liquidity of investor demand, especially at high frequencies.

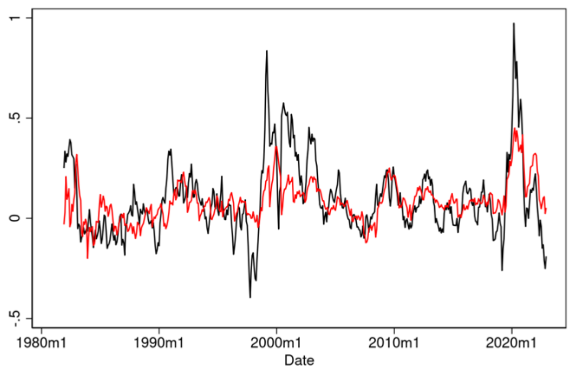
As argued in Section 2.2, the imperfect correlation between EBRs and returns suggests that analysis do not mechanically infer expectations from prices, at least not exclusively so. Price inference also implies that measures of fundamental news, such as contemporaneous cash flow growth and/or forecast errors, should not add explanatory power for EBRs controlling for returns. As we show in the Appendix (Table B7), this is also rejected: fundamental news substantially increases explanatory power for EBRs compared to contemporaneous returns. Another possible reason for the imperfect correlation between EBRs and returns is that analyst forecasts may be a noisy proxy for market beliefs (e.g. they only cover horizons up to 5 years out, causing EBRs to miss longer term variation in market expectations).[17] In Appendix A.3 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 conservative.

We conclude this section by plotting, in Figure 1, the time series correlation between EBRs and returns for the five Fama-French long minus short portfolios.

Panel A: HML



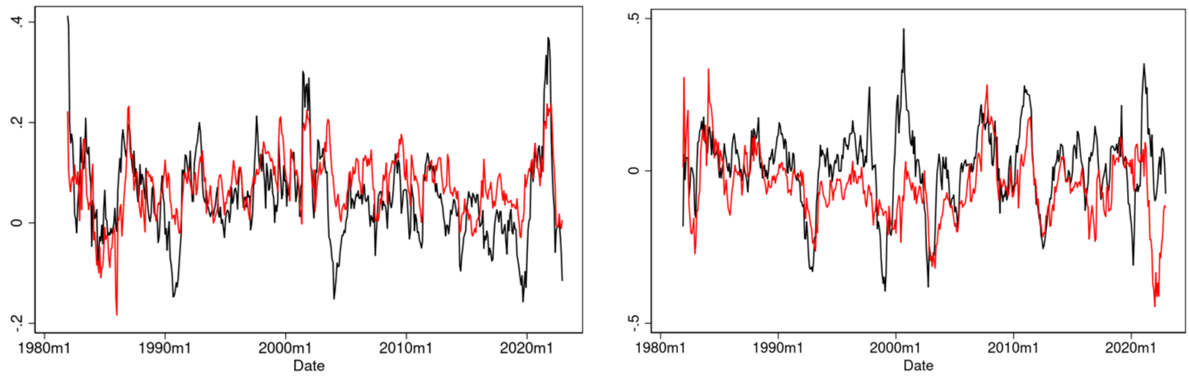
Panel B: SMB



Panel C: CMA

Panel D: RMW

[17] Measured analyst beliefs may also depart from market beliefs due to (unobserved) disagreement between the marginal investor and the analyst consensus.



Panel E: WML

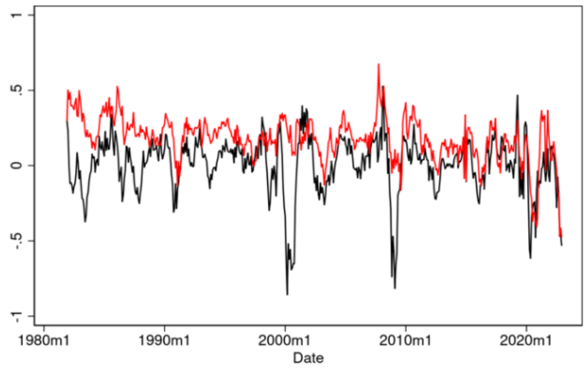


Figure 1. Actual and expectations-based return spreads

Note: Panel A plots one-year log returns (black line) and expectations-based returns (EBRs, red line) for the long-short value minus growth (HML) portfolio, where EBRs are computed following Equation (13) in the text. Panel B plots one-year log returns (black line) and EBRs (red line) for the long-short small minus big (SMB) portfolio. Panel C plots one-year log returns (black line) and EBRs for the long-short conservative minus aggressive investment (CMA) portfolio. Panel D plots one-year log returns (black line) and EBRs (red line) for the long-short robust minus weak profitability (RMW) portfolio. Panel E plots one-year log returns (black line) and EBRs (red line) for the long-short winners minus losers (WML) portfolio. The sample period is December 1981 to December 2023. See Table 1 for a description of the portfolios.

Return spreads are more volatile than EBR spreads, which may again be due either to our expectations proxy being imperfect 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 assess the extent to which the variation in EBRs can account both for average return spreads, Section 3, and their predictable time variation, Section 4. To do so, we use EBRs to conduct the market efficiency tests developed in Section 2. Studying these 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}.

3. The Value Premium and Market Efficiency with time-constant required returns

In Section 3.1, we estimate Equation (11) by using EBRs constructed for the value minus growth long-short portfolios starting from analyst prices, as in Section 2.3.[18] In Section 3.2 we estimate the same Equation by using analyst forecast errors and revisions (the components of EBRs) as separate regressors, following Equation (3). Both methods show that, after accounting for EBRs, there is little systematic variation in the value spread left for risk premia to explain. Forecast errors play a key role in this result. We extend the analysis to the other Frama and French long-short portfolios in Section 5.

3.1 EBRs Explain the Value Premium

Table 3, Panel A reports regressions of the actual value long-short portfolio spreads $r_{LMS,t+h}$ on $\text{EBR}_{LMS,t+h}$, as specified in Equation (11), for various horizons h. In this regression, the constant term captures the required return spread (given that analysts do not fully infer forecasts from prices). If the value premium is due to required returns, the regression constant should be positive and significant.[19] To deal with overlapping observations for horizons $h > 1$ month, we correct standard errors using the Newey-West (1987) procedure.

Table 3
EBRs and returns and for the HML portfolio

Note: Panel A presents univariate regression results of log returns for the long-short value minus growth (HML) portfolio on expectations-based returns for that portfolio, $\text{EBR}_{HML,t,t+h}$. Separate regressions are estimated for

[18] In Appendix A, we further correct the estimated for the possibility that analyst expectations may be a noisy proxy of market beliefs about earnings growth. Our empirical analysis shows that such discrepancies indeed exist but that accounting for them usually makes little difference for the entailed required return differential $r_L - r_S$.

[19] 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.

horizons h of one-month, three-month, one-year, three-year, and five-year. Panel B extends the analysis by adding EBRs 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 from December 1981 to December 2023. See Table 1 for a description of the portfolios Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Panel A
Regression of HML returns on HML EBRs

	Dep. variable: $r_{HML,t,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$\text{EBR}_{HML,t,t+h}$	0.5067[a]	0.8313[a]	1.0274[a]	1.1719[a]	1.1723[a]
	(0.1527)	(0.1638)	(0.1156)	(0.2274)	(0.1842)
Constant	0.0032[c]	0.0052	-0.0055	-0.0269	-0.0308
	(0.0018)	(0.0046)	(0.0140)	(0.0462)	(0.0424)
Obs	504	502	493	469	445
Adj R^2	4%	16%	46%	45%	50%

Panel B
Regression of HML returns on HML EBRs and market EBRs

	Dep. variable: $r_{HML,t,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$\text{EBR}_{HML,t+h}$	0.3965[b]	0.7954[a]	1.0884[a]	1.2214[a]	1.1425[a]
	(0.1561)	(0.1732)	(0.1452)	(0.2369)	(0.2260)
$\text{EBR}_{Mkt,t+h}$	0.3438[c]	0.0824	-0.1408	-0.2122	-0.3558[b]
	(0.1813)	(0.1511)	(0.1374)	(0.1629)	(0.1676)
Constant	0.0002	0.0032	0.0065	0.0344	0.1591[c]
	(0.0026)	(0.0059)	(0.0167)	(0.0661)	(0.0893)
Obs	504	502	493	469	445
Adj R^2	6%	16%	47%	47%	57%

EBRs and actual returns are significantly positively correlated for the HML portfolio, especially at longer horizons. The R^2 also sharply rises with the horizon: it is 4% 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 (11), the estimated coefficient on EBRs is close to one, and statistically indistinguishable from it, at most horizons. It is smaller than one at the monthly horizon, arguably because analyst forecasts are a noisier proxy for market beliefs at higher frequencies.

Crucially, the regression constant capturing the HML premium is small in magnitude and statistically indistinguishable from zero at all horizons over one month. After accounting for the average difference in "perceived cash flow news" with EBRs, there is no systematic value spread left for risk to explain. Our first test rejects the joint hypothesis of market efficiency and constant required returns. This conclusion holds even if we allow for analysts to form expectations in part by inferring their forecasts from prices.

BGLS (2024) show that the HML spread is also predicted by lagged aggregate optimism, as measured by high expectations of long-term aggregate earnings growth, high LTG_t, suggesting that at least part of the value spread is driven by a predictable, aggregate expectations 'factor'. In Table 3 Panel B we run a horse race between the market-level contemporaneous expectations-based return $\text{EBR}_{Mkt,t,t+h}$ and the portfolio one $\text{EBR}_{HML,t,t+h}$, in accounting for the contemporaneous return spread.[20] The results indicate that the value premium is largely due to cross sectional cycles in expectations, as proxied by $\text{EBR}_{HML,t,t+h}$. Expectational boom and bust patterns obtain in the cross section, just as they do in the aggregate stock market (BGLS 2024).

3.2 Decomposing EBRs into Forecast Errors and Revisions

Instead of computing EBRs, we can use Equation (3) to 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 different belief components (which are linked to distinct belief biases). Second, it relaxes the parametric restrictions embedded in our computation of EBRs and the assumptions about long-term expectations. This constitutes a robustness check and allows forecasts (available up to 5 years ahead) to capture correlated variation in unmeasured longer-term beliefs.

[20] 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.

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 $FE_{i,t+h} = \ln\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right) - \ln\left(\widetilde{\mathbb{E}}_t^a\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right)\right)$. At five-year horizons, we compute the forecast error using LTG as $FE_{i,t+60} = \ln\left(\frac{EPS_{i,t+60}}{EPS_{i,t}}\right)/5 - LTG_{i,t}$. We compute revisions of short-term growth forecasts, $h = 12{,}24$, as $\Delta_h STG_{i,t+h} = \left(\widetilde{\mathbb{E}}_{t+h}^a - \widetilde{\mathbb{E}}_t^a\right) ln\left(\frac{EPS_{i,t+h+12}}{EPS_{i,t+h}}\right)$ and revisions of long-term forecasts as $\Delta_h LTR_{i,t+h} = LTG_{i,t+h} - LTG_{i,t}$.

Under rational expectations, forecast errors $FE_{i,t+h}$ and revisions $\Delta_h STG_{i,t+h}$ are mean zero (formally, a martingale difference), so they meet the assumptions for testing market efficiency with Equation (11). The case of LTG revisions is slightly more complicated because LTG captures forecasts with over horizons of 3 to 5 years. Thus, $\Delta_h LTG_{i,t+h}$ includes a component of revisions for fixed horizons (up to year $t + 4$ when considering yearly LTG changes), but also a shift in the forecast horizon (year $t + 5$). The shift in the forecast horizon implies that even under rational expectations $\Delta_h LTG_{i,t+h}$ is not strictly a martingale difference sequence, at least when considering revisions over one year or more (the problem is arguably limited at the monthly horizon). Therefore, as a robustness check we estimate Equation (11) omitting LTG revisions. Our results are robust because much explanatory power comes from pure forecast errors and revisions (see Appendix B2).[21]

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 analogously for forecast revisions. We use the differences in these aggregated forecast errors and revisions between

[21] This concern does not apply to the analysis on return predictability in Section 4.

the long and short portfolios as explanatory variables for contemporaneous long minus short return spreads.[22] Table 4 shows the results.

Table 4
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) portfolio for horizons (h) of one-year, three-years, and five-years, $r_{HML,t,t+h}$. The independent variables include: (a) spreads in forecast errors between t and $t+h$, $FE_{HML,t+h}$ (defined in the text), (b) spreads in forecast revisions for short-term earnings growth between t and $t+h, \Delta_h STG_{HML,t+h}$ (defined in the text), and (c) spreads in changes in long-term growth forecasts between t and $t+h$, $\Delta_h LTG_{HML,t+h}$. All independent variables have unit standard deviation. 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.

	Dep. variable: $r_{HML,t,t+h}$		
	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)
$FE_{HML,t+h}$	0.1318[a]	0.1743[a]	0.1413[a]
	(0.0134)	(0.0261)	(0.0376)
$\Delta_h STG_{HML,t+h}$	0.0808[a]	0.0575[b]	-0.0240
	(0.0127)	(0.0248)	(0.0289)
$\Delta_h LTG_{HML,t+h}$	0.0276[a]	0.0182	0.0665[c]
	(0.0094)	(0.0207)	(0.0346)
Constant	-0.1111[a]	-0.2847[a]	-0.3308[a]
	(0.0213)	(0.0634)	(0.0872)
Obs	493	469	445
Adj R^2	53%	48%	45%

The expectation components are strongly predictive of HML spreads, with forecast errors playing a dominant role in the multivariate regression (note that the regressors have unit standard deviation). The pre-eminence of forecast errors for explaining HML suggests that returns reflect a disappointment in the short arm compared to the long arm, i.e. of growth stocks compared to value stocks. Portfolio expectations revisions also play a role, so part of the HML spread is accounted for by 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 our previous findings, the constant terms in Table 4 suggest that the average returns spreads are entirely explained by expectations. The regression constants in

[22] 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 $EPS_{i,t+h}$.

Table 4 are all negative, and often statistically significant, contrary to the hypothesis of a positive time-constant required return spread between value and growth stocks.

In sum, the systematic return spreads on the Fama French HML factor are explained by EBRs, and in particular by systematic differences in forecast errors and revisions (Tables 3 and 4). Taking these into account, there is little evidence of systematic differences in required returns. Growth stocks do worse when optimism for them falls relative to that for value stocks, and they do worse on average because such optimism happens more than its reverse, in line with La Porta et al (1997). In principle, the expectations mechanism implies that the value spread may be predictably negative when the prospects of high growth firms are underestimated relative to those of value firms. We later assess this possibility.

4. The Value Premium and Market Efficiency under time varying required returns

We next allow required returns to vary over time. This case is important because spreads exhibit dramatic variation over time, as shown in Figure 1. This variation has led many researchers to compare time variation in required returns with systematic mispricing (Merton 1973, Lettau Ludvigson 2001, Lewellen Nagel 2006, Lochstoer and Tetlock 2020, Campbell, Giglio, and Polk 2023). Second, and crucially, previous work has shown that analysts' forecasts of earnings growth are not rational: lagged forecasts predict future errors and revisions (BGLS 2019, 2024). We can then assess whether nonrationality of analyst forecasts can account for predictable time variation in return spreads using the tests developed in Section 2.2.

We test the efficient markets null by estimating Equation (12), which amounts to running a horse race between *current* measured expectations $\sum_{s\geq 0}\rho^s\widetilde{\mathbb{E}}_t^a\left(g_{HML,t+1+s}\right)$ and price dividend ratio $pd_{LMS,t} = pd_{H,t} - pd_{L,t}$ in predicting *future* return spreads $r_{HML,t+1}$. If the market is efficient and analysts partially infer expectations from prices under a wrong required return model, the price dividend ratio should strongly negatively predict future return spreads, while analyst expectations should positively predict future returns. The test rejects market efficiency if either of these restrictions fails.

As described in Section 2, we capture measured expectations with short-term growth forecasts $STG_{HML,t}$ and long-term growth forecasts, $LTG_{HML,t}$ (expectations for longer horizons are not available). Following BGLS (2024), we decompose these expectations in terms of lagged levels, $LTG_{HML,t-h}$, and current revisions $\Delta_h LTG_{HML,t} = LTG_{HML,t} - LTG_{HML,t-h}$, which provide two measures of optimism, and similarly for short-term forecasts, $STG_{HML,t-h}$ and $\Delta_h STG_{i,t} = \left(\widetilde{\mathbb{E}}_t^a - \widetilde{\mathbb{E}}_{t-h}^a\right) ln\left(\frac{EPS_{i,t+h+1}}{EPS_{i,t+h}}\right)$. We also include lagged forecast errors, $FE_{HML,t-h,t}$, which are serially correlated under either over- or under-reaction.[23] Finally, we also include other measures of current optimism, namely levels and revisions of forecasts of aggregate growth, which BGLS (2024) show predict future HML spreads. Independent variables have unit standard deviations to allow for a quantitative comparison of explanatory power.

Table 5 presents the results for HML long minus short portfolio. The last row of Table 5 presents the R^2 of a univariate regression of $r_{HML,t,t+h}$ on $pd_{HML,t}$.

Table 5
Predicting future return spreads from expectations data

Note: This table presents regressions of log returns for the long-short value minus growth (HML) portfolio $r_{HML,t,t+h}$. Separate regressions are estimated for horizons (h) one-month, one quarter, and one, three and five years. The set of independent variables includes: (a) the portfolio log price-dividend ratio at time t, $pd_{HML,t}$, (b) the change in the portfolio forecast for long-term earnings growth between $t-h$ and t, $\Delta_h LTG_{HML,t}$, (c) the lagged portfolio forecast for long-term earnings growth at $t-h$, $LTG_{HML,t-h}$, (d) the change in the portfolio forecast for short-term earnings growth between $t-h$ and t, $\Delta_h STG_{HML,t}$ (defined in the text), (e) the lagged portfolio forecast for short-term earnings growth at $t-h$, $STG_{HML,t-h}$, (f) the portfolio forecast error in earnings between $t-h$ and t, $FE_{HML,t-h,t}$ (defined in the text), (g) the change in the aggregate forecast for long-term aggregate earnings growth between $t-h$ and t, $\Delta_h LTG_{Mkt,t}$, and (h) the forecast for long-term aggregate earnings growth at $t-h$, $LTG_{Mkt,t-h}$. All independent variables have unit standard deviation. The last row reports the R^2 from a univariate regression of $r_{LMS,t+h}$ on $pd_{LMS,t}$. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period is 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.

---

[23] Lagged forecast errors $FE_{HML,t-h,t}$ are defined analogously to the contemporaneous forecast errors $FE_{HML,t+h}$ defined above Table 4 (the notation makes the forecasted period explicit).

	Dep variable: $r_{HML,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$pd_{HML,t}$	0.0057	-0.0301	-0.0256	-0.0273	-0.3057[b]
	(0.0187)	(0.0454)	(0.1778)	(0.2592)	(0.1451)
$\Delta_h LTG_{HML,t}$	0.0045[c]	0.0093	-0.0864[b]	-0.0556[c]	0.0438
	(0.0026)	(0.0074)	(0.0350)	(0.0332)	(0.0320)
$LTG_{HML,t-h}$	-0.0190	-0.0395	-0.2779[c]	-0.4239	-0.2472
	(0.0143)	(0.0364)	(0.1669)	(0.2733)	(0.1980)
$\Delta_h STG_{HML,t}$	-0.0022	0.0017	0.0459[a]	0.0808[a]	0.1061[a]
	(0.0023)	(0.0075)	(0.0171)	(0.0256)	(0.0339)
$STG_{HML,t-h}$	0.0158	0.0676[b]	0.2209[a]	0.4318[a]	0.2701[c]
	(0.0105)	(0.0341)	(0.0645)	(0.0994)	(0.1382)
$FE_{HML,t-h,t}$			0.0424[c]	0.0329	-0.1039[a]
			(0.0236)	(0.0422)	(0.0304)
$\Delta_h LTG_{Mkt,t}$	-1.3527[c]	0.6327	1.1462	2.4281	8.1230[a]
	(0.7412)	(0.9656)	(1.3897)	(1.6632)	(1.6512)
$LTG_{Mkt,t-h}$	0.1895[c]	0.5294	1.4495	6.5183[a]	6.2778[a]
	(0.1003)	(0.3283)	(1.3405)	(1.8266)	(2.1769)
Constant	-0.0401[b]	-0.1147[b]	-0.4006[b]	-1.0748[a]	-1.2718[a]
	(0.0164)	(0.0533)	(0.1875)	(0.3408)	(0.3522)
Obs	444	442	433	409	385
Adjusted R^2	5%	7%	18%	26%	69%
Univariate R^2	0%	1%	3%	1%	5%

Contrary to the efficient markets hypothesis, and accounting for "price-based inference", the price dividend ratio is mostly insignificant when controlling for expectations. The sign is often negative, consistent with Equation (12), but is only marginally significant at the 5 years horizon. This implies that some required return variation may exist, but it is too small to meaningfully contribute to the predictability of the portfolio return spreads. Expectations have instead a strong predictive power for the future HML return spread, even at the short 1-month horizon, a challenging test, and increasing for longer horizons.

The dividend price ratio's explanatory power in univariate regressions is negligible, as seen by the univariate R^2 shown in the last row. Adding expectations leads to a much higher adjusted R^2 in all specifications, for example increasing from 3% to 18% at 1-year horizon, and from 5% to 69% at a 5-year horizon. These results are robust to using other valuation ratios, such as book to market (Appendix C).[24, 25]

Importantly, and again inconsistent with market efficiency, long-term growth forecasts negatively predict future returns. This is consistent with overreaction, in line with previously documented analyst long-term overreaction to news (BGLS 2019, 2024): given good news on the long arm of the portfolio compared to the short one, analysts become too optimistic about the long-term growth of the HML portfolio, which inflates the current portfolio price and generates predictably low returns as expectations are corrected.[26]

Short-term expectations instead predict returns positively. A positive coefficient is not a symptom of belief distortions per se (see the positive coefficient on expectations in Equation (12)) but, given the lack of predictive power of the dividend price ratio, it is suggestive of a belief distortion: specifically, under-reaction of short-term growth forecasts. After receiving positive cash flow news for the HML portfolio, analysts revise their near-term expectations too little, which predicts upward expectations revisions and hence positive HML returns going forward. Short-term under-reaction is also consistent with prior independent evidence from Bouchaud et al. (2019), who document underreaction of short-term analyst expectations and use it to account for the profitability anomaly.

---

[24] These results are 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, as well as with the book-to-market ratio. The results are very similar, see Table C3 in Appendix C.

[25] Nagel (2024) raises the possibility that 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, and therefore also implausible for EBRs.

[26] Note that the specification in Table 5 is very demanding: whether or not market efficiency holds, and for any level of price inference, the *pd* variable also reflects measured expectations. It is therefore not surprising that coefficients on expectation variables are not statistically significant in all specifications. In Table C1 in the Appendix, we run the same specification of Table 5 but omit the pd ratio. The evidence on overreaction of LTG is considerably strengthened, and in particular the lagged level of LTG significantly, and negatively, predicts returns at all horizons.

Taken together, our results show that returns are predictable from 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, which would entail negative return predictability from the price dividend ratio.

If market efficiency is rejected, how do current expectations predict future returns? The previous discussion suggests that return spreads may be explained, at least in part, by the unfolding of predictable market expectations errors as proxied by predictable forecast errors and revisions, for both short- and long-term forecasts, incorporated into EBRs. To directly assess this mechanism, we perform a two-step exercise. In the first stage, we use the variables in Table 5 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 $\widehat{\text{EBR}}_{HML,t\to t+1}$, to explain contemporaneous returns, again controlling for the price dividend ratio:

$$r_{H,t+1} - r_{L,t+1} = \alpha + \beta \cdot \widehat{\text{EBR}}_{HML,t\to t+1} + \gamma \cdot pd_{HML,t} + v_{t+1}$$

Compared to Equation (12), this test ties return differentials to error predictability, the hallmark of non-rationality, and allows a quantitative assessment of predicted EBRs.

The first stage R^2 is presented in the last row of Table 6 (see Appendix C, Table C4 for the full first stage results). EBRs are strongly predicted by lagged expectations, reflecting systematic predictability of forecast revisions and errors. As in Table 5, these capture overreaction of beliefs of long-term growth and under-reaction of beliefs of short-term growth, especially for horizons of a year and longer. Second stage results appear in Table 6.

Table 6
Non-rational expectations and predictable returns

Note: This table presents second-stage regressions for the long-short value minus growth (HML) portfolio, $r_{HML,t,t+h}$. The first-stage regressions –reported in Table C4-- predict long-short EBRs using the independent variables in Table 5. In the second stage we regress log returns on predicted EBRs and the log price-dividend ratio ($pd_{HML,t}$). The price dividend ratio $pd_{HML,t}$ has unit standard deviation. Separate regressions are estimated for each horizon: one-month, one quarter, one-, three- and five years. The last row of the table presents the R^2s from the first stage regressions. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{HML,t,t+h}$				
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$
	(1)	(2)	(3)	(4)	(5)
$\widehat{EBR}_{HML,t,t+h}$	1.3267[a]	0.9367[a]	1.0656[a]	1.5689[a]	1.6707[a]
	(0.4452)	(0.3548)	(0.2814)	(0.4504)	(0.3362)
$pd_{HML,t}$	-0.0131	-0.0494[c]	-0.1343[b]	-0.1231	-0.2950[b]
	(0.0097)	(0.0293)	(0.0651)	(0.1405)	(0.1195)
Constant	-0.0039	-0.0145	-0.0577[b]	-0.1333[c]	-0.2169[a]
	(0.0040)	(0.0125)	(0.0254)	(0.0795)	(0.0720)
Obs	444	442	433	409	385
Adjusted R^2	4%	6%	18%	24%	55%
1[st] stage R^2	14%	28%	32%	34%	42%

Table 6 shows that actual spreads load strongly on predicted EBRs, with all coefficients highly significant and indistinguishable from 1 for horizons up to a year. Coefficients are above one for longer horizons (suggesting that for such horizons current expectations have stronger predictive power for future returns than that captured in predicted EBRs). The loadings on the price dividend ratio are instead small compared to those of EBRs and often insignificant or marginally significant. Intercepts for actual return spreads are negative and mostly insignificant, confirming that EBRs, and in particular their predictable component, can account for the observed spreads in actual returns.[27]

These results reject a market efficiency explanation of the value premium, consistent with the view that return spreads capture, at least in part, the unfolding of predictable errors

[27] Appendix C extends the analysis in Table 6 to the firm level (Table C.4): we predict firm level EBRs using the expectation variables in Table 5, 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. This firm-level analysis adds to the evidence that expectations are not spuriously capturing information about required returns.

of market expectations, as proxied by predictable EBRs: excess optimism about a portfolio is systematically disappointed in the future, leading to low predictable returns. Consistent with this view, Engelberg et al (2018) show that such cross-sectional returns accrue mainly on cash flow news events when forecast errors materialize, in line with measured expectations capturing market expectations and being systematically surprised. [28]

While this mechanism generates outperformance of value portfolios after periods of high optimism about growth stocks, it can also produce the reverse; at times of strong relative optimism about value stocks, or equivalently pessimism about growth, future predicted EBRs can be negative. This finding is inconsistent with any model of a risk averse marginal investor in which characteristics capture required returns, even of time varying magnitude.

In sum, the analysis of expectations data and their role in return predictability is hard to square with the hypothesis of efficient markets with time varying required returns, even accounting for "price-based inference." Measured expectations contain genuine information about non-rational market beliefs that affect prices and help predict future returns.

5. Other standard factors

We repeat the analysis of Sections 3 and 4 for the size, investment, profitability and momentum anomalies. Table 7 presents our estimates of equation (8) for these factors.

Table 7
Actual and expectations based long short portfolio return spreads

Note: This table presents univariate regression results for log returns, $r_{LMS,t,t+h}$, against expectations-based returns (EBRs) for four distinct long-short (LMS) portfolios. The portfolios examined are: (1) SMB, which is long stocks in the bottom decile of market capitalization and short stocks in the top decile, (2) CMA, which is long stocks in lowest quintile of one-year asset growth and short stocks in highest quintile, (3) RMW, which is long stocks in the highest quintile of operating profitability and short stocks in the lowest quintile, and (4) WML, which is long stocks in the top quintile of returns during period $t-11$ through $t-1$ and short stocks in

[28] 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 (3). 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.

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.

	Dep. variable: $r_{LMS,t,t+1}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
Size (SMB)	(1)	(2)	(3)	(4)	(5)
$EBR_{LMS,t,t+h}$	0.5698[a]	0.7813[a]	1.1470[a]	1.1858[a]	1.0915[a]
	(0.1304)	(0.1559)	(0.1779)	(0.1787)	(0.2321)
Constant	0.0027	0.0092[c]	0.0177	0.0365	0.0698
	(0.0017)	(0.0049)	(0.0134)	(0.0291)	(0.0485)
Adj R^2	5%	12%	34%	55%	49%
Investment (CMA)	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
$\text{EBR}_{LMS,t,t+h}$	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)	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
$\text{EBR}_{LMS,t,t+h}$	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)	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
$\text{EBR}_{LMS,t,t+h}$	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, EBRs returns have strong explanatory power for actual spreads across these portfolios. The slope coefficients are large, statistically significant, and increase with the holding horizon. For SMB, CMA and RMW, magnitudes are comparable to those obtained for HML returns: for size and 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.

These patterns are confirmed in the EBR decomposition exercise, reported in Appendix B (Table B6): 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 and momentum, as well as for HML as shown in Table 4).

Finally, we perform our analysis of return predictability, i.e. testing the null of market efficiency with time varying required returns, on the size, investment, profitability and momentum following Equation (12). The results mirror those for HML (Tables 5 and 6): current expectations also have overwhelming predictive power for future return spreads on these portfolios, while the portfolios price-dividend ratio has no predictive power, except for size and momentum at the 5-year horizon (see Table C.1; Table C.2 repeats the analysis using bm and lagged returns as proxies for required returns). Table 8 shows the horse race between the portfolios' price dividend ratio and predicted EBRs, as in Table 6. For brevity, we focus on the one-month and one-year horizons.

Table 8
Non-rational expectations and predictable returns

Note: This table presents two-stage regressions of log returns for for four distinct long-short (LMS) portfolios, $r_{LMS,t,t+h}$. The portfolios examined are: (1) SMB, which is long stocks in the bottom decile of market capitalization and short stocks in the top decile, (2) CMA, which is long stocks in lowest quintile of one-year asset growth and short stocks in highest quintile, (3) RMW, which is long stocks in the highest quintile of operating profitability and short stocks in the lowest quintile, and (4) 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. The first stage regressions – reported in Table C4 – predict long-short EBRs using the independent variables in Table 5. In the second stage we regress returns on predicted EBRs and the log price-dividend ratio, $pd_{LMS,t}$. The price dividend ratio $pd_{LMS,t}$ has unit standard deviation. Separate regressions are estimated for one-month, and one- year horizons. The last row of the table presents the R^2s from the first stage regressions. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{SML,t,t+h}$		$r_{CMA,t,t+h}$		$r_{RMW,t,t+h}$		$r_{WML,t,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{EBR}_{LMS,t,t+h}$	1.4470[a]	0.8375[b]	0.9758[a]	1.3952[a]	0.9479	-0.1958	1.1135[b]	0.8337[a]
	(0.4084)	(0.3716)	(0.2722)	(0.3535)	(0.7012)	(0.2713)	(0.5053)	(0.2897)
$pd_{LMS,t}$	-0.0002	-0.0588	-0.0120	-0.1073[b]	-0.0135	0.0127	-0.0058	-0.0746
	(0.0088)	(0.0813)	(0.0079)	(0.0427)	(0.0119)	(0.0635)	(0.0077)	(0.0577)
Constant	0.0035	0.0762	-0.0065[c]	-0.1038[a]	0.0015	0.0021	-0.0397[c]	-0.1582[a]
	(0.0070)	(0.0681)	(0.0037)	(0.0333)	(0.0033)	(0.0286)	(0.0205)	(0.0595)
Obs	444	433	444	433	444	433	444	433
Adjusted R^2	2%	12%	4%	24%	0%	0%	2%	18%
1st stage R^2	9%	46%	14%	17%	5%	30%	7%	51%

As in Table 6, for all portfolios except profitability, the coefficients on the predicted $\widehat{EBR}_{LMS}$ are highly significant and indistinguishable from 1, while those on the price to dividend ratios are small and insignificant. The size, investment, and momentum puzzles are also 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.

6. Characteristics and return predictability at the firm level

If portfolio returns do not reflect systematic risk exposure, the rationale for portfolio level analysis diminishes because non-rational expectations can drive return predictability also at the firm level. Here we take the first steps in this direction. We ask two questions: first, if firm level returns are driven by idiosyncratic firm level beliefs, why should standard firm characteristic predict returns? Second, given that forecast errors and revisions – and thus future returns – are predictable from current news, can we tie return predictability to cash flow news?

One answer to the first question is that characteristics may capture non-rational beliefs rather than an elusive connection to risks. In this case, characteristics should also predict future firm level EBRs. To assess this hypothesis, Table 9 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 may absorb the role of the measured characteristics themselves.

Table 9
Characteristics predict firm level EBRs

Note: This table presents regressions of firm level log expectations-based returns at horizons (h) of one-month, three-months, one-year, three-years, and five-years, $EBR_{i,t,t+h}$. The independent variables include: (a) log book-to-market ratio ($bm_{i,t}$) 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, $size_{i,t}$, (d) operating profitability at time t, $op_{i,t}$, and (e) returns between periods $t-12$ and t, $r_{i,t-12,t}$. All specifications have firm fixed effects. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. Except for $r_{i,t-12,t}$, all independent variables have unit standard deviation. 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.

	Dep. variable: $EBR_{i,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(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)
$r_{i,t-12,t}$	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.[29] 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.

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). We find that, to a large extent, characteristics predict returns precisely because they capture distorted expectations (Table D1). Following BGLS (2024), we also 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 for longer horizons, the strong predictability of $bm_{i,t}$ is entirely captured by expectations (Table D2).

Finally, since expectations are not rational, future forecast errors and revisions – and therefore future returns – should be predictable from current news. These could include intangible news, as in Daniel and Titman (2006), but also tangible news such as growth in earnings or sales, as in BGLS (2024). The latter possibility is particularly informative because, by directly linking future returns to current news, it constitutes clear evidence of the expectations channel. To assess it, we regress firm level returns on lagged news, proxied by

[29] 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.

growth in sales (Table 10, panel A) or growth in earnings (panel B), crucially controlling for a proxy for required returns such as lagged book-to-market or lagged return.[30]

Table 10

Tangible news predict future returns at the firm level

Note: This table presents firm-level regressions of log returns across holding horizons (h) of one month, one year, and five years. Panel A includes as independent variables: (a) log sales-per-share growth between t-h and t $\Delta_h sps_{i,t-h,t}$, (b) log firm-level returns between t-h and t, $r_{i,t-h,t}$, and (c) log book-to-market ratio, $bm_{i,t}$. Panel B replaces $\Delta_h sps_{i,t-h,t}$ with log earnings-per-share growth between t-h and t, $\Delta_h eps_{i,t-h,t}$, as an independent variable, while retaining log firm-level returns and the log book-to-market ratio. All specifications include firm and time fixed effects. Except for lagged returns, all independent variables have unit standard deviation. Standard errors are corrected for overlapping observations and cross-sectional dependence using the Driscoll and Kraay (1998) procedure. The sample includes all firms on CRSP and COMPUSTAT during the period spanning December 1981 to December 2023. Superscripts indicate significance levels: [a] at 1%, [b] at 5%, and [c] at 10%.

Panel A. Lagged growth in sales per share

	Dep. variable: $r_{i,t,t+h}$								
	$h = 1$			$h = 12$			$h = 60$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta_h sps_{i,t-h,t}$	0.0027[a]	0.0027[a]	0.0030[a]	-0.0202[a]	-0.0168[a]	-0.0103[a]	-0.1045[a]	-0.0828[a]	-0.0735[a]
	(0.0004)	(0.0004)	(0.0004)	(0.0030)	(0.0029)	(0.0024)	(0.0085)	(0.0110)	(0.0106)
$r_{i,t-h,t}$		-0.0049[a]			-0.0334[a]			-0.1199[a]	
		(0.0014)			(0.0113)			(0.0138)	
$bm_{i,t}$			0.0078[a]			0.0837[a]			0.2811[a]
			(0.0013)			(0.0125)			(0.0400)
Obs	1,210,205	1,210,205	1,206,866	1,086,584	1,085,718	1,085,492	629,062	625,433	628,507
Adj R^2	-1%	-1%	-1%	-1%	0%	1%	1%	3%	6%

Panel B. Lagged growth in earnings per share

	Dep. variable: $r_{i,t,t+h}$								
	$h = 1$			$h = 12$			$h = 60$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta_h eps_{i,t-h,t}$	0.0633[a]	0.0659[a]	0.0664[a]	0.0666[a]	0.0700[a]	0.0718[a]	-0.0006	0.0126[b]	0.0157[a]
	(0.0027)	(0.0027)	(0.0026)	(0.0034)	(0.0035)	(0.0032)	(0.0060)	(0.0057)	(0.0051)
$r_{i,t-h,t}$		-0.0422[a]			-0.0335[a]			-0.0561[a]	
		(0.0071)			(0.0083)			(0.0168)	
$bm_{i,t}$			0.0096[a]			0.0263[a]			0.0912[a]
			(0.0010)			(0.0029)			(0.0116)
Obs	1,007,395	1,007,395	1,005,149	985,406	985,326	985,087	871,936	871,443	871,273
Adj R^2	-1%	0%	0%	0%	0%	1%	-1%	0%	1%

[30] The results carry through when controlling for the dividend price ratio, but since firms do not always pay dividends, this variable has lower coverage at the firm level.

Table 10 confirms that strong performance of a firm, as measured by sales growth, negatively predicts returns. Linking returns to lagged tangible news provides direct evidence of an overreaction mechanism, consistent with La Porta (1996) and BGLS (2024).

7.Taking stock

We started the paper with a simple question: does understanding the cross-section of stock returns by firm characteristics require exotic risk factors, first introduced by Fama and French (1993)? The evidence we presented says no. We develop tests of the efficient market hypothesis with both constant and time varying required returns, and use analyst forecasts to tease out the role of risk and beliefs in driving prices in characteristic-based cross sections of stocks. The risk premia identified by Fama and French appear to reflect corrections of measurable and predictable errors in expectations of earnings growth. We see this result as a victory for financial economics, because it points to its ability to explain the data without the need for exotic risks.

Our evidence shows that characteristics-based return spreads arise because expectations of future growth of firms in the short arm of the portfolios are systematically too optimistic relative to those of firms in the long arm, so that the long portfolio outperforms the short one as expectation errors are corrected. 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 (van Binsbergen et al 2023, Frey 2023, Cho and Polk 2024). Crucially, the very same mechanism and expectations data account for aggregate stock return predictability (BGLS 2024), reconciling longstanding cross sectional and time series puzzles based on measurable analyst forecasts. These findings have significant implications not just for 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 generate 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 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, and Christian Opp, 2019, Real anomalies. *Journal of Finance* 74 (4): 1659-1706.

van Binsbergen, Jules, Martijn Boons, Christian Opp, and Andrea Tamoni, 2023a, 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* 38: 311-347.

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar, 2019, Sticky expectations and the profitability anomaly. *Journal of Finance* 74 (2): 639-674.

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.

Chen, Xiaohong, Lars Peter Hansen, and Peter Hansen, 2020, Robust identification of investor beliefs. *Proceedings of the National Academy of Sciences* 117(52): 33130-33140.

Chinco, Alex, Samuel Hartzmark, and Abigail Sussman, 2022, A new test of risk factor relevance." *Journal of Finance* 77 (4): 2183-2238.

Cho, Thummim, and Christopher Polk, 2024, Putting the price in asset pricing. *Journal of Finance* 3943 - 3984.

Cochrane, John, 2011, Presidential address: Discount rates. *Journal of Finance* 66 (4): 1047-1108.

Da, Zhi, and Mitch Warachka, 2011, The disparity between long-term and short-term forecasted earnings growth. *Journal of Financial Economics* 100 (2): 424-442.

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.

Decaire, Paul, and John Graham, 2024, Valuation fundamentals. *Working paper*.

Dechow, Patricia, and Richard Sloan, 1997, Returns to contrarian investment strategies: Tests of naive expectations hypotheses. *Journal of Financial Economics* 43 (1): 3-27.

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.

Gabaix, Xavier, and Ralph Koijen. 2023, In search of the origins of financial fluctuations: The inelastic markets hypothesis. NBER w28967.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer, 2016, Expectations and investment. *NBER Macroeconomics Annual* 30: 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, 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.

Lee, Charles, Eric So, and Charles Wang, 2020, Evaluation firm-level expected-return proxies: implications for estimating treatment effects. *Review of Financial Studies* 34 (4): 1907 – 1951.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns. *Journal of Finance 56*(3): 815-849.

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.

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics* 82 (2): 289–314

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

Lou, Dong, 2012, A flow-based explanation for return predictability. *Review of Financial Studies* 25 (12): 3457-3489.

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

Merton, Robert, 1973, An intertemporal capital asset pricing model. *Econometrica* 41 (5): 867-887.

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.

Pavlova, Anna, and Taisiya Sikorskaya, 2023, Benchmarking intensity. *Review of Financial Studies* 36 (3): 859-903.

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

Sarkar, Suproteem, 2024, Economic Representations, *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* 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.

Sloan, Richard G., and Annika Wang, 2923, Predictable EPS growth and the performance of value investing. *Review of Accounting Studies*, https://doi.org/10.1007/s11142-023-09812-6

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

ONLINE APPENDIX

Appendix A

A.1 Tests of the Efficient Markets Hypothesis with inference from prices

Consider the case with asset specific, time-invariant required returns. Returns are:

$$r_{i,t+1} = (r_i - r) + r + \sum_{s\geq 0} \rho^s \big(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\big)\big(g_{i,t+1+s}\big)$$

Here $\widetilde{\mathbb{E}}_t$ denote market expectations of cash flow growth. By differentiating this expression across portfolios, we obtain the test in Equation (6) in the case of no inference from prices. To account for inference, note that for each asset i:

$$\sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t\big(g_{i,t+1+s}\big) = pd_{i,t} - \frac{k - r_i}{1-\rho},$$

Thus, EBR_{it} for asset i at time t can be written as:

$$r + \sum_{s\geq 0} \rho^s \big(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\big)\big(g_{i,t+1+s}\big) = g_{t+1} + k + (r - r_i) + \rho \cdot pd_{i,t+1} - pd_{i,t}$$

Now, our assumption of price inference is:

$$\sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t^a\big(g_{i,t+1+s}\big) = (1-\varphi)\lambda \sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t\big(g_{i,t+1+s}\big) + \varphi\left(pd_t - \frac{k-r}{1-\rho}\right)$$

We use this assumption to write analyst EBR^as in terms of market EBRs. Consider first the case of full inference, $\varphi = 1$. Then:

$$\sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t^a\big(g_{i,t+1+s}\big) = pd_{i,t} - \frac{k-r}{1-\rho}$$

And so analyst EBRs can be written as:

$$EBR_{i,t+1}^a = g_{t+1} + k + \rho \cdot pd_{i,t+1} - pd_{i,t}$$

where the key difference to market EBRs is that the return term $(r - r_i)$ cancels out due to the analysts' erroneous assumption about required returns. Taking the difference across portfolios, we find:

$$EBR_{LMS,t+1}^a = EBR_{LMS,t+1} + r_{LMS}$$

For general φ:

$$EBR_{LMS,t+1}^a = (1-\varphi)EBR_{LMS,t+1} + \varphi\big[EBR_{LMS,t+1} + r_{LMS}\big] = EBR_{LMS,t+1} + \varphi \cdot r_{LMS}$$

Introducing this in the expression for returns, we find Equation (11):

$$r_{i,t+1} = (1-\varphi)r_{LMS} + EBR_{LMS,t+1}^a$$

So the constant depends on the true return differential and on the extent of inference. If there is full inference, the constant is zero even if there is a required return r_{LMS}.

Consider next the case of time-varying required returns. From Equation (7), we obtain:

$$\widetilde{\mathbb{E}}_t(r_{i,t+1}) = k_1 - (1-\rho\eta)\cdot pd_{i,t} + (1-\rho\eta)\cdot \sum_{s\geq 0}\rho^s \widetilde{\mathbb{E}}_t(g_{i,t+1+s}) \qquad (A1)$$

from which the test (10) follows in the case of no inference from prices (constant terms are gathered in k_1). Under inference from prices, the last term reads:

$$\sum_{s\geq 0}\rho^s \widetilde{\mathbb{E}}_t(g_{i,t+1+s}) = \frac{1}{(1-\varphi)\lambda}\sum_{s\geq 0}\rho^s \widetilde{\mathbb{E}}_t^a(g_{i,t+1+s}) - \frac{\varphi}{(1-\varphi)\lambda}\left(pd_{i,t} - \frac{k-r}{1-\alpha}\right)$$

Inserting this in the previous equation yields:

$$\widetilde{\mathbb{E}}_t(r_{i,t+1}) = k_2 - (1-\rho\eta)\left(1+\frac{\varphi}{(1-\varphi)\lambda}\right)\cdot pd_{i,t} + \frac{1-\rho\eta}{(1-\varphi)\lambda}\cdot \sum_{s\geq 0}\rho^s \widetilde{\mathbb{E}}^a{}_t(g_{i,t+1+s}) \quad (A2)$$

where constant terms are gathered in k_2. We test the null that expectations are rational, which yields Equation (12).

A.2 Return predictability with non-rational growth expectations

We next generalize the previous analysis by allowing for market expectations of returns, $\widetilde{\mathbb{E}}_t \sum_{s\geq 0}\rho^s r_{i,t+1+s}$, and of growth rates, $\widetilde{\mathbb{E}}_t \sum_{s\geq 0}\rho^s g_{i,t+1+s}$, that need not be rational. We keep for simplicity the functional form assumption on market expectations of returns:

$$\widetilde{\mathbb{E}}_t \sum_{s\geq 0}\rho^s r_{i,t+1+s} = \frac{\widetilde{\mathbb{E}}_t(r_{i,t+1})}{\psi} \qquad (A3)$$

for $\psi > 0$. This nests the case of rational expectations about required returns described by Equation (4), considered above, in which case $\psi = 1-\rho\eta$. A benchmark of interest is that of i.i.d. required returns, with $\eta = 0$ and $\psi = 1$. Note that for such i.i.d. required returns, several non-rational models of expectations (such as diagnostic expectations or rational inattention) would also yield $\psi = 1$. As in the text, we allow for analysts to infer growth expectations from prices following equation (10). To simplify notation, we omit the asset index i.

With these beliefs, the Campbell-Shiller decomposition implies:

$$\widetilde{\mathbb{E}}_t \sum_{s\geq 0} \rho^s r_{t+1+s} = \frac{k'}{1-\rho} - pd_t + \widetilde{\mathbb{E}}_t \sum_{s\geq 0} \rho^s g_{t+1+s}$$

which, using (A2) and (10), can be written as:

$$\widetilde{\mathbb{E}}_t(r_{t+1}) = \frac{\psi k_i'}{1-\rho} - \psi\left(1 + \frac{\varphi}{(1-\varphi)\lambda}\right) pd_t + \frac{\psi}{(1-\varphi)\lambda} \widetilde{\mathbb{E}}_t^a \sum_{s\geq 0} \rho^s g_{t+1+s} \quad (A3)$$

which generalises Equation (A.2). Because we now allow for non-rational expectations, we need to explicitly account for future news. By the Campbell Shiller return decomposition, the realized return at $t+1$ is given by:

$$r_{t+1} = \widetilde{\mathbb{E}}_t(r_{t+1}) - r_{t+1}^n + g_{t+1}^n \quad (A4)$$

where $g_{t+1}^n = \left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\right) \sum_{s\geq 0} \rho^s g_{t+1+s}$ and $r_{t+1}^n = \left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\right) \sum_{s\geq 1} \rho^s r_{t+1+s}$ are news, relative to time-t market expectations, about growth rates and required returns respectively.

Inserting (A3) into (A4), we get the process for the future realized returns in terms of current analyst expectations of cash flow growth, current price dividend ratio, and future news:

$$r_{t+1} = c_0 - c_1 \cdot pd_t + c_2 \cdot \widetilde{\mathbb{E}}_t^a \sum_{s\geq 0} \rho^s g_{t+1+s} - r_{t+1}^n + g_{t+1}^n \quad (A4)$$

with $c_0 = \frac{\psi k_i'}{1-\rho}$, $c_1 = \psi\left(1 + \frac{\varphi}{(1-\varphi)\lambda}\right)$, and $c_2 = \frac{\psi}{(1-\varphi)\lambda}$.

As in Section 2, we study this process with the regression specification:

$$r_{t+1} = \beta_0 + \beta_1 \cdot pd_{i,t} + \beta_2 \cdot \widetilde{\mathbb{E}}_t^a \sum_{s\geq 0} \rho^s g_{t+1+s} + u_{t+1} \quad (A5)$$

where the term u_{t+1} is orthogonal to the regressors pd_t and $\sum_{s\geq 0} \rho^s \widetilde{\mathbb{E}}_t(g_{t+1+s})$. Under market efficiency, the news terms r_{t+1}^n and g_{t+1}^n are orthogonal to information available at t, including the regressors above, so the regression (A5) yields unbiased coefficients $\beta_1 = -c_1$ and $\beta_2 = c_2$. We are back to the test of EMH given by Equation (12).

When expectations are not rational, the news terms may be predictable from information at time t and therefore yield biased coefficient estimates. To see this, consider first an illustrative case:

- cash flow news are given by:

$$g_{t+1}^{n} = -\theta \cdot \sum_{s \geq 0} \rho^{s} \tilde{\mathbb{E}}_{t}(g_{t+1+s}) + \upsilon_{t+1}, \qquad (A6)$$

  where υ_{t+1} is the truly unpredictable component of cash flow news. Expectations of cash flow overreact when $\theta > 0$, so that high optimist today predicts bad news in the future, and they underreact is $\theta < 0$, so that high optimism today predicts good news in the future. This is consistent with the evidence in BGLS (2024) as well as with our evidence that analyst EBRs are predictable.

- required returns are given by $\tilde{\mathbb{E}}_{t}(r_{t+1}) = r + \omega_{t}$, where $\omega_{i,t}$ is i.i.d. with mean zero, and the market holds rational expectations about required returns (though not about realized returns). Thus, the term r_{t+1}^{n} continues to be unpredictable at time t, and all predictability stems from the current realized required return $\tilde{\mathbb{E}}_{t}(r_{t+1})$ and the expectations term.

The process for realized returns, Equation (A4), thus becomes:

$$r_{t+1} = c_0 - c_1 \cdot pd_t + (c_2 - \theta) \cdot \tilde{\mathbb{E}}_{t}^{a} \sum_{s \geq 0} \rho^{s} g_{t+1+s} - r_{t+1}^{n} + \upsilon_{t+1} \qquad (A7)$$

where $-r_{t+1}^{n} + \upsilon_{t+1}$ is indeed a news term, but the regression coefficients in (A5) are now:

$$\beta_1 = -c_1, \;\; \gamma = \; c_2 - \theta$$

Overreacting expectations can generate negative predictability of returns, $\gamma < 0$, consistent with the evidence in BGLS (2024). In particular, replacing c_2 above yields:

$$\gamma = \frac{\psi}{(1-\varphi)\lambda} - \theta.$$

Thus, for a given degree of belief overreaction θ, γ is negative if investors' expected returns are sufficiently persistent (high η, thus lower $\psi = 1 - \rho\eta$), and if analyst beliefs capture or even amplify variations in investors' beliefs, low φ and high λ (for instance because the marginal investor is less biased than the average analyst, due to the presence of some arbitrage forces). Note that, given that news are the forecast revision, $g_{t+1}^n = \widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t$, the extent of overreaction is given by $\theta = \left[var\left(\widetilde{\mathbb{E}}_t\right) - cov\left(\widetilde{\mathbb{E}}_{t+1}, \widetilde{\mathbb{E}}_t\right)\right]/var\left(\widetilde{\mathbb{E}}_t\right)$. This increases when $cov\left(\widetilde{\mathbb{E}}_{t+1}, \widetilde{\mathbb{E}}_t\right)$ is low, which calls for a sufficiently strong belief reversal. Over a longer timescale, $cov\left(\widetilde{\mathbb{E}}_{t+1}, \widetilde{\mathbb{E}}_t\right)$ can even be negative, $\theta > 1$, due to optimism reverting to pessimism as fundamentals mean revert, but $\theta > 1$ is not required to obtain $\gamma < 0$.

We now turn to the general case. The coefficient of interest in (A5) is β_2 and its OLS estimate is (summarizing the news terms $-r_{t+1}^n + g_{t+1}^n$ as n_{t+1}):

$$\beta_2 = c_2 + \frac{cov\left(n_{t+1}, \widetilde{\mathbb{E}}_t^a\right)var(pd_t) - cov(n_{t+1}, pd_t)cov\left(pd_t, \widetilde{\mathbb{E}}_t^a\right)}{var(pd_t)var\left(\widetilde{\mathbb{E}}_t^a\right) - cov\left(pd_t, \widetilde{\mathbb{E}}_t^a\right)^2} \qquad (A8)$$

where $\widetilde{\mathbb{E}}_t^a$ stands for $\widetilde{\mathbb{E}}_t^a \sum_{s\geq 0} \rho^s g_{i,t+1+s}$. To assess the role of overreaction, we wish to write the correction term in terms of our primitives, the market expectations of returns $\tilde{R}_t \equiv \widetilde{\mathbb{E}}_t \sum_{s\geq 0} \rho^s r_{t+1+s}$ and of growth rates, $\widetilde{\mathbb{E}}_t = \widetilde{\mathbb{E}}_t \sum_{s\geq 0} \rho^s g_{t+1+s}$, and in particular in terms of the extent to which they predict future news, namely $cov\left(n_{t+1}, \widetilde{\mathbb{E}}_t\right)$ and $cov\left(n_{t+1}, \tilde{R}_t\right)$.

The denominator in the correction term is positive. In the numerator, we use the Campbell-Shiller decomposition to rewrite the pd_t terms and Equation (10) on price inference to rewrite the $\widetilde{\mathbb{E}}_t^a$ terms. The numerator becomes proportional to:

$$cov\left(n_{t+1}, \widetilde{\mathbb{E}}_t\right)\left[var\left(\tilde{R}_t\right) - cov\left(\widetilde{\mathbb{E}}_t, \tilde{R}_t\right)\right] + cov\left(n_{t+1}, \tilde{R}_t\right)\left[var\left(\widetilde{\mathbb{E}}_t\right) - cov\left(\widetilde{\mathbb{E}}_t, \tilde{R}_t\right)\right].$$

Overreaction of cash flow expectations now corresponds to $cov\left(n_{t+1}, \widetilde{\mathbb{E}}_t\right) < 0$. In the plausible case where cash flow expectations and return expectations are not positively related, $cov(\widetilde{\mathbb{E}}_t, \tilde{R}_t) \leq 0$, overreaction lowers β_2 below c_2. If overreaction is strong enough, it generates $\beta_2 < 0$.

If expectations about future required returns also overreact, $cov(n_{t+1}, \tilde{R}_t) > 0$, then β_2 increases. Thus, our negative estimate of β_2 points not only to market inefficiency, but also to overreaction of cash flow expectations that dominate the predictability of returns, even in the presence of price inference and overreacting expectations about required returns.

In the simpler case in which required returns are i.i.d., and return expectations are not distorted – i.e. $\widetilde{\mathbb{E}}_t(r_{t+1}) = r + \omega_t$, e.g. due to diagnostic expectations or rational inattention – the numerator of the correction becomes proportional to:

$$cov(n_{t+1}, \widetilde{\mathbb{E}}_t) \cdot var(\tilde{R}_t),$$

so $\beta_2 < c_2$ if and only if market expectations overreact, $cov(n_{t+1}, \widetilde{\mathbb{E}}_t) < 0$ (and $var(\tilde{R}_t) > 0$). In this case, the denominator also admits a simple expression, and we obtain $\beta_2 = c_2 - cov(n_{t+1}, \widetilde{\mathbb{E}}_t)/var(\widetilde{\mathbb{E}}_t)$, as in our earlier example. Coefficient β_2 falls short of c_2 due to the negative predictability of the forecast revision from past forecasts.

A.3 Adjusting for measurement noise in analysts' expectations

Here we present a method for recovering the expected return differential $r_L - r_S$ from Equation (6), 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:

$$r_{i,t+1} = r_i + [g_{i,t+1} - \widetilde{\mathbb{E}}_t(g_{i,t+1})] + \sum_{s\geq 1} \rho^s (\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t)(g_{i,t+1+s}) \qquad (A9)$$

where $\widetilde{\mathbb{E}}_t$ are market expectations, and second, that measured analysts' expectations $\widetilde{\mathbb{E}}_t^a$ relate to $\widetilde{\mathbb{E}}_t$ as:

$$\widetilde{\mathbb{E}}_t^a(g_{i,t+1}) = \beta + \tau \cdot \widetilde{\mathbb{E}}_t(g_{i,t+1}) + \sigma \cdot \varepsilon_{it}.$$

where ε_{it} is an iid white noise shock (possibly stock specific), scaled by a volatility parameter $\sigma > 0$. This specification allows for three distortions: $\beta \geq 0$ may capture analysts' systematic over-optimism relative to the market, which may be due to agency problems, and 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 $\tau < 1$.

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

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

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

$$g_{i,t+1} - \widetilde{\mathbb{E}}^a_{i,t}\left(g_{i,t+1}\right) = g_{i,t+1} - \beta - \tau \cdot \widetilde{\mathbb{E}}_{it}\left(g_{i,t+1}\right) - \sigma \cdot \varepsilon_{i,t}$$

so the measured EBR for firm i is equal to:

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

We can then aggregate Equation (A.7) at the level of portfolio p:

$$\begin{aligned}r_{p,t+1} &= r_p + \left[g_{p,t+1} - \widetilde{\mathbb{E}}_t\left(g_{p,t+1}\right)\right] + \sum_{s\geq 1} \rho^s \left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\right)\left(g_{p,t+1+s}\right) \\ &= \left(r_p - r\right) + \text{EBR}_{p,t+1} \qquad (A10)\end{aligned}$$

Where $\text{EBR}_{p,t+1}$ are true EBRs. Then, EBR^m for portfolio p is equal to:

$$\begin{aligned}\text{EBR}^a_{p,t+1} &= r + g_{p,t+1} - \beta - \tau \cdot \widetilde{\mathbb{E}}_{pt}\left(g_{i,t+1}\right) - \sigma \cdot \varepsilon_{p,t} \\ &\quad + \sum_{s\geq 1} \rho^s \left(\tau \cdot \left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\right)\right)\left(g_{p,t+1+s}\right) + \rho^s \sigma \cdot \left(\varepsilon_{pt+1} - \varepsilon_{pt}\right) \\ &= r - \beta + (1-\tau) g_{p,t+1} + \tau\left[g_{p,t+1} - \widetilde{\mathbb{E}}_{pt}\left(g_{p,t+1}\right)\right] - \sigma \cdot \varepsilon_{p,t} \\ &\quad + \tau \sum_{s\geq 1} \rho^s \left(\widetilde{\mathbb{E}}_{t+1} - \widetilde{\mathbb{E}}_t\right)\left(g_{p,t+1+s}\right) + \rho^s \sigma \cdot (\varepsilon_{it+1} - \varepsilon_{it}) \\ &= r(1-\tau) - \beta + (1-\tau) g_{p,t+1} + \tau \text{EBR}_{p,t+1} - \sigma \cdot \varepsilon_{p,t} + \rho^s \sigma \\ &\quad \cdot (\varepsilon_{it+1} - \varepsilon_{it}) \qquad (A11)\end{aligned}$$

We can rewrite equations (A10) and (A11) as

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

where $v_{p,t+1} = \text{EBR}_{p,t+1} - r$, that is $v_{p,t+1}$ is just the expectations component of $\text{EBR}_{p,t+1}$ (without the aggregate required return r), and

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

Next, denote by $\overline{r_p}$ the (sample) average of the realized portfolio return $r_{p,t+1}$. The estimated portfolio regression gives:

$$\overline{r_p} = \kappa + \gamma \cdot \overline{\text{EBR}^a_p},$$

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

$$r_p = \kappa + \gamma \cdot \overline{\mathrm{EBR}_p^a} - \overline{v_p}$$

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

$$\overline{\mathrm{EBR}_p^a} = r - \beta + (1-\tau)\overline{g_p} + \tau \cdot \overline{v_p} \qquad (A12)$$

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

$$\left[g_{p,t+1} - \tilde{\mathbb{E}}_t^a(g_{p,t+1})\right] = g_{p,t+1} - \beta - \tau \cdot \mathbb{E}_t(g_{p,t+1}) - \sigma \cdot \varepsilon_{pt},$$

on average obey (again, denoting sample averages with upperbars):

$$\overline{\left[g_{p,t+1} - \tilde{\mathbb{E}}_t^a(g_{p,t+1})\right]} = \overline{g_p} - \beta - \tau \cdot \overline{g_p},$$

So that:

$$\beta = (1-\tau) \cdot \overline{g_p} - \overline{e_p},$$

where $\overline{e_p}$ is the average forecast error for portfolio p in our sample. We can then plug this expression for β in equation (A4) and obtain:

$$\overline{\mathrm{EBR}_p^a} = r + \overline{e_p} + \tau \cdot \overline{v_p}$$

To recover τ note that:

$$cov(\mathrm{EBR}_p^a, g_p) = (1-\tau) \cdot var(g_p) + \tau \cdot cov(v_p, g_p)$$

where we know $cov(\mathrm{EBR}_p^a, g_p)$ and $var(g_p)$ but not $cov(v_p, g_p)$. Under our assumptions on the true required returns, the latter can be obtained from:

$$cov(v_p, g_p) = cov(r_p, g_p).$$

This yields:

$$\tau = \frac{cov(\mathrm{EBR}_p^a, g_p) - var(g_p)}{cov(r_p, g_p) - var(g_p)}$$

This allows us to back up $\overline{v_p}$ from $\overline{\mathrm{EBR}_p^a}$, and thus the required return spread:

$$r_L - r_S = \kappa_L - \kappa_S + \overline{\mathrm{EBR}_L^a - \mathrm{EBR}_S^a}\left[\gamma - \frac{cov(r_{LMS}, g_{LMS}) - var(g_{LMS})}{cov(\mathrm{EBR}_{LMS}^a, g_{LMS}) - var(g_{LMS})}\right] + \frac{cov(r_{LMS}, g_{LMS}) - cov(\mathrm{EBR}_{LMS}^a, g_{LMS})}{cov(\mathrm{EBR}_{LMS}^a, g_{LMS}) - var(g_{LMS})}(\overline{g}_L - \overline{g}_S). \qquad (A13)$$

Table A1 presents the corrections to the estimated intercepts of Tables 3 and 7 in the text.

Table A1
Expectation based estimates of long-short portfolios required return spread

Note: This table presents estimates of the required return premia (adjusted κ) for the long-short value minus growth (HML), small minus big market capitalization (SMB), conservative minus aggressive investment (CMA), robust minus weak profitability (RMW), and winners minus losers momentum (WML) portfolios. The adjustment allows for three distortions in expectations-based returns (EBRs) as described in Equation (A13). As benchmarks, we report (in the first row) the sample long-short annual spreads for the relevant portfolios for horizons h 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 annualized 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.

	Dep. variable: $r_{LMS,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
Value (HML)	(1)	(2)	(3)	(4)	(5)
Average HML	0.0522[b]	0.0471[b]	0.0395[b]	0.0316[b]	0.0237[c]
	(0.0204)	(0.0195)	(0.0199)	(0.0150)	(0.0122)
Average κ	0.0374[c]	0.0197	-0.0067	-0.0112	-0.0074
	(0.0209)	(0.0179)	(0.0138)	(0.0153)	(0.0085)
Adjusted κ	0.0252	0.0142	-0.0104	-0.0049	-0.0027
Size (SMB)	(1)	(2)	(3)	(4)	(5)
Average SMB	0.0271	0.0308	0.0376[c]	0.0285	0.0261[c]
	(0.0212)	(0.0201)	(0.0208)	(0.0178)	(0.0150)
Average κ	0.0332	0.0390[b]	0.0220	0.0167[c]	0.0185[c]
	(0.0208)	(0.0187)	(0.0135)	(0.0097)	(0.0098)
Adjusted κ	0.0377	0.0403	0.0178	0.0273	0.0383
Investment (CMA)	(1)	(2)	(3)	(4)	(5)
Average CMA	0.0584[a]	0.0568[a]	0.0515[a]	0.0339[a]	0.0254[a]
	(0.0123)	(0.0112)	(0.0105)	(0.0081)	(0.0068)
Average κ	0.0358[b]	0.0159	-0.0113	-0.0030	-0.0071
	(0.0140)	(0.0124)	(0.0127)	(0.0142)	(0.0056)
Adjusted κ	-0.0208	-0.0240	-0.0253	-0.0064	-0.0055
Profitability (RMW)	(1)	(2)	(3)	(4)	(5)
Average RMW	0.0225	0.0217	0.0160	0.0043	-0.0047
	(0.0167)	(0.0142)	(0.0139)	(0.0063)	(0.0062)
Average κ	0.0420[b]	0.0464[a]	0.0396[a]	0.0226[a]	0.0161[a]
	(0.0165)	(0.0133)	(0.0106)	(0.0055)	(0.0042)
Adjusted κ	0.0480	0.0239	-0.0136	0.0224	0.0038
Momentum (WML)	(1)	(2)	(3)	(4)	(5)
Average WML	0.0596[b]	0.0501[b]	0.0019	-0.0088	-0.0083
	(0.0270)	(0.0230)	(0.0214)	(0.0119)	(0.0083)
Average κ	0.0007	-0.1972[a]	-0.1513[a]	-0.0429[a]	-0.0221[a]
	(0.0528)	(0.0613)	(0.0412)	(0.0106)	(0.0066)
Adjusted κ	-0.2827	-0.2915	-0.1727	-0.0557	-0.0330

The findings in the text 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.

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: This 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 expectations-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) market capitalization, (5) the lagged one-year return, (6) the earnings-to-price ratio, (7) the ratio of the three-year moving average of earnings per share to price (CAPE ratio), (8) the dividend to price ratio, and (9) the average number of observations. For each variable, we report sample means for the two samples, their difference, the standard error of the difference, the z-stat, and the significance of the difference between the two means (p).

	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 $\frac{\widetilde{\mathbb{E}}_t^a DPS_{i,t+12}}{\widetilde{\mathbb{E}}_t^a EPS_{i,t+12}}$ in our sample for those firms which paid dividends that year, $DPS_{i,t+12} > 0$, which equals 0.41. In the text we assumed a constant payout ratio, which has a correlation with the above specification of over 97%.

Table B2
EBRs with Alternative Measures of Payout Ratio

Note: Panel A presents sample means of log expectations-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 (13) in the text with the payout ratio set to: (1) the ratio of expected dividends-to-earnings in period *t+12*, $\widetilde{\mathbb{E}}_t^a DPS_{i,t+12}/\widetilde{\mathbb{E}}_t^a EPS_{i,t+12}$, 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, $\widetilde{\mathbb{E}}_t^a DPS_{i,t+12}/\widetilde{\mathbb{E}}^a_{t\,t} EPS_{i,t+12,}$ in our sample for those firms which paid dividends in the previous year but do not have $\widetilde{\mathbb{E}}_t^a DPS_{i,t+12}$. Portfolio EBRs are equally weighted with monthly rebalancing. Panel B shows pairwise correlations between log returns and EBR for high and low portfolios of stocks formed on book-to-market, investment, size, profitability and momentum sorts over holding horizons ranging from one month to five years. See Table 1 for portfolio definitions. The sample period spans 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 Table 3 is robust to various choices regarding the construction of EBRs. In Table B3, we consider two alternative specifications of the terminal value in the analyst price (Equation 14). Panel A considers the case where LTG forecasts are assumed to last until 10 years out (compared to 5 years out in our baseline specification), while Panel B considers the case where forecasted growth rates gradually converge from firm specific LTG five years out towards a common terminal growth rate g ten years out. The Table shows the results are very robust, consistent with the fact that correlations of the above specifications with our baseline specification (e.g. at the 1 month horizon) are of 93.5% and 99.9% respectively.

Table B3
Alternative Measures of Expectation based returns and the HML

Note: This table presents univariate regression results of log returns for the long-short value minus growth (HML) portfolio on two alternative measures of expectations-based returns, $EBR_{HML,t+h}$. Both measures of EBRs define the analyst price ($\tilde{P}^{a}_{i,t}$) as the present value of expected earnings between years one and ten, plus a terminal value based on a perpetuity growth rate g (which is set to match the average stock price across all firms and months in our sample). The first alternative EBR measure assumes that the expected earnings growth rate during the period from $t+60$ to $t+120$ is $LTG_{i,t}$. The second alternative EBR measure assumes that the expected earnings growth rate declines linearly from $LTG_{i,t}$ at period $t+60$ to g at period to $t+120$. Separate regressions are estimated for holding horizons (h) of one month, three months, one year, three years, and five years. 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: EBR based on expected earnings growth of $LTG_{i,t}$ during the period $t+60$ to $t+120$

	(1)	(2)	(3)	(4)	(5)
$EBR_{HML,t+h}$	0.3910[a]	0.6787[a]	0.8704[a]	1.0352[a]	0.9594[a]
	(0.1139)	(0.1299)	(0.0915)	(0.1471)	(0.1263)
Constant	0.0016	-0.0031	-0.0457[a]	-0.1395[a]	-0.1644[a]
	(0.0020)	(0.0053)	(0.0155)	(0.0493)	(0.0354)
Obs	504	502	493	469	445
Adj R^2	4%	16%	46%	48%	51%

Panel B: EBR based on expected earnings growth declining from $LTG_{i,t}$ to g during the period $t+60$ to $t+120$

	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)

$EBR_{HML,t+h}$	0.5132[a]	0.8358[a]	1.0275[a]	1.1537[a]	1.1428[a]
	(0.1490)	(0.1590)	(0.1158)	(0.2255)	(0.1793)
Constant	0.0033[c]	0.0052	-0.0059	-0.0283	-0.0321
	(0.0018)	(0.0046)	(0.0139)	(0.0457)	(0.0425)
Obs	503	502	493	469	445
Adj R^2	5%	17%	46%	45%	49%

Next, we show our baseline analysis in Table 3 is robust to value weighting portfolio EBRs. The correlation of this specifications of EBRs with our baseline specification ranges from 70% to 95% at the yearly horizon.

Table B4
Value weighted EBRs and Returns

Note: Panel A presents univariate regression results of log value-weighted returns for the long-short value minus growth (HML) portfolio on expectations-based returns (EBRs) for that portfolio in columns 1 to 5 and for the long-short small minus big (SMB) portfolio in columns 6 to 10. Separate regressions are estimated for horizons h of one month, three months, one year, three years, and five years horizon. Panel B repeats the analysis for the long-short portfolio robust minus weak profitability (RMW) portfolio in columns 1 to 5 and the long-short portfolio conservative minus aggressive investment (CMA) portfolio in columns 6 to 10. Panel C shows analogous results for the long-short winners minus losers (WML) 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

	Dep. variable: $r_{HML,t,t+h}$					Dep. variable: $r_{SMB,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$EBR_{LMS,t+h}$	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

	Dep. variable: $r_{RMW,t,t+h}$					Dep. variable: $r_{CMA,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$EBR_{LMS,t+h}$	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

	Dep. variable: $r_{WML,t,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$EBR_{LMS,t+h}$	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%

B.2 Decomposition of EBR spreads and the role of news

We next expand on the EBR decomposition analysis presented in Table 4 for the HML portfolio. We first complement Table 4 by including only expectation revisions (i.e. we drop changes in long-term growth forecasts, $\Delta_h LTG_{t+h}$, which entails a shift in horizon for the 5-year out forecast). The results are very similar.

Table B5
Portfolio level forecast errors and revisions predict spreads

Note: This table presents multivariate regressions of log returns for the long-short value minus growth portfolio, $r_{HML,t,t+h}$. Separate regressions are estimated for holding horizons (h) of one year, three years, and five years. The independent variables include: (a) forecast errors between t and $t+h$, $FE_{HML,t+h}$ (defined in the text), and (b) forecast revisions for short-term earnings growth between t and $t+h, \Delta_h STG_{HML,t+h}$ (defined in the text). All independent variables have unit standard deviation. 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.

	Dep. variable: $r_{HML,t,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$FE_{HML,t+h}$			0.1456[a]	0.1814[a]	0.1737[a]
			(0.0134)	(0.0258)	(0.0369)
$\Delta_h STG_{HML,t+h}$	0.3731[a]	0.4313[a]	0.0896[a]	0.0631[b]	-0.0002
	(0.0566)	(0.0679)	(0.0128)	(0.0250)	(0.0299)
Constant	0.0055[a]	0.0119[b]	-0.0894[a]	-0.2600[a]	-0.2439[a]
	(0.0017)	(0.0047)	(0.0201)	(0.0554)	(0.0821)
Obs	476	474	493	469	445
Adj R^2	8%	13%	50%	48%	41%
F-stat	43.3	40.2	59.0	25.4	11.1

We next reproduce the EBR decomposition analysis of Table 4 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 B6
Portfolio level forecast errors and revisions predict spreads

Note: Panel A presents multivariate regressions of log returns for the long-short value minus growth (HML) portfolio and the small minus big (SMB) portfolio for horizons (h) of one year, three years, and five years. The independent variables include: (a) spreads in forecast errors in earnings between t and $t+h$, $FE_{LMS,t+h}$ (defined in the text), (b) spreads in forecast revisions for short-term earnings growth between t and $t+h$, $\Delta_h STG_{LMS,t+h}$ (defined in the text), (c) spreads in changes in long-term earnings growth forecasts between t and $t+h$, $\Delta_h LTG_{LMS,t+h}$, and (d) the log price-dividend ratio, $pd_{LMS,t}$. Panel B presents analogous results for the conservative minus aggressive investment (CMA) portfolio and the robust minus weak profitability (RMW) portfolio. Panel C shows results for the winners minus losers (WML) portfolio. All independent variables have unit standard deviation. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans 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

	Dep. variable: $r_{HML,t,t+h}$					Dep. variable: $r_{SMB,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$FE_{LMS,t+h}$			0.1305[a]	0.1744[a]	0.1351[a]			0.0648[a]	0.0736[a]	0.1023[a]
			(0.0137)	(0.0267)	(0.0376)			(0.0089)	(0.0126)	(0.0199)
$\Delta_h STG_{LMS,t+h}$			0.0802[a]	0.0569[b]	-0.0251			0.0495[a]	0.0399[a]	0.0332[b]
			(0.0128)	(0.0248)	(0.0284)			(0.0081)	(0.0123)	(0.0134)
$\Delta_h LTG_{LMS,t+h}$	-0.0002	0.0141[a]	0.0278[a]	0.0197	0.0487	0.0032[a]	0.0115[a]	0.0149[b]	0.0428[a]	-0.0207
	(0.0017)	(0.0043)	(0.0094)	(0.0212)	(0.0345)	(0.0010)	(0.0025)	(0.0065)	(0.0137)	(0.0202)
$pd_{LMS,t}$	-0.0024	-0.0065	-0.0021	0.0044	-0.0502	0.0006	-0.0018	-0.0123[c]	0.0133	-0.0197
	(0.0017)	(0.0049)	(0.0107)	(0.0263)	(0.0351)	(0.0010)	(0.0027)	(0.0067)	(0.0135)	(0.0185)
Constant	0.0005	-0.0102	-0.1144[a]	-0.2830[a]	-0.3671[a]	-0.0039[c]	-0.0019	0.0215	0.0449	0.0524
	(0.0034)	(0.0097)	(0.0259)	(0.0661)	(0.0916)	(0.0023)	(0.0062)	(0.0157)	(0.0319)	(0.0515)
Obs	504	502	493	469	445	504	502	493	469	445
Adj R^2	0%	5%	53%	48%	48%	2%	7%	46%	47%	39%

Panel B: Investment, Profitability

	Dep. variable: $r_{CMA,t,t+h}$					Dep. variable: $r_{RMW,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$FE_{LMS,t+h}$			0.0490[a]	0.0651[a]	0.0143			0.1037[a]	0.1295[a]	0.0382
			(0.0155)	(0.0173)	(0.0167)			(0.0204)	(0.0367)	(0.0486)
$STG_{LMS,t+h}$			0.0226	0.0074	0.0213			0.0337[c]	0.0110	-0.0340
			(0.0158)	(0.0173)	(0.0164)			(0.0184)	(0.0262)	(0.0351)
$\Delta_h LTG_{LMS,t+h}$	0.0030[b]	0.0014	-0.0200	-0.0135	-0.0096	0.0003	-0.0016	-0.0280[b]	0.0175	0.0799[b]
	(0.0014)	(0.0035)	(0.0127)	(0.0147)	(0.0217)	(0.0018)	(0.0049)	(0.0139)	(0.0291)	(0.0351)
$pd_{LMS,t}$	0.0011	0.0011	0.0129	0.0345[b]	0.0101	-0.0031[c]	-0.0117[b]	-0.0251[c]	-0.0239	-0.0663
	(0.0014)	(0.0038)	(0.0132)	(0.0164)	(0.0255)	(0.0018)	(0.0050)	(0.0148)	(0.0342)	(0.0485)
Constant	0.0041	0.0084	0.0323	0.0439	0.0594	-0.0114[b]	-0.0403[a]	-0.1415[a]	-0.1896[c]	-0.2954[c]

	(0.0028)	(0.0074)	(0.0245)	(0.0274)	(0.0460)	(0.0055)	(0.0156)	(0.0442)	(0.1069)	(0.1545)
Obs	504	502	493	469	445	504	502	493	469	445
Adj R^2	0.0065	-0.003	11%	20%	4%	0%	2%	34%	40%	26%

Panel C: Momentum

	Dep. variable: $ret_{WML,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$FE_{LMS,t+h}$			0.1317[a]	0.1794[a]	0.0705[b]
			(0.0270)	(0.0292)	(0.0342)
$STG_{LMS,t+h}$			0.0821[a]	0.0662[a]	0.0237
			(0.0228)	(0.0235)	(0.0256)
$\Delta_h LTG_{LMS,t+h}$	0.0004	0.0134[a]	0.0193	0.0428[c]	0.0012
	(0.0023)	(0.0052)	(0.0176)	(0.0249)	(0.0393)
$pd_{LMS,t}$	-0.0038[c]	-0.0091	-0.0142	0.0329	-0.0600
	(0.0023)	(0.0056)	(0.0165)	(0.0237)	(0.0371)
Constant	0.0052[c]	0.0033	-0.1582[a]	-0.0898[b]	0.0144
	(0.0030)	(0.0076)	(0.0379)	(0.0352)	(0.0478)
Obs	504	502	493	469	445
Adj R^2	0%	3%	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 B7 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 B7
EBRs and contemporaneous news

Note: This table presents portfolio-level regressions of expectations-based returns (EBRs) at horizons (h) of one month, three months, one year, three years, and five years. Panel A presents results for the long-short value minus growth (HML) portfolio and the small minus big (SMB) portfolio. Panel B presents results for the long-short conservative minus aggressive investment (CMA) portfolio and the robust minus weak profitability (RMW) portfolio. Panel C presents results for the long-short winners minus losers (WML) portfolio. Each panel reports results for two sets of regressions. The independent variables in the first set of regressions are: (a) log return between t and $t+h$, $r_{LMS,t+h}$, (b) earnings growth between t and $t+h$, $\Delta_h e_{LMS,t+h}$, and (c) the (log) forecast error for earnings growth between t and $t+h$, $FE_{LMS,t+h}$ (defined in the text). The independent variable in the second set of regressions is the log return between t and $t+h$. Except for $r_{LMS,t+h}$,all independent variables have unit standard deviation. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans 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

Dep. variable: $EBR_{HML,t+h}$ Dep. variable: $EBR_{SMB,t+h}$

	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t+h}$	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)
$\Delta_h e_{LMS,t+h}$	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)
$FE_{LMS,t+h}$			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

	Dep. variable: $EBR_{CMA,t+h}$					Dep. variable: $EBR_{RMW,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t+h}$	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)
$\Delta e_{LMS,t+h}$	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)
$FE_{LMS,t+h}$			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%

	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t+h}$	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)

Dep. variable: $EBR_{WML,t+h}$

	$h=1$ (1)	$h=3$ (2)	$h=12$ (3)	$h=36$ (4)	$h=60$ (5)
$r_{LMS,t+h}$	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)
$\Delta e_{LMS,t+h}$	0.9235[b]	0.4931[a]	0.1464[c]	-0.0318	0.1481
	(0.4523)	(0.1677)	(0.0749)	(0.1073)	(0.2198)
$FE_{LMS,t+h}$			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%

	$h=1$ (1)	$h=3$ (2)	$h=12$ (3)	$h=36$ (4)	$h=60$ (5)
$r_{LMS,t+h}$	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 6), 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 Equation (3) for realized returns holds 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 4) that the components into of EBRs – forecast errors and revisions – explain returns. Finally, we repeat the analysis in Table B7, 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 Equation (6), we similarly report the explanatory power of contemporaneous firm level EBR for actual firm level returns.

Table B8
Expectation based returns and actual returns

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

Dep. variable: $r_{i,t,t+h}$

	$h = 1$		$h = 3$		$h = 12$		$h = 36$		$h = 60$	
	(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 4, which decomposes the link between EBRs and returns into EBR's forecast error and forecast revision components.

Table B9
Forecast errors and revisions predict returns at the firm level

Note: This table presents firm-level multivariate regressions of log returns, $r_{i,t,t+h}$, over holding horizons (h) of one year, three years, and five years. The independent variables include: (a) the (log) forecast error for earnings growth between t and $t+h$, $FE_{i,t+h}$ (defined in the text), (b) forecast revisions for the short-term growth forecast between t and $t+h$, $\Delta_h STG_{i,t+h}$, and (c) changes in the forecast for long-term earnings growth between t and $t+h$ $\Delta_h LTG_{i,t+h}$. Regressions include time and firm fixed effects in columns (1), (4), and (7), only time fixed effects in columns (2), (5), and (8), and no fixed effects in columns (3), (6), and (9). All independent variables have unit standard deviation. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

Dep. variable: $r_{i,t,t+h}$

	$h = 12$			$h = 36$			$h = 60$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$FE_{i,t+h}$	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)
$\Delta_h STG_{i,t+h}$	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 the explanatory power of expectations – both the coefficients and R2 – remains unchanged. In sum, Tables B7 and B9 place constraints on the role of required returns in the Campbell Shiller decomposition. The results are consistent with the decomposition (4) with the persistent component of required return being a constant r_{ih} that scales with horizon h, as captured by the constant term in columns 3, 6, and 9.[31]

Finally, we replicate Table B7 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. All specifications have firm fixed effects.

Table B10
Expectations based returns and contemporaneous news

Note: Panel A presents firm-level univariate regressions of expectations-based returns, $EBR_{i,t+h}$, at horizons (h) of one month, three months, one year, three years, and five years on the log return between t and $t+h$, $r_{i,t+h}$. Panel B presents firm-level multivariate regressions of $EBR_{i,t+h}$ at horizons (h) of one month, three months, one year, three years, and five years. The independent variables include: log returns between t and $t+h$, $r_{i,t+h}$, (b) earnings growth between t and $t+h$, $\Delta_h e_{i,t+h}$, and (c) the forecast error for earnings growth between t and $t+h$, $FE_{i,t+h}$ (defined in the text). All specifications include firm fixed effects. Except for $r_{i,tt+h}$, all independent

[31] 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).

variables have unit standard deviation. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans 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: Dependent variable $EBR_{i,t,+h}$

	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(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: Dependent variable $EBR_{i,t,+h}$

	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$r_{i,t,t+h}$	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)
$\Delta_h e_{i,t+h}$	0.0098[a]	0.0319[a]	0.0219[a]	-0.0170[b]	0.0059
	(0.0002)	(0.0014)	(0.0026)	(0.0071)	(0.0183)
$FE_{i,t+h}$			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 B10 shows that stock returns are not mechanically incorporated into expectations.

B.4 Log versus raw returns

As discussed in Section 2.3, our analysis uses log, not raw, returns. A concern with the logarithmic transformation is that it may cause a violation in the martingale-difference property, due to the differential volatility of the long and short portfolio EBR components.

By Jensen's inequality, such differential volatility may cause the portfolio EBR to be different from zero *on average*. Here, we show that our key tests are not affected if we use log returns and a Jensen inequality term is allowed for. We also show that our empirical analysis changes little if we use raw returns rather than log returns (but we argue below that, in line with standard practice, when compounding over multiple periods the use of log returns is more appropriate).

Consider first the issue of allowing for Jensen's inequality term when using log returns. We discuss it in the context of the class of time-constant required return models, given that this is the class of models where the issue is most pressing (constant volatility differences between the EBRs in the long and short portfolio would not matter for the return predictability exercise). The raw gross return on the long minus short portfolio based on characteristic i is:

$$r_{it}^{G,raw} = r_{i}^{G,raw} + \mathrm{EBR}_{it}$$

where r_i is the constant required return on the portfolio and EBR_{it} is its expectations-based return. This decomposition follows from the definition of return $r_{it}^{G,raw} = (P_{i,t} + D_{i,t})/P_{i,t-1}$ and the assumption that the price is the NPV of rationally expected dividends under a constant gross discount rate r_i^G. In this equation, EBR_{it} is a function of forecast revisions of dividends scaled by $P_{i,t-1}$ (in particular, it is not the specification used in Equation (5) which obtained under log returns and the Campbell-Shiller decomposition). The current definition preserves the key properties that EBRs are mean zero and unpredictable.

Assume that, under the rational expectations null, EBR_{it} has mean zero and variance σ_i^2 (which is the sum of the variances of the EBRs of the long and short portfolios minus twice the covariance between them). Taking logs and defining the net actual and required returns as $r_{it}^{raw} = r_{it}^{G,raw} - 1$ and $r_{i}^{raw} = r_{i}^{G,raw} - 1$ we find that, either exactly (if EBRs are lognormal) or based on a second order approximation, the log gross return obeys:

$$\mathrm{E}[\ln(1+r_{it}^{raw})] = \mathrm{E}[\ln(1+r_i^{raw}+\mathrm{EBR}_{it})] \approx \ln E[1+r_i^{raw}+\mathrm{EBR}_{it}] - k\cdot\sigma_i^2$$
$$= r_i^{raw} - k\cdot\sigma_i^2 \approx r_i^{raw} + E[\ln(\mathrm{EBR}_{it})].$$

This equation simply reflects the fact that, under the log transformation, the mean of the EBRs of portfolio i is no longer zero, but equal to $-k\cdot\sigma_i$. When regressing the log return of portfolio i on the log EBRs of the same portfolio, the Jensen's inequality adjustment would be captured by the mean observed log EBR. As a result, under the rational null the constant α_i from regressing the log return of portfolio i on its log EBR is equal to:

$$\alpha_i = \mathrm{E}[\ln(r_{it})] - E[\ln(\mathrm{EBR}_{it})] = r_i.$$

It is therefore the case that, even if we allow for a Jensen's inequality correction, the regression constant identifies the risk premium in the rational expectations null. This, in turn, implies that if we take the difference of the log return of a long and short portfolio and we regress it on the difference in the logs of their EBRs, by the linearity of the log regression the constant will be equal to the required return difference, $\alpha_{LMS} = r_L - r_S$. Thus, our interpretation of Equation (6) is valid under the log return approximation, which we have assumed to derive it, as well as under raw returns and raw EBRs. The logic for the validity of the log transformation is that, notwithstanding the Jensen's-inequality term, this term appears both on the average return spread (the explanandum) and on the EBR (the explanatory variable). Differencing out the two, which is what the regression constant does, leaves us with the risk premium.[32]

We now come to the second point, namely that the results do not change appreciably when we use raw returns rather than log returns. This is important, because whether taking

[32] The implications carry through if we allow for an i.i.d. shock to affect the required return of different portfolios (due for instance to liquidity or other transient factors), which would be captured by the error term in Equation (6). This return shock would reduce the average log return of each portfolio, due to a Jensen's inequality adjustment. Being i.i.d., it would cancel out when taking long minus short differences. Even in the more general case in which such return shock has different variances for the long and the short portfolio, the difference in the corresponding Jensen's inequality term would be a stable feature of log return differences α_i. Thus, under the efficient market null, variation in log EBRs would not explain it away. We see the opposite in the data, rejecting the efficient market hypothesis' prediction that differences in raw or average log returns only reflect stable or idiosyncratic required return variation across portfolios.

logs affects or not our results is also and crucially an empirical matter. Specifically, Table B11 below reproduces Table 1 in the paper, by computing average *raw* expectations-based returns on long versus short portfolios and compares them to the corresponding *raw* realized returns. Levels of raw returns are higher as expected by the Jensen inequality, but the resulting spreads are essentially identical to those in Table 2 at horizons of 1 year and above. For shorter horizons, particularly at 1 month, we run into the problem of averaging over raw returns that can be very large (as well as systematically larger when due to a price increase rather than a price decrease).

Table B11. Average raw returns and raw EBRs of portfolios

Note: Panel A presents sample means of portfolio raw returns over holding horizons *h* ranging from one month to five years. Portfolios are formed independently based on quintiles. Results are displayed for the following five quintile portfolios: (1) book-to-market, *Growth* stocks in the bottom quintile and *Value* stocks in the top quintile, (2) investment, *Aggressive* stocks in the top quintile and *Conservative* ones in the bottom 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 expectations-based raw returns (EBRs) following Equation (13) 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 (raw) returns

Holding Horizon	Growth	Value	Aggr.	Cons.	Big	Small	Weak	Robust	Losers	Winners
1 Month	44.9%	60.4%	52.9%	48.2%	60.5%	31.8%	59.8%	45.0%	102.1%	47.3%
3 Months	21.4%	30.7%	28.0%	21.3%	32.2%	17.9%	26.7%	23.7%	33.7%	25.9%
1 Year	14.1%	19.9%	18.8%	12.9%	20.5%	14.0%	15.8%	16.5%	18.4%	16.1%
3 Years	13.3%	17.3%	16.2%	12.3%	17.0%	13.5%	14.4%	14.8%	15.4%	14.1%
5 Years	13.4%	16.4%	15.6%	12.7%	16.2%	13.1%	14.8%	14.1%	14.9%	13.8%

Panel B. Average portfolio (raw) EBRs

Holding Horizon	Growth	Value	Aggr.	Cons.	Big	Small	Weak	Robust	Losers	Winners
1 Month	12.7%	18.5%	20.3%	9.2%	13.8%	13.4%	20.3%	11.5%	-11.2%	41.5%
3 Months	11.3%	16.6%	18.3%	7.9%	12.3%	12.5%	18.0%	10.6%	-10.2%	36.3%
1 Year	10.1%	16.0%	16.6%	7.5%	13.2%	11.2%	16.1%	10.3%	0.8%	22.6%
3 Years	10.2%	14.5%	14.5%	8.8%	12.4%	11.1%	14.7%	10.6%	9.1%	13.9%
5 Years	10.6%	13.8%	14.2%	9.8%	12.2%	11.2%	14.8%	10.7%	10.5%	12.8%

Comparing Table B11 to Table 1 reveals that raw EBR spreads are similar to the log EBR spreads which in turn are similar to the raw return spreads. Our regression results are also robust to using raw rather than log returns. Table B12 below reproduces Table 3 in the paper, regressing raw return spreads on raw expectations-based returns, again with very similar results, both for HML (shown below) and for the other portfolios (available upon request).

Table B12: EBRs and the HML spreads

Note: Panel A presents univariate regression results of raw returns for the long-short value minus growth, $R_{HML,t,t+h}$, portfolio on expectations-based raw returns $EBR_{HML,t,t+h}$, for that portfolio. Separate regressions are estimated for horizons h of one-month, three-month, one-year, three-year, and five-year horizon. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period is 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.

	Dep. variable: $R_{HML,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$EBR_{HML,t+h}$	0.4635[a]	0.7850[a]	0.9944[a]	1.1931[a]	1.3523[a]
	(0.1523)	(0.1613)	(0.1053)	(0.2084)	(0.3151)
Constant	0.0036[b]	0.0066	-0.0008	-0.0244	-0.0384
	(0.0017)	(0.0046)	(0.0164)	(0.0707)	(0.0783)
Obs	504	502	493	469	445
Adj R^2	4%	15%	40%	40%	41%

In these regressions, the constant term –the estimate for the required return spread – is very small and statistically insignificant. Finally, Table B13 below reproduces Table 4 in the paper, which regresses return spreads – here replaced with raw realized spreads – on the components of expectations-based returns. The results are again very robust.

Table B13
Portfolio level forecast errors and revisions predict HML returns

Note: This table presents multivariate regressions of log returns for the long-short value minus growth (HML) portfolio for horizons (h) of one-year, three-years, and five-years. The independent variables include: (a) spreads in forecast errors between t and $t+h$, $FE_{HML,t+h}$ (defined in the text), (b) forecast revisions for the short-term growth forecast between t and $t+h$, $\Delta_h STG_{HML,t+h}$ (defined in the text), and (c) spreads in changes in long-term growth forecasts between t and $t+h$, $\Delta_h LTG_{HML,t+h}$. All independent variables have unit standard deviation. 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.

	Dep. variable: $R_{HML,t,t+h}$		
	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)
$FE_{HML,t+h}$	0.1491[a]	0.2356[a]	0.2023[b]

	(0.0195)	(0.0513)	(0.0914)
$\Delta_h STG_{HML,t+h}$	0.0868[a]	0.0813[c]	-0.0174
	(0.0180)	(0.0457)	(0.0681)
$\Delta_h LTG_{HML,t+h}$	0.0361[a]	0.0350	0.2105[b]
	(0.0132)	(0.0354)	(0.0829)
Constant	-0.1213[a]	-0.3350[a]	-0.5406[a]
	(0.0293)	(0.1127)	(0.1956)
Obs	465	441	417
Adj R^2	44%	35%	38%

In sum, using raw rather than log returns does not materially change our results. We continue to reject the presence of significant required returns spreads after variation in expectations is accounted for.

Appendix C: Return predictability

C1. Portfolio level results.

In the return predictability analysis of Table 5, return spreads are predicted from lagged expectations variables and from the price to dividend ratio (as a proxy for required returns). As discussed in Footnote 24, this is very demanding since, whether or not market efficiency holds and for any level of price inference, the *pd* variable also reflects measured expectations. Table C1 repeats the analysis omitting the pd ratio.

Table C1. Predicting future return spreads from expectations data alone

Note: the table presents regressions of log returns for portfolios that are long value and short growth stocks, $r_{HML,t,t+h}$. Separate regressions are estimated for horizons (h) one-month, one quarter, and one, three and five years. The set of independent variables includes: (a) the change in the portfolio forecast for long-term earnings growth between $t-h$ and t, $\Delta_h LTG_{HML,t}$, (b) the lagged portfolio forecast for long-term earnings growth at $t-h$, $LTG_{HML,t-h}$, (c) forecast revisions for the short-term growth forecast between t and $t+h$, $\Delta_h STG_{HML,t+h}$, (d) the lagged portfolio short-term growth forecast at $t-h$, $STG_{HML,t-h}$ (defined in the text), (e) the forecast error in portfolio earnings between $t-h$ and t, $FE_{HML,t-h,t}$ (defined in the text), (f) the change in the portfolio forecast for long-term aggregate earnings growth between $t-h$ and t, $\Delta_h LTG_{Mkt,t}$, and (f) the lagged portfolio forecast for long-term aggregate earnings growth at $t-h$, $LTG_{Mkt,t-h}$. All independent variables are standardized. The last row reports the R^2 from a univariate regression of $r_{LMS,t+h}$ on $pd_{LMS,t}$. 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.

	Dep variable: $r_{HML,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$\Delta_h LTG_{HML,t}$	0.0045[c]	0.0089	-0.0875[b]	-0.0583[c]	-0.0115

	(0.0026)	(0.0072)	(0.0344)	(0.0303)	(0.0446)
$LTG_{HML,t-h}$	-0.0163[c]	-0.0546[b]	-0.2928[a]	-0.4428[a]	-0.4735[a]
	(0.0089)	(0.0247)	(0.0903)	(0.1435)	(0.1694)
$\Delta_h STG_{HML,t}$	-0.0023	0.0019	0.0457[a]	0.0795[a]	0.0883[a]
	(0.0023)	(0.0076)	(0.0173)	(0.0306)	(0.0307)
$STG_{HML,t-h}$	0.0170[c]	0.0619[c]	0.2130[a]	0.4220[a]	0.2378[b]
	(0.0094)	(0.0338)	(0.0514)	(0.1226)	(0.1135)
$FE_{HML,t-h,t}$			0.0422[c]	0.0326	-0.1061[a]
			(0.0232)	(0.0423)	(0.0325)
$\Delta_h LTG_{Mkt,t}$	-0.0045[c]	0.0038	0.0238	0.0557	0.2505[a]
	(0.0025)	(0.0068)	(0.0288)	(0.0341)	(0.0488)
$LTG_{Mkt,t-h}$	0.0040[c]	0.0124[c]	0.0317	0.1265[a]	0.1324[a]
	(0.0022)	(0.0068)	(0.0282)	(0.0343)	(0.0394)
Constant	-0.0389[b]	-0.1223[b]	-0.4069[b]	-1.0812[a]	-1.4077[a]
	(0.0155)	(0.0497)	(0.1664)	(0.3043)	(0.3348)
Obs	444	442	433	409	385
Adj R^2	5%	7%	18%	27%	67%
R^2 univariate $pd_{HML,t}$	0%	1%	3%	1%	5%

Relative to Table 5, the evidence on overreaction of LTG is considerably strengthened, and in particular the lagged level of LTG significantly, and negatively, predicts returns at all horizons.

Table C2 extends the return predictability exercise of Table 5 to the other factors. As in the text, we adopt the following notation for forecasts of short-term growth, $STG_{LMS,t-h} = \widetilde{\mathbb{E}}^a_{t-h}\Delta_{12}e_{LMS,t-h+24}$ and $\Delta_h STG_{LMS,t} = \widetilde{\mathbb{E}}^a_t\Delta_{12}e_{HML,t+24} - \mathbb{E}_{t-h}\Delta_{12}e_{HML,t-h+24}$.

Table C2

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 month, three months, one year, three years, and five years. Panel A presents results for the long-short small minus big (SMB) portfolio and the conservative minus aggressive investment (CMA) portfolio. Panel B presents results for the long-short robust minus weak profitability (RMW) portfolio. Panel C presents results for the long-short winners minus losers (WML) portfolio. Each panel presents two sets of regressions The independent variables in the first set of regressions includes: (a) the portfolio log price-dividend ratio at time t, $pd_{HML,t}$, (b) the change in the portfolio forecast for long-term earnings growth between $t-h$ and t, $\Delta_h LTG_{HML,t}$, (c) the lagged portfolio forecast for long-term earnings growth at $t-h$, $LTG_{HML,t-h}$, (d) the change in the portfolio forecast for short-term earnings growth between $t-h$ and t, $\Delta_h STG_{HML,t}$ (defined in the text), (e) the lagged portfolio forecast for short-term earnings growth at $t-h$, $STG_{HML,t-h}$, (f) the portfolio forecast error in earnings between $t-h$ and t, $FE_{HML,t-h,t}$ (defined in the text), (g) the change in the aggregate forecast for long-term aggregate earnings growth between $t-h$ and t, $\Delta_h LTG_{Mkt,t}$, and (h) the forecast for long-term aggregate earnings growth at $t-h$, $LTG_{Mkt,t-h}$. The independent variable in the second set of regression is the log price-dividend ratio for the relevant long-short portfolio at time t ($pd_{LMS,t}$). All independent variables have unit standard deviation. 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} 10% level.

Panel A: Size and investment

			$r_{SMB,t,t+h}$					$r_{CMA,t,t+h}$		
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$pd_{LMS,t}$	-0.0077	-0.0641^{c}	-0.1188	-0.0984	-0.6169^{a}	0.0066	0.0396	-0.0652	0.2482^{c}	0.0273
	(0.0120)	(0.0365)	(0.0902)	(0.0941)	(0.1418)	(0.0115)	(0.0296)	(0.0721)	(0.1340)	(0.1611)
$\Delta_h LTG_{LMS,t}$	0.0060^{b}	-0.0020	0.0055	-0.0147	-0.0195	0.0038^{b}	0.0051	0.0062	-0.0730^{a}	-0.0514
	(0.0025)	(0.0104)	(0.0275)	(0.0519)	(0.0580)	(0.0016)	(0.0045)	(0.0169)	(0.0155)	(0.0365)
$LTG_{LMS,t-h}$	0.0052	0.0280	0.0341	-0.2033^{c}	0.1689	-0.0098^{c}	-0.0386^{a}	-0.1432^{a}	-0.2839^{b}	-0.1183
	(0.0068)	(0.0172)	(0.0700)	(0.1038)	(0.1482)	(0.0057)	(0.0138)	(0.0515)	(0.1164)	(0.1301)
$\Delta_h STG_{LMS,t}$	0.0020	0.0028	0.0455^{a}	0.0435^{a}	0.0192	-0.0007	0.0044	0.0138	0.0013	0.0207
	(0.0026)	(0.0079)	(0.0158)	(0.0150)	(0.0177)	(0.0017)	(0.0043)	(0.0112)	(0.0229)	(0.0238)
$STG_{LMS,t-h}$	0.0304^{b}	0.0407	0.1332^{c}	0.0193	-0.1077	0.0081	0.0197	0.1228^{a}	-0.0146	0.1096
	(0.0132)	(0.0261)	(0.0756)	(0.1358)	(0.1442)	(0.0092)	(0.0243)	(0.0461)	(0.1181)	(0.1522)
$FE_{LMS,t-h,t}$			0.0015	-0.0298	0.0078			-0.0013	-0.0048	0.0248
			(0.0157)	(0.0413)	(0.0279)			(0.0085)	(0.0138)	(0.0175)
$\Delta_h LTG_{Mkt,t}$	2.4655^{a}	0.1018	-0.1510	4.2048^{a}	5.4256^{a}	-0.6611^{c}	0.3459	0.8748	0.2023	1.6040^{b}
	(0.8304)	(1.4134)	(1.2195)	(1.0503)	(1.2486)	(0.3377)	(0.4043)	(0.5567)	(0.7860)	(0.7338)
$LTG_{Mkt,t-h}$	0.1402	0.2952	2.2585^{b}	5.9413^{a}	3.6553^{a}	0.1758^{a}	0.5494^{a}	1.2001^{a}	0.8635	2.1723^{b}
	(0.1461)	(0.3174)	(0.9535)	(1.2544)	(1.3941)	(0.0624)	(0.1595)	(0.4383)	(0.9310)	(0.9002)
Constant	-0.0307	-0.0314	-0.2567	-0.4361^{b}	0.0057	-0.0255^{a}	-0.0722^{a}	-0.2518^{a}	-0.1022	-0.2077
	(0.0271)	(0.0542)	(0.1590)	(0.1699)	(0.2301)	(0.0086)	(0.0223)	(0.0687)	(0.1825)	(0.1337)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	4%	3%	23%	43%	58%	6%	11%	26%	19%	14%

$pd_{LMS,t}$	-0.0093	-0.0408^{c}	-0.1597^{b}	-0.3820^{a}	-0.4896^{a}	0.0009	-0.0180	-0.1425^{a}	-0.0563	-0.0865
	(0.0080)	(0.0230)	(0.0718)	(0.1471)	(0.1462)	(0.0070)	(0.0200)	(0.0423)	(0.0765)	(0.0561)
Constant	0.0101	0.0413^{b}	0.1707^{a}	0.4201^{a}	0.5668^{a}	0.0052^{b}	0.0089	0.0056	0.0617^{c}	0.0614^{b}
	(0.0064)	(0.0187)	(0.0634)	(0.1413)	(0.1514)	(0.0024)	(0.0070)	(0.0161)	(0.0336)	(0.0277)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	0%	2%	7%	19%	26%	0%	0%	8%	0%	1%

Panel B: Profitability and momentum

			$r_{RMW,t+h}$					$r_{WML,t+h}$		
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$pd_{LMS,t}$	-0.0176	-0.0686	-0.1065	0.2879	0.1332	-0.0088	-0.0093	-0.0696	-0.0814	-0.3051^{a}
	(0.0164)	(0.0420)	(0.1525)	(0.2586)	(0.1522)	(0.0121)	(0.0237)	(0.0696)	(0.0650)	(0.0753)

$\Delta_h LTG_{LMS,t}$	0.0014	-0.0026	-0.0055	0.0045	0.0160	0.0031	0.0018	-0.0279	-0.0151	0.0122
	(0.0023)	(0.0043)	(0.0161)	(0.0205)	(0.0216)	(0.0030)	(0.0082)	(0.0231)	(0.0357)	(0.0617)
$LTG_{LMS,t-h}$	0.0034	0.0210	0.0818	-0.1707	0.3391[a]	-0.0134[c]	-0.0372[b]	-0.1462[b]	-0.1798[a]	-0.0042
	(0.0080)	(0.0172)	(0.0733)	(0.1126)	(0.0958)	(0.0075)	(0.0149)	(0.0571)	(0.0511)	(0.1136)
$\Delta_h STG_{LMS,t}$	0.0039	0.0052	0.0022	0.0013	-0.0228[c]	0.0075[b]	0.0199[b]	0.0692[b]	0.0960[a]	0.0871[b]
	(0.0025)	(0.0047)	(0.0117)	(0.0196)	(0.0134)	(0.0032)	(0.0093)	(0.0268)	(0.0248)	(0.0399)
$STG_{LMS,t-h}$	0.0234[c]	0.0410[c]	0.0076	0.0161	-0.1541	0.0558[a]	0.1102[a]	0.2517[a]	0.4210[a]	0.4535
	(0.0139)	(0.0238)	(0.0674)	(0.1349)	(0.1259)	(0.0176)	(0.0387)	(0.0707)	(0.1448)	(0.3202)
$FE_{LMS,t-h,t}$			-0.0175[c]	-0.0066	-0.0102			-0.0116	-0.0088	-0.0240
			(0.0090)	(0.0321)	(0.0386)			(0.0350)	(0.0394)	(0.0445)
$\Delta_h LTG_{Mkt,t}$	-2.3797[a]	-0.9571	1.8280	-0.0563	1.5699[c]	0.6559	0.0205	0.3969	-3.4682[a]	-3.6590[a]
	(0.6772)	(0.8804)	(1.2404)	(1.1091)	(0.9360)	(0.7280)	(0.9215)	(0.8756)	(1.3153)	(1.1701)
$LTG_{Mkt,t-h}$	0.1201	0.4940[c]	1.6938	-1.6826	6.6079[a]	0.1297	0.3554	-0.9316	-3.0757[a]	-3.2093[c]
	(0.1266)	(0.2719)	(1.2554)	(1.8974)	(1.9104)	(0.1257)	(0.2729)	(1.3530)	(1.1357)	(1.8777)
Constant	-0.0102	-0.0579[c]	-0.1951	0.2645	-0.7383[a]	0.0027	0.0056	0.2329	0.4724[a]	0.4727[c]
	(0.0170)	(0.0311)	(0.1532)	(0.2384)	(0.2486)	(0.0160)	(0.0335)	(0.1818)	(0.1622)	(0.2592)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	5%	4%	9%	7%	28%	7%	16%	26%	36%	23%

$pd_{LMS,t}$	-0.0001	-0.0131	0.0038	0.1155	0.3221[b]	-0.0055	-0.0163	-0.0992[c]	-0.1340	-0.1338[b]
	(0.0081)	(0.0198)	(0.0618)	(0.0793)	(0.1282)	(0.0077)	(0.0163)	(0.0524)	(0.0818)	(0.0678)
Constant	0.0010	-0.0023	0.0097	0.0511	0.0928[b]	0.0040	0.0093	-0.0068	-0.0487	-0.0726[b]
	(0.0033)	(0.0086)	(0.0255)	(0.0317)	(0.0434)	(0.0025)	(0.0067)	(0.0216)	(0.0336)	(0.0363)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	0%	0%	0%	3%	13%	0%	0%	4%	4%	3%

The results in Table C2 generalize those of Table 5 to other factors: 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 price to dividend ratio with the book-to-market ratio or lagged return spread, shown in Table C3.

Table C3
Robustness of return predictability from expectations data

Note: The table presents regressions of log returns for the long-short value minus growth (HML) portfolio, the small minus big (SMB) portfolio, the conservative minus aggressive investment (CMA) portfolio, the robust minus weak profitability (RMW) portfolio, and the winners minus losers (WML) portfolio. Separate regressions are estimated for horizons (h) one month and one year. The set of independent variables includes (a) the change in the portfolio forecast for long-term earnings growth between $t-h$ and t, $\Delta_h LTG_{HML,t}$, (b) the lagged portfolio forecast for long-term earnings growth at $t-h$, $LTG_{HML,t-h}$, (c) the change in the portfolio forecast for short-term earnings growth between $t-h$ and t, $\Delta_h STG_{HML,t}$ (defined in the text), (d) the lagged portfolio forecast for short-term earnings growth at $t-h$, $STG_{HML,t-h}$, (e) the portfolio forecast error in earnings between $t-h$ and t, $FE_{HML,t-h,t}$ (defined in the text), (f) the change in the aggregate forecast for long-term aggregate earnings growth between $t-h$ and t, $\Delta_h LTG_{Mkt,t}$, (g) the forecast for long-term aggregate earnings growth at $t-h$, $LTG_{Mkt,t-h}$, and (h) the log book-to-market ratio, $bm_{LMS,t}$ for the relevant long-short portfolio at time t in Panel A, (i) and the log lagged portfolio return between $t-h$ and t, $r_{LMS,t-h,t}$, in Panel B. 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). All independent variables have unit standard deviation. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] 10% level.

Panel A: Controlling for book-to-market ratio spreads

	$r_{HML,t,t+h}$		$r_{SMB,t,t+h}$		$r_{CMA,t,t+h}$		$r_{RMWt,t+h}$		$r_{WML,t,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$bm_{LMS,t}$	-0.0144[c]	0.0733	-0.0080[b]	0.0305	-0.0096[b]	0.0053	-0.0123[b]	0.1435[a]	-0.0076[b]	0.0291
	(0.0079)	(0.0748)	(0.0037)	(0.0289)	(0.0039)	(0.0265)	(0.0059)	(0.0436)	(0.0037)	(0.0258)
$\Delta_h LTG_{LMS,t+h}$	0.0046[c]	-0.0872[a]	0.0043[a]	0.0005	0.0017	-0.0110	0.0063[b]	-0.0158	0.0036	-0.0238
	(0.0027)	(0.0311)	(0.0016)	(0.0165)	(0.0024)	(0.0179)	(0.0026)	(0.0257)	(0.0032)	(0.0227)
$LTG_{LMS,t-h}$	-0.0058	-0.3722[a]	-0.0110[b]	-0.1508[a]	-0.0002	0.0346	0.0016	0.0082	-0.0227[a]	-0.1354[b]
	(0.0083)	(0.1224)	(0.0049)	(0.0419)	(0.0057)	(0.0354)	(0.0056)	(0.0488)	(0.0074)	(0.0550)
$\Delta_h STG_{LMS,t}$	-0.0025	0.0462[a]	-0.0010	0.0156	0.0042[c]	0.0048	0.0021	0.0364[b]	0.0084[a]	0.0703[a]
	(0.0024)	(0.0167)	(0.0017)	(0.0104)	(0.0024)	(0.0130)	(0.0026)	(0.0166)	(0.0032)	(0.0244)
$STG_{LMS,t-h}$	0.0119	0.2449[a]	0.0078	0.1056[b]	0.0221	0.0058	0.0338[a]	0.1355[c]	0.0620[a]	0.1992[a]
	(0.0091)	(0.0662)	(0.0088)	(0.0456)	(0.0141)	(0.0686)	(0.0125)	(0.0716)	(0.0172)	(0.0746)
$FE_{LMS,t-h,t}$		0.0374[b]		0.0057		-0.0168[c]		0.0122		0.0037
		(0.0181)		(0.0092)		(0.0097)		(0.0136)		(0.0358)
$\Delta_h LTG_{Mkt,t}$	-0.0044[c]	0.0151	-0.0024[b]	0.0223[c]	-0.0081[a]	0.0333	0.0085[a]	-0.0414[c]	0.0018	0.0086
	(0.0024)	(0.0283)	(0.0012)	(0.0120)	(0.0023)	(0.0223)	(0.0029)	(0.0250)	(0.0025)	(0.0188)
$LTG_{Mkt,t-h}$	0.0050[b]	0.0231	0.0028[c]	0.0303[a]	0.0029	0.0273	0.0059	0.0044	0.0010	-0.0095
	(0.0022)	(0.0293)	(0.0015)	(0.0095)	(0.0026)	(0.0247)	(0.0036)	(0.0186)	(0.0028)	(0.0308)
Constant	-0.0132	-0.5367[b]	-0.0184[c]	-0.2734[a]	-0.0272	-0.1168	-0.0398	-0.1857	0.0029	0.1927
	(0.0176)	(0.2136)	(0.0102)	(0.0666)	(0.0175)	(0.1652)	(0.0268)	(0.1332)	(0.0162)	(0.1815)
Obs	442	433	442	433	442	433	442	433	442	433
Adjusted R^2	6%	19%	8%	27%	6%	8%	5%	27%	9%	27%
Univariate R^2	2%	0%	2%	1%	1%	0%	0%	16%	0%	9%

Panel B: Controlling for lagged return spreads

	$r_{HML,t,t+h}$		$r_{SMB,t,t+h}$		$r_{CMA,t,t+h}$		$r_{RMW,t,t+h}$		$r_{WML,t,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t-h,t}$	0.1728[a]	0.0336	0.1691[a]	0.0016	0.1897[a]	-0.0657[b]	0.2056[a]	-0.0570[c]	0.1352[a]	0.0257
	(0.0261)	(0.0296)	(0.0208)	(0.0276)	(0.0268)	(0.0281)	(0.0203)	(0.0292)	(0.0233)	(0.0270)
$\Delta_h LTG_{LMS,t+h}$	0.0051[b]	-0.0938[a]	0.0024[c]	0.0035	-0.0020	-0.0087	0.0023	-0.0093	0.0017	-0.0185
	(0.0026)	(0.0340)	(0.0013)	(0.0170)	(0.0020)	(0.0153)	(0.0021)	(0.0252)	(0.0027)	(0.0273)
$LTG_{LMS,t-h}$	-0.0118	-0.2815[a]	-0.0064	-0.1640[a]	-0.0002	0.0480	0.0023	-0.0104	-0.0129[c]	-0.1881[a]
	(0.0082)	(0.0857)	(0.0044)	(0.0385)	(0.0046)	(0.0343)	(0.0045)	(0.0510)	(0.0068)	(0.0530)
$\Delta_h STG_{LMS,t}$	0.0012	0.0439[b]	0.0005	0.0138	0.0025	0.0102	0.0039[b]	0.0510[a]	0.0050	0.0770[a]
	(0.0022)	(0.0183)	(0.0014)	(0.0112)	(0.0020)	(0.0141)	(0.0019)	(0.0154)	(0.0040)	(0.0266)
$STG_{LMS,t-h}$	0.0163[c]	0.1769[a]	0.0090	0.1080[b]	0.0088	0.0225	0.0228[b]	0.1637[b]	0.0315[c]	0.2085[a]
	(0.0089)	(0.0607)	(0.0080)	(0.0427)	(0.0119)	(0.0709)	(0.0102)	(0.0740)	(0.0165)	(0.0693)
$FE_{LMS,t-h,t}$		0.0366		-0.0017		-0.0107		0.0261		0.0001
		(0.0252)		(0.0092)		(0.0103)		(0.0219)		(0.0340)
$\Delta_h LTG_{Mkt,t}$	0.0005	0.0272	-0.0002	0.0193	-0.0006	0.0188	0.0030	-0.0065	0.0019	-0.0030
	(0.0022)	(0.0266)	(0.0012)	(0.0120)	(0.0022)	(0.0230)	(0.0026)	(0.0247)	(0.0024)	(0.0196)
$LTG_{Mkt,t-h}$	0.0029	0.0299	0.0023[c]	0.0261[a]	0.0015	0.0265	0.0021	0.0505[b]	0.0021	-0.0354
	(0.0019)	(0.0279)	(0.0013)	(0.0097)	(0.0022)	(0.0233)	(0.0025)	(0.0201)	(0.0026)	(0.0331)
Constant	-0.0289[b]	-0.3836[b]	-0.0174[b]	-0.2421[a]	-0.0057	-0.1190	-0.0241	-0.2960[b]	0.0012	0.2272
	(0.0137)	(0.1684)	(0.0078)	(0.0671)	(0.0131)	(0.1263)	(0.0208)	(0.1427)	(0.0151)	(0.1834)
Obs	444	433	444	433	444	433	444	433	444	433
Adjusted R^2	23%	19%	24%	26%	25%	13%	29%	24%	24%	26%
Univariate R^2	21%	2%	23%	1%	26%	9%	29%	2%	21%	0%

Table C4 presents the first stage of the exercise of Tables 6 and 8, in which EBRs spreads are predicted from lagged expectations variables and from the price to dividend ratio.

Table C4

Note: The table presents portfolio-level regressions of log portfolio expectations-based returns (EBRs) at horizons (h) of one month, three months, one year, three years, and five years. Panel A presents results for the long-short value minus growth (HML) portfolio and the small minus big (SMB) portfolio. Panel B presents results for the long-short conservative minus aggressive investment (CMA) portfolio and the robust minus weak (RMW) portfolio. Panel C presents results for long-short winners minus losers (WML) portfolio. In each panel, the independent variables in the regression include: (a) the portfolio log price-dividend ratio at time t, $pd_{HML,t}$, (b) the change in the portfolio forecast for long-term earnings growth between $t-h$ and t, $\Delta_h LTG_{HML,t}$, (c) the lagged portfolio forecast for long-term earnings growth at $t-h$, $LTG_{HML,t-h}$, (d) the change in the portfolio forecast for short-term earnings growth between $t-h$ and t, $\Delta_h STG_{HML,t}$ (defined in the text), (e) the lagged portfolio forecast for short-term earnings growth at $t-h$, $STG_{HML,t-h}$, (f) the portfolio forecast error in earnings between $t-h$ and t, $FE_{HML,t-h,t}$ (defined in the text), (g) the change in the aggregate forecast for long-term aggregate earnings growth between $t-h$ and t, $\Delta_h LTG_{Mkt,t}$, and (h) the forecast for long-term aggregate earnings growth at $t-h$, $LTG_{Mkt,t-h}$.All independent variables have unit standard deviation. The last row reports the R^2 from a 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 December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] 10% level.

Panel A: Value and size

	$EBR_{HML,t+h}$					$EBR_{SMB,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$pd_{LMS,t}$	0.0078	0.0308	0.1739[b]	0.1467	-0.0028	0.0033	-0.0051	0.0314	-0.0132	-0.3141[a]
	(0.0066)	(0.0189)	(0.0882)	(0.1374)	(0.0723)	(0.0042)	(0.0130)	(0.0451)	(0.0609)	(0.0566)
$\Delta_h LTG_{LMS,t}$	0.0033[a]	0.0125[a]	-0.0671[a]	-0.0471[c]	-0.0169	0.0011	0.0029	0.0032	-0.0331	-0.0548[b]
	(0.0010)	(0.0037)	(0.0193)	(0.0244)	(0.0202)	(0.0008)	(0.0031)	(0.0098)	(0.0238)	(0.0266)
$LTG_{LMS,t-h}$	-0.0052	-0.0267	-0.3003[a]	-0.3747[b]	-0.1416	-0.0040[c]	-0.0098	-0.0642[a]	-0.2041[a]	-0.0554
	(0.0049)	(0.0166)	(0.0940)	(0.1608)	(0.0951)	(0.0022)	(0.0066)	(0.0198)	(0.0428)	(0.0529)
$\Delta_h STG_{LMS,t}$	-0.0039[a]	-0.0094[a]	0.0417[a]	0.0553[a]	0.0608[a]	-0.0004	-0.0020	0.0329[a]	0.0623[a]	0.0587[a]
	(0.0009)	(0.0025)	(0.0112)	(0.0116)	(0.0122)	(0.0008)	(0.0029)	(0.0071)	(0.0122)	(0.0075)
$STG_{LMS,t-h}$	-0.0000	0.0121	0.1612[a]	0.2695[a]	0.2152[b]	0.0170[a]	0.0363[a]	0.1224[a]	0.2175[a]	0.3404[a]
	(0.0041)	(0.0119)	(0.0421)	(0.0681)	(0.0940)	(0.0045)	(0.0133)	(0.0267)	(0.0553)	(0.0715)
$FE_{LMS,t-h,t}$			0.0371[b]	0.0053	-0.0356[b]			0.0250[a]	-0.0112	0.0173
			(0.0176)	(0.0145)	(0.0156)			(0.0070)	(0.0288)	(0.0284)
$\Delta_h LTG_{Mkt,t}$	-0.0001	0.0028	0.0277[b]	0.0318[b]	0.0961[a]	0.0014	0.0007	0.0181[b]	0.0409[b]	0.0269
	(0.0008)	(0.0032)	(0.0113)	(0.0137)	(0.0207)	(0.0010)	(0.0032)	(0.0083)	(0.0159)	(0.0193)
$LTG_{Mkt,t-h}$	0.0024[a]	0.0072[a]	0.0277[b]	0.0537[a]	0.0170	0.0000	0.0007	0.0212[a]	0.0142	-0.0321[a]
	(0.0008)	(0.0026)	(0.0131)	(0.0133)	(0.0174)	(0.0007)	(0.0023)	(0.0059)	(0.0138)	(0.0111)
Constant	-0.0174[a]	-0.0588[a]	-0.3236[a]	-0.4989[a]	-0.2083	-0.0087	-0.0122	-0.0920[c]	0.0511	0.4416[a]
	(0.0055)	(0.0210)	(0.0939)	(0.1550)	(0.1557)	(0.0067)	(0.0192)	(0.0557)	(0.1294)	(0.1579)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	14%	28%	32%	34%	56%	9%	16%	46%	56%	68%
R^2 univariate	1%	2%	0%	0%	0%	1%	6%	16%	30%	46%

Panel B: Investment and profitability

	$EBR_{CMA,t,t+h}$					$EBR_{RMW,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$pd_{LMS,t}$	0.0114[c]	0.0434[a]	-0.0049	0.0569	-0.1896[b]	-0.0035	-0.0180	-0.0394	0.0195	-0.0503
	(0.0060)	(0.0167)	(0.0462)	(0.0836)	(0.0874)	(0.0076)	(0.0195)	(0.0633)	(0.1002)	(0.1510)
$\Delta_h LTG_{LMS,t}$	0.0010	0.0014	0.0132	-0.0388[a]	-0.0103	0.0012	0.0078[a]	0.0188[b]	0.0011	0.0404[c]
	(0.0007)	(0.0036)	(0.0104)	(0.0143)	(0.0169)	(0.0009)	(0.0022)	(0.0089)	(0.0122)	(0.0216)
$LTG_{LMS,t-h}$	-0.0039	-0.0156[b]	-0.0790[b]	-0.1263[c]	-0.0366	0.0062[c]	0.0178[b]	-0.0337	-0.0954[c]	0.0889
	(0.0035)	(0.0071)	(0.0356)	(0.0732)	(0.0575)	(0.0032)	(0.0081)	(0.0394)	(0.0487)	(0.0852)
$\Delta_h STG_{LMS,t}$	-0.0045[a]	-0.0029	0.0205[b]	0.0148	0.0508[a]	-0.0009	0.0006	0.0183[b]	0.0270[b]	0.0291[c]
	(0.0009)	(0.0023)	(0.0096)	(0.0091)	(0.0089)	(0.0009)	(0.0023)	(0.0087)	(0.0125)	(0.0157)
$STG_{LMS,t-h}$	-0.0058	0.0065	0.1136[a]	0.1580[a]	0.3033[a]	0.0055	0.0229[c]	0.1453[a]	0.2332[a]	0.1429
	(0.0048)	(0.0124)	(0.0336)	(0.0454)	(0.0489)	(0.0047)	(0.0123)	(0.0400)	(0.0682)	(0.1055)

$FE_{LMS,t-h,t}$			0.0006	-0.0158	0.0074			-0.0158[c]	-0.0521	-0.0208
			(0.0071)	(0.0104)	(0.0149)			(0.0081)	(0.0345)	(0.0339)
$\Delta_h LTG_{Mkt,t}$	-0.0003	0.0014	0.0118	0.0029	0.0194	-0.0012	-0.0025	-0.0132	-0.0020	-0.0123
	(0.0007)	(0.0026)	(0.0088)	(0.0106)	(0.0134)	(0.0008)	(0.0025)	(0.0094)	(0.0141)	(0.0259)
$LTG_{Mkt,t-h}$	0.0018[a]	0.0071[a]	0.0169[a]	-0.0069	0.0407[a]	0.0007	0.0024	-0.0143	-0.0056	0.0152
	(0.0007)	(0.0021)	(0.0064)	(0.0093)	(0.0091)	(0.0009)	(0.0026)	(0.0142)	(0.0205)	(0.0386)
Constant	-0.0049	-0.0202	-0.0967[b]	0.1110	-0.1992[b]	-0.0035	-0.0131	0.0920	0.0446	-0.1949
	(0.0046)	(0.0138)	(0.0488)	(0.0904)	(0.0771)	(0.0059)	(0.0149)	(0.0778)	(0.1394)	(0.2534)
Obs	444	442	433	409	385	444	442	433	409	385
Adjusted R^2	14%	15%	17%	17%	30%	5%	13%	30%	20%	22%
R^2 univariate	3%	4%	1%	0%	0%	3%	5%	1%	2%	7%

Panel C: Momentum

	$EBR_{WML,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$pd_{LMS,t}$	0.0097[c]	0.0203	-0.0119	0.0335	-0.1154[a]
	(0.0057)	(0.0134)	(0.0387)	(0.0389)	(0.0444)
$\Delta_h LTG_{LMS,t}$	-0.0009	-0.0052	-0.0745[a]	-0.0771[a]	-0.0305
	(0.0011)	(0.0035)	(0.0149)	(0.0189)	(0.0257)
$LTG_{LMS,t-h}$	-0.0097[a]	-0.0342[a]	-0.0989[a]	-0.2397[a]	-0.0764
	(0.0032)	(0.0091)	(0.0278)	(0.0293)	(0.0723)
$\Delta_h STG_{LMS,t}$	-0.0005	-0.0025	0.0441[a]	0.0774[a]	0.0804[a]
	(0.0011)	(0.0026)	(0.0090)	(0.0094)	(0.0108)
$STG_{LMS,t-h}$	0.0059	0.0333[b]	0.2025[a]	0.4334[a]	0.4122[a]
	(0.0045)	(0.0146)	(0.0499)	(0.1076)	(0.1199)
$FE_{LMS,t-h,t}$			0.0363[b]	0.0422[c]	-0.0040
			(0.0176)	(0.0225)	(0.0237)
$\Delta_h LTG_{Mkt,t}$	-0.0008	0.0068[c]	0.0336[a]	-0.0067	0.0250[a]
	(0.0012)	(0.0037)	(0.0119)	(0.0085)	(0.0092)
$LTG_{Mkt,t-h}$	0.0049[a]	0.0151[a]	0.0318[a]	-0.0174[b]	-0.0106
	(0.0011)	(0.0026)	(0.0116)	(0.0081)	(0.0149)
Constant	0.0142[b]	0.0259	0.0589	0.2978[a]	0.2472[b]
	(0.0071)	(0.0167)	(0.0745)	(0.0776)	(0.1079)
Obs	444	442	433	409	385
Adjusted R^2	7%	25%	51%	53%	32%
R^2 univariate	0%	0%	1%	4%	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 (Table C5 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 C5
Predicted EBRs explain 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: The set of independent variables includes: (a) the log price-dividend ratio at time t, $pd_{i,t}$, (b) the change in the forecast for long-term earnings growth between $t-h$ and t, $\Delta_h LTG_{i,t}$, (c) the lagged forecast for long-term earnings growth at $t-h$, $LTG_{i,t-h}$, (d) the change in the forecast for short-term earnings growth between $t-h$ and t, $\Delta_h STG_{i,t}$ (defined in the text), (e) the lagged forecast for short-term earnings growth at $t-h$, $STG_{i,t-h}$, (f) the forecast error in earnings between $t-h$ and t, $FE_{i,t-h,t}$ (defined in the text), (g) the change in the aggregate forecast for long-term aggregate earnings growth between $t-h$ and t, $\Delta_h LTG_{Mkt,t}$, and (h) the forecast for long-term aggregate earnings growth at $t-h$, $LTG_{Mkt,t-h}$. Except for the last two variables, all variables are defined at the firm level. All independent variables have unit standard deviation. The dependent variable in Panel B is firm-level returns over the same holding horizons h, $r_{i,t+h}$. The independent variables in Panel B include: (a) $pd_{i,t}$ and (b) the EBRs predicted by the Panel A regressions ($\widehat{EBR}_{i,t+h}$). 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 December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] 10% level.

Panel A: Predicting firm level EBRs

	Dependent Variable: $EBR_{i,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$pd_{i,t}$	0.0043[a]	0.0132[a]	0.0235[a]	0.0255[a]	0.0264[c]
	(0.0005)	(0.0016)	(0.0053)	(0.0070)	(0.0140)
$\Delta_h LTG_{i,t}$	0.0019[a]	0.0010	-0.0519[a]	-0.1380[a]	-0.1577[a]
	(0.0003)	(0.0012)	(0.0058)	(0.0109)	(0.0218)
$LTG_{i,t-h}$	-0.0050[a]	-0.0221[a]	-0.1058[a]	-0.1505[a]	-0.1686[a]
	(0.0007)	(0.0027)	(0.0099)	(0.0144)	(0.0257)
$\Delta_h STG_{i,t}$	-0.0107[a]	-0.0343[a]	-0.0025	-0.0096	-0.0065
	(0.0007)	(0.0045)	(0.0140)	(0.0196)	(0.0266)
$STG_{i,t-h}$	-0.0117[a]	-0.0197[a]	0.0241[c]	0.0158	0.0554[b]
	(0.0010)	(0.0049)	(0.0143)	(0.0179)	(0.0279)
$FE_{i,t-h,t}$			0.0255[a]	-0.2916[a]	-0.5641[a]
			(0.0090)	(0.0354)	(0.0664)
$\Delta_h LTG_{Mkt,t}$	0.0040[a]	0.0136[a]	-0.0061	-0.0588[a]	-0.0205
	(0.0008)	(0.0032)	(0.0076)	(0.0175)	(0.0291)
$LTG_{Mkt,t-h}$	0.0002	-0.0025	-0.0308[a]	-0.0005	-0.0692[c]
	(0.0006)	(0.0022)	(0.0091)	(0.0244)	(0.0372)
Observations	501,342	467,941	443,350	355,688	273,440

Adj R^2	0%	2%	2%	5%	5%
F-Stat	69	41	34	28	21

Panel B: Predicted EBRs and actual returns

	Dependent Variable: $r_{i,t+h}$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
$pd_{i,t}$	-0.0014	-0.0057	-0.0301[a]	-0.0625[a]	-0.1195[a]	-0.0015	-0.0070[c]	-0.0317[a]	-0.0540[a]	-0.0826[a]
	(0.0014)	(0.0037)	(0.0101)	(0.0211)	(0.0273)	(0.0016)	(0.0037)	(0.0099)	(0.0177)	(0.0184)
$\widehat{EBR}_{i,t+h}$						0.0363	0.1370	0.4556[a]	0.5797[a]	0.8565[a]
						(0.1650)	(0.1171)	(0.1057)	(0.1246)	(0.1748)
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%	1.6%	4.4%
F-stat	1.0	2.3	8.8	8.8	19.1	0.5	2.3	10.6	13.5	15.0

APPENDIX D: Characteristics, expectations and market efficiency

We complement our analysis in a text with a mediation exercise (MacKinnon 2012), whose goal is 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:

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

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 9 in the text):

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

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, $r_{i,t,t+h}$ at horizons (h) of one month, three months, one year, three years, and five years. The independent firm-level variables include: (a) log expectations-based returns at the same horizons, $EBR_{i,t,t+h}$, (b) log book-to-market ratio at time t, $bm_{i,t}$, (c) log market value of equity at time t, $\ln size_{i,t}$, (d) one-year growth in assets between $t-12$ and t, $inv_{i,t}$, (e) operating profitability at time t, $op_{i,t}$, and (f) returns between periods $t-12$ and t, $r_{i,t-12,t}$. All specifications have firm fixed effects. All independent variables have unit standard deviation. Standard errors are corrected for overlapping observations and cross-correlations using the Driscoll and Kraay (1998) procedure. The sample period spans December 1981 to December 2023. Panel B shows the share of predictability of log firm-level returns at each horizon h accounted for by $\ln bm_{i,t}$, $\ln size_{i,t}$, $Inv_{i,t}$ $op_{i,t}$, and $r_{i,t-12,t}$ as detailed in Equations (D1, D2) 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

	Dep. variable: $r_{i,t,t+h}$				
	$h=1$	$h=3$	$h=12$	$h=36$	$h=60$
	(1)	(2)	(3)	(4)	(5)
$\mathrm{EBR}_{i,t,t+h}$	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)
$bm_{i,t}$	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)
$r_{i,t-12,t}$	-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

$bm_{i,t}$	113%	70%	46%	38%	43%
$\ln size_{i,t}$	17%	24%	22%	16%	10%
$inv_{i,t}$	-58%	-54%	-253%	385%	237%
$op_{i,t}$	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 B.6. 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 9, 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. Controlling for other characteristics, there is no explanatory power for profitability. These results offer direct evidence that analyst expectations help explain the documented predictive power of firm characteristics for future returns.

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 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 $j = 1, \ldots, h$. 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
Residualization exercise

Note: In Panel A, we regress firm-level log returns, $r_{i,t,t+h}$, at horizons (h) of one month, three months, one year, three years, and five years on the log book-to-market ratio at time t, $bm_{i,t}$. In Panel B, we regress residualized log returns—derived from separate regressions (not shown) of returns at the same horizons h on forecast errors and revisions in both long-term and short-term growth between t and $t+h$—on log book-to-market at time t. $bm_{i,t}$ has unit standard deviation. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans 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: $r_{i,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$bm_{i,t}$	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 $r_{i,t+h}$				
	$h = 1$	$h = 3$	$h = 12$	$h = 36$	$h = 60$
	(1)	(2)	(3)	(4)	(5)
$bm_{i,t}$	0.0043[a]	0.0074[a]	0.0165[b]	0.0115	0.0085
	(0.0012)	(0.0028)	(0.0065)	(0.0125)	(0.0179)
Constant	-0.0023	-0.0040	-0.0086	-0.0060	-0.0043
	(0.0015)	(0.0032)	(0.0083)	(0.0177)	(0.0231)
Obs	477,256	445,657	401,537	352,841	302,816
Adj R^2	0%	0%	0%	0%	0%
F-stat	13.3	7.0	6.4	0.8	0.2

The Table shows that $bm_{i,t}$ 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.