Finance without (exotic) risk

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

July 1 2024

Abstract

We address the joint hypothesis problem in cross-sectional asset pricing by using measured analyst expectations of earnings growth. We construct a firm-level measure of Expectations Based Returns (EBR) 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 spreads related to value, investment, size, profitability, and momentum. Second, predictable time variation in EBRs due to non-rational expectations explains predictable time variation in the spreads. This result holds even after controlling for price-scaled variables, which may capture discount-rate differentials. Third, disappointment of expectations (and hence of EBRs) is predicted by firm characteristics typically viewed as capturing risk. Predictable movements in non-rational expectations hold significant explanatory power for spreads typically attributed to exotic risk factors.

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

1. Introduction.

The capital asset pricing model (CAPM, Sharpe 1964) represents a major contribution to the understanding of financial markets. Starting from mean variance preferences and rational expectations, it offers a theory of returns that became the cornerstone of the efficient markets hypothesis (EMH): average return differences across stocks are determined by their differential exposures to market risk. Subsequent evidence on return differentials not tied to such exposure challenged the CAPM, casting doubt on the EMH (Basu 1977, 1983; Rosenberg et al. 1985; Banz 1981). However, as pointed out by Fama (1970) and Fama and French (1993), this evidence is not necessarily a rejection of the EMH, but possibly of the CAPM model of risk. This came to be known as the joint hypothesis problem: without observing expectations or risk, any test of market efficiency is also a test of a model of risk.

Mainstream finance dealt with this problem by keeping rational expectations and introducing new risk factors (Fama and French 1993, 2015), an assumption that has driven research in cross-sectional asset pricing for the past 30 years. It has however proved challenging to link these risk factors to tangible risks such as bankruptcy and distress (La Porta et al 1997). Behavioral finance has instead relaxed rational expectations, allowing for belief extrapolation, over- and underreaction, or other biases (e.g., Lakonishok et al. 1994, Barberis, Shleifer, and Vishny 1998, Hong and Stein 1999, Jegadeesh and Titman 2011, Daniel, Hirshleifer 2015, Barberis et al 2018, Kozak, Nagel, and Santosh 2018). Even in this approach, returns are matched to a model of beliefs, not to measured expectations.

We deal with the joint hypothesis problem by using measured expectations about the future earnings growth of US listed firms, so that the rationality of these expectations and their ability to explain cross sectional patterns in returns can be assessed directly. In recent years, it has become possible to measure such expectations using surveys of market analysts and other investors, which has proved useful in a variety of domains (e.g., La Porta 1996, Greenwood and Shleifer 2014, Giglio and Kelly 2018, Bordalo, Gennaioli, La Porta, Shleifer 2024, de la O and Myers 2021, Nagel Xu 2022, Jiang et al 2022, Bianchi, Ludvigson, and Ma 2024). We tie cross sectional returns to measured expectations about firm level fundamentals, which allow us to trace not only changes in market wide optimism (see Bordalo, Gennaioli, La Porta, and Shleifer, BGLS, 2024), but also co-movement across different portfolios of stocks. We then ask: can leading cross-sectional return spreads typically attributed to risk factors come from expectations? We show that, to a significant extent, the answer is yes.

In Section 2, using the Campbell-Shiller decomposition and measured expectations, we construct firm level *expectations based returns* (EBRs), which attribute all variation in returns to observed belief errors and revisions, while shutting down any cross sectional and time series variation in required return. In our key test, we show that variation in EBRs quantitatively explains most of the contemporaneous returns of the long-short high minus low book-to-market and size portfolios (HML and SMB, Fama French 1993). That is, the residual spread after accounting for expectations – the new target for risk-based stories – is close to zero and insignificant. HML and SMB are puzzles of expectations, with little room for risk.

We then extend the analysis to the other standard factors used in the literature, namely investment and profitability (making up the Fama French (2015) five factor model), as well as momentum (Jegadeesh Titman 1993). Here as well, after accounting for variation in EBRs there is little left to explain, with the only possible exception of profitability (where we however face the empirical challenge of attrition of low profitability firms with negative earnings). These results hold even after accounting for aggregate market optimism, showing that EBRs capture systematic cross-sectional variation that affects cross sectional spreads.

Why do EBRs explain cross sectional spreads? The most intuitive mechanism entails non-rational expectations: 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. Prima facie consistent with this mechanism, we show that realized spreads have large and significant loadings on forecast errors, to an extent that leaves little room for risk to explain observed average return differentials.

One challenge to this interpretation is a concern about measured beliefs: analysts may mechanically extract expectations from prices, erroneously interpreting risk-driven price movements as informative about future expected growth. BGLS (2024) address this concern for the aggregate market and firm level expectations with two key pieces of evidence: i) aggregate and firm level expectations are revised based on earnings news, not past returns, and ii) they predict aggregate and firm level stock returns even after controlling for the market's dividend price ratio. This second test is crucial, for it rejects the hypothesis that expectations are redundant proxies for the required returns embedded in prices.

In this paper we offer a version of these tests connected to the relevant objects of our analysis, cross sectional EBRs, to rule out the role of "price contamination" along this dimension. Section 4 shows that firm level EBRs are driven by earning news even after controlling for contemporaneous returns. EBRs are not mechanically tied to returns.

The more powerful evidence comes from two return predictability tests in Section 6. The first test directly addresses the joint hypothesis problem: we predict future cross-sectional spreads and EBRs using current portfolio-level expectations *and* price-scaled variables. Because the latter capture both market expectations and required returns, in an efficient market they should be the only source of predictability (including if measured expectations are mechanically inferred from prices). In the data, however, expectations predict future returns and EBRs to an extent that dwarfs price-based predictability. Market inefficiency appears to play a key role in producing cross sectional return spreads.

In the second test we predict future firm level EBRs using firm characteristics such as book to market, investment, etc. These characteristics are typically viewed as risk factors but have not been reliably connected to tangible risks. We show that, in contrast, the same characteristics are strongly tied to non-rational beliefs: high book to market, strong investment, etc predict future belief disappointment and downward forecast revisions, in the form of low firm level EBRs. Market inefficiency reconciles predictable returns and the role of characteristics, proving empirically more plausible than risk-based required returns.

Our results have significant implications for the classic risk-based analysis of cross-sectional return spreads but also for more recent attempts to introduce 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, Giglio, and Polk 2023). These papers do not offer direct measures of risk, so one cannot exclude that their horizon-dependent and time-varying return patterns may be generated by non-rational expectations, whose biases have indeed been shown to be horizon-dependent and time varying (e.g. BGMS 2019, BGLS 2024). In fact, our return predictability tests in Section 6 show that cross sectional spreads are time varying: they are particularly high at times in which analysts are very optimistic about the portfolio's short arm and low otherwise. Future work may study this time variation further.

We also connect to fast-growing research showing that measured expectations allow an empirically disciplined and theoretically structured approach to asset prices that helps solve the joint hypothesis problem. La Porta (1996) showed that LTG predicts low returns due to its overreacting to good news and subsequently correcting (Bordalo, Gennaioli, La Porta, and Shleifer 2019). Aggregate LTG is also excessively volatile, leading to sufficiently volatile valuations to quantitatively explain Shiller's (1981) excess volatility puzzle (Bordalo et al. 2024). De la O and Myers (2021) show that volatility of valuation ratios reflects short term earnings expectations. BGLS (2024) show that overreaction leads to departures from market efficiency: forecast errors of aggregate LTG are predictable, and these errors and their correction account for essentially all predictability of market level returns from valuation ratios such as the aggregate price dividend ratio (see also Nagel and Xu 2022, Adam and Nagel 2023). Frey (2023) analyses a large number of potential factors and shows that many predict convergence of earnings growth forecasts between the long and short arms of the factor. Expectations account for puzzles about returns on the aggregate market without the need to invoke time varying required returns. Here we offer a new method using EBRs to develop a systematic analysis of classical cross-sectional return spreads.

2. Concepts and Methods

2.1 Risk in Efficient Markets Finance

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

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

where $p_{i,t}$ is the firm's log price at t, $d_{i,t}$ is its log dividend, $g_{i,t+1} = d_{i,t+1} - d_{i,t}$ is its dividend growth between t and $t+1$, while k and α are constants that depend on the mean log price dividend ratio (we equalize k and α across firms as in Campbell and Mei 1993). By iterating Equation (1) forward and ruling out "bubbles" we obtain the ex-post identity:

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

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

$$p_{i,t}^e - d_{i,t} = \frac{k - r_i}{1-\alpha} + \sum_{s\geq 0} \alpha^s \widetilde{\mathbb{E}}_t\left(g_{i,t+1+s}\right). \tag{3}$$

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

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

which is higher if the firm is riskier, the required return r_i is higher, or if good news are received on the firm, either as a positive dividend growth surprise (the term in square brackets), or as an upward revision of expectations of future growth (the sum).

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

$$\overline{r}_{i,1} = r_i. \tag{5}$$

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

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

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

2.2 Expectations Based Returns (EBRs)

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

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

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

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

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

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

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

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

2.3 Data and Construction of EBRs

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

We collect median forecast data on dividends for the upcoming fiscal year from IBES, and use them to compute the stock's expected payout ratio (see Section 2.3). 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.

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

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

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

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

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

We derive expected dividends per share from expected earnings per share, as follows. We proxy expected earnings per share with analyst short term earnings expectations $\widetilde{\mathbb{E}}_t EPS_{i,t+12s}$ up to the second fiscal year; starting with the last non-missing positive EPS forecast and up to five years out, analysts expect EPS_{it} to grow at the rate LTG_{it}. To translate expected earnings into expected dividends, we use the expected payout ratio implied from analysts' expectations of dividends and earnings. Specifically, we assume a constant ratio equal to the average expected payout ratio $\frac{\widetilde{\mathbb{E}}_t DPS_{i,t+12}}{\widetilde{\mathbb{E}}_t EPS_{i,t+12}}$ in our sample for those firms which paid dividends that year, $DPS_{i,t+12} > 0$, which equals 0.41.[2]

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

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

The firm level raw EBR is extended to the monthly raw return of a portfolio π using an equal weighted average, $\mathrm{EBR}^{\mathrm{m,r}}_{\pi,t,t+1} = \frac{1}{|\pi|}\sum_{i\in\pi} \mathrm{EBR}^{\mathrm{m,r}}_{i,t,t+1}$.[3] Our results also hold when using value weighted portfolios (see Appendix). To compute (log) returns over longer horizons t to $t+h$, we compute 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 raw returns. We then rebalance portfolios at the end of each month j and compound monthly EBRs to obtain the (log) return:[4]

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

For example, the one month log EBR equals $\ln\left(\mathrm{EBR}^{\mathrm{r}}_{\pi,t,t+1}\right)$ which, under the Campbell Shiller approximation of Equation (9), takes the form of Equation (5). Finally, we obtain EBRs for factor portfolios as the difference between the returns of the portfolio's long and short arms. Specifically, we compute EBRs for the following long-short portfolios:

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, $\mathrm{EBR}_{HML,t,t+h} = \mathrm{EBR}_{V,t,t+h} - \mathrm{EBR}_{G,t,t+h}$.
2. 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). $\mathrm{EBR}_{CMA,t,t+h} = \mathrm{EBR}_{C,t,t+h} - \mathrm{EBR}_{A,t,t+h}$.
3. 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). $\mathrm{EBR}_{RMW,t,t+h} = \mathrm{EBR}_{R,t,t+h} - \mathrm{EBR}_{W,t,t+h}$.

[3] IBES surveys analysts in the middle of each month (i.e. the Thursday before the third Friday of every month, see IBES Unadjusted US Summary Statistics file). We use CRSP daily file to compute actual returns over the same periods as EBRs. Results are similar if we compute actual returns using calendar months, but the correlation between one- and three-months EBRs and returns is slightly stronger when using IBES.
[4] Though we construct expectation based returns from measures of analyst expectations, for simplicity we use the same notation as for EBRs based on market beliefs, Equation (6).

4. 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), *i.e.* $\text{EBR}_{SMB,t,t+h} = \text{EBR}_{S,t,t+h} - \text{EBR}_{B,t,t+h}$.
5. Winners Minus Losers momentum (WML): EBRs of a portfolio that is long winning stocks ($\pi = W$, top quintile prior returns between periods $t - 11$ and $t - 1$) and short losing stocks ($\pi = L$, bottom quintile), i.e. $\text{EBR}_{WML,t,t+h} = \text{EBR}_{W,t,t+h} - \text{EBR}_{L,t,t+h}$.

Our sample consists of monthly firm level observations from 1981 to 2023 for which LTG_t and LTG_{t+h} exist.[5] This requirement restricts our sample from the 2 million observations in the CRSP/Compustat database to about 1.3 million observations for $h = 1$ and 1.1 million for $h = 12$. The sample drops firms that tend to be smaller in market cap, but the samples are comparable in characteristics such as book to market (0.8 in the full sample, 0.6 in our samples) and investment (0.18 in the full sample, 0.19 in our samples). 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 very similar to our sample in all characteristics. Appendix B provides descriptive statistics of the samples and shows that our results hold when imposing added restrictions on the dependent variable.

2.4 Raw Portfolio EBRs and Correlations with Actual Returns

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

Table 1. Average returns and EBRs of portfolios

Note: Panel A presents sample means of log portfolio returns over holding horizons h ranging from one month to five years, following the methodology outlined in the website of Ken French. Portfolios are formed independently based on quintiles. Results are displayed for the following five quintile portfolios: (1) book-to-market, with *Growth* stocks in bottom quintile and *Value* stocks in the top quintile, (2) investment, *Aggressive* stocks in the bottom quintile and *Conservative* ones in the top quintile, (3) size, *Big* stocks in the top quintile and *Small* ones in the bottom quintile, (4) Profitability, *Weak* profitability in the bottom quintile and *Robust*

[5] 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).

profitability in the top quintile), and (5) Momentum, *Losers* stocks in the bottom quintile and *Winners* stocks in the top quintile. Panel B presents sample means of log expectation based returns (EBR) returns computed following equation 9 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. For all horizons, the last observation is based on returns from December 2018 onwards. Please see text for details

Panel A. Average portfolio returns

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

Panel B. Average portfolio expectation based returns

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

In line with existing work, Panel A shows that portfolios in the long arm exhibit higher average returns than those in the short arm, at both long and short horizons. An assessment of the statistical significance of the long-short return spreads (Tables 4 and 8 below) reveals that the value and investment spreads are large and significant at all horizons in our sample, with annualized spreads between 3 and 5%. Momentum spreads are large and significant at horizons of under a year, with annualized spreads of around 5%. These are key targets of our analysis. The size spreads are instead not significant in our sample, which is 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). Average profitability spreads are also not significant in our sample period.[6]

[6] 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.

Panel B shows that EBRs display quantitatively similar patterns for HML, CMA, SMB and WML. For these portfolios, target and explanatory variables have similar magnitudes and go in the same direction, suggesting that i) expectation errors and revisions captured by in EBRs differ systematically across portfolios, and ii) systematic differences in EBRs align with average return differences. The exception is profitability for which we observe the reverse spread, which may again be driven by attrition of unprofitable firms with negative earnings, whose expectations based returns are unmeasured by may be low.

Average return spreads are only one dimension of comparison between EBRs and true returns. For EBRs to be a good proxy for the latter, the two must also be positively correlated. The "perfect proxy" benchmark of Equation (8) implies that the correlation coefficient should be one, because the variance in realized return is entirely driven by EBRs. We compute the correlation between $EBR_{\pi,t+h}$ and $r_{\pi,t+h}$ for the long and short portfolios of HML, CMA, SMB, RMW and WML. We consider horizons h of $\{1,3,12,36,60\}$ months, covering the short horizons typical of the cross-sectional analysis (e.g., Fama and French 1993, 2015) as well as longer horizons typical of reversal anomalies (de Bondt and Thaler 1985) and aggregate stock market variation (BGLS 2014). Table 2 reports the results.

Table 2. Portfolio level correlations for actual and expectation based returns.

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

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, showing that EBRs constitute good proxies for news. In all cases, though, the correlation coefficient is less than one, the theoretical "perfect proxy" benchmark. There are two main reasons to expect EBRs to constitute an imperfect proxy. First, analyst beliefs may depart from market beliefs due to (unobserved) disagreement between the marginal investor and the analyst consensus. Second, analyst beliefs only cover horizons up to 5 years out, so our proxy for EBRs surely misses longer term variation in market expectations. A third possible reason for the imperfect correlation in Table 2 is that the required return spread in Equation (8) is time varying, so it yields a distinct source of return variation not captured by earnings growth expectations. This seems unlikely given that both fundamental risk and investor preferences toward it arguably vary more at lower frequencies, but correlations between returns and EBRs are uniformly and counterfactually stronger at lower frequencies.[7]

The noise in our proxy for EBRs implies that we should treat them as only a proxy for market beliefs, so we should not expect EBR to perfectly mimic actual returns and average spreads. We therefore cannot assess the explanatory power of EBRs by looking at their implied average return spread, but rather by regressing contemporaneous returns on contemporaneous EBRs, so that the estimated coefficients optimally extract the information contained in EBRs, partially correcting for measurement error. Of course, the measurement error contaminates the regression constant in (8), which is our measure of risk-based spread. In the next section we present one method to adjust the estimated constant under specific assumptions about measurement error in EBRs. Looking at the regression constant also offers a quantitative assessment of any need for time varying required returns to account for average return spreads after belief variation is controlled for.

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

To study whether EBRs help account for cross sectional spreads we proceed as follows. In Section 3, we assess the roles of expectations versus risk by constructing EBRs for the value minus growth and small minus big long-short portfolios and using them to estimate Equation (8). We also assess the contribution of different components of EBR by separately introducing into (8) the realized forecast errors and forecast revisions at different horizons, which offers a direct test of Equation (4) without making the parametric restrictions embedded in the construction of the analyst price. We then allow analyst expectations to serve as an imperfect proxy for market beliefs about future earnings growth and develop a method that allows us to correct the coefficients obtained from estimating Equation (8) for measurement error. Our empirical analysis shows that such discrepancies indeed exist but that accounting for them usually makes little difference for the entailed required return differential $r_L - r_S$. In Section 4, we repeat the analysis for the other long-short portfolios.

3. EBRs and the Value Premium

In Section 3.1 we implement our test in Equation (8) on the book to market and size long short portfolios in the Fama French (1993) three factor model. In Section 3.2 we present a second, less structured, way to deal with the joint hypothesis in Equation (8) by regressing the value and size spreads on the contemporaneous portfolio forecast errors and forecast revisions at different horizons. After accounting for EBRs, there is little systematic variation in the value and size spreads left for risk to explain.

3.1 EBRs Explain the Value and Size Premia

Table 3, Panel A reports regressions of the actual value and size long-short portfolio spreads $r_{LMS,t+h}$ on $\text{EBR}_{LMS,t+h}$, as specified in Equation (8), for various horizons h. In this regression, the constant term represents a first estimate of the required return spread, under the assumption that analyst beliefs are a perfect proxy for market beliefs.

Table 3
Expectation based returns and the HML and SMB spreads

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

Panel A

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

Panel B

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

EBRs and actual returns are significantly positively correlated for the HML and SMB portfolios, especially at longer horizons. The R^2 also sharply rises with the horizon: it is 6% at one month, and 40% or more at horizons of one to five years. Expectations thus contain substantial information about the news perceived by the market. Consistent with Equation

(8), the estimated coefficient on EBRs is close to one, and is statistically indistinguishable from one, for most horizons. It is smaller than one at the monthly horizon, suggesting that analyst forecasts are a noisier proxy for market beliefs at higher frequencies.

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

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

To answer this question, we compute the market-level expectations-based return $\mathrm{EBR}_{Mkt,t,t+h} = \frac{1}{|M|}\sum_{i\in Mkt}\mathrm{EBR}_{i,t,t+h}$, and run a horse race between the aggregate EBR, $\mathrm{EBR}_{Mkt,t,t+h}$, and the portfolio ones, $\mathrm{EBR}_{HML,t,t+h}$ and $\mathrm{EBR}_{SMB,t,t+h}$, in accounting for the contemporaneous observed return spread, Table 3 Panel B. The question here is not only whether the portfolio EBR survives in the regression, but also the extent to which adding it affects the regression R^2 and the estimated regression constant compared to Panel A.

The results of Panel B indicate that the cross-sectional value spread is mostly accounted for by cross sectional movement in expectations. The estimated coefficients of $\mathrm{EBR}_{HML,t+h}$ and $\mathrm{EBR}_{SMB,t+h}$ remain similar to those in Panel A, and are always statistically significant. The proxy for market-wide growth in optimism, $\text{EBR}_{Mkt,t+h}$ plays some role at the one month and 5 years horizons, it adds a bit of explanatory power in terms of R^2, and it causes the regression constant to shrink, further reducing the need for required return differentials. Broadly speaking, the value and size premia are largely due to cross sectional cycles in expectations, as proxied by $\text{EBR}_{HML,t,t+h}$ and $\text{EBR}_{SMB,t,t+h}$. This finding is consistent with the results in BGLS (2019), and suggests that understanding the drivers of systematic cross sectional expectations is important for future work.

As seen in Table 3, EBRs are a good proxy for market expectations, consistent with a growing body of work showing their explanatory power for prices and returns (De la O and Myers 2024, Bordalo et al 2024, BGLS 2024), but not a perfect measure, especially at short horizons. We next allow analyst expectations to constitute an imperfect proxy of market beliefs and develop a correction for the estimated required return differential that takes measurement error into account. Our correction builds on the following affine-stochastic discrepancy between analyst forecasts $\tilde{\mathbb{E}}_{it}^m$ and market forecasts $\tilde{\mathbb{E}}_{it}$ at time t about stock i:

$$\tilde{\mathbb{E}}_{it}^m = \beta + \tau \cdot \tilde{\mathbb{E}}_{it} + \sigma \cdot \varepsilon_{it}. \tag{10}$$

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

As we show in the appendix, given the measurement error structure in (10), we can use the estimated constant κ and slope γ in Table 3 and other known moments in the data to recover the distortion parameters and adjust the original estimate of the required return spread κ for measurement error. Out of parameters (β, τ, σ) the analysts' systematic bias β is not relevant for the spread because it cancels out when comparing different portfolios. By contrast, noisier expectations, higher σ, or excess reaction to news by analysis, $\tau > 1$, both dampen the regression slope γ. This, in turn, affects the spread estimate κ.

Parameters σ and τ can be recovered using two moments of the data. The first is the regression slope γ, and in particular its deviation from one. The second is the gap between the covariance of observed earnings growth g_{it} with $EBR_{i,t}$ and its covariance with actual returns $r_{i,t}$.[8] Under the maintained assumption that the required return is constant, the two covariances should be equal. This yields a data moment to back out an additional parameter distorting analyst expectations compared to market ones. This estimation only uses the restriction that the required return is constant but allows for any possible risk model satisfying that restriction, including conventional unconditional factor models of risk premia.[9]

Applying this method, we find that the average τ is 1.26 so that analysts tend to respond to news more strongly than the market. We can use these estimated parameters to adjust the estimated constant κ of the required return premium of the HML and SMB portfolios that account for measurement error in analyst expectations. Table 4 compares the actual return spread between the value and growth portfolios, and the small and big firm portfolios, to the adjusted required return measure (both measures are annualized).

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

Note: the table presents estimates of the required return premia for the portfolio that is long value and short growth (HML, Columns 1 to 5) and for the portfolio that is long small firms and short big firms (SMB, Columns 6 to 10). The adjustment allows for three distortions in expectation based returns (EBR) as described in Equation (10). As benchmarks, we report (in the first row) the sample long-short spreads for the relevant portfolios for horizons of one-month, three-month, one-year, three-year, and five-year horizons. The second row

[8] The adjusted estimate is equal to (where variables capture differences between value and growth portfolios):

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

If there is no measurement error, not only $\gamma = 1$, but the covariance of actual return and cash flow growth is identical to the covariance of EBRs and cash flow growth, $cov(r_t, g_t) = cov(\mathrm{EBR}_t, g_t)$, so the estimate for the required risk premium is equal to the estimated regression constant, $r = \kappa$, as in Equation (7).

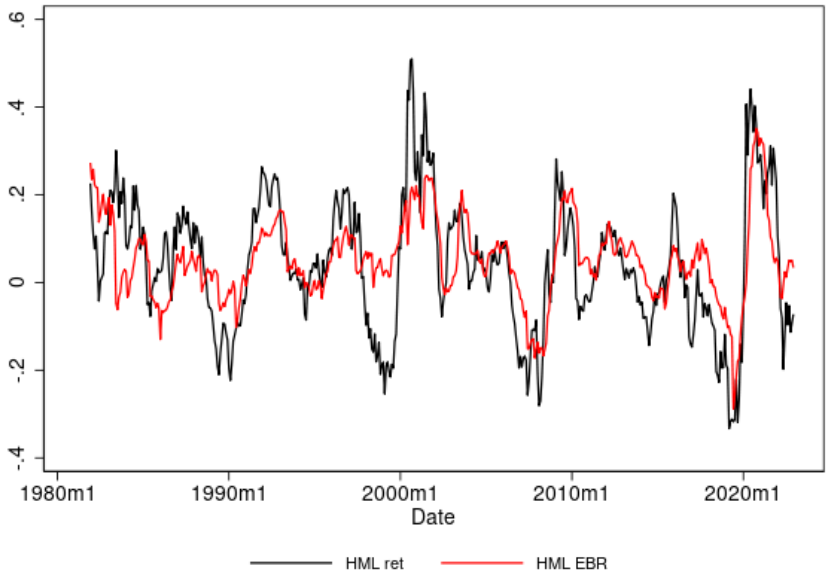
[9] This exercise also allows to estimate the extent to which time variation in required returns unrelated to time variation in non-rational analyst expectations may be needed to account for the return spread observed on average. A direct horse race between possibly time varying required returns and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

reports the intercept from a univariate regression of annualized log returns of relevant long-short portfolio and horizon h on their EBRs. The last row reports annualized estimates of the required risk premia. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period extends from December 1981 to December 2018. For all horizons, the last observation is based on returns from December 2018 onwards. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

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

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

In sum, the evidence is consistent with the hypothesis that the observed cross-sectional differences in expectations explain HML and SMB spreads. Growth stocks do worse when optimism about them drops relative to that about value stocks, and similarly for big stocks relative to small ones. In Figure 1, growth stocks do worse on average because this happens more than its opposite, rather than because growth stocks are less risky. But in principle, it is entirely possible that at particular times the value and size spreads may be negative, when the prospects of high growth (resp. big) firms are underestimated while those of value (resp. small) firms are overestimated. In this sense, time variation in the spread is more fundamental than its average value. Understanding the predictable time variation in analyst expectations and its connection to the return spreads is thus an important way forward, which we tackle in Section 6 when we perform our spread-predictability analysis.



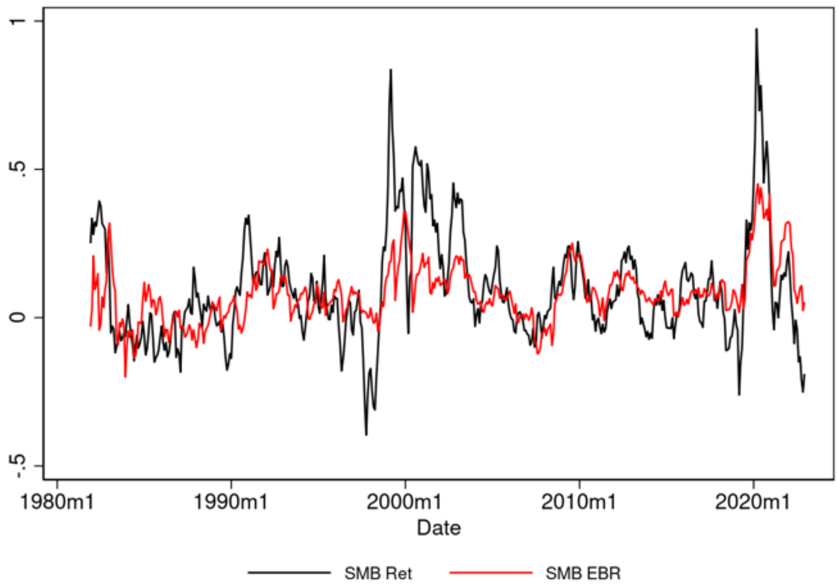


Figure 1. Time series of actual and expectation based HML and SMB returns.

3.3 An Alternative Joint-Hypothesis Strategy: Decomposing EBRs

Rather than computing a proxy for EBRs, we can test for the return Equation (4) by directly regressing realized returns on contemporaneous forecast errors and revisions at different horizons. This alternative strategy is informative for two reasons. First, it relaxes the parametric restrictions embedded in our computation of EBRs, allowing for measured expectations to also capture correlated unmeasured variation in beliefs about longer horizons.

Second, and related, it allows to separately assess the explanatory of different components of EBRs: the forecast error and forecast revisions at different horizons.

We perform the decomposition at the yearly horizon or above, because the forecast error cannot be computed at the monthly and quarterly levels. The shortest horizon at which expectations are formed is one year.

For one and three years horizons, $h = 12, 36$, we compute the firm level forecast error as the difference between realized one or three year earnings growth and the growth expected one or three years prior $FE_{i,t+h} = \ln\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right) - \ln\left(\mathbb{E}_t\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right)\right)$. At five year horizons, we compute the forecast error using LTG as $FE_{i,5} = \ln\left(\frac{EPS_{i,t+h}}{EPS_{i,t}}\right)/5 - LTG_{i,t}$. We then compute forecast revisions as follows. Revisions of short term growth forecast, namely forecasts at horizons $h = 12{,}24$ are computed as $\Delta_h STG_{i,t+h} = (\mathbb{E}_{t+h} - \mathbb{E}_t) ln\left(\frac{EPS_{i,t+h+1}}{EPS_{i,t+h}}\right)$ and those of long term forecasts by $\Delta_h LTR_{i,t+h} = LTG_{i,t+h} - LTG_{i,t}$.

We aggregate each measure of forecast error and revision at the portfolio level, e.g. for forecast error we compute $FE_{\pi,t+h} = \frac{1}{|\pi|}\sum_{i\in\pi} FE_{i,t+h}$ and we analogously aggregate forecast revisions. We use the differences in these aggregated forecast errors and revisions between the long and short portfolios as explanatory variables for contemporaneous long minus short return spreads.[10] Table 5 shows the results.

Table 5

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

[10] Earnings are published quarterly so we refrain to compute forecast errors at a 1 month or 3 month frequency. 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 $EPS_{i,t+h}$.

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

Two results stand out. First, the expectation components are strongly predictive of HML spreads, with forecast errors playing a dominant role in the multivariate regression (note the regressors are standardized). The pre-eminence of forecast errors for both factors is consistent with the view that returns reflect a disappointment of the short arm compared to the long arm, i.e. of growth stocks compared to value stocks for HML. Portfolio expectations revisions also play a role, so part of the HML spread is accounted for by 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 5 show that the average returns spreads are almost entirely explained away when we account for expectations. In fact, the regression constants here for both HML and SMB are negative and statistically significant, but we cannot interpret this finding as cleanly as in Table 3 given that this test is less theoretically disciplined.[11]

In sum, estimating Equation (7) shows that the systematic return spreads on the Fama French HML and SMB factors are explained by EBRs, and in particular by systematic

[11] We consider the robustness of these results when controlling for the price dividend ratio $pd_{LS,t} = pd_{L,t} - pd_{S,t}$, as a proxy for the difference in required returns across portfolios. Appendix D shows that the expectation components continue to predict returns, while $pd_{LS,t}$ are weak, negative predictors of returns (and statistically insignificant at 1 year horizon). Interestingly, the coefficient on forecast errors drops dramatically, consistent with the view that prices are capturing expectations as well.

differences in forecast errors and revisions (Tables 3 and 5). Taking these into account, there is little evidence of systematic differences in required returns (Table 4).

4. Other FF factors

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

Table 6
Actual and expectations based long short portfolio return spreads

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

	$r_{LMS,t+1}$	$r_{LMS,t+3}$	$r_{LMS,t+12}$	$r_{LMS,t+36}$	$r_{LMS,t+60}$
Investment (CMA)	(1)	(2)	(3)	(4)	(5)
$\text{EBR}_{LMS,t,t+h}$	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)					
$\text{EBR}_{LMS,t,t+h}$	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					
$\text{EBR}_{LMS,t,t+h}$	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, expectation-based returns have strong explanatory power for actual returns. The slope coefficients are large, statistically significant, and increase with the holding horizon. For CMA and RMW, their magnitudes are comparable to those obtained for HML and SMB returns; in particular, for longer horizons coefficients are close to, or statistically indistinguishable from, the benchmark value of 1 for investment. For momentum, and particularly for profitability, they are lower than 1 although still substantial throughout.

Turning to our main test, EBRs account for the average return spread for all factors except profitability, in the sense that the intercepts are either small and statistically indistinguishable from zero, or negative, with the exception of short horizons for investment and size. The negative coefficients for momentum may be consistent with a negative spread in required returns whereby winners are deemed safer than losers.

For profitability, intercepts are positive and significant. This is in line with the finding in Table 1 that the average EBR is higher for low profitability firms, while actual returns are directionally higher for high profitability firms. Recall, however, that comparing average EBR spreads to average return spreads, as in Equation (8), yields a upward-biased measure of required returns when the estimated slope γ is smaller than 1 (see footnote 6), which is particularly pronounced for profitability. We therefore adjust the estimated required return for measurement error across portfolios, and present the results in Table 7.

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

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

	$r_{LMS,t+1}$	$r_{LMS,t+3}$	$r_{LMS,t+12}$	$r_{LMS,t+36}$	$r_{LMS,t+60}$
Investment (CMA)	(1)	(2)	(3)	(4)	(5)

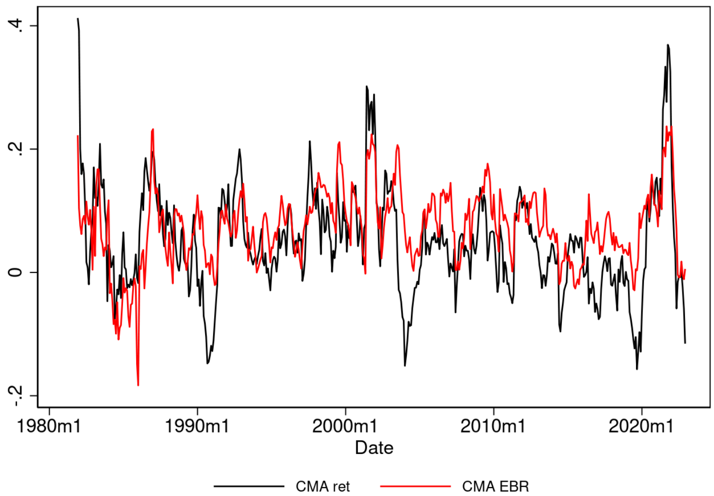
Average spread	0.0588[a] (0.0124)	0.0571[a] (0.0113)	0.0515[a] (0.0106)	0.0329[a] (0.0081)	0.0249[a] (0.0068)
Constant κ	0.0361[b] (0.0140)	0.0162 (0.0123)	-0.0116 (0.0130)	-0.0047 (0.0137)	-0.0085 (0.0056)
Adjusted κ	-0.0212	-0.0240	-0.0264	-0.0076	-0.0068
Profitability (RMW)					
Average spread	0.0221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

The adjustment broadly confirms the message of the estimated regression constant. For investment the correction proves relatively important for spreads at short rather than long horizons. Regarding profitability and momentum, Table 6 suggests EBRs have more noise and the corrections are accordingly larger. For profitability the estimated required return spreads decrease, particularly at longer horizons. The corrections for momentum are in line with the earlier interpretation that firms in the long portfolio (winners) are if anything viewed as safer than those in the short portfolio.

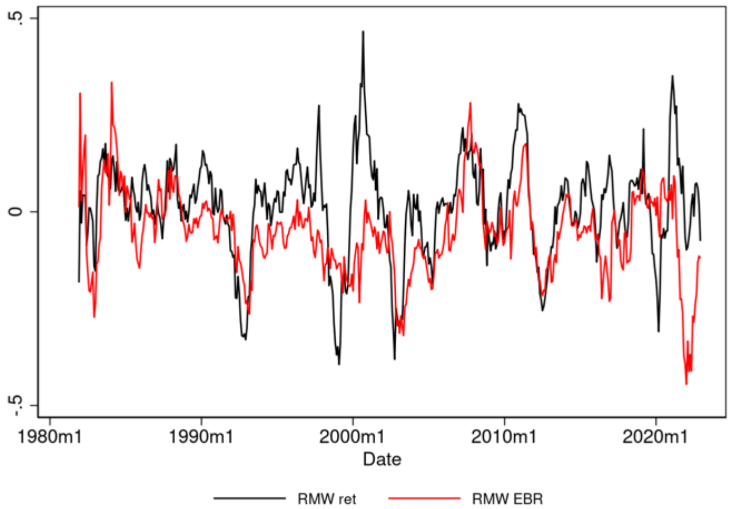
These patterns are confirmed in the EBR decomposition exercise, which is reported in the Appendix: spreads in forecast errors and revisions positively and significantly predict return spreads and the intercepts are either small and insignificant -- for investment and profitability -- or negative, for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

In sum, the investment, size and momentum puzzles are solved with expectations. As with HML, the market does not see conservative firms as riskier than aggressive ones, nor winners as riskier than losers. Instead, analysts and the market appear to have systematically bullish expectations about firms in the short portfolios, compared to firms in the long portfolio, and the former do worse on average because that relative optimism systematically decreases. Our evidence suggests the same may be true for profitability. Figure 2 plots the time series of actual and expectation based spreads of these long-short portfolios.

Panel A: Investment



Panel B: Profitability



Panel C: Momentum

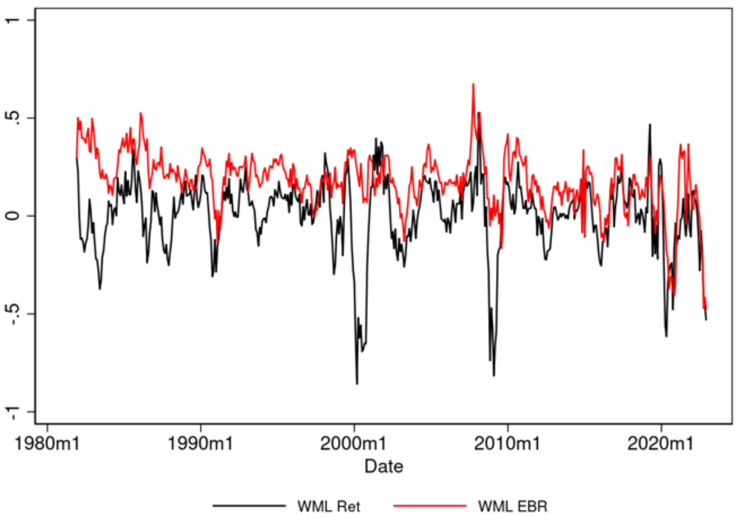


Figure 2. Actual and expectation based return spreads, CMA, RMW and WML.

Relative to Table 6, Figure 2 further shows that EBR spreads track actual spreads well across time. As for book to market and size (Figure 1), EBRs systematically capture periods of high spreads but also correctly account for periods of negative spreads. The ability to account for the time variation on spreads, and in particular negative spreads, is a key implication of the expectations mechanism. EBRs exhibit less extreme fluctuations than actual returns despite the fact that the slope coefficients are often in the neighbourhood of 1. This may correspond to periods where expectations for growth beyond 5 years matter and correlate with LTG (variation in such expectations is currently shut down in EBRs), or where the distribution of expectations beyond that of the median analyst plays a role.

5. Validity of expectations data

A possible critique of our analysis concerns the validity of analyst forecasts as a proxy for expectations. We already discussed measurement noise, which works against finding a systematic association between EBRs and realized returns. Another concern, in the opposite direction, is that analysts may infer earnings growth expectations from market prices under the erroneous assumption of constant, and perhaps CAPM-based, returns. If so, analysts could erroneously incorporate in cash flow expectations priced information about required returns.[12] BGLS (2024) present two approaches that reject this hypothesis.

First, revisions in LTG at both the market and the firm levels are more reliably explained by past earnings growth than by past stock returns. Changes in measured beliefs respond to realized fundamentals, and do not appear to mechanically respond to prices. The lack of a mechanical correlation between beliefs and prices is also evident from our previous analysis, which shows that realized returns and EBRs are imperfectly correlated (R^2s are below one in Tables 3 and 7, see also Figures 1 and 2).

Second, BGLS (2024) show that measures of expectations predict future returns both in the aggregate and at the firm level, while controlling for current price scaled variables such as the (aggregate or firm-level) price dividend ratio. This is a key test: if expectations surreptitiously captured discount rate variation embedded in market prices, their predictive power in a horse race with the latter would be zero. But this is strongly not the case, indicating that measure expectations contain genuine information about non-rational market beliefs that affects prices and helps predict future returns.

Here we expand this analysis to our EBRs proxy and to the cross-sectional portfolios. Compared to the firm level analyses in BGLS (2024), portfolios aggregate many stocks so idiosyncratic variation in analyst forecasts is less likely to matter. The exercise also checks for the possibility that "price contamination", while not relevant for aggregate and firm level expectations, is relevant for our portfolios. Finally, the return predictability exercise is of independent interest because it allows us to assess whether market inefficiency produces systematic time variation in cross-sectional spreads, just as is produces systematic return variation in the aggregate market as shown in BGLS (2024).

[12] In particular, if equity analysts infer cash flow expectations from prices while correctly adjusting for the true model of risk, then they would in fact recover market expectations about earnings growth.

For the first exercise, Table 8 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. We present results for 1 month and 1 year horizons, with other horizons presented in the Appendix.

Table 8
Expectation based portfolio returns and contemporaneous news

Note: Panel A presents portfolio-level univariate regressions of expectation based returns (EBR) at horizons (h) of one-month and one-year. The independent variable is: (a) log returns between t and $t+h$. Panel B presents portfolio-level multivariate regressions of expectation based returns (EBR) at horizons (h) of one-month and one-year. The independent variables include: (a) log returns between t and $t+h$, (b) earnings growth between t and $t+h$, and (c) the forecast error for earnings growth between t and $t+h$. 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

	$EBR_{HML,t+h}$		$EBR_{SMB,t+h}$		$EBR_{CMA,t+h}$		$EBR_{RMW,t+h}$		$EBR_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t,t+h}$	0.7452^{a}	0.4476^{a}	0.6280^{a}	0.2982^{a}	0.7165^{a}	0.4534^{a}	0.6498^{a}	0.3335^{a}	0.7611^{a}	0.3904^{a}
	(0.1312)	(0.0498)	(0.0982)	(0.0426)	(0.1372)	(0.0581)	(0.1430)	(0.0821)	(0.1269)	(0.0627)
Constant	0.0418^{a}	0.0270^{a}	0.0127	0.0041	0.0750^{a}	0.0553^{a}	-0.0500^{a}	-0.0545^{a}	0.2016^{a}	0.2045^{a}
	(0.0121)	(0.0086)	(0.0104)	(0.0080)	(0.0074)	(0.0068)	(0.0116)	(0.0108)	(0.0161)	(0.0142)
Obs	493	493	493	493	493	493	493	493	493	493
Adj R^2	8%	46%	9%	34%	6%	36%	5%	16%	8%	29%

Panel B

	$EBR_{HML,t+h}$		$EBR_{SMB,t+h}$		$EBR_{CMA,t+h}$		$EBR_{RMW,t+h}$		$EBR_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$r_{LMS,t,t+h}$	0.6313^{a}	0.2819^{a}	0.6313^{a}	0.2819^{a}	0.6677^{a}	0.3262^{a}	0.5551^{a}	0.1457^{a}	0.7864^{a}	0.2403^{a}
	(0.1111)	(0.0492)	(0.1111)	(0.0492)	(0.1323)	(0.0528)	(0.0901)	(0.0349)	(0.1248)	(0.0456)
$\Delta_h e_{LMS,t+h}$	1.7713^{a}	0.2801^{a}	1.7713^{a}	0.2801^{a}	0.4480	0.0280	1.2244^{a}	0.0126	0.9235^{b}	0.1464^{c}
	(0.3345)	(0.0686)	(0.3345)	(0.0686)	(0.2841)	(0.0521)	(0.3545)	(0.0685)	(0.4523)	(0.0749)
$(1-E_t)$ $\Delta_h e_{LMS,t+h}$		0.1342^{b}		0.1342^{b}		0.2402^{a}		0.3297^{a}		0.4035^{a}
		(0.0639)		(0.0639)		(0.0470)		(0.0535)		(0.0553)
Constant	0.0338^{a}	-0.0097	0.0338^{a}	-0.0097	0.0675^{a}	0.0318^{a}	0.0135	0.0355^{a}	0.1360^{a}	0.0634^{a}
	(0.0106)	(0.0131)	(0.0106)	(0.0131)	(0.0087)	(0.0069)	(0.0098)	(0.0098)	(0.0357)	(0.0190)
Obs	493	493	493	493	493	493	493	493	493	493
Adj R^2	24%	57%	24%	57%	8%	51%	17%	58%	12%	62%

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

6. Predictable returns and market efficiency.

The central puzzle that motivates our paper is the predictability of cross-sectional return spreads from firm characteristics, which are sometimes interpreted are proxies for risk. Informed by our previous analysis, we now ask two questions that circle back to such predictability. First, can cross-sectional returns be predicted using measured expectations, including after controlling for the current stock prices? Second, do standard firm characteristics predict future EBRs (and hence future forecast errors and revisions)?

An affirmative answer to the first question offers evidence for market inefficiency: non-rational expectations (not required returns embedded in prices) are the source of return predictability. An affirmative answer to the second question ties market inefficiency to firms' characteristics, as these characteristics predict the forecast errors that produce return spreads. Given the notorious difficulty of showing that characteristics are connected to tangible risks, this evidence would offer a direct mechanism for why characteristics matter.

6.1 Predicting Portfolio EBR and Return Spreads from Expectations

We analyze return predictability by testing first whether current expectations predict future EBRs, and second whether the same expectations *similarly* predict actual returns. In both tests, we also control for current price scaled variables. The first test assesses the non-rationality of expectations, because EBRs, as a combination of forecast errors and forecast revisions, should not be predictable if expectations are rational. The second test assesses market inefficiency, because it asks whether expectations predict future returns in the same way in which they predict future EBRs.

Controlling in both tests for current prices is critical: if measured expectations exhibit predictive power conditional on prices, then measured expectations are not merely extracted from prices. That is, they contain genuine information about market expectations, and do not surreptitiously capture required returns, which are fully incorporated in the price control itself. This test is very demanding because current prices also contain information about longer term unmeasured expectations and because analyst expectations are only a proxy for market expectations embedded in prices.

Following BGLS (2024), we regress future EBR based spreads $\text{EBR}_{LMS,t,t+h}$ of the long short portfolio LMS on current and lagged expectations proxies including: the current LTG revision $\Delta_{12}LTG_{LMS,t}$, the lagged LTG portfolio level $LTG_{LMS,t-12}$, the current portfolio short term growth forecast $\mathbb{E}_t[\ln EPS_{LMS,t+24} - \ln EPS_{LMS,t+12}]$, lagged ranked forecast errors, $FE_{LMS,t-l,t}$ for $l = 12, 36,$ and 60 months (to account for potential persistence in errors), as well as the current revision $\Delta_{12}LTG_{M,t}$ and lagged value $LTG_{M,t}$ of the aggregate market portfolio. We also control for the difference in book to market ratio between the value and growth portfolios, $bm_{LMS,t} = \ln BM_{L,t} - \ln BM_{S,t}$. Table 9 presents the results for the 1-month and 1 year horizon for each factor, with other horizons reported in the Appendix.

Table 9
Predicting future expectation based spreads from expectations data

Note: the table presents regressions of log expectations-based returns (EBRs) for portfolios that are long value and short growth stocks (HML), long small and short big stocks (SMB), long conservative and short aggressive investment stocks (CMA), and long winner and short momentum stocks (WML) on expectations based returns (EBR) for that portfolio. Separate regressions are estimated for horizons (h) one-month and one year.. The set of independent variables includes: (a) the ranked forecast error in portfolio earnings between $t-12$ and t, (b) the ranked forecast error in portfolio earnings between $t-36$ and t, (c) the ranked forecast error in portfolio earnings between $t-60$ and t, (d) the lagged portfolio forecast for long-term growth in earnings at $t-12$, (e) the change in the portfolio forecast for long-term growth in earnings between $t-12$ and t, (f) the portfolio forecast for one-year growth in earnings at t, (g) the portfolio forecast for one-year growth in earnings at $t-12$, (h) the aggregate forecast for long-term growth in earnings at *t*, (i) the forecast error in aggregate earnings between $t-12$ and t, (j) the forecast error in aggregate earnings between $t-36$ and t, (k) the forecast error in aggregate earnings between $t-60$ and t, and (l)) portfolio log book-to-market ($\ln bm_{LMS,t}$) at time *t*. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile. 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

	$EBR_{HML,t+h}$		$EBR_{SMB,t+h}$		$EBR_{CMA,t+h}$		$EBR_{RMW,t+h}$		$EBR_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$LTG_{LMS,t-12}$	-0.0545	-2.3218^{b}	-0.0353	-0.9355^{a}	-0.0163	-0.6924	-0.0093	-0.5065	-0.1372^{b}	-2.3154^{a}
	(0.0573)	(1.0372)	(0.0441)	(0.3429)	(0.0437)	(0.4211)	(0.0351)	(0.3985)	(0.0544)	(0.4469)
$\Delta_{12}LTG_{LMS,t}$	-0.1091	-3.4667^{a}	-0.0029	0.0845	-0.0287	0.3995	0.1730^{a}	1.9333^{a}	-0.2267^{b}	-2.8157^{a}
	(0.0956)	(1.0807)	(0.0794)	(0.6332)	(0.0683)	(0.4797)	(0.0511)	(0.4829)	(0.0883)	(0.5599)
$E_t[\ln EPS_{LMS,t+24} - \ln EPS_{LMS,t+12}]$	0.0427^{a}	0.4177^{a}	0.0325^{a}	0.2330^{a}	0.0190^{b}	0.1509^{b}	0.0193^{a}	0.1870^{a}	0.0209^{a}	0.3730^{a}
	(0.0072)	(0.0948)	(0.0076)	(0.0836)	(0.0090)	(0.0756)	(0.0060)	(0.0496)	(0.0074)	(0.0660)
$(1-E_t)$ $\Delta_{12}e_{LMS,t}\ rank$	0.0689^{a}	0.4445^{a}	0.0688^{a}	0.2387^{c}	0.0245	-0.0423	-0.0055	-0.3572^{a}	0.0615^{a}	0.2247
	(0.0156)	(0.1477)	(0.0185)	(0.1345)	(0.0155)	(0.1271)	(0.0133)	(0.1246)	(0.0229)	(0.1541)
$(1-E_t)$ $\Delta_{36}e_{LMS,t}\ rank$	-0.0033	-0.1512	0.0026	0.3365^{a}	-0.0396^{b}	-0.5302^{a}	0.0231	0.4271^{b}	-0.0267	-0.0495
	(0.0204)	(0.2308)	(0.0200)	(0.1204)	(0.0201)	(0.1708)	(0.0182)	(0.1716)	(0.0234)	(0.1932)
$(1-E_t)$ $\Delta_{36}e_{LMS,t}rank$	0.0092	0.0845	0.0029	-0.2964^{a}	0.0527^{b}	0.6495^{a}	-0.0069	-0.3289^{b}	0.0291	-0.0114
	(0.0132)	(0.1683)	(0.0168)	(0.0944)	(0.0211)	(0.1532)	(0.0141)	(0.1576)	(0.0218)	(0.1656)
$\Delta_{12}LTG_{Mkt,t}$	0.1824^{a}	1.0871^{b}	0.0348	0.4380	0.0644	0.1857	-0.0021	-0.7208	0.1492^{c}	1.5358^{a}
	(0.0467)	(0.4424)	(0.0610)	(0.3461)	(0.0446)	(0.4073)	(0.0474)	(0.4888)	(0.0809)	(0.5149)
$LTG_{Mkt,t-12}$	0.2028^{a}	1.1110^{c}	0.0728	0.7704^{b}	0.0692^{c}	0.2603	-0.0246	-0.5333	0.1699^{b}	0.7402
	(0.0444)	(0.6586)	(0.0582)	(0.3665)	(0.0386)	(0.3165)	(0.0531)	(0.6809)	(0.0757)	(0.6187)
$bm_{LMS,t}$	-0.0078^{b}	0.0296	-0.0123^{b}	0.0157	-0.0062	-0.0496	-0.0074	-0.1052^{c}	-0.0193^{a}	-0.1288^{a}
	(0.0038)	(0.0582)	(0.0051)	(0.0386)	(0.0061)	(0.0448)	(0.0058)	(0.0607)	(0.0061)	(0.0410)
Constant	-0.0182^{b}	-0.3360^{b}	-0.0024	-0.0816^{c}	0.0007	0.0418	-0.0044	-0.0041	-0.0069	0.0532
	(0.0089)	(0.1412)	(0.0066)	(0.0428)	(0.0071)	(0.0597)	(0.0097)	(0.0944)	(0.0105)	(0.0876)
Obs	442	433	442	433	442	433	442	433	442	433
Adjusted R^2	19%	27%	16%	44%	5%	26%	7%	31%	13%	53%

On average, analyst differential optimism about long term growth of the short arm – as measured by low $LTG_{LMS,t-12}$ and $\Delta_{12}LTG_{LMS,t}$, negatively predicts subsequent forecast errors and revisions, and hence high EBRs, in line with previously documented analyst overreaction to news (BGLS 2019, 2024). Analyst differential optimism about short term growth instead positively predicts future spreads. Controlling for these measures, aggregate long term optimism tends to predict higher EBR spreads, as in BGLS (2024), but not uniformly so. The coefficient on book to market is generally insignificant, and if anything has the wrong sign.

We next regress the future realized long shot spreads themselves on the same expectation proxies and $bm_{LMS,t}$. Table 10 presents the results for each of the long-short portfolio spreads, focusing again on the 1 month and 1 year horizons.

Table 10
Predicting future return spreads from expectations data

Note: the table presents regressions of log returns for portfolios that are long value and short growth stocks (HML), long small and short big stocks (SMB), long conservative and short aggressive investment stocks (CMA), and long winner and short momentum stocks (WML) on expectations based returns (EBR) for that portfolio. Separate regressions are estimated for horizons (h) one-month and one year.. The set of independent variables includes: (a) the ranked forecast error in portfolio earnings between $t-12$ and t, (b) the ranked forecast error in portfolio earnings between $t-36$ and t, (c) the ranked forecast error in portfolio earnings between $t-60$ and t, (d) the lagged portfolio forecast for long-term growth in earnings at $t-12$, (e) the change in the portfolio forecast for long-term growth in earnings between $t-12$ and t, (f) the portfolio forecast for one-year growth in earnings at t, (g) the portfolio forecast for one-year growth in earnings at $t-12$, (h) the aggregate forecast for long-term growth in earnings at t, (i) the forecast error in aggregate earnings between $t-12$ and t, (j) the forecast error in aggregate earnings between $t-36$ and t, (k) the forecast error in aggregate earnings between $t-60$ and t, and (l)) portfolio log book-to-market ($\ln bm_{LMS,t}$) at time t. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{HML,t+h}$		$r_{SMB,t+h}$		$r_{CMA,t+h}$		r_{RMWt+h}		$r_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$LTG_{LMS,t-12}$	-0.2339[c]	-4.1287[a]	0.0304	0.3144	-0.1449[b]	-1.1618[b]	0.0173	0.7473	-0.5119[a]	-2.4700[a]
	(0.1290)	(1.4216)	(0.1374)	(1.0971)	(0.0710)	(0.5756)	(0.0893)	(0.4872)	(0.1608)	(0.7104)
$\Delta_{12}LTG_{LMS,t}$	0.0240	-5.0531[a]	-0.0479	-1.6284	0.0958	-0.1462	0.0905	0.2717	-0.0718	-1.5429[b]
	(0.2574)	(1.8146)	(0.3027)	(1.6915)	(0.1378)	(0.7903)	(0.1392)	(0.8660)	(0.2044)	(0.7064)
$E_t[\ln EPS_{LMS,t+24} - \ln EPS_{LMS,t+12}]$	0.0723[a]	0.4593[a]	0.0507[b]	0.2205	0.0198	0.0123	0.0230	-0.0627	0.1110[a]	0.5256[a]
	(0.0188)	(0.1255)	(0.0233)	(0.1686)	(0.0168)	(0.1047)	(0.0165)	(0.0906)	(0.0322)	(0.1650)
$(1-E_t)$ $\Delta_{12}e_{LMS,t}\ rank$	-0.0275	0.3257	0.0651	-0.0281	0.0252	0.1131	-0.0552[c]	-0.3216[c]	0.0634	-0.2228
	(0.0444)	(0.2587)	(0.0499)	(0.3332)	(0.0270)	(0.1591)	(0.0285)	(0.1920)	(0.0458)	(0.2494)
$(1-E_t)$ $\Delta_{36}e_{LMS,t}\ rank$	0.0409	-0.0959	0.0682	0.3692[c]	-0.1218[a]	-0.7962[a]	0.1098[a]	0.4694[c]	0.0787	0.2768

	(0.0528)	(0.4418)	(0.0510)	(0.2153)	(0.0423)	(0.2735)	(0.0417)	(0.2421)	(0.0518)	(0.3445)
$(1-E_t)$ $\Delta_{36}e_{LMS,t}rank$	0.0649[c]	0.1756	-0.0716[c]	-0.1790	0.1465[a]	0.6997[a]	-0.0588[c]	-0.5343[b]	-0.0622	0.1348
	(0.0339)	(0.3737)	(0.0413)	(0.2901)	(0.0432)	(0.2306)	(0.0329)	(0.2348)	(0.0513)	(0.2313)
$\Delta_{12}LTG_{Mkt,t}$	0.1543	0.5711	0.1379	-2.9559[a]	-0.0329	0.7372	0.0594	1.7799[c]	0.0376	0.3750
	(0.1331)	(1.3447)	(0.2645)	(1.0076)	(0.0843)	(0.5542)	(0.1433)	(0.9264)	(0.1546)	(0.8586)
$LTG_{Mkt,t-12}$	0.3214[a]	1.1467	0.4461[b]	-0.3726	0.0791	0.7758[c]	0.1934	1.5432	0.0378	-0.2508
	(0.1207)	(1.3559)	(0.2062)	(0.8365)	(0.0717)	(0.4114)	(0.1424)	(1.0512)	(0.1976)	(1.4271)
$bm_{LMS,t}$	-0.0213[b]	0.0779	-0.0386[a]	0.3213[a]	-0.0265[a]	0.0073	-0.0377[a]	-0.0020	-0.0282[c]	0.0983
	(0.0107)	(0.0873)	(0.0144)	(0.0906)	(0.0102)	(0.0691)	(0.0139)	(0.0826)	(0.0146)	(0.0863)
Constant	-0.0080	-0.5230[b]	-0.0476	-0.1173	-0.0043	-0.1223	-0.0495[b]	-0.1421	-0.0337	0.1805
	(0.0245)	(0.2176)	(0.0301)	(0.1273)	(0.0133)	(0.0745)	(0.0214)	(0.1717)	(0.0242)	(0.2061)
Obs	442	433	442	409	442	433	442	433	442	433
Adjusted R^2	6%	17%	4%	47%	8%	31%	2%	14%	10%	27%

As in Table 9, lagged expectations have predictive power for the future HML return spread, even at the short 1 month horizon (a challenging test) and increasing at the 1 year horizon. This is driven in particular by the role of lagged optimism, for both short and long term growth, as is also the case in Table 9. These patterns are confirmed, and in fact get stronger at longer horizons, reported in the Appendix. Taken together, these results are consistent with future return spreads capturing, at least in part, the unfolding of predictable market expectations errors as proxied by predictable EBRs.

To study more directly the role of departures from rationality for return predictability, we connect the predictability of EBR to average return differences. We do so by performing a two-stage analysis. We use Table 10 as the first stage to construct at each t a predicted EBR of portfolios $\pi = L, S$ at time $t+1$, which we denote by $\widehat{\mathrm{EBR}}_{\pi,S,t\to t+1}$. In the second stage, we test the ability of the predicted EBR differential to predict portfolio returns:

$$r_{L,t+1} - r_{S,t+1} = \beta_0 + \beta_1 \cdot \left(\widehat{\mathrm{EBR}}_{L,t\to t+1} - \widehat{\mathrm{EBR}}_{S,t\to t+1}\right) + v_{t+1}$$

Compared to Equation (8), this test ties return differentials to error predictability, the hallmark of non-rationality, but also allows us to study time variation in return spreads. That is, the long minus short return spread may be high following periods when beliefs about firms in the short arm are particularly bullish, and low otherwise. Table 11 reports second stage results from Equation (8), focusing on regressions at 1 month and 1 year horizons here as in Tables 10 and 11.

Table 11
Instrumented expectation based returns explain actual spreads

Note: This table resents univariate instrumental variable regressions of log returns for the portfolio that is long value and short growth (HML, columns 1 and 2), long conservative and short aggressive investment (CMA, in column 3 and 4), long robust and short weak profitability (RMW, columns 5 and 6), and long winners and short loser momentum stocks (RMW, columns 7 and 8) on expectations based returns (EBR) for that portfolio. Separate regressions are estimated for one-month and one-year horizons. The set of instrumental variables includes: (a) the ranked forecast error in portfolio earnings between $t-12$ and t, (b) the ranked forecast error in portfolio earnings between $t-36$ and t, (c) the ranked forecast error in portfolio earnings between $t-60$ and t, (d) the lagged portfolio forecast for long-term growth in earnings at $t-12$, (e) the change in the portfolio forecast for long-term growth in earnings between $t-12$ and t, (f) the forecast for growth in portfolio earnings between $t+12$ and $t+24$ at t, (h) the forecast for long-term growth in aggregate earnings at $t-12$, and (i) the change in the forecast for long-term growth in aggregate earnings between $t-12$ and t. Portfolio forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). We present Kleinbergen Paap (KP) and Montiel Pflueger (MP) F-statistics. Standard errors are corrected for overlapping observations using the Newey-West (1987) procedure. The sample period spans from December 1981 to December 2023. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	$r_{HML,t+h}$		$r_{SMB,t+h}$		$r_{CMA,t+h}$		$r_{RMW,t+h}$		$r_{WML,t+h}$	
	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$	$h=1$	$h=12$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\widehat{EBR}_{LMS,t,t+h}$	1.1261[a]	1.1423[a]	1.2027[a]	0.8840[a]	1.4193[a]	1.3163[a]	0.1028	-0.0538	1.2837[a]	0.9145[a]
	(0.3466)	(0.2160)	(0.3327)	(0.2494)	(0.4412)	(0.2421)	(0.4756)	(0.2537)	(0.4667)	(0.2791)
Constant	0.0008	-0.0168	0.0032	0.0265[c]	-0.0064[c]	-0.0655[a]	0.0017	0.0064	-0.0464[b]	-0.1730[a]
	(0.0020)	(0.0134)	(0.0019)	(0.0142)	(0.0034)	(0.0205)	(0.0030)	(0.0185)	(0.0191)	(0.0546)
Obs	444	433	444	433	444	433	444	433	444	433
KP F-stat	10.4	6.3	8.5	13.2	4.2	4.6	4.8	7.3	7.9	8.9
MP F-stat	10.5	3.8	7.6	9.2	3.5	4.0	3.7	3.7	6.1	10.3

Actual spreads load strongly on predicted expectations-based spreads, with all coefficients strongly significant and indistinguishable from 1, except for profitability where EBRs do not predict returns. To assess the strength of our expectation-based instruments, we present Kleinbergen Paap (2006) and Montiel Pflueger (2013) F-statistics. In line with previous results, the predictability from lagged expectations is stronger for HML, SMB and WML, with instruments in about half the specifications exceeding the heuristic of an F-stat of 10. For CMA and RMW, instruments are weak. These instruments are stronger for horizons above one year, with 8 (5) out of 10 KP (MP) F-statistics exceeding 10. Overall, these results are consistent with the fact that future return spreads capture, at least in part, unfolding of predictable market expectations errors, where the latter are proxied by predictable EBRs.

Intercepts for actual return spreads are mostly small and insignificant, and in some cases negative, confirming that EBRs, and in particular their predictable component, can account for the totality of the observed HML and CMA spread in actual returns. This further evidence is in line with the hypothesis that these spreads reflect market inefficiency rather than risk. We next assess directly the extent to which characteristics reflect non-rational beliefs.

Finally, we can use the results in Table 10 to revisit the hypothesis that analysts erroneously infer growth expectations from movements in prices due to required returns. In that case, the predictability of future returns should be absorbed by current prices themselves, contrary to our findings. To confirm these results at the firm level, we predict firm level EBRs using the expectation variables in Table 10, and then run a horse race between predicted future $\widehat{EBR}_{i,t+12}$ and current characteristics $bm_{i,t}$ to explain future firm level returns $r_{i,t+12}$. Table 12 shows the results.

Table 12
Predicted EBRs versus book to market

Note: This table presents regressions of firm level log returns at horizons (*h*) of one-month, three-months, one-year, three-years, and five-years. The independent variables include: (a) predicted expectation based returns ($\widehat{EBR}_{i,t\to t+h}$) between t and $t+h$, and (b) log book-to-market ($bm_{i,t}$) at time *t* for firm i. $\widehat{EBR}_{i,t\to t+h}$ is generated from separate regressions –which are not shown-- using the following predictors: (a) the time $t-12$ long-term growth in earnings ($LTG_{i,t-12}$), (b) the change in long-term growth in earnings between time $t-12$ and *t* ($\Delta_{12}LTG_{i,t}$), (c) the time-*t* forecast for growth in earnings between $t+12$ and $t+24$ ($E_t[e_{t+24}-e_{t+12}$), (d) the forecast error in growth in earnings between $t-12$ and t ($(1-E_t)\Delta_{12}e_{i,t}\ rank$), (e) the forecast error in growth in earnings between $t-36$ and t ($(1-E_t)\Delta_{36}e_{i,t}\ rank$), (f) the forecast error in growth in earnings between $t-60$ and t ($(1-E_t)\Delta_{60}e_{i,t}\ rank$), (g) the time $t-12$ forecast for long-term growth in earnings of the market, and the change in long-term growth in earnings of the market between $t-12$ and t ($\Delta_{12}LTG_t$), Forecast errors are ranked from 0 (lowest percentile) to 1 (top percentile). 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.

	(1)	(2)	(3)	(4)	(5)
	$r_{i,t+1}$	$r_{i,t+3}$	$r_{i,t+12}$	$r_{i,t+36}$	$r_{i,t+60}$
$\widehat{EBR}_{i,t\to t+h}$	0.5269[a]	0.4519[a]	0.4173[a]	0.4093[a]	0.4221[a]
	(0.1574)	(0.1324)	(0.1140)	(0.1163)	(0.1567)
$bm_{i,t}$	0.0002	0.0013	0.0135	0.0373[c]	0.0647[b]

	(0.0011)	(0.0029)	(0.0090)	(0.0204)	(0.0320)
Constant	0.0025	0.0102	0.0588[b]	0.1924[a]	0.3213[a]
	(0.0034)	(0.0090)	(0.0264)	(0.0543)	(0.0858)
Obs	518,470	506,485	490,328	397,044	323,939
R^2	0%	1%	1%	1%	1%
Adj R^2	0%	1%	1%	1%	1%
F-stat	5.6	6.5	8.0	6.3	9.3
Time FE	N	N	N	N	N
Firm FE	N	N	N	N	N

At the firm level, predicted EBRs again explain future returns at all horizons and firm book to market has little predictive power except at long horizons, confirming that expectations are not spuriously capturing information about required returns.

6.2 Characteristics and EBRs.

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

Table 13 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. Columns 6 to 10 control for the expectation variables used in Table 9 to predict portfolio level returns, which allows us to assess whether characteristics hold predictive power over and above measured expectations, and vice versa. We do not include firm fixed effects here, because they would potentially absorb the role of the measured characteristics themselves.

Table 13
Characteristics predict firm level expectation based returns

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

	$EBR_{i,t,t+1}$	$EBR_{i,t,t+3}$	$EBR_{i,t,t+12}$	$EBR_{i,t,t+36}$	$EBR_{i,t,t+60}$
	(1)	(2)	(3)	(4)	(5)
$bm_{i,t}$	0.0056[a]	0.0160[a]	0.0573[a]	0.1113[a]	0.1434[a]
	(0.0003)	(0.0019)	(0.0081)	(0.0138)	(0.0202)
$size_{i,t}$	0.0082[a]	0.0115[a]	-0.0717[a]	-0.2916[a]	-0.3975[a]
	(0.0002)	(0.0017)	(0.0055)	(0.0207)	(0.0306)
$Inv_{i,t}$	-0.0067[a]	-0.0215[a]	-0.0705[a]	-0.0854[a]	-0.0869[a]
	(0.0004)	(0.0022)	(0.0056)	(0.0095)	(0.0107)
$op_{i,t}$	-0.0014[b]	-0.0048[c]	-0.0050	0.0034	-0.0098
	(0.0006)	(0.0025)	(0.0110)	(0.0213)	(0.0313)
$r_{i,t-12\to t-1}$	0.0176[a]	0.0502[a]	0.0806[a]	0.0357[b]	0.0305[b]
	(0.0002)	(0.0046)	(0.0150)	(0.0155)	(0.0143)
Obs	878,185	818,858	775,234	596,792	475,520
Adj R^2	0%	1%	2%	10%	17%

On their own, characteristics have strong and highly significant predictive power for forecast errors and revisions (columns 1 to 5). Low book to market, high investment and low returns predict subsequent disappointment and low EBRs at all horizons, consistent with the average spread of the corresponding factors.[13] 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.

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

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

large extent, characteristics predict returns precisely because they capture distorted expectations. These tests quantify the extent to which characteristic based predictability reflects the expectations channel, and thus market inefficiency, as in Equation (8) and shed new light on time variation in average return spreads.

Specifically, we run the regression:

$$r_{i,t+h} = a_r + b \cdot \mathrm{EBR}_{i,t,t+h} + c_{bm} \cdot bm_{i,t} + c_{inv} \cdot inv_{i,t} + c_{size} \cdot size_{i,t} + c_{prof} \cdot op_{i,t} + c_{mom} \cdot r_{i,t-12\to t-1} + \epsilon_{t+h}, \qquad (11)$$

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 12, columns 1 to 5):

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

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. Note that this exercise offers a lower bound on the role of expectations: 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 14 shows the empirical results, reporting Equation (11) in Panel A.

Table 13

Return predictability from characteristics is mediated by expectations

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

Panel A: Explaining Returns

	$r_{i,t+1}$ (1)	$r_{i,t+3}$ (2)	$r_{i,t+12}$ (3)	$r_{i,t+36}$ (4)	$r_{i,t+60}$ (5)
$\text{EBR}_{i,t,t+h}$	0.1579[a]	0.2168[a]	0.4077[a]	0.5222[a]	0.5589[a]
	(0.0010)	(0.0065)	(0.0164)	(0.0227)	(0.0267)
$\ln bm_{i,t}$	-0.0080[a]	-0.0135[a]	-0.0195[b]	-0.0043	0.0064
	(0.0004)	(0.0032)	(0.0086)	(0.0109)	(0.0134)
$\ln size_{i,t}$	0.0040[a]	-0.0295[a]	-0.1296[a]	-0.2229[a]	-0.2831[a]
	(0.0002)	(0.0026)	(0.0103)	(0.0255)	(0.0418)
$Inv_{i,t}$	0.0115[a]	0.0135[a]	0.0199[a]	0.0470[a]	0.0764[a]
	(0.0003)	(0.0028)	(0.0074)	(0.0149)	(0.0179)
$op_{i,t}$	0.0093[a]	0.0175[a]	0.0457[a]	0.0720[a]	0.0856[a]
	(0.0006)	(0.0035)	(0.0103)	(0.0213)	(0.0272)
$r_{i,t-12\to t-1}$	-0.0058[a]	-0.0124[a]	-0.0596[a]	-0.0774[a]	-0.0824[a]
	(0.0002)	(0.0030)	(0.0109)	(0.0184)	(0.0194)
Obs	878,185	818,858	775,211	596,734	475,452
Adj R^2	2%	7%	31%	52%	59%

Panel B: Share of predictability from characteristics via expectations

bm	7%	20%	54%	55%	51%
$size$	24%	-9%	18%	41%	44%
inv	11%	26%	60%	92%	115%
op	-2%	-5%	-5%	2%	-6%
$r_{i,t-12\to t-1}$	-125%	-975%	-126%	-32%	-26%

Panel A shows that EBR has substantial explanatory power: conditioning on characteristics, b is large and significant, consistent with Table 12. 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 14 but the correct sign in Table 13, suggesting all of the predictability is captured by EBR.

In Panel C we compute a lower bound for the expectation-channel share of the predictability of characteristic $\chi = bm, size, inv, prof, mom$ as $\frac{b \cdot c_\chi}{b \cdot c_\chi + d_\chi}$, which are reported in Table 14 panel B. At horizons of 1 year or longer, most predictability from bm and inv, and a substantial share of predictability from $size$, works through the expectations channel.[14] There is no explanatory power for profitability, in line with the result from Table 13 that, controlling for other characteristics, op does not predict EBRs. These result offers direct evidence that analyst expectations help explain the documented predictive power of firm characteristics for future returns.

7.Taking stock

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

Our evidence shows that spreads are generated because expectations about future growth of firms in the short arm of the portfolios are systematically too optimistic, and those about firms in the long arm too pessimistic, so that the long portfolio outperforms the short one as expectation errors are corrected in the future. Characteristics such as book to market or investment predict returns at least in part because they predict differential optimism and forecast errors. Notably, the same mechanism helps account for momentum. Predictability

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

from other characteristics may also work through expectations. This, of course, has significant implications not just for cross sectional (and time series) asset pricing, but also for firm investment policies, financial policies, and other decisions. We conclude by highlighting two follow-up questions.

The first question concerns the structure of expectations. Analyst beliefs can reproduce return co-movements across firms sharing similar characteristics because expectations themselves comove within groups of 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 particular groups of firms or sectors, but co-movement in beliefs may also reflect 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 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 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, 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.

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.

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

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

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

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.

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Nagel, Stefan, and Xu, Zhengyang, 2019, Asset pricing with fading memory. *Review of Financial Studies* 35 (5): 2190-2245.

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

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

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

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

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

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

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