Belief Overreaction and Stock Market Puzzles

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

November 2022, First Draft October 2019

Abstract

We construct an index of long term expected earnings growth for S&P500 firms and show that it has remarkable power to jointly predict future errors in these expectations and stock returns, in both the aggregate market and the cross section. The evidence supports a mechanism whereby good news cause investors to become too optimistic about long term earnings growth, for the market as a whole but especially for a subset of firms. This leads to inflated stock prices and, as beliefs are systematically disappointed, to subsequent low returns in the aggregate market and for the subset of firms. Overreaction of long term expectations helps resolve major asset pricing puzzles without time series or cross-sectional variation in required returns.

[1] The authors are from Oxford Said Business School, Università Bocconi, Brown University, and Harvard University, respectively. Gennaioli thanks the Italian Ministry of Education for Financial Support (PRIN Grant). We are grateful to Nick Barberis, Francesca Bastianello, John Campbell, Kent Daniel, Paul Fontanier, Spencer Kwon, Yueran Ma, Peter Maxted, Dev Patel, Jesse Shapiro, Adi Sunderam, and the referees and editor of this journal for extremely helpful comments. Julien Manili provided outstanding research assistance.

1. Introduction

In the textbook asset pricing model, the price of a stock is the rational expectation of future dividends discounted by a time-invariant required return. That required return is higher for stocks that are riskier, in the sense of being more exposed to aggregate market movements. Over the past four decades, this approach has been challenged by two key findings. First, the return on the aggregate stock market is predictably low following periods of high valuations, as measured for instance by a high aggregate price to dividend ratio (Campbell and Shiller 1988). This fact is inconsistent with the assumed time-invariance of required returns. Second, large cross-sectional average return differences are traceable to firm characteristics, not to market exposure. For instance, high market-to-book stocks have lower average returns than low market-to-book ones (Fama and French 1993). Ultimately, the key stock market puzzles concern excessive return predictability, both in the time series and the cross section.

The conventional approach to these puzzles enriches the theory of required returns while maintaining rational expectations of future dividends. In the time series, required returns are assumed to vary due to changes in risk preference (e.g., Campbell and Cochrane 1999), or in long run or disaster risk (Rietz 1988, Bansal and Yaron 2004, Barro 2006). In the cross section, required returns are assumed to vary due to exposure to characteristics-based "risk factors" (Fama and French 1993). A challenge for this approach is that investors should rationally expect low future returns during a stock market boom. In survey expectations of returns, however, the opposite is the case (Greenwood and Shleifer 2014). A deeper problem is that changes in risk preference and risk are hard to measure, and cross sectional risk factors remain a black box.

In this paper we try to address these puzzles by pursuing an orthogonal approach: keep required returns constant and relax rational expectations. In this approach, return predictability arises from the eventual correction of systematic pricing errors caused by non-rational beliefs. Using data on analyst expectations of future earnings growth of listed firms, we empirically characterize belief errors and connect them to realized returns. We show that errors in expectations of aggregate long-term earnings growth, LTG, offer a promising source of return predictability in both the time series and the cross section, helping to reconcile key anomalies.

In the first part of the paper we study survey expectations and return predictability in the time series. Section 2 shows that high expected aggregate long-term earnings growth predicts sharply lower future aggregate stock returns. The predictive power of LTG is robust to controlling for the current price dividend ratio and other prominent macroeconomic predictors of returns. Expectations of short-term earnings growth, in contrast, do not predict future returns.

Section 3 studies the mechanism linking beliefs and return predictability, documenting three facts. First, LTG overreacts: upward LTG revisions predict future disappointment of growth forecasts. Second, such predicted disappointment is associated with low returns. Third, systematic unwinding of aggregate LTG forecast errors accounts for a large share of the link between the price dividend ratio and future returns. These findings point to a mechanism in which overreaction to good news causes excess optimism and inflated stock prices. Going forward, systematically disappointing aggregate earnings growth causes a price reversal and hence low returns.

In Section 4 we consider cross sectional return differences. We ask whether variation in the aggregate LTG, which captures systematic belief biases, can also produce cross sectional return comovement and average return spreads. We first revisit the return spread earned by stocks with low LTG compared to stocks with high LTG (La Porta 1996). We find that this spread varies systematically with aggregate LTG: current optimism about aggregate fundamentals is followed by lower returns and more disappointing forecast errors for high LTG stocks than low LTG ones. This evidence is consistent with a mechanism in which high LTG firms exhibit stronger overreaction to aggregate good news, perhaps because these firms belong to the "hot sector" of the moment. Remarkably, we find that a similar mechanism also sheds light on the well-known book-to-market, profitability and investment factors (Fama and French 1993). The short arm in these factors disappoints more sharply, both in returns and in realized earnings growth, after periods of high aggregate optimism, again measured using aggregate LTG.

Our evidence indicates that high aggregate LTG captures periods when the aggregate market and specific stocks are overvalued, in the sense of having subsequent disappointing returns. In the language of standard finance, these periods look like moments of low risk aversion, when the price of the aggregate market and particularly of risky firms is elevated. One concern is that perhaps survey expectations spuriously capture time varying risk aversion. This could occur if analysts mechanically infer expectations about long term growth by fitting stock prices. Our analysis, however, shows that LTG is a genuine proxy for expectations, and that return predictability to a substantial degree reflects belief overreaction. First, consistent with our interpretation and inconsistent with expectations inferred from prices, our results are robust to controlling for price ratios. Second, excess optimism in LTG arises as an overreaction to news: we find that both LTG revisions and subsequent errors are predictable using news about fundamentals, even after controlling for stock returns. Finally, and crucially, we show that much of the predictability of aggregate returns attributed to price ratios comes from predictable reversals in LTG. After controlling for these reversals, price ratios have little predictive power for returns.

A few recent papers study stock market puzzles using measured expectations.[2] Bordalo et al. (BGLS 2019) account for the La Porta (1996) LTG spread through belief over-reaction, but do not connect the spread in returns and forecast errors to systematic belief biases.

Using analysts' forecasts of short-term earnings growth, De la O and Myers (2020) construct a dividend discount index and show that it strongly correlates with the aggregate price to earnings ratio. This exercise showcases the usefulness of expectations data, but does not shed light on return predictability: unlike LTG, short term expectations do not predict stock returns. Nagel and Xu (2019) show that past aggregate dividend growth correlates negatively with future aggregate returns and positively with earnings growth expectations. However, they do not directly

[2] Frankel and Lee (1998), Lee, Myers, and Swaminathan (1999), Lee and Swaminathan (2000), Bachetta et al (2009), and Koijen and Nieuwerburgh (2011) also use beliefs data to study asset prices. Cutler, Poterba, Summers (1990), DeLong et al (1990b), Barberis et al 2015, Adam, Beutel and Marcet (2017) study price extrapolation, which is also consistent with returns expectations data (Greenwood and Shleifer 2014, Giglio et al 2021).

connect expectations to forecast errors and returns, and therefore do not show that return predictability is driven by belief overreaction. In fact, the growth of past dividends might affect required returns through consumption as in Campbell and Cochrane (1999). And to the extent that past dividend growth correlates with expectations, it does so only partially: beliefs and stock prices may overreact to other news, such as the arrival of new technologies.

More broadly, we are the first to show that a parsimonious mechanism of belief overreaction throws new light on both aggregate return predictability and cross sectional return differentials by characterizing the joint behavior of returns and forecast errors.

Our work offers a new angle on macro volatility. In macroeconomics, departures from rational expectations typically take the form of rational inattention (Sims 2003, Woodford 2003, Gabaix 2019), or overconfidence (Kohlhas and Walther 2021). These mechanisms generate rigidity in consensus beliefs and prices (Mankiw and Reis 2002). We document the importance of the opposite phenomenon of belief overreaction. Compared to Bordalo, Gennaioli, Ma, and Shleifer (BGMS, 2020), who find overreaction by individual professional forecasters, we find overreaction in consensus expectations and connect it to excess stock market volatility. Our analysis points to belief volatility as a source of macro-financial volatility, in line with recent work in macroeconomics (Bianchi et al. 2021, Bordalo et al. 2021, L'Huillier et al. 2021).

2. Predictability of Aggregate Stock Returns: Data and Basic Facts

We gather monthly data on analyst forecasts for firms in the S&P 500 index from the IBES Unadjusted US Summary Statistics file. We focus on median forecasts of a firm's earnings per share (EPS_{it}) and long-term earnings growth (LTG_{it}). IBES defines LTG 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."[3] Data coverage starts on 3/1976 for EPS_{it} and

[3] It is not obvious whether LTG captures $g = \sqrt[T]{\widetilde{\mathbb{E}}[(1+g_1)\ldots(1+g_T)} - 1$, or the average point estimate $g = (\hat{g}_1 + \cdots + \hat{g}_T)/T$. We take the former interpretation, but the distinction is not key for studying return predictability.

12/1981 for LTG_{it}. (Data on dividend forecasts starts in 2002 and uses shorter horizons.) We fill in missing forecasts by linearly interpolating EPS_{it} at horizons ranging from 1 to 5 years (in one-year increments). Beyond the second fiscal year we assume that analysts expect EPS_{it} to grow at the rate LTG_{it} starting with the last non-missing positive EPS forecast.

Analysts may distort their forecasts due to agency conflicts. As showed in BGLS (2019), this is unlikely to affect the time series variation in forecasts, which is key here. Furthermore, all brokerage houses typically cover S&P 500 firms, so investment banking relationships and analyst sentiment are unlikely to influence the decision to cover firms in the S&P 500.[4] Our focus on median forecasts further alleviates these concerns, reducing the impact of outliers.

We aggregate the earnings forecasts of S&P 500 firms into an index of aggregate beliefs. We multiply each forecast EPS_{it} by the number of shares outstanding in month t and sum these forecasts across all S&P 500 firms. We then divide this aggregate earnings forecast by the total number of shares in the S&P 500 index to obtain the expected earning per share EPS_t. (Log) earnings growth one or two-years ahead are computed based on EPS_t.[5]

We aggregate LTG forecasts by value-weighting firm level forecasts:

$$LTG_t = \sum_{i=1}^{S} LTG_{i,t} \frac{P_{i,t} \cdot Q_{i,t}}{\sum_{i=1}^{S} P_{i,t} \cdot Q_{i,t}}$$

where S is the number of firms in the S&P 500 index with IBES data on LTG_{it}, P_{it} is the stock price of firm i at time t, and $Q_{i,t}$ is the number of shares outstanding of firm i at time t.[6]

Figure 1 plots one year ahead and long term expected earnings growth. LTG_t is more persistent than expected short term growth. In particular, it does not exhibit short run reversals

[4] For example, in December of 2018, nineteen analysts followed the median S&P500 firm, while four analysts followed the median firm not in S&P500. Analysts are also less likely to rate as "buy" firms in the S&P500 index.

[5] The number of shares in the index (what S&P refers to as the divisor) is the ratio of the market capitalization of S&P 500 and the S&P 500 index. It is 100 in the base year and it is adjusted due to shares outstanding, the index composition, and corporate actions. We compute growth forecasts using aggregate earnings because many firm-level observations have zero or very low current earnings. We set an observation in a given month to missing if the market cap of firms for which we have forecasts at a given horizon is less than 90% of the market cap of the index.

[6] Nagel and Xu (2019) weigh $LTG_{i,t}$ using firm level earnings forecasts. The correlation between their index and our LTG_t is 95.44%. Since stocks with high LTG often have negative earnings, our preferred measure is LTG_t.

such as the expected short term growth peak in 2009. As we show later, the persistence of LTG_t is crucial, for it allows it to capture the low frequency predictability of returns.

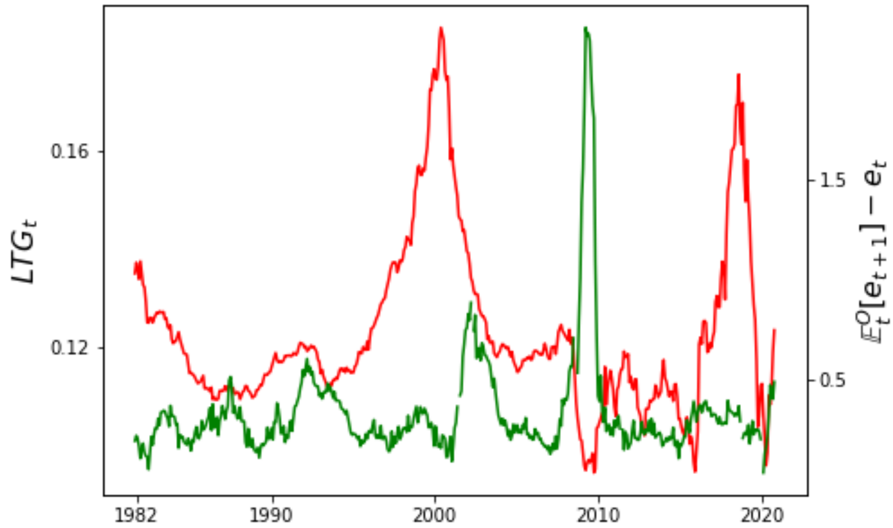

Figure 1. We plot the expected short- and long-term growth in earnings ($\mathbb{E}_t^O[e_{t+1}] - e_t$ in green and LTG_t in red, respectively, where $e_t = \log EPS_t$ and $\mathbb{E}_t^O$ represents measured expectations). The scale for short-term earnings ($\mathbb{E}_t^O[e_{t+1}] - e_t$) is on the right. The sample period is 12/1981 to 12/2020.

De la O and Myers (2021, 2022) use measured expectations of one year ahead earnings growth to construct a discounted expected stock market index. They show that this index is highly correlated with the actual price earnings ratio. One issue is whether the correlation arises because expectations of earnings track current earnings or because they capture stock price anomalies (Adam and Nagel 2022). From the viewpoint of market efficiency, the key question is whether beliefs produce excessive price variation and hence return predictability.

To address this issue, we regress future cumulative raw aggregate stock returns over 1, 3 and 5 years on our three measures of expected earnings growth: at one and two years, and long term. Table 1 reports the results. We also run a horse race between the different expectations measures. In this and other tests, we focus on raw returns but the results are very similar if we use excess returns, see Appendix B, Table B.1.

Table 1
Return Predictability and Expectations of Earnings Growth

We examine the association between earnings growth forecasts and returns at different horizons. The dependent variables are the (log) one-year return in column [1] and the discounted value of the cumulative 3- and 5-year return in columns [2] and [3], respectively. Here $\alpha = 1/(1+e^{dp})$ where dp is the average price dividend ratio in our sample ($\alpha = 0.9779$). The independent variables are the forecast for earnings growth: (a) in the long run LTG_t, (b) one-year ahead, $\mathbb{E}_t^O[e_{t+1} - e_t]$, and (c) between year $t+1$ and $t+2$, $\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$. All variables are standardized and intercepts are not shown. The sample period is 1981:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (number of lags ranges from 12 in the first column to 60 in the last one). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	(1)	(2)	(3)
	r_{t+1}	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
		Panel A: Returns and LTG	
LTG_t	-0.2389[b]	-0.4019[a]	-0.4349[a]
	(0.0928)	(0.0944)	(0.0831)
Observations	409	409	409
Adj R^2	9%	24%	25%
		Panel B: Returns and growth forecast for year 1	
$E_t[e_{t+1}-e_t]$	-0.0335	0.0467	0.1556[a]
	(0.1027)	(0.0716)	(0.0587)
Observations	404	404	404
Adj R^2	0%	0%	3%
		Panel C: Returns and growth forecast for year 2	
$\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$	-0.0527	0.0408	0.2113
	(0.0885)	(0.1556)	(0.1686)
Observations	404	404	404
Adj R^2	0%	0%	6%

High current expectations of long term earnings growth strongly predict low future returns. LTG_t accounts for 25% of variation in realized returns over the following five years.[7] In contrast,

[7] It is well known that the OLS estimator in predictive regressions using lagged stochastic regressors, such as LTG_t, may be biased (Stambaugh 1999). The bias arises because the disturbances in the regression for returns may be correlated with future values of LTG_t. We follow the methodology of Kothari and Shanken (1997): we use simulations to compute the coefficient that we would estimate under the null of no predictability and bootstrap a p-value for the OLS value in Table 1. See Appendix E for details of the methodology, and Table E.1 for the results. We find that, under the null that the LTG_t coefficient is zero, the p-values for the LTG coefficients in Table 1: 3.24% in column 1, 3.84% in column 2, and 3.48% in column 3. Moreover, a mechanical link between return disturbances and future LTG_t is less of a concern here: the correlation between the residuals of regressions of the univariate regressions of r_{t+1} and LTG_{t+1} on LTG_t is -0.05. Similarly, the correlation between one-month returns and $\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$ is negligible (i.e., 0.02). In contrast, correlation between shocks to one-month returns and $\mathbb{E}_t^O[e_{t+1} - e_t]$ is -0.41. To put these numbers in perspective, Kothari and Shanken report that the correlation between shocks to book-to-market and to annual returns is -0.80.

expectations of short term earnings growth do not predict returns or have a very weak explanatory power (one year ahead earnings expectations only account for 3% of five years ahead return variation). To our knowledge, this is the first time series evidence of strong return predictability using measured expectations of fundamentals. The lack of predictive power of short-term growth expectations suggests that this proxy likely captures earnings variation rather than mispricing.[8]

We next assess two questions. First, does the predictive power of LTG actually reflect non-rational market beliefs? Second, how does it compare to stock return predictors studied in previous work? To address these questions, we present several tests in Table 2.

With respect to the first question, the concern is that LTG_t may spuriously reflect time varying required returns. This could happen if analysts estimate LTG_t by fitting the growth rate of earnings that justifies the current stock price, while erroneously assuming that required returns are constant. To assess this possibility, columns (1) and (2) assess the predictive power of LTG_t controlling for the current price dividend and price earnings ratios, respectively. If LTG_t is reverse engineered from stock prices, these controls should eliminate its explanatory power. If instead LTG_t retains some explanatory power, it must be because it captures market expectations about long term fundamentals (with price ratios instead capturing the independent role of expectations at other horizons as well as, possibly, variation in required returns). Note that these are challenging tests: prices incorporate market expectations while LTG_t is a noisy proxy for them. Even if all price variation was due to expectations as opposed to required returns, controlling for market prices may overshadow the predictive power of LTG_t.

To further assess the ability of LTG_t to capture beliefs, columns (3), (4) and (5) control for the three leading proxies of time varying required returns: surplus consumption (Campbell and Cochrane 1999), the consumption wealth ratio (cay, Lettau and Ludvigson 2001), and $SVIX^2$ (Martin 2017). The first proxies for fluctuations of marginal utility in habit formation models, the

[8] De la O and Myers (2022) show that short term earnings expectations predict returns at a very long (10 year) horizon, consistent with Table 1, panel B. However, this relationship disappears once we control for LTG (see also Table 2).

second for required returns in a large class of rational expectations models, and the third for the required return of a rational log utility investor fully invested in the market.

Table 2
Return Predictability, Expectations and Measures of Required Returns

We study the association between realized returns, ex-ante proxies for required returns and macroeconomic predictors of returns. The dependent variable is the discounted value of the cumulative return between year t and $t+5$. All regressions include the forecast for earnings growth in the long run LTG_t, In Panel A the additional independent variables are: (a) the price to dividend ratio, pd_t, in column [1], (b) the price to earnings ratio, pe_t, in column [2], (c) the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [3], (d) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [4], (e) the Martin (2013) expected one-year return on the market, $SVIX_t^2$, in column [5], and (f) the forecast for one-year ahead, $\mathbb{E}_t^O[e_{t+1}-e_t]$, in column [6]. In Panel B the additional independent variables are: (a) the term spread defined as the log difference between the gross yield of 10-year and 1-year US government bonds from the St. Louis Fed in column [1], (b) the credit spread defined as the log difference between the gross yield of BAA and AAA bonds from the St. Louis Fed in column [2], (c) the Baker et al. (2016) economic policy uncertainty index in column [3], the Kelly and Pruitt (2013) optimal forecast of aggregate equity market returns in column [4], the forecast for CPI inflation in year *t+1* by the Survey of Professional Forecasters, $\mathbb{E}_t^O[\pi_{t+1}]$, in column [5] , and (12) the Nagel and Xu (2021) experienced dividend growth from, $exp(d_t)$. The sample period is 1981:12-2015:12. Data is quarterly in column [4] of Panel A and column [6] of Panel B, and monthly elsewhere. All variables are standardized and intercepts are not shown. The sample period is 1981:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
LTG_t	-0.2350[b]	-0.4675[a]	-0.4522[a]	-0.5569[a]	-0.3946[a]	-0.4881[a]
	(0.1162)	(0.1081)	(0.1033)	(0.1179)	(0.1016)	(0.1036)
X_t	-0.5826[a]	-0.1803	-0.1387	0.1894	0.3852[b]	0.1001
	(0.1397)	(0.1662)	(0.1035)	(0.1766)	(0.1782)	(0.0851)
Obs	409	409	409	137	193	404
Adj R^2	52%	28%	27%	28%	47%	26%
X_t	p_t-d_t	p_t-e_t	spc_t	cay_t	$SVIX_t^2$	$\mathbb{E}_t^O[e_{t+1}-e_t]$

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
LTG_t	-0.4345[a]	-0.4682[a]	-0.4761[a]	-0.7542[a]	-0.5052[a]	-0.3450[a]
	(0.1031)	(0.1217)	(0.1198)	(0.2648)	(0.1002)	(0.0483)
X_t	0.1672	0.1945	0.2297	-0.2800	0.3848[a]	-0.6142[a]
	(0.1365)	(0.1994)	(0.1875)	(0.2708)	(0.1346)	(0.1177)
Obs	409	409	372	13400%	409	137
Adj R^2	27%	29%	37%	27%	40%	59%
X_t	Term $Spread_t$	Credit $Spread_t$	Uncertainty $Index_t$	Kelly Pruitt MRP_t	$\mathbb{E}_t^O[\pi_{t+1}]$	$\exp(d_t)$

Columns (1)-(5) in Panel A show that LTG_t is unlikely to proxy for required returns. First, its explanatory power is robust to controlling for prices (columns (1) and (2)). This suggests that the price dividend ratio may at least in part predict returns due to its ability to capture market expectations. In Section 3 we present a test based on our theory to assess this possibility.

The explanatory power of LTG_t is also robust to controlling for proxies of time varying required returns (columns (3), (4), and (5)). The coefficient on LTG_t is fairly stable between -0.4 and -0.5 and highly statistically significant. The spc and cay proxies are themselves insignificant and do not add explanatory power. $SVIX^2$ adds explanatory power, but in a way orthogonal to LTG_t: the R^2 of $SVIX^2$ alone is 19%. Overall, then, columns (1)-(5) in Panel A validate LTG_t as a measure of beliefs and confirm its high predictive power for returns.

Coming to the second question, we compare the explanatory power of LTG_t to that of determinants or predictors of stock prices/returns from previous work. In Panel A, column (6) we control for short term earnings growth expectations. The predictive power of LTG_t is robust to introducing this control which is itself insignificant, consistent with Table 1. In Panel B, columns (1)-(4), we control for well-established macroeconomic predictors of stock returns: the term spread, the credit spread, Bloom's uncertainty index, and the Kelly Pruitt factor (Kelly and Pruitt 2013). None of these predictors is statistically significant once we control for LTG_t and the gain in R^2 compared to Table 1 is modest.

Finally, we consider the role of expected long term inflation and past dividend growth. De la O and Myers (2022) view expected long term inflation as a determinant of beliefs about real fundamentals: excessively high (low) expected inflation should be associated with excess pessimism (optimism) about future earnings, predicting high (low) future returns. In column (5) of Panel B the predictive power of LTG_t is shown to be robust to controlling for expected long term inflation, confirming that it captures significant variation in real expected fundamentals. [9]

[9] Based on the predictive role of long term inflation expectations, De La O and Myers (DM 2022) argue that expectations about real short-term earnings growth is what predicts returns, not LTG. This conclusion is however

We control for past dividend growth based on Nagel and Xu (2021), who see it as causing excess optimism about future dividend growth, in turn leading to low future returns. Column (6) of Panel B show that LTG_t is robust to this control as well. This evidence strengthens the link between beliefs and return predictability: past dividend growth may affect returns by also changing preferences and hence discount rates. In addition, even if past dividends only affect expectations, the predictive power of LTG_t shows that beliefs do not just reflect past performance, but also news about the future (Daniel and Titman 2006). This resonates with Kindleberger's (1979) idea that new technologies help inflate asset bubbles.[10,11]

If LTG_t predicts returns, what determines its evolution? An analysis of this issue also provides useful input into the rest of our study. Table 3 reports, in column (1), the regression of the one year revision ΔLTG_t on lagged beliefs, LTG_{t-1}, and on earnings surprises relative to cyclically adjusted earnings, $e_t - cae_{t-5}$. The coefficient on LTG_{t-1} provides information on the persistence of beliefs, the coefficient on $e_t - cae_{t-5}$ on whether beliefs respond to sustained earnings growth (which is more relevant to assessing long term fundamentals than temporary growth episodes). Of course, because LTG_t may also be updated based on news about the future, we should not expect past fundamentals to account for 100% of its revisions.

Column (2) presents an additional test that LTG_t is not mechanically set to fit market prices or required returns by controlling for stock returns in the past year and for one-year ahead expected

---

flawed for three reasons. First, the predictive role of long-term inflation expectations may be spurious, for it reflects the high inflation of the 70s, which was followed by low inflation and high stock returns in the 80s. Second, DM's analysis does not address the basic fact that the predictive power of short-term earnings growth expectations for returns is weak (Table 1, Panel B), and disappears when one controls for LTG_t (Table 2, Panel A, column 6). On a related note, short term inflation expectations would seem to be more relevant than long term ones for their emphasis on short term real earnings growth. Finally, DM propose a test that is sufficient, but not necessary, for return predictability and use an incorrect specification LTG_t as growth between years 3 and 5. When using the correct definition of cumulative growth over the next 3 to 5 years, LTG_t passes the test. Specifically, future LTG_t errors, both actual and predicted using the model in Table 4 are negatively predicted by the current price dividend ratio (univariate regression coefficients of -0.2883 ($p = 0.077$) and -0.3915 ($p = 0.055$) respectively in our main sample). In Section 3 we perform a more systematic "horse race" to assess the extent to which the predictive power of pd_t for returns is due to predictable future LTG_t revisions and errors.

[10] The results of Table 2 hold at a 3 year horizon and when including other predictors (Appendix B, Table B.2).

[11] Hillenbrand and McCarthy (2022) regress the price earnings ratio on measured beliefs and on required return proxies. The R^2 of the regression using measured beliefs is 77%, which increases to 84% when proxies for required returns are added. In this analysis, consistent with our results, LTG_t is the variable with largest explanatory power.

return from the CFO Survey[12]. In columns (3), (4), (5) and (6) we control for the dividend price ratio and the proxies for discount rates we used in Table 2. If market prices move with news about future fundamentals, LTG_t revisions will correlate with contemporaneous returns as well as with price ratios. A key aspect of this exercise is to check whether theory-based drivers of expectations, such as recent growth in fundamentals, predict revisions even after controlling for prices.

Table 3
Determinants of LTG revisions

We study the association between one-year changes in the forecast for growth in the long run and predictors of returns (empirical and theoretical). The dependent variable is the change in the forecast for growth in earnings in the long run LTG_t between year t and t-1, ΔLTG_t. The independent variables are: (a) the one-year lagged value of LTG_t, (b) log of earnings for the S&P500 in year t relative to cyclically-adjusted earnings in year $t-5$, $e_t - cae_{t-5}$, (c) the (log) return on the S&P500 between year $t-1$ and t, r_{t-1}, (d) the forecast for the S&P500's one-year return from the Graham and Harvey survey, $\mathbb{E}_t^O[r_{t+1}]$, (e) the price to dividend ratio, pd_t, in column [3] (f) the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [4], (g) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [5] and (h) the Martin (2013) expected return on the market, SVIX^2, in column [6]. Data is monthly (quarterly) in columns [1], [3], [4] and [6] ([2] and [5]). All variables are standardized and intercepts are not shown. The sample period is 1981:12-2020:12. Newey-West standard errors are shown in parentheses (with 12 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: ΔLTG_t					
LTG_{t-1}	-0.4349[a]	-0.4624[a]	-0.5451[a]	-0.4393[a]	-0.3232[a]	-0.3338[b]
	(0.1616)	(0.1090)	(0.1489)	(0.1429)	(0.1187)	(0.1510)
$e_t - cae_{t-5}$	0.3938[a]	0.3006[a]	0.3409[a]	0.3274[a]	0.3883[a]	0.4663[a]
	(0.0827)	(0.0561)	(0.0570)	(0.0770)	(0.0889)	(0.1173)
r_{t-1}		0.0572				
		(0.1023)				
$E_t^O[r_{t+1}]$		0.0858				
		(0.0959)				
X_t			0.2828[a]	0.2291[a]	-0.0928	0.1459
			(0.0945)	(0.0655)	(0.1214)	(0.1754)
Observations	457	76	457	457	148	193
Adj R^2	31%	38%	37%	36%	31%	52%
X_t			$p_t - d_t$	spc_t	cay_t	$SVIX_t^2$

In column (1), the coefficient on LTG_{t-1} is negative and smaller than one in magnitude, showing that LTG_t is quite persistent but tends to mean revert. The positive coefficient on $e_t -$

[12] Available at https://www.richmondfed.org/cfosurvey.

cae_{t-5} further suggests that LTG_t is revised upward after periods of sustained earnings growth. These two forces alone account for roughly one third of the variation in *LTG* revisions.

None of these conclusions change materially when we control for past and expected returns, the price dividend ratio, and the required return proxies. The evidence confirms that the change in LTG_t reflects genuine belief revisions about future fundamentals. Two out of four controls are insignificant, and they mostly only marginally improve explanatory power.[13]

Overall, we showed that LTG_t strongly predicts future aggregate stock returns and that it offers a good proxy for market expectations of long-term fundamentals. How are beliefs, as measured by LTG_t, and returns connected? We study this question next.

3. Expectations and Stock Returns

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

$$r_{t+1} = \alpha p_{t+1} + (1-\alpha)d_{t+1} - p_t + k, \tag{1}$$

where p_t is log stock price at t, d_{t+1} is the log dividend at $t+1$, while $k > 0$ and $\alpha \in (0,1)$ are constants. Iterating Equation (1) forward and imposing the transversality condition, we obtain:

$$p_t - d_t = \frac{k}{1-\alpha} + \sum_{s\geq 0} \alpha^s g_{t+1+s} - \sum_{s\geq 0} \alpha^s r_{t+1+s}, \tag{2}$$

where $g_{t+s+1} \equiv d_{t+s+1} - d_{t+s}$ is dividend growth between $t+s$ and $t+s+1$.

The average firm in the economy, which we call "the market," has dividend growth:

$$g_{t+1} = \mu g_t + v_{t+1}, \tag{3}$$

where is v_{t+1} is an i.i.d. Gaussian shock with mean zero and variance σ_v^2 and $\mu \in [0,1]$. In BGLS (2020) we showed that our key results hold under a general covariance stationary process. The shock v_{t+1} captures tangible news arriving at $t+1$ such as earnings news, proxied for instance

[13] Table B.4 in Appendix B shows the results are robust to controlling for further measures of required returns, as well as for lagged 5 year returns (as a proxy for expectations of returns, Greenwood and Shleifer (2014)).

by the measure $e_t - cae_{t-5}$, but it can also capture intangible news learned at t but affecting future earnings, such as the introduction of a new technology. We write $v_{t+1} = \tau_{t+1} + \eta_t$, where τ_{t+1} is tangible news, η_t intangible news, and the variance of v_{t+1} reflects the two components $\sigma_v^2 = \sigma_\tau^2 + \sigma_\eta^2$. By using expectations data we can capture both tangible and intangible news. Table 2 shows that intangible news is important: expectations data have considerable explanatory power even controlling for past fundamentals.[14]

In Equation (2), the variation in the current price to dividend ratio is due to expected variation in future dividend growth (captured by the g_{t+1+s} terms), required returns (captured by the r_{t+1+s} terms), or both. Rational expectations theories of return predictability rely only on the second source of variation. In these theories, expectations of fundamentals $\mathbb{E}_t(g_{t+s+1})$ are formed by optimally using Equation (3), while rational expectations of future returns $\mathbb{E}_t(r_{t+s+1})$ are also formed using the true model of required returns (which we do not need to specify). Under rational expectations, the realized stock return between t and $t+1$ is then given by:

$$r_{t+1} = \mathbb{E}_t(r_{t+1}) + \sum_{s\geq 0} \alpha^s (\mathbb{E}_{t+1} - \mathbb{E}_t)(g_{t+1+s}) - \sum_{s\geq 1} \alpha^s (\mathbb{E}_{t+1} - \mathbb{E}_t)(r_{t+1+s}), \qquad (4)$$

so that realized returns are driven by three components: the required return between t and $t+1$, $\mathbb{E}_t(r_{t+1})$, rational belief revisions about future dividends $(\mathbb{E}_{t+1} - \mathbb{E}_t)(g_{t+1+s})$, and rational belief revisions about future returns, $(\mathbb{E}_{t+1} - \mathbb{E}_t)(r_{t+1+s})$. Because rational belief revisions reflect news arriving at $t+1$, they are unpredictable at time t. As a result, under rational expectations return predictability is only due to variation in $\mathbb{E}_t(r_{t+1})$.

In our approach to predictability, in contrast, required and hence expected returns are constant at r, but beliefs about future fundamentals are formed using a distorted operator $\widetilde{\mathbb{E}}_t(g_{t+1+s})$, not by optimal forecasts using Equation (3). Here realized returns are given by:

---

[14] We perform a systematic analysis of tangible (i.e. measured in terms of fundamentals) versus intangible news in Appendix D. We find that predictive power of past fundamentals is typically economically smaller and statistically less significant than that of measured beliefs, suggesting an important role for intangible news.

$$r_{t+1} = r + \sum_{s \geq 0} \alpha^s \left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)(g_{t+1+s}). \quad (5)$$

Critically, the belief revision $\left(\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t\right)(g_{t+1+s})$ occurring at $t+1$ is no longer pure "news". It is also shaped by systematic belief distortions prevailing at t. These distortions, embedded in the time t forecast $\tilde{\mathbb{E}}_t(g_{t+1+s})$, are the source of return predictability in our approach.

To characterize the predictions from (5), we lay out a reduced form model of beliefs that nests the leading departures from rationality studied in macroeconomic and finance: overreaction to news, as in models of diagnostic expectations (Bordalo, Gennaioli, and Shleifer 2018, BGLS 2019), but also as in earlier models (e.g., Barberis et al. 1998), and underreaction to news, as in models of rational or non-rational inattention (Sims 2003, Gabaix 2013, Huang and Liu 2007, Bouchaud et al. 2019). The model highlights the distinctive predictions of these theories with respect to the forecast errors and their link to return predictability.

3.1 Non-Rational Beliefs and their Empirical Predictions

We model departures from rationality as a time varying distortion ϵ_t whose impact on beliefs decays with the forecast horizon according to the true persistence μ of fundamentals:

$$\tilde{\mathbb{E}}_t(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \mu^{s-1}\epsilon_t, \quad (6)$$

where $s \geq 1$ and $\mathbb{E}_t(g_{t+s}) = \mu^{s-1}(\mu g_t + \eta_t)$ is the rational forecast based on (3).

The distortion ϵ_t follows an AR(1) process, $\epsilon_t = \rho\epsilon_{t-1} + u_t$, where $\rho \in [0,1]$ and u_t is an expectations shock. Parameter ρ captures the observed persistence in LTG_t. We impose $\rho < \mu$ to reproduce one key fact in Table 3: the negative correlation between LTG_t revisions and lagged forecast LTG_{t-1}, i.e., $cov\left[\tilde{\mathbb{E}}_{t+1}(g_{t+s}) - \tilde{\mathbb{E}}_t(g_{t+s}), \tilde{\mathbb{E}}_t(g_{t+s})\right] < 0$. This implies that excess optimism or pessimism gradually yet systematically revert over time.

The over- vs under-reaction in beliefs is incorporated into the expectations shock u_t. We assume that u_t is proportional to news, captured by the rational belief revision at t. Formally, $u_t = \theta(\mu\tau_t + \eta_t)$. If $\theta = 0$, expectations are rational. If $\theta > 0$, investors overreact, exaggerating the impact of news on expectations. If $\theta < 0$, investors underreact, dampening the effect of news on expectations. We assume $\theta > -1$, which ensures that good news are not viewed as bad and vice versa. Appendix A shows that, for $\theta > 0$, Equation (6) is a special case of the diagnostic expectations model (Bordalo et al. 2018).[15]

Equations (5) and (6) yield our two empirical tests, one on the predictability of forecast errors and another on predictability of returns from predictable forecast errors. The first test detects whether beliefs over or under react ($\theta \lesseqgtr 0$) by using the predictability of future forecast errors based on current expectations revisions.

Proposition 1. *Under Equation (6), the forecast error predictability regression:*

$$g_{t+s} - \widetilde{\mathbb{E}}_t(g_{t+s}) = \beta_0 + \beta_1\left[\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s})\right] + \beta_2\widetilde{\mathbb{E}}_{t-1}(g_{t+s}) + z_{t+s}, \qquad (7)$$

has $\beta_1 < 0$ *if and only if beliefs overreact to news,* $\theta > 0$. $\theta > 0$ *also implies* $\beta_2 < 0$.

Consistent with Bordalo, Gennaioli, Ma, and Shleifer (2020), who build on Coibion and Gorodnichenko (2015), a negative association $\beta_1 < 0$ between the current forecast revision $\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s})$ and the future forecast error $g_{t+s} - \widetilde{\mathbb{E}}_t(g_{t+s})$ is indicative of overreaction to *current* news. After good news (i.e., a positive revision) beliefs become too optimistic, predicting future disappointment (i.e., a negative error). Under-reaction entails the opposite association. Proposition 1 additionally says that, in our model, if beliefs over-react, then the lagged forecast $\widetilde{\mathbb{E}}_{t-1}(g_{t+s})$ also negatively predicts forecast errors. The reason is that the belief distortion ϵ_t is persistent, so high lagged forecasts incorporate over-reaction to past news, which also leads to systematic future disappointment.

Our second, and key, test links systematic forecast errors in earnings growth to return predictability. It is obtained using Equations (3), (5), and (6).

[15] Our model rules out non-fundamental noise (Black 1986, DeLong et al. 1990a). It can be easily introduced in the analysis to capture an extreme form of overreaction, in which beliefs react to wholly irrelevant factors.

Proposition 2 *The realized return at t+1 is given by:*

$$r_{t+1} = r + \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\mathbb{E}_t\left[g_{t+1} - \widetilde{\mathbb{E}}_t(g_{t+1})\right] + \omega_{t+1} \tag{8}$$

where $\omega_{t+1} = \left(\frac{1+\alpha\theta}{1-\alpha\mu}\right)\tau_{t+1} + \alpha\left(\frac{1+\theta}{1-\alpha\mu}\right)\eta_{t+1}$ *is a combination of tangible and intangible news.*

The realized return depends on news arriving at $t+1$, captured by ω_{t+1}, but also on the $t+1$ forecast error predictable using information available at t. This positive association between future returns and predictable forecast errors connects our two tests.

Prediction 1 LTG_t *overreacts,* $\theta > 0$*, if and only if forecast errors in earnings growth and future stock returns are both negatively predicted by the current* LTG_t *revision. If* $\theta > 0$*, then lagged* LTG_t *also negatively predicts both forecast errors and returns.*

If upward LTG revisions and high lagged LTG predict stronger disappointment of earnings growth expectations (i.e. more negative forecast errors), then beliefs overreact. In our theory, then, upward LTG revisions and high lagged LTG additionally imply a currently inflated stock market, in turn predicting lower future stock returns. Using expectations data, we can now test for this joint predictability of forecast errors and returns.

3.2 Predictability of Aggregate Stock Returns

Table 4 tests Prediction 1, combining Propositions 1 and 2. Column (1) tests Equation (7) from Proposition 1: it predicts the forecast error in the five years ahead earnings growth using the one year LTG_t revision and the lagged forecast, LTG_{t-1}. Column (2) uses the same explanatory variables to predict five year ahead returns. Column (3) performs an IV strategy testing Equation (8) from Proposition 2: in the first stage we predict forecast errors using the model in Column (1), in the second stage we use the fitted forecast errors to predict returns. In Column (4) we perform a robustness test: we add the current price to dividend ratio as a regressor.

Table 4
Predictability of Forecast Errors and Returns

This table links aggregate forecast errors and market returns. We report regressions using as dependent variable the error in forecasting five-year growth in aggregate earnings in column [1] and the discounted value of the cumulative market return between year t and $t+5$ in columns [2] and [3 Five-year cumulative market returns ($\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$) are computed using monthly data and run from $t+1/12$ through $t+60/12$. We define the forecast error as the difference between (a) the annual growth in earnings per share between year t and $t+5$, $\Delta_5 e_{t+5}/5$, and (b) the expected long term growth in earnings, LTG_t. The independent variables are the one-year change in LTG_t, ΔLTG_t, the lagged forecast, LTG_{t-1}, and the predicted forecast error, $\Delta_5 e_{t+5}/5 - LTG_t$. We assume that earnings are reported with a with a 3-month lag (i.e. we define e_t as earnings for the calendar period $t-1/4$). We report OLS estimates in columns [1] and [2], and second-stage IV results in column [3]. The instrumental variables are ΔLTG_t and LTG_{t-1}. Except for ΔLTG_t, all variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	Dependent Variable:			
	$\Delta_5 e_{t+5}/5 - LTG_t$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
ΔLTG_t	-0.8407[a]	-0.6403[a]		
	(0.1533)	(0.0764)		
LTG_{t-1}	-0.2157	-0.5252[a]		
	(0.1369)	(0.0864)		
$\Delta_5 e_{t+5}/5 - LTG_t$			0.8460[a]	
			(0.2501)	
$\widehat{\Delta_5 e_{t+5} - LTG_t}$				0.3853[a]
				(0.1198)
pd_t				-0.6377[a]
				(0.1773)
Obs	397	397	397	397
Adj R^2	25%	31%		48%
Montiel-Pflueger F-stat			10,9	
Instrument	.	.	ΔLTG_t, LTG_{t-1}	

Column 1 shows that beliefs overreact, $\theta > 0$. Upward LTG_t revisions predict future disappointment, suggesting that beliefs become too optimistic when good news arrives. This confirms, at the level of the S&P 500 index, the belief overreaction documented by BGLS (2019) at the level of stock portfolios. Here we find overreaction at the consensus level (remember we are using the median LTG forecast). This is a strong result: analysts' overreaction here overwhelms information frictions. The latter can in fact produce consensus underreaction even if individual forecasters overreact (Woodford 2003, BGMS 2020). Also in line with overreaction, a higher lagged forecast LTG_{t-1} is associated with a lower forecast error. This association is not significant at conventional levels, but our other results show statistical significance for LTG_{t-1}.[16]

Column (2) connects belief overreaction to return predictability. Upward LTG_t revisions and higher lagged forecast LTG_{t-1} predict sharply lower future stock returns. This finding is consistent with our mechanism. Overreaction to current news causes excessive upward LTG_t revisions, high ϵ_t, and hence an excessive stock market boom at t. This is followed by belief disappointment, a downward price correction and hence low returns r_{t+1}. Higher lagged forecast LTG_{t-1} also predicts low future returns for the same reason.

Column (3) links predictable forecast errors to future returns. Consistent with Equation (8), periods of excess pessimism in which future forecast errors are systematically high (growth is above expectations) are on average followed by high stock returns. Conversely, periods of excess optimism in which future forecast errors are systematically low (growth is below expectations) are on average followed by low returns. Column (4) shows that, as in Table 2, this holds even after controlling for the price dividend ratio, confirming that the link between forecast errors and returns is unlikely to be due the reverse engineering LTG from stock prices. In this exercise, LTG_t proxies for beliefs at a specific horizon, while $p_t - d_t$ captures beliefs at other horizons (which may be affected by independent factors), and required return variation.

Quantitatively, the effects are sizable. In column (2) a one standard deviation higher revision ΔLTG_t (equal to 0.62) is associated with a roughly 0.4 of a standard deviation lower future return, and a one standard deviation higher lagged forecast LTG_{t-1} (equal to 1) is associated with a roughly 0.5 of a standard deviation lower future return. These effects imply reductions in 5-year log returns of 0.13 and 0.17, respectively. Since the average monthly log return is 0.007, this corresponds to losing 19 to 25 months' worth of returns over five years.

[16] One may worry that measurement error in LTG_t may create a spurious negative correlation between forecast errors and ΔLTG_t. However, i) forecast errors are predictable from LTG revisions instrumented by past growth in fundamentals (Table B.6 in Appendix B), and ii) revisions of LTG_t negatively predict returns, consistent with overreaction (Table 4 column 2). Table B.5 shows the results in Table 4 follow through at the 3 year horizon.

The explanatory power of expectations is also high in terms of R^2: the model in column (2) accounts for 31% of return variation at a five year horizon. The explanatory power of expectations is much higher than that of past fundamentals. Our measure of earnings growth in the past five years, $e_t - cae_{t-5}$, negatively predicts returns but with an R^2 of only 13%.

Do beliefs about long term growth help account for the predictive power of the price dividend ratio? We can address this question using the exact relationship between returns and news in Equations (4) and (5). Under rational expectations, when Equation (4) holds, $p_t - d_t$ predicts future returns by acting as an inverse measure of the required return $\mathbb{E}_t(r_{t+1})$. In this case, $p_t - d_t$ is orthogonal to proxies for the news affecting r_{t+1}. Under belief over-reaction and constant required returns, in contrast, returns are described by Equation (5). Now high $p_t - d_t$ signals excess optimism, so it predicts low returns by predicting systematically disappointing future "news", as measured by systematic expectations revisions $\left(\widetilde{\mathbb{E}}_{t+k} - \widetilde{\mathbb{E}}_t\right)(g_{t+1+s}) < 0$. A key implication follows: under rational expectations, the explanatory power of $p_t - d_t$ is unaffected if future returns r_{t+1} are purified from expectation revisions at $t + 1$. Under belief over-reaction, in contrast, the explanatory power of $p_t - d_t$ is reduced, because part of the expectation revisions in $t + 1$ capture systematic reversal of overreaction at t.

We test this implication using a two stage test. In the first stage, we regress realized returns r_{t+k} at the one, three, and five years horizons, $k = 1,3,5$, on the LTG_t news occurring at the same horizon. We proxy news by the one year LTG revisions and by the LTG forecast errors occurring between $t + 1$ and $t + k$. The return residuals $\tilde{r}_{t+k}$ from this regression purify the return r_{t+k} from variation due to future LTG news. In the second stage, we use $p_t - d_t$ to predict residuals $\tilde{r}_{t+k}$. Under rationality, the coefficient on $p_t - d_t$ should not change when predicting raw returns or residuals. Under overreaction, the coefficient's magnitude should be smaller in the latter case.

Table 5 reports the results of this exercise. For brevity, we report only the second stage result (the first stage is in Appendix B, Table B.7 and shows, intuitively, that better news entail higher returns). Columns (1), (3) and (5) report the coefficient from regressing raw returns r_{t+k} on $p_t - d_t$. Columns (2), (4) and (6) report the coefficient for residualized returns $\tilde{r}_{t+k}$.

Table 5
Unbundling Return Predictability from Price Dividend Ratio

This table examines why the price-dividend ratio ($p_t - d_t$) predicts stock market returns. We proceed in two steps. In the first step, we regress realized returns r_{t+k} at the one, three, and five years horizons ($k = 1,3,5$) on one-year revisions in long term growth in earnings occurring between $t+1$ and $t+k$ (i.e. ΔLTG_{t+1} through ΔLTG_{t+k}) and long-term forecast errors occurring between $t+1$ and time $t+k$ (i.e. $e_{t+1} - LTG_{t-4}$ through $\Delta_5 e_{t+k} - LTG_{t+k-5}$), presented in Table B.7. We generate return residuals $\tilde{r}_{t+k}$ from these first-stage regression. In the second step, we use $p_t - d_t$ to predict $r_{t+\mathrm{k}}$ in columns [1], [3], and [5] and $\tilde{r}_{t+k}$ in columns [2], [4], and [6]. Except for ΔLTG_t, all variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 12 lags in columns [1]-[2], 36 lags in column [3]-[4], and 60 lags in columns [5]-[6]). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	r_{t+1}		$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$		$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	
	raw	residual	raw	residual	raw	residual
$p_t - d_t$	-0.3742[a]	-0.3017[a]	-0.6219[a]	-0.2537[a]	-0.8204[a]	-0.1670[c]
	(0.1446)	(0.0956)	(0.2012)	(0.0898)	(0.2115)	(0.0983)
Obs	361	361	361	361	361	361
R^2	10%	10%	28%	13%	48%	10%
Adj R^2	10%	9%	27%	13%	48%	10%

Overreaction in the expectations of long term growth, reflected in systematic LTG revisions and forecast errors, accounts for a large chunk of price dividend ratio's predictive power at medium to long term horizons. At the one year horizon, LTG matters little in the sense that the estimated coefficient in column (1) is indistinguishable from that in column (2). At the 3 and 5 years horizons, in contrast, the effect is dramatic: after removing returns that are due to LTG news, the magnitude of the estimated coefficient for $p_t - d_t$ drops by a factor of 2.5 at the three years horizon and more than 5 at a five year horizon.[17] This evidence suggests that the bulk of the price dividend ratio's predictive power is due to its ability to capture non-rational beliefs, in particular systematic reversals of overreaction, as proxied by LTG news.

[17] Results are similar if we proxy for LTG news using only forecast errors (Appendix B, Table B.8).

3.3 Predictability of Firm Level Stock Returns

The results in Table 4 might be influenced by a few outlier episodes, such as the internet bubble. One way to address this concern is to test Prediction 1 at the firm level, since this allows us to control for time dummies, purging the effects of common shocks (including shocks to required returns). We can also include firm fixed effects, which control for constant differences in average returns across firms.

Table 6 performs this firm level exercise, controlling for year dummies and firm fixed effects. Column (1) predicts forecast errors for a firm's five years ahead earnings growth using the one year changes in a firm's forecast ΔLTG_{it} and the lagged forecast $LTG_{i,t-1}$. Column (2) uses the same regressors to predict the firm's stock returns over the next five years. Column (3) uses the errors fitted in column (1) as instruments to predict returns.[18] In line with Table 4, in column (4) we control for the firm's price dividend ratio (restricting to the observations where dividends are paid), and in column (5) we control for the price earnings ratio (using observations with positive earnings). To assess the role of the episodes in our data with the greatest returns, in column (6) we exclude the years 1998-2002 and 2007-2009.

Table 6
Firm-Level Results

We present firm-level regressions for all US firms in the IBES sample. We define firm-level forecast errors as the difference between (a) the growth in firm i's earnings per share between year t and $t+5$, $\Delta_5 e_{i,t+5}/5$, and (b) the expected long term growth in firm i's earnings, $LTG_{i,t}$. In column [1] we perform an OLS regression of the error in forecasting the five-year earnings growth on: (a) the one year revision of the forecast for a firm's long-term earnings growth, $\Delta LTG_{i,t}$ and (b) the lagged forecast $LTG_{i,t-1}$. In column [2] we perform an OLS regression of discounted the cumulative (log) return for firm i between year t and $t+5$, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$ on the same two independent variables. In column [3] we perform an IV regression of stock returns, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$, on the forecast errors fitted in column [1]. In column [4] we perform an OLS regression of five-year returns, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$, repeating the regression on the forecast errors fitted in column [1] controlling for the price dividend ratio in column [4] (for observations with positive dividends) and the price earnings ratio in column [5] (for observations with positive earnings). In column [6], we return to the benchmark specification in column [3] but exclude from the sample the years 1998-2002 and 2007-2009. Except $\Delta LTG_{i,t}$, all variables are standardized. Regressions include time- and firm-fixed effects, which we do not report. Except in column [6], the sample period is 1982:12-2015:12. We report Driscoll–Kraay standard errors with autocorrelation of up to 60 lags. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

[18] Following BGLS (2019), here we consider all domestic common stocks in the IBES Unadjusted US Summary Statistics file, which includes stocks listed on major U.S. stock exchanges (i.e., NYSE, AMEX, and NASDAQ) except for closed-end funds and REITs. From the IBES Detail History Tape file we obtain analyst earnings forecasts.

	(1)	(2)	(3)	(4)	(5)	(6)
	$\frac{\Delta_5 e_{i,t+5}}{5} - LTG_{i,t}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$
$\Delta LTG_{i,t}$	-0.3286[a]	-0.1773[a]				
	(0.0248)	(0.0409)				
$LTG_{i,t-1}$	-0.3626[a]	-0.2163[a]				
	(0.0256)	(0.0446)				
$\widehat{\Delta_5 e_{i,t+5}/5 - LTG_t}$			0.5768[a]	0.3647[a]	0.4868[a]	0.4594[a]
			(0.0919)	(0.0554)	(0.1056)	(0.0665)
$pd_{i,t}$				-0.4460[a]		
				(0.1109)		
$pe_{i,t}$					-0.2963[a]	
					(0.0677)	
Observations	371.571	371.571	371.571	259,734	371,567	304,256
Adj R^2	4%	1%		3%	3%	
KP F-stat	.	.	101,8	.	.	187.7
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Instrument	.	.	$LTG_{i,t-1}$, $\Delta LTG_{i,t}$			$LTG_{i,t-1}$, $\Delta LTG_{i,t}$

Column (1) again shows strong evidence of overreaction. Upward firm level LTG_{it} revisions predict future disappointment (negative forecast errors), and the same does a high lagged forecast $LTG_{i,t-1}$, in line with the aggregate results. Column (2) confirms, at the firm level, the result on return predictability: higher firm level forecast revisions $\Delta LTG_{i,t}$ and lagged forecast $LTG_{i,t-1}$ are associated with sharply lower returns. The R^2 in column (2) is lower than that for the aggregate market, perhaps because there are many sources of idiosyncratic and unpredictable variation in firm level returns. Still, coefficient magnitudes are sizable: a one standard deviation higher forecast revision (equal to 0.53) or of lagged forecast (equal to 1) are followed by a 0.09 (respectively 0.22) of a standard deviation lower return at the firm level.

Column (3) confirms the direct link between predictable disappointment and predictable returns: periods in which beliefs about a firm are over-pessimistic (over-optimistic), in the sense that they are systematically followed by earnings growth predictably above (below) expectations, are also periods in which the firm's stock return is higher (lower). Columns (4) and (5) show that the results are robust to controlling for the price dividend ratio (for the observations that pay dividends) and for the price earnings ratio, providing further evidence that LTG is not inferred from prices. In Column (6) we exclude the years of the internet bubble and the financial crisis: our results are not driven by the episodes with largest returns in the data.[19]

In sum, measured expectations display strong overreaction to news and boom-bust stock price dynamics: good news lead to excessive optimism, which is associated with an inflated stock price and a future price reversal when over-optimism is disappointed. The same mechanism operates for both the aggregate market and individual firms, indicating its generality.

4. Return Predictability in the Cross Section

Decades of asset pricing research have unveiled puzzling differences in average returns across stocks grouped based on observed characteristics such as the book to market ratio, profitability, etc. Such predictability is systematic. For instance, high book to market stocks tend to do poorly together, compared to low book to market stocks, and likewise for other characteristics (Fama and French 1993). Some scholars view such cross sectional return predictability as reflecting differential exposure to systematic risk factors (Fama and French 1993, Cochrane 2011). Other scholars argue instead that it reflects systematic psychological factors (DeLong et al. 1990, Lakonishok et al. 1994, Kozak, Nagel and Santosh 2018).

Expectations data allow to empirically assess this debate. We just showed that expectations about *aggregate* long term earnings growth, LTG_t, capture systematic overreaction in market beliefs. Can such systematic overreaction also shed light on the comovement of returns in the cross section? In Section 4.1 we address this question by focusing on the cross-sectional return spread between high and low LTG firms (La Porta 1996). Section 4.2 broadens the analysis to consider the returns of Fama-French (1993) factors.

[19] In Appendix B we extend the results of Table 6 by predicting returns at a 3 year horizon (Table B.9), as well as by instrumenting revisions using fundamental growth (Table B.10).

4.1 LTG and time variation in the LTG Spread

La Porta (1996) showed that firms in the top LTG decile have predictably lower stock returns than firms in the bottom LTG decile. BGLS (2019) show that a model in which beliefs about a firm's long-term earnings growth overreact can account for this finding. Here we ask a new question: do the returns of stocks in the top (bottom) LTG decile comove with aggregate LTG_t, causing systematic variation in the LTG cross sectional spread? Addressing this question is a key first step to understanding whether systematic belief biases shape cross sectional mispricing.

To make progress, in Table 7 we regress the five year log return of portfolios of stocks sorted based on LTG on our proxies for aggregate over-optimism, namely the forecast revision ΔLTG_t and the lagged forecast LTG_{t-1}. We also add the contemporaneous market return, which captures the CAPM co-movement based on the fundamental risk exposure. Column (1) reports the regression results for the low LTG portfolio ($LLTG$), defined as the bottom decile of stocks based on their median LTG. Column (2) presents the same regression for the high LTG portfolio ($HLTG$) defined as the top decile of stocks based on their median LTG. Column (3) estimates the same model for the return on the low minus high LTG portfolio. We call this portfolio "Pessimism minus Optimism" LTG, or PMO, adopting the Fama-French convention of forming a portfolio whose long arm is the group of firms earning a higher average return, $LLTG$ in our case. Columns (4) through (7) add to column (3) regressions the three conventional proxies for discount rates.

Table 7
Market Return and LTG portfolio returns

We predict the return for portfolios formed by the forecast for long-term growth in earnings for firm *i*, $LTG_{i,t}$ using expectations about earnings growth for the market. On each month between December 1982 and December 2015, we form decile portfolios based on $LTG_{i,t}$ and report regression results for the five-year cumulative (log) returns on: (a) the lowest decile ($LLTG$) in column [1], (b) the highest decile ($HLTG$) in column [2], and (c) the difference between the two ($PMO = LLTG - HLTG$) in columns [3]-[7]. The independent variables are: (a) the one year forecast revision for long term growth in aggregate earnings, ΔLTG_t, (b) the one-year lagged forecast, LTG_{t-1}, (c) the (log) five-year return of CRSP's value-weighted index between t and $t+5$, $\ln(Mkt_{t,t+5})$, and: (d) the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [4], (d) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [5], (e) the Martin (2013) expected return on the market $SVIX^2$ in column [6], and (f) the price dividend ratio, pd_t, in column [7]. Except for ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: (Log) Five-year Return						
	LLTG	HLTG	PMO	PMO	PMO	PMO	PMO
ΔLTG_t	0.2335[c]	-0.7946[a]	0.9878[a]	0.9207[a]	0.9119[a]	1.0709[a]	1.0545[a]
	(0.1414)	(0.1779)	(0.1991)	(0.1895)	(0.2045)	(0.2928)	(0.2098)
LTG_{t-1}	0.3445[c]	-0.4515[a]	0.6408[a]	0.5990[a]	0.5767[a]	0.5150[a]	0.7299[a]
	(0.1831)	(0.0851)	(0.1462)	(0.1356)	(0.1403)	(0.1157)	(0.1071)
$\ln(Mkt_{t,t+5})$	0.8461[a]	0.5177[a]	-0.2691[a]	-0.2451[b]	-0.3241[a]	-0.2396	-0.4176[a]
	(0.1037)	(0.0958)	(0.0973)	(0.0958)	(0.1228)	(0.1962)	(0.1001)
X_t				0.1875	0.1937	0.0370	-0.3027[a]
				(0.1371)	(0.1294)	(0.1623)	(0.1093)
Observations	397	397	397	397	133	193	397
Adjusted R^2	69%	83%	70%	71%	71%	73%	74%
X_t				spc_t	cay_t	$SVIX_t^2$	$p_t - d_t$

There is a strong systematic variation in the LTG spread. $HLTG$ stocks appear to do worse in bad times than $LLTG$ stocks (the PMO return loads negatively on the market).[20] Thus, the LTG spread cannot be explained by the fact that $LLTG$ stocks are riskier in the standard CAPM sense. A more promising avenue is to look at the LTG_t proxies: $HLTG$ stocks are more exposed to waves of aggregate optimism compared to $LLTG$ stocks. In columns (1) and (2), good news about long term earnings growth, reflected in high LTG_t revisions, is followed by higher returns for $LLTG$ stocks and by lower returns for $HLTG$ stocks. The same holds when the lagged forecast LTG_{t-1} is high. Thus, in column (3), the PMO spread is higher after periods of aggregate optimism, due to the good performance of $LLTG$ stocks, the long arm of the portfolio, and the poor performance of $HLTG$ stocks, the short arm of the portfolio.

Differential co-movement with aggregate optimism sheds new light on the origins of the LTG spread. The model in column (3) accounts for 70% of the time variation in the LTG spread, compared to only 25% of the market return alone. Measures of discount rates play no role in explaining the data (see columns (4), (5) and (6)). Furthermore, the magnitude of the PMO spread explained by aggregate expectations is very significant: relative to a long term average of 5.5%

[20] The market loading of the PMO portfolio is not the difference between the *LLTG* loading in column (1) and the *HLTG* loading in column (2) because the variables are standardized, and *LLTG* has lower variance than *HLTG*.

per year, the spread increases to 23.5% per year as LTG_{t-1} goes from its average of 12.2% to 15.5% (a two standard deviations increase). This suggests that the average aggregate LTG captures some level of over-optimism, corresponding to the over-valuation of the *HLTG* firms.[21]

To study how aggregate optimism can create cross sectional co-movement, we introduce firm heterogeneity in our model. For simplicity, we abstract from intangible news by setting $\eta_t = 0$, but this is not critical (see footnote 24). Each firm i exhibits AR(1) dividend growth:

$$g_{i,t+1} = \mu g_{i,t} + v_{i,t}, \tag{9}$$

As in Equation (6), expected growth at horizon $s \geq 1$ for firm i is believed to be:

$$\widetilde{\mathbb{E}}_t(g_{i,t+s}) = \mathbb{E}_t(g_{i,t+s}) + \mu^{s-1}\epsilon_{i,t}. \tag{10}$$

The firm specific belief distortion continues to follow an AR(1) process $\epsilon_{i,t} = \rho\epsilon_{i,t-1} + u_{i,t}$ with persistence $\rho \in [0,1]$, where $u_{i,t}$ is a firm level expectations shock.

As in standard cross-sectional asset pricing, firm level and aggregate shocks are connected. The firm level fundamental shock is the product $v_{i,t} = v_i * v_t$ of the aggregate fundamental shock v_t and a parameter $v_i > 0$ capturing the firm's exposure to it. This is the standard CAPM exposure to fundamental risk, which varies across firms. Similarly, the firm level expectations shock $u_{i,t}$ can be written as the aggregate expectation shock u_t times a firm specific exposure to it. Think of it as a firm specific degree of belief overreaction θ_i, so that $u_{i,t} = \theta_i * v_t$. This key new aspect creates differential exposure of firms to aggregate optimism and pessimism.[22] The firm level belief distortion is then proportional to the aggregate one, $\epsilon_{i,t} = (\theta_i/\theta)\epsilon_t$.

A firm may be more exposed to aggregate optimism because it belongs to the "hot" sectors of the moment, or because it is similar enough to firms in such sectors (as in Bordalo et al 2021).

[21] The results of Table 7 hold when controlling for the contemporaneous price dividend ratio (Appendix C, Table C.1) and at a 3 year horizon (Table C.2, Panel A). BGLS (2019) show that the average (value weighted) PMO at 1-year horizon is 6.4%, similar to 5.5% in Table 7. We use value weighted portfolios due to our focus on the S&P500 index.
[22] The assumption that firms' fundamentals and beliefs perfectly comove with the market makes the model tractable. As we show in Appendix A, this assumption is not necessary. Enriching the model with idiosyncratic fundamental shocks would yield the additional implication that the LTG spread reflects also the overreaction of $HLTG$ firms to good firm-level fundamental news, as in BGLS (2019), not only the average degree of maker over-optimism ϵ_t.

For instance, optimism about the aggregate market may be due to the rapid growth of some high-tech firms. Such optimism may contaminate other high-tech firms due to higher fundamentals (high v_i), but also due to mere similarity, which increases θ_i for given v_i. The distinction between these two effects is key for understanding returns.

The realized return for firm i at $t+1$ is isomorphic to Equation (1). Using Equations (9) and (10) we can show (see Appendix A) that the realized return for firm i is given by:

$$r_{i,t+1} - r_i = \left(\frac{v_i + \alpha\theta_i}{1+\alpha\theta}\right)(r_{t+1} - r) - \left(\frac{1-\alpha\rho}{1-\alpha}\right)\left(\frac{\theta_i - \theta v_i}{\theta + \alpha\theta^2}\right)\epsilon_t, \tag{11}$$

where θ is market overreaction, r_i is the firm-specific required return, while $(r_{t+1} - r)$ is the realized market return in excess of the required one.[23] The firm's realized return depends on the excess market return $(r_{t+1} - r)$ and on past excess optimism ϵ_t according to firm specific coefficients. If a firm's exposure to aggregate optimism is only shaped by its exposure to market fundamentals, $\theta_i = \theta v_i$, the model boils down to the CAPM. The return of firm i loads with coefficient v_i on the market return, which is the only source of comovement. Even though the aggregate market displays excess volatility and return predictability, the cross section is correctly priced in terms of market exposure. Thus, the case $\theta_i = \theta v_i$ cannot explain Table 7.

If instead firms overreact more or less than warranted by their exposure to fundamentals, $\theta_i \neq \theta v_i$, the CAPM breaks down. Now the realized market return captures the firm's reaction to current aggregate shocks, while aggregate excess optimism ϵ_t captures the firm's relative overreaction to past shocks. Firms that overreact more than the average, $\theta_i - \theta v_i > 0$, exhibit a stronger comovement with the market, due to stronger overreaction to current news v_{t+1}. Critically, they are also more inflated during periods of high aggregate optimism ϵ_t. Thus, they

[23] Appendix A shows that under our assumptions, if investors have mean variance preferences and are naïve about ϵ_t, the required return r_i can be endogenized and is determined as in the CAPM: $r_i = r_f + v_i(r - r_f)$.

exhibit a stronger reversals in the future, in beliefs and in returns. The reverse holds for firms that overreact less than the average firm, $\theta_i - \theta v_i < 0$.[24]

We can examine whether this mechanism is at play by considering forecast errors. In our model the belief distortion for firm i (which is inversely related to the forecast error) is given by:

$$\epsilon_{i,t} = \left(\frac{\theta_i}{\theta}\right)\epsilon_t. \tag{12}$$

Excess optimism $\epsilon_{i,t}$ about firm i at time t is proportional to aggregate excess optimism ϵ_t, with a proportionality coefficient that increases in the extent θ_i to which beliefs about firm i overreact compared to beliefs about the market θ.

Denote by (v_H, θ_H) the exposure to fundamental risk and to belief overreaction of high LTG firms and by (v_L, θ_L) the exposures of low LTG ones. We obtain the following result.

Prediction 2 *The beliefs about $LLTG$ firms overreact to aggregate news less than those of HLTG firms, $\theta_L < \theta_H$, if and only if forecast errors in earnings growth for the PMO portfolio are positively predicted by the current revision ΔLTG_t and the lagged forecast LTG_{t-1}. If $\theta_L < \theta_H$ and in addition the two portfolios are similarly exposed to fundamental risk, $v_L \approx v_H$, the same LTG proxies predict a higher PMO spread.*

This prediction allows to assess whether high LTG firms overreact more than low LTG ones, $\theta_L < \theta_H$, by testing whether high LTG firms disappoint more than low LTG ones after periods of aggregate optimism. If so, then, Prediction 2 also offers a way to explain the result of Table 7: the PMO spread widens after periods of high aggregate optimism. As in the case of time series predictability, overreaction characterizes the joint behavior of returns and errors.

Table 8 studies forecast errors. Column (1) regresses the forecast errors for five years ahead earnings growth for the $LLTG$ portfolio on the current forecast revision ΔLTG_t and lagged forecast

---

[24] The presence of intangible news simply adds to Equation (11) a third factor capturing contemporaneous aggregate intangible news η_{t+1}. For simplicity, we omit this factor.

LTG_{t-1}. Column (2) does the same for forecast errors in the $HLTG$ portfolio, and column (3) for the PMO portfolio.

Table 8
Forecast Errors of LTG Portfolios

This table predicts forecast errors for portfolios formed on the basis of expected long-term growth in earnings for firm i, $LTG_{i,t}$ using beliefs about aggregate earnings growth. On each month between December 1982 and December 2015, we form decile portfolios based on $LTG_{i,t}$ and report regressions for the forecast errors in predicting earnings growth between t and $t+5$ of the following three portfolios: (a) the lowest decile ($LLTG$) in column [1], (b) the highest decile ($HLTG$) in column [2], and (c) the difference between the two ($PMO = LLTG - HLTG$) in column [3]. We define portfolio errors as the mean forecast error of the firms in the relevant LTG portfolio, i.e. the time t average difference between: (1) the annual growth in firm i's earnings per share between year t and $t+5$, $\Delta_5 e_{i,t+5}/5$, and (2) the expected long term growth in firm i's earnings, $LTG_{i,t}$. The independent variables are: (a) the one year forecast revision for aggregate earnings, ΔLTG_t, and (b) the lagged one-year forecast, LTG_{t-1}. Except ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)
	Dependent variable: Five-Year Forecast Error		
	LLTG	HLTG	PMO
ΔLTG_t	-0.3595[c]	-0.8597[a]	0.7970[a]
	(0.1905)	(0.1389)	(0.1217)
LTG_{t-1}	0.0911	-0.7493[a]	0.7910[a]
	(0.1687)	(0.0759)	(0.0891)
Observations	397	397	397
Adjusted R^2	6%	52%	55%

The results indeed point to stronger overreaction to aggregate news for $HLTG$ than for $LLTG$ firms, $\theta_H > \theta_L$. Higher aggregate forecast revisions ΔLTG_t predict belief disappointment in the $LLTG$ portfolio (column 1), but even stronger disappointment in the $HLTG$ portfolio (column 2). Likewise, higher lagged forecast LTG_{t-1} predicts disappointment for the $HLTG$ portfolio, but not for the $LLTG$ one. Due to these patterns, the PMO LTG portfolio exhibits systematically positive earnings growth surprises after periods of aggregate optimism, captured by the positive coefficients in column (3). These positive surprises reflect lower disappointment in the long arm of the portfolio, $LLTG$, compared to the short arm, $HLTG$.[25]

[25] In Table 8 we focus on predictability at the 5 year horizon to match Table 7. The results are robust to shorter horizons (Appendix C, Table C.2 panel B). BGLS (2019) document that average forecast errors of portfolios with high v low LTG at the one year horizon are positive. The results are also robust to the prediction of value-weighted, as opposed to equal weighted, forecast errors (Table C.3).

We can connect Tables 7 and 8 using our model. The positive predictability of PMO forecast errors in Table 7 points to excess pessimism about $LLTG$ firms compared to $HLTG$ ones in good times, $(\theta_H - \theta_L) > 0$. The positive predictability of PMO returns in Table 7 suggests that in the same good times $LLTG$ firms are undervalued compared to $HLTG$ ones, $(\theta_H - \theta_L) > \theta(v_H - v_L)$. The two conditions are met if the fundamental exposure of $HLTG$ is not much larger than that of $LLTG$ firms. In fact, the two conditions are identical if these firms are similarly exposed, $(v_H - v_L) \approx 0$. In this case, Tables 7 and 8 are two sides of the same coin.

Note also that in Table 7 the return of the PMO portfolio loads negatively on the market factor. Equation (11) accounts for this fact provided $\alpha(\theta_H - \theta_L) > (v_L - v_H)$. Similar fundamental exposure by high and low LTG firms, $(v_H - v_L) \approx 0$, guarantees this result as well.[26] Here we do not try to measure the exposures of $HLTG$ and of $LLTG$ firms to fundamental risk, but the message is clear: Differential overreaction of firms to aggregate news offers a parsimonious account of co-movement of forecast errors and returns in the cross section, even absent any differential exposure to aggregate risk. This approach is able to account for the PMO spread, and once again underscores the importance of using beliefs as predictors of returns.

4.2 LTG and the Fama-French risk factors

In a series of influential papers Fama and French (1993, 2015) show that, over and beyond the standard market factor, cross sectional spreads and return co-movement are to a large extent explained by other return factors constructed using firm characteristics such as book to market, size, profitability, and investment. The efficient markets explanation for these findings is that these factors reflect sources of risk to which firms are differentially exposed. Attempts to directly

[26] When $v_H - v_L \approx 0$, not only our model reconciles Tables 7 and 8, but the PMO spread is entirely due to overreaction. In this case, the contemporaneous market return in Table 6 captures the excess overreaction of $HLTG$ stocks to contemporaneous news, whereas the beginning of period LTG proxies capture the excess overreaction of the same firms to past news. Compared to $LLTG$ firms, contemporaneous overreaction drives up the return of $HLTG$ firms, while disappointment of past overreaction drives it down.

measure these risks have however proved elusive, leading some researchers to argue that these factors can at least in part capture relative under-valuation of stocks in the long arm of the factor-return portfolio due to systematic belief biases (Lakonishok, Shleifer, and Vishny 1994). Our previous analysis of the LTG spread suggests that aggregate expectations LTG_t can be regarded as a proxy for such systematic biases. This raises the question of whether this proxy, in line with the logic of Equation (11), can shed light on the Fama-French factors.

We conclude by showing that this connection may be promising. Table 9 below regresses the five-year returns (Panel A) and forecast errors (Panel B) of the Fama-French (2015) factor portfolios, including book to market (HML), profitability (RMW), investment (CMA), and size (SMB) on our measures of aggregate excess-optimism: the aggregate LTG_t revision and lagged forecast LTG_{t-1}. For returns, we also use the contemporaneous market return as a control.

Table 9
Predictability of factor returns and forecast errors

This table links beliefs about growth in earnings to Fama-French factor returns (Panel A) and forecast errors (Panel B). The dependent variables in Panel A are the compounded (log) return between year t and $t+5$ of the following 4 factors: (a) high-minus-low book-to market (HML) in column [1], (b) robust-minus-weak profitability factor (RMW) in column [2], (c) conservative-minus-aggressive investment (CMA) in column [3], and (d) small-minus-big factor (SMB) in column [4]. The dependent variables in Panel B are the forecast errors in predicting the growth in earnings between *t* and *t+5* for the: (1) HML, (2) RMW, (3) CMA, and (4) SMB portfolios. We define portfolio errors as the mean forecast error of the firms in the relevant *LTG* portfolio, i.e. the time *t* average difference between: (1) the annual growth in firm i's earnings per share between year t and $t+5$, $\Delta_5 e_{i,t+5}/5$, and (2) the expected long term growth in firm i's earnings, $LTG_{i,t}$.In Panel A, the independent variables are: (a) the one-year revision in aggregate earnings growth forecast, ΔLTG_t, (b) the one-year lagged forecast, LTG_{t-1}, (c) the (log) five-year return of CRSP's value-weighted index between t and $t+5$, $\ln(Mkt_{t,t+5})$. In Panel B, the independent variables are ΔLTG_t and LTG_{t-1}. Except ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A: Returns and forecasts about growth

	(1)	(2)	(3)	(4)
	Dependent Variable: Five-year (log) Return of			
	HML	RMW	CMA	SMB
ΔLTG_t	0.9881[a]	0.3311[c]	0.9283[a]	0.2985
	(0.1478)	(0.1738)	(0.2623)	(0.1901)
LTG_{t-1}	0.9150[a]	0.1925	0.6744[a]	0.5304[a]
	(0.1280)	(0.1815)	(0.1388)	(0.1059)
$\ln(Mkt_{t,t+5})$	0.4247[a]	-0.3954[b]	0.0354	-0.4051[a]
	(0.1369)	(0.1693)	(0.1662)	(0.1171)

Observations	397	397	397	397
Adjusted R^2	62%	30%	52%	59%

Panel B: Forecast errors and forecasts about growth

	(1)	(2)	(3)	(4)
	Dependent Variable: Five-year Forecast Error			
	HML	RMW	CMA	SMB
ΔLTG_t	0.2564[c]	0.0847	0.6449[a]	-0.3126[b]
	(0.1358)	(0.1141)	(0.1500)	(0.1329)
LTG_{t-1}	0.4508[a]	0.2922[a]	-0.0117	-0.2413[b]
	(0.1107)	(0.1040)	(0.0979)	(0.1056)
Observations	397	397	397	397
Adjusted R^2	16%	6%	14%	6%

The coefficients on the LTG proxies in the return regression are all positive, suggesting that part of the cross-sectional return differentials may indeed reflect undervaluation of the long arm of the portfolio during times of aggregate optimism (compared to the short arm). The explanatory power of LTG is high: using the market factor alone accounts for only 1% of the $HML\ BM$ return, 27% of the RMW return, 14% of the CMA return, and 40% of the SMB return (see Table C.6 in Appendix C). Aggregate optimism helps explain cross sectional co-movement.[27]

In line with the analysis of the PMO spread, we next ask whether co-movement is due to weaker belief overreaction for stocks in the long arm of the portfolio compared to those in the short arm. To address this question in Panel B we study forecast error predictability. Consider the $HML\ BM$ portfolio in column (1). Higher aggregate optimism predicts positive surprises (less belief disappointment) in long term earnings growth for high book to market stocks compared to low book to market ones. This points to weaker overreaction for high BM stocks compared to low BM ones, $\theta_{HBM} < \theta_{LBM}$. The undervaluation of high BM stocks during periods of aggregate

[27] The predictive power of LTG_t in Panel A is robust to a shorter, 3 year horizon (Appendix C, Table C.4) as well as to including proxies for required returns. In particular, in the spirit of Lettau and Ludvigson's (2001) we can include in the regression cay alone and cay interacted with the contemporaneous market return (Table C.5). This causes the LTG revision to become insignificant in the RMW regression, but modestly improves the regression R^2, which becomes 66% for HML, 36% for RMW (for which cay is itself insignificant), 66% for CMA, and 65% for SMB.

optimism in Panel A can thus reflect their weaker over-reaction. The mechanisms for the $PMO\ LTG$ and $HML\ BM$ return spreads are similar.

The same message holds for the RMW and CMA factors: columns (2) and (3) in panel B show that, after times of aggregate optimism, firms that are highly profitable and invest conservatively exhibit less belief disappointment than firms that are less profitable and invest aggressively, respectively. This is also consistent with the fact that, during the same times, profitable or conservative firms are relatively undervalued, as captured by columns (2) and (3) in Panel A. In terms of Proposition 2 and Equation (11), the weaker overreaction of the portfolios' long arms (Panel B) and their relative undervaluation (Panel A) can be jointly explained if the short arm of the portfolio is not much more exposed to fundamental risk than the long arm, as in the case of the $PMO\ LTG$ spread.[28]

The findings for the SMB factor are not as clear. In Panel B, small firms experience sharper belief disappointment than big firms, suggesting $\theta_S > \theta_B$, and yet they appear to be undervalued compared to big firms during times of excess optimism (column (4), Panel A). There is no direct connection between return and forecast error predictability for the size factor. Equation (11) is consistent with the results in Table 9 if small firms are sufficiently more exposed to fundamental risk than large firms $v_S > v_L$. Small firms may then be undervalued in good times because they display small overreaction compared to their market exposure $\theta_S - \theta v_S < \theta_L - \theta v_L$, and yet disappoint after good times because they display larger absolute overreaction $\theta_S > \theta_L$.[29] The SMB factor is not easily accommodated by our model.

Overall, our results bring together return predictability for the aggregate stock market and in the cross section in terms of a common mechanism of overreacting expectations. High aggregate LTG captures overvaluation of stocks, in the aggregate but particularly stocks with low book to

[28] According to (11), the non-negative loadings on the market factor for $HML\ BM$ and CMA in Panel A additionally require a sufficiently stronger exposure to fundamentals of the long arm of the portfolio compared to the short arm.
[29] Specifically, one needs that $0 < (\theta_S - \theta_B) < \max[\theta(v_S - v_B), (v_B - v_S)/\alpha]$.

market, low profitability, and aggressive investing, which overreact more to aggregate conditions. Such overvaluation leads to lower subsequent returns, both for the market and for firms with those characteristics. Over-optimism acts similarly to lower risk aversion, if the short arms of the factors were indeed riskier along the lines of Fama French (1993).[30] But the expectations approach goes farther: it explains that over-optimism arises in response to news, that forecast errors about factors are systematically predictable from aggregate optimism, and that those predictable errors in turn help explain factor returns.

5. Conclusion

Measured expectations of fundamentals throw new light on leading aggregate and cross-sectional stock market puzzles, even assuming that required returns are constant in the time series and in the cross section, and no price extrapolation. The main takeaway is that overreaction of expectations of long-term fundamentals unveils a common, parsimonious mechanism behind the anomalies. Good news cause investors to become too optimistic about long term fundamentals of the average firm or of particular firms. This inflates both the market and individual firm valuations, leading to predictably low returns in the future, in absolute terms or compared to other firms, as earnings expectations are disappointed. The mechanism is empirically confirmed by the joint predictability of returns and forecast errors, in both the aggregate market and in the cross section.

A skeptic may argue that measured long term expectations surreptitiously incorporate variation in discount rates. We consider this possibility, but do not find support for it. In particular,

[30] There is little evidence this is the case in terms of market beta and volatility, either for HML (LaPorta et al 1997) or for PMO (BGLS 2019). In Appendix C, we show that high beta stocks are also more exposed to aggregate overreaction, which helps explain their underperformance (Table C.7). However, recent work has shown that the standard risk factors load on stocks whose cashflows are relatively more concentrated in the short term, and for which long term growth expectations are also lower (Weber 2018, Gormsen and Lazarus 2021). Gormsen and Lazarus (2021) propose aversion to short term cash flow variation as a risk-based explanation for these factors' average returns. The fact that returns on the factors are (partially) linked to errors in long term growth forecasts, which are in turn predictable from fundamental aggregate shocks helps explain the negative correlation between market returns and factor returns documented in Gormsen (2021), under the unifying mechanism of overreacting expectations. We leave it to future work to evaluate in a systematic way the ability of overreacting beliefs to account for conventional cross sectional return anomalies.

beliefs about long term growth have remarkable predictive power for aggregate returns even when we control for leading proxies for required returns and for the price dividend ratio. At the firm level, these beliefs predict a firm's future return even after introducing time fixed effects, which controls for common shocks to required returns. Finally, revisions in measured beliefs are in good part driven by earnings news, and not by past stock returns or expected stock returns. These results further strengthen our overreacting expectations interpretation of the evidence.

In his well-known Asset Pricing book, Cochrane (2001) writes about the possibility that price movements may reflect irrational exuberance (Shiller 2000): "Perhaps, but is it just a coincidence that this exuberance comes at the top of an unprecedented economic expansion, a time when the average investor is surely feeling less risk averse than ever, and willing to hold stocks despite historically low risk premia?" At a most basic level, our analysis shows that this fact is not a coincidence, but obtains for a different reason: at the top of an unprecedented expansion the average investor is *more optimistic*, rather than less risk averse. Our analysis of LTG_t strongly supports this possibility, which is also confirmed by a growing body of evidence using survey expectations of corporate managers, professional forecasters, and individual investors (Bordalo et al. 2022). The data suggest that belief overreaction holds significant promise for explaining many macro-financial puzzles.

REFERENCES

Abel, Andrew, 1990, Asset prices under habit formation and catching up with the joneses, *American Economic Review Papers and Proceedings* 80, 38-42.

Adam, Klaus, Albert Marcet, and Johannes Beutel, 2017, Stock price booms and expected capital gains, *American Economic Review* 107, 2352-2408.

Adam, Klaus, and Stefan Nagel, 2022, Expectations data in asset pricing, Mimeo, Chicago Booth.

Afrouzi, Hassan, Spencer Kwon, Augustin Landier, Yueran Ma and David Thesmar, 2020, Overreaction and working memory, NBER w27947.

Alti, Aydogan, and Paul Tetlock, 2014, Biased beliefs, asset prices, and investment: a structural approach, *Journal of Finance* 69, 325-361.

Azeredo da Silveira, Rava, Yeji Sung, and Michael Woodford, 2020, Optimally imprecise memory and biased forecasts, working paper.

Bacchetta, Philippe, Elmar Mertens, and Eric van Wincoop, 2008, Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance* 28, 406-426.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481-1509.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2010, Long run risks, the macroeconomy, and asset prices, *American Economic Review* 100, 542-46.

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

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

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, Extrapolation and bubbles, *Journal of Financial Economics* 129, 203-227.

Barro, Robert, 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823-866.

Barsky, Robert, and J. Bradford DeLong, 1993, Why does the stock market fluctuate*, Quarterly Journal of Economics* 108, 291-312.

Ben-David, Itzhak, John Graham, and Campbell Harvey, 2013, Managerial miscalibration, *Quarterly Journal of Economics* 128, 1547-1584.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo, 2021, Implications of diagnostic expectations: theory and applications, NBER w28604.

Black, Fischer, 1986, Noise, *Journal of Finance* 41, 528-543.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2018, Diagnostic expectations and credit cycles, *Journal of Finance* 73, 199-227.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen Terry, 2021, Real credit cycles, *NBER* wp28416

Bordalo, Pedro, Nicola Gennaioli, Spencer Kwon, and Andrei Shleifer, 2020, Diagnostic bubbles, *Journal of Financial Economics*, 141, 1060 – 1077.

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

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2020, Expectations of fundamentals and stock market puzzles, *NBER* wp 27283.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748-2782.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, Frederik Schwerter, and Andrei Shleifer, 2020b, Memory and representativeness, *Psychological Review*, 128, 71-85.

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

Campbell, John, and John Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behaviour, *Journal of Political Economy* 107, 205-251.

Campbell, John, and Robert Shiller, 1987, Cointegration and tests of present value models, *Journal of Political Economy* 95, 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, 195-228.

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

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

Constantinides, George, 1990, Habit formation: a resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519-543.

Cummins, Jesse, Kevin Hassett, and Stephen Oliner, 2006, Investment behaviour, observable expectations, and internal funds, *American Economic Review* 96, 796 – 810.

Cutler, David, James Poterba, and Lawrence Summers, 1990, Speculative dynamics and the role of feedback traders, *American Economic Review* 80, 63–68.

Daniel, Kent, David Hirshleifer, and Avanidhar Subramanyam, 1998, Investor psychology and security market under- and over-reaction, *Journal of Finance* 53, 1839 – 1885.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, A., 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Finance* 56, 921-965.

Daniel, Kent and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

D'Arienzo, Daniele, 2019, Excess volatility from increasing overreaction, Working paper.

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

De la O, Ricardo, and Sean Myers, 2022, When do subjective expectations explain asset prices, *working paper.*

De Bondt, Werner, and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793 - 805.

De Bondt, Werner, and Richard Thaler, 1990, Do security analysts overreact? *American Economic Review* 80, 52-57.

De Long, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann, 1990a, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703-738.

De Long, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann, 1990b, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 379-395.

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

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

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

Gabaix, Xavier, 2012, Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645-700.

Gabaix, Xavier, 2019, Behavioral inattention. In *Handbook of Behavioral Economics: Applications and Foundations* 1 (Vol. 2), 261-343. North-Holland.

Giglio, Stefano, and Bryan Kelly, 2017, Excess volatility: Beyond discount rates, *Quarterly Journal of Economics* 133, 71-127.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021, Five facts about beliefs and portfolios, *American Economic Review* 111, 1481-1522.

Gilchrist, Simon and Egon Zakrajšek, 2012, Credit spreads and business cycle fluctuations. *American Economic Review* 102, 1692–1720.

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

Gormsen, Niels, and Eben Lazarus, 2021, Duration-driven returns, *Journal of Finance, forthcoming.*

Greenwood, Robin, and Samuel Hanson, 2015, Waves in ship prices and investment, *Quarterly Journal of Economics* 131, 55-109.

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

Guo, Hongye, and Jessica Wachter, 2020, 'Superstitious' Investors, NBER w25603.

Hillenbrand, Sebastian, and Odhrain McCarthy, 2022, Heterogeneous investors and stock market fluctuations, NYU, mimeo.

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

Huang, Lixin, and Hong Liu, 2007, Rational inattention and portfolio selection, *Journal of Finance* 62, 1999-2040.

Jagannathan, Ravi, and Binying Liu, 2019, Dividend dynamics, learning, and expected stock index returns, *Journal of Finance* 74, 401-448.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jin, Lawrence J., and Pengfei Sui, 2019, Asset pricing with return extrapolation, *Journal of Financial Economics* 145 (2), 273-295.

Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of present values, *Journal of Finance* 68, 1721-1756.

Kohlhas, Alexandre, and Ansgar Walther, 2021, Asymmetric attention, *American Economic Review* 111, 2879-2925.

Koijen, Ralph S J, and Stijn Van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467-491.

Kothari, S.P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time-series analysis, *Journal of Financial Economics* 44, 169–203.

Kothari, S. P., Jonathan Lewellen, and Jerold Warner, 2006, Stock returns, aggregate earnings surprises, and behavioral finance, *Journal of Financial Economics* 79, 537-568.

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

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

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

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

Lee, Charles, James N. Myers, and Bhaskaran Swaminathan, 1999, What is the intrinsic value of the Dow? *Journal of Finance* 54, 1693-1741.

Lee, Charles, and Bhaskaran Swaminathan, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017-2069.

LeRoy, Stephen, and Richard Porter, 1981, The present-value relation: Tests based on implied variance bounds, *Econometrica* 49, 555-574.

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

L'Huillier, Jean-Paul, Sanjay Singh, and Donghoon Yoo, 2021, Diagnostic expectations and macroeconomic volatility, working paper SSRN 3778231.

Mankiw, Gregory and Ricardo Reis, 2002, Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve, *Quarterly Journal of Economics* 117, 1295–1328.

Marsh, Terry, and Robert Merton, 1986, Dividend variability and variance bounds tests for the rationality of stock market prices, *American Economic Review* 76, 483-498.

Martin, Ian, 2017, What is the expected return on the market?, *Quarterly Journal of Economics* 132, 367–433.

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, A simple, positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Odean, Terrance, 1998, Volume, volatility, price, and profit when all traders are above average, *Journal of Finance* 53, 1887-1934.

Ortu, Fulvio, Andrea Tamoni, and Claudio Tebaldi, 2013, Long-run risk and the persistence of consumption shocks, *Review of Financial Studies* 26, 2876–2915.

Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam, 2020, Financial market risk perceptions and the macroeconomy, *Quarterly Journal of Economics* 135, 1443-1491.

Rietz, Thomas A, 1988, The equity risk premium: A solution, *Journal of Monetary Economics* 22, 117-131.

Shiller, Robert, 1981, The use of volatility measures in assessing market efficiency, *Journal of Finance* 36, 291-304.

Shiller, Robert, 2014, Speculative asset prices, *American Economic Review* 104, 1486-1517.

Sims, Christopher, 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–690.

Stambaugh, Robert, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.

Timmermann, Allan, 1993, How learning in financial markets generates excess volatility and predictability in stock prices, *Quarterly Journal of Economics* 108, 1135-1145.

Wachter, Jessica, 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance* 68, 987-1035.

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

Woodford, Michael, 2003, Imperfect common knowledge and the effects of monetary policy, In *Knowledge, Information, and Expectations in Modern Macroeconomics* (Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, eds.), Princeton University Press.

Appendix A. Proofs

Equivalence between Equation (6) and Diagnostic Expectations when $\boldsymbol{\theta} > \mathbf{0}$. In their internet Appendix, Bordalo et al. (2018) define a generalized slow-moving diagnostic distribution as:

$$f^{\theta}(g_{t+s}) \propto f(g_{t+s}|g_t, \eta_t) R_{s,t}^{\theta},$$

$$R_{s,t} = \prod\nolimits_{n\geq 1} \left[\frac{f(g_{t+s}|\mathbb{E}_{t+1-n}(g_t), \eta_{t+1-n})}{f(g_{t+s}|\mathbb{E}_{t-n}(g_t), \eta_{t-n})}\right]^{\gamma_n},$$

where $f(g_{t+s}|g_t, \eta_t)$ is the true conditional distribution, which depends on the current state g_t and news shock η_t, while $R_{s,t}$ is the representativeness of realization g_{t+s} at time t. In the standard memoryless DE model, $\gamma_1 = 1$ and $\gamma_n = 0$ for $n > 1$. A sluggish DE model features $\gamma_n > 0$ for some $n > 1$.

A special case of a sluggish DE model is one where $\gamma_n = \gamma^{n-1}$, where $\gamma \in (0,1)$ parameterizes the speed of the decay of overreaction to past news. In this case we can write:

$$R_{s,t} = \prod\nolimits_{n\geq 1} \left[\frac{f(g_{t+s}|\mathbb{E}_{t+1-n}(g_t), \eta_{t+1-n})}{f(g_{t+s}|\mathbb{E}_{t-n}(g_t), \eta_{t-n})}\right]^{\gamma^{n-1}}$$

In this case, applying the BGS (2018) formula, the diagnostic expectation is equal to:

$$\mathbb{E}_t^{\theta}(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \theta \sum\nolimits_{n>1} \gamma^{n-1} [\mathbb{E}_{t+1-n}(g_{t+s}) - \mathbb{E}_{t-n}(g_{t+s})].$$

We now show that this equation is equivalent to (6) for a suitable choice of the parameter γ.

Iterating backward the belief distortion ϵ_t, we can write Equation (6) as:

$$\widetilde{\mathbb{E}}_t(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \mu^{s-1} \sum\nolimits_{n>1} \rho^{n-1} u_{t+1-n}.$$

Plugging the expression for the expectations shock we obtain:

$$\widetilde{\mathbb{E}}_t(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \theta\mu^{s-1} \sum\nolimits_{n>1} \rho^{n-1} (\mu\tau_{t+1-n} + \eta_{t+1-n}).$$

In the AR(1) process of Equation (3), the shock $(\mu\tau_{t+1-n} + \eta_{t+1-n})$ is equivalent to the expectation revision for next period dividend growth, namely $(\mu\tau_{t+1-n} + \eta_{t+1-n}) =$

$\mathbb{E}_{t+1-n}(g_{t+2-n}) - \mathbb{E}_{t-n}(g_{t+2-n})$. In turn, Equation (3) implies that $\mathbb{E}_{t+1-n}(g_{t+1}) = \mu^{n-1}\mathbb{E}_{t+1-n}(g_{t+2-n})$ and $\mathbb{E}_{t-n}(g_{t+1}) = \mu^{n-1}\mathbb{E}_{t-n}(g_{t+2-n})$. As a result, we can write:

$$\widetilde{\mathbb{E}}_t(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \theta\mu^{s-1}\sum_{n>1}\left(\frac{\rho}{\mu}\right)^{n-1}[\mathbb{E}_{t+1-n}(g_{t+1}) - \mathbb{E}_{t-n}(g_{t+1})].$$

Equation (3) also implies that $\mathbb{E}_{t-k}(g_{t+s}) = \mu^{s-1}\mathbb{E}_{t-k}(g_{t+1})$. In turn, this implies that:

$$\widetilde{\mathbb{E}}_t(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \theta\sum_{n>1}\left(\frac{\rho}{\mu}\right)^{n-1}[\mathbb{E}_{t+1-n}(g_{t+s}) - \mathbb{E}_{t-n}(g_{t+s})].$$

As a result, for $\theta > 0$, $\widetilde{\mathbb{E}}_t(g_{t+s})$ in Equation (6) is equivalent to a slow-moving diagnostic expectation with decay parameter $\gamma = \rho/\mu$.

Proof of Proposition 1. The goal is to compute the regression coefficients β_1 and β_2 for Equation (7) and to characterize their sign as a function of θ. Using Equation (6), the forecast error and the forecast revision are equal to:

$$\begin{aligned} g_{t+s} - \widetilde{\mathbb{E}}_t(g_{t+s}) &= -\mu^{s-1}\epsilon_t, \\ \widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s}) &= \mu^{s-1}[(1+\theta)(\mu\tau_t + \eta_t) - (\mu-\rho)\epsilon_{t-1}]. \end{aligned}$$

Using these expressions, after some algebra we obtain the following moments:

$$\begin{aligned}
\gamma_{11} &\equiv \mathrm{Var}\left[\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s})\right] \\
&\qquad = \mu^{2(s-1)}\sigma_v^2\left[(1+\theta)^2 + (\mu-\rho)^2\frac{\theta^2}{(1-\rho^2)}\right]; \\
\gamma_{21} &\equiv \mathrm{Cov}\left[\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s}), \widetilde{\mathbb{E}}_{t-1}(g_{t+s})\right] \\
&\qquad = -\mu^{2s-1}(\mu-\rho)\theta\sigma_v^2\left[\frac{1}{1-\mu\rho} + \frac{\theta}{(1-\rho^2)}\right]; \\
\gamma_{22} &\equiv \mathrm{Var}\left[\widetilde{\mathbb{E}}_{t-1}(g_{t+s})\right] \\
&\qquad = \mu^{2s}\sigma_v^2\left[\frac{1}{1-\mu^2} + \frac{2\theta}{1-\mu\rho} + \frac{\theta^2}{1-\rho^2}\right]; \\
\gamma_{Y1} &\equiv \mathrm{Cov}\left[\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s}), g_{t+s} - \widetilde{\mathbb{E}}_t(g_{t+s})\right] \\
&\qquad = -\mu^{2(s-1)}\theta\sigma_v^2\left[1 + \theta\left(1 - \rho\frac{(\mu-\rho)}{(1-\rho^2)}\right)\right]; \\
\gamma_{Y2} &\equiv \mathrm{Cov}\left[\widetilde{\mathbb{E}}_{t-1}(g_{t+s}), g_{t+s} - \widetilde{\mathbb{E}}_t(g_{t+s})\right] \\
&\qquad = -\mu^{2s-1}\rho\theta\sigma_v^2\left[1 + \frac{\theta}{1-\rho^2}\right],
\end{aligned}$$

where $\sigma_v^2 = \left(\mu^2\sigma_\tau^2 + \sigma_\eta^2\right)$ is the variance of the fundamental shock. Then we have:

$$\beta_1 = \frac{\gamma_{22}\gamma_{1Y} - \gamma_{12}\gamma_{2Y}}{\gamma_{11}\gamma_{22} - \gamma_{21}^2}, \quad \beta_2 = \frac{-\gamma_{12}\gamma_{1Y} + \gamma_{11}\gamma_{2Y}}{\gamma_{11}\gamma_{22} - \gamma_{21}^2}$$

The sign of β_1 is equal to the sign of $\gamma_{22}\gamma_{1Y} - \gamma_{12}\gamma_{2Y}$, where

$$\gamma_{22}\gamma_{1Y} - \gamma_{12}\gamma_{2Y} = -\mu^{4s-2}\sigma_v^4 \times \theta G(\theta),$$

where

$$\begin{aligned} G(\theta) \quad &= \left((1+\theta) - \rho(\mu-\rho)\frac{\theta}{1-\rho^2}\right) \times \left(\frac{1}{1-\mu^2} + \frac{2\theta}{1-\mu\rho} + \frac{\theta^2}{1-\rho^2}\right) \\ &+\theta\rho(\mu-\rho)\left(1 + \frac{\theta}{1-\rho^2}\right) \times \left(\frac{1}{1-\mu\rho} + \frac{\theta}{1-\rho^2}\right). \end{aligned}$$

Note that the condition $\beta_1 < 0$ is equivalent to $\theta > 0$ if and only if $G(\theta) > 0$ for all θ.

We now show that $G(\theta) > 0$ for all θ. We proceed in two steps. First, we show $G(\theta) > 0$ for all $\theta \in [-1, -(1-\rho^2)]$. Next, we show $G'(\theta) > 0$ for all $\theta \geq -(1-\rho^2)$, which guarantees that $G(\theta) > 0$ in this range given that $G\left(-(1-\rho^2)\right) > 0$. For the first step, note that $G(\theta) > 0$ can be equivalently written as:

$$\rho(\mu-\rho)\theta - (1-\rho^2)(1+\theta) < \rho(\mu-\rho) \cdot \theta \cdot \frac{(1-\rho^2+\theta)\left(\frac{1}{(1-\mu\rho)} + \frac{\theta}{(1-\rho^2)}\right)}{\left(\frac{1}{1-\mu^2} + \frac{\theta^2}{(1-\rho^2)} + \frac{2\theta}{(1-\mu\rho)}\right)}.$$

When $\theta < 0$ the LHS is always negative, since it is a sum of negative terms. On the other hand, if $\theta \leq -(1-\rho^2)$ the RHS is greater or equal than zero. This is due to the facts that: i) $\rho(\mu-\rho) > 0$, ii) it can be easily verified that $\frac{1}{(1-\mu\rho)} + \frac{\theta}{(1-\rho^2)} > 0$, iii) in the range $\theta \leq -(1-\rho^2)$ the terms θ and $(1-\rho^2+\theta)$ are non-positive so their product is non-negative, and iv) the denominator is positive. Point iv) follows from the fact that the sign of the denominator is equal to the sign of the following quadratic polynomial (letting $\bar{\theta} \equiv -\theta > 0$):

$$(1-\mu^2)(1-\mu\rho)\bar{\theta}^2 - 2(1-\mu^2)(1-\rho^2)\bar{\theta} + (1-\rho^2)(1-\mu\rho),$$

whose determinant is

$$\Delta \quad = 4(1-\mu^2)(1-\rho^2)[-(\mu-\rho)^2] \leq 0,$$

which is negative when $\rho < \mu$ and zero when $\rho = \mu$. Noting that for $\theta = 0$ the polynomial is positive, it follows that the polynomial is positive for all θ. So, we have shown that $G(\theta) > 0$ for all $\theta \leq -(1-\rho^2)$.

Consider now the second step of the proof, which is to show that $G'(\theta) \geq 0$ for all $\theta \geq -(1-\rho^2)$. First, write $G(\theta)$ as follows:

$$\begin{aligned} G(\theta) \quad &= \left(\frac{1}{1-\mu^2} + \frac{\theta^2}{(1-\rho^2)} + \frac{2\theta}{(1-\mu\rho)}\right) \\ &\quad +\theta\left(\frac{1-\mu\rho}{1-\rho^2}\right)\left(\frac{1}{1-\mu^2} + \frac{\theta^2}{1-\rho^2} + \frac{2\theta}{1-\mu\rho}\right) \\ &\quad +\theta\rho(\mu-\rho)\left[\frac{1}{1-\mu\rho} + \frac{\theta^2}{(1-\rho^2)^2} + \frac{\theta}{1-\rho^2}\left(1 + \frac{1}{1-\mu\rho}\right)\right]. \end{aligned}$$

Note, then, that $G(\theta)$ has the structure $G(\theta) = f(\theta) + \theta h(\theta)$, so that

$$\begin{aligned} G'(\theta) \quad &= f'(\theta) + h(\theta) + \theta h'(\theta) \\ &= [f'(\theta) - h'(\theta)] + h(\theta) + (1+\theta)h'(\theta). \end{aligned}$$

We now show that for $\theta \geq -(1-\rho^2)$ the following sufficient condition for $G'(\theta) \geq 0$ holds: (a) $h'(\theta) \geq 0$; (b) $f'(\theta) - h'(\theta) \geq 0$; and (c) $h(\theta) \geq 0$. Consider first that:

$$h'(\theta) \quad = \left[1 - \frac{\rho(\mu-\rho)}{1-\rho^2}\right]\left(\frac{2\theta}{1-\rho^2} + \frac{2}{1-\mu\rho}\right) + \frac{\rho(\mu-\rho)}{1-\rho^2} \cdot \left(\frac{2\theta}{1-\rho^2} + 1 + \frac{1}{1-\mu\rho}\right) \geq 0,$$

where the inequality holds because we have the sum of two positive terms. The second term is positive because it is positive at the lowest admissible value $\theta = -(1-\rho^2)$. Next note that:

$$f'(\theta) - h'(\theta) = \frac{\rho(\mu-\rho)}{1-\rho^2} \cdot \left(\frac{2\theta}{1-\rho^2} + \frac{2}{1-\mu\rho}\right) - \frac{\rho(\mu-\rho)}{1-\rho^2} \cdot \left(1 + \frac{1}{1-\mu\rho} + \frac{2\theta}{1-\rho^2}\right)$$

$$= \frac{\rho(\mu-\rho)}{1-\rho^2} \cdot \left(\frac{1}{1-\mu\rho} - 1\right) \geq 0.$$

Finally, it is easy to show that $h(\theta) \geq 0$ when $\theta \geq -(1-\rho^2)$. Since $h'(\theta) \geq 0$ for all $\theta \geq -(1-\rho^2)$, it suffices to show that $h\big(-(1-\rho^2)\big) \geq 0$, which is easily verified. This completes the proof that $G(\theta) \geq 0$ for all admissible θ, so $\beta_1 < 0$ if and only if $\theta > 0$.

We conclude by showing that $\beta_2 < 0$ if $\theta > 0$. The sign of β_2 is equal to the sign of $-\gamma_{12}\gamma_{1Y} + \gamma_{11}\gamma_{2Y}$, where

$$-\gamma_{12}\gamma_{1Y} + \gamma_{11}\gamma_{2Y} = -\mu^{4(s-1)}\sigma_v^4\theta Q(\theta),$$

where

$$\begin{aligned} Q(\theta) \quad &= (\mu-\rho)\theta\left(\frac{1}{1-\mu\rho} + \frac{\theta}{1-\rho^2}\right) \times \left(1 + \theta\frac{1-\mu\rho}{1-\rho^2}\right) \\ &+\mu\rho\left(1 + \frac{\theta}{1-\rho^2}\right) \times \left(1 + \theta^2 + (\mu-\rho)^2\frac{\theta^2}{1-\rho^2}\right). \end{aligned}$$

The desired claim holds if $Q(\theta) > 0$ for all $\theta > 0$. But this is easily seen to be true.

Proof of Proposition 2. Equation (1) can be rewritten as:

$$r_{t+1} = \alpha(p_{t+1} - d_{t+1}) + g_{t+1} - (p_t - d_t) + k.$$

Plugging in this equation the expressions for $p_{t+1} - d_{t+1}$ and $p_t - d_t$ derived from Equation (3) we obtain:

$$r_{t+1} = r + \alpha\sum_{s\geq 0}\alpha^s\widetilde{\mathbb{E}}_{t+1}(g_{t+2+s}) + g_{t+1} - \sum_{s\geq 0}\alpha^s\widetilde{\mathbb{E}}_t(g_{t+1+s}),$$

which is in turn equivalent to:

$$r_{t+1} = r + \sum_{s \geq 1} \alpha^s \big[\tilde{\mathbb{E}}_{t+1}(g_{t+1+s}) - \tilde{\mathbb{E}}_t(g_{t+1+s})\big] + \big[g_{t+1} - \tilde{\mathbb{E}}_t(g_{t+1})\big].$$

Using Equation (6), the forecast revision is equal to:

$$\tilde{\mathbb{E}}_{t+1}(g_{t+1+s}) - \tilde{\mathbb{E}}_t(g_{t+1+s}) = \mathbb{E}_{t+1}(g_{t+1+s}) - \mathbb{E}_t(g_{t+1+s}) + \mu^{s-1}(\epsilon_{t+1} - \mu\epsilon_t),$$

which is in turn equal to:

$$\tilde{\mathbb{E}}_{t+1}(g_{t+1+s}) - \tilde{\mathbb{E}}_t(g_{t+1+s}) = \mu^{s-1}(\mu\tau_{t+1} + \eta_{t+1})(1+\theta) - \mu^{s-1}(\mu - \rho)\epsilon_t.$$

Plugging this expression in the equation for returns and using the definition of the current forecast error we obtain:

$$r_{t+1} = r + \frac{\alpha}{1-\alpha\mu}[(\mu\tau_{t+1} + \eta_{t+1})(1+\theta) - (\mu-\rho)\epsilon_t] + (\tau_{t+1} - \epsilon_t),$$

which can be rewritten as:

$$r_{t+1} = r + \frac{1+\alpha\mu\theta}{1-\alpha\mu}\tau_{t+1} + \frac{\alpha(1+\theta)}{1-\alpha\mu}\eta_{t+1} - \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\epsilon_t,$$

which proves the proposition.

Firm level analysis, including stock returns in the Cross Section

There are many firms i, each of which exhibits AR(1) dividend growth $g_{i,t+1} = \mu g_{i,t} + v_{i,t}$, with firm level fundamental shock $v_{i,t} = v_i * v_t + \varphi_{it}$, where φ_{it} is a firm specific Gaussian shock uncorrelated with v_t, and the firm specific belief distortion follows $\epsilon_{i,t} = \rho\epsilon_{i,t-1} + u_{i,t}$, with firm level expectations shock $u_{i,t} = \theta_i * (v_t + \varphi_{it})$. Note that, given these assumptions, we can write $\epsilon_{i,t} = \left(\frac{\theta_i}{\theta}\right)\epsilon_t + \varepsilon_{it}$, where ε_{it} is the idiosyncratic component of the distortion.

In analogy with the aggregate market return, the realized return on firm i is equal to:

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

The forecast revision is equal to:

$$\widetilde{\mathbb{E}}_{t+1}(g_{i,t+1+s}) - \widetilde{\mathbb{E}}_t(g_{i,t+1+s}) = \mu^s v_{i,t+1} + \mu^s\theta(v_{t+1} + \varphi_{i,t+1}) - \mu^{s-1}(\mu - \rho)\epsilon_{i,t}.$$

Plugging this expression in the equation for returns and using the definition of the current forecast error we obtain:

$$r_{i,t+1} = r_i + \frac{\alpha}{1-\alpha\mu}\left[\mu v_{i,t+1} + \mu\theta_i(v_t + \varphi_{it}) - (\mu - \rho)\epsilon_{i,t}\right] + (v_{i,t+1} - \epsilon_{i,t}),$$

which can be rewritten as:

$$r_{i,t+1} = r_i + \frac{v_i + \alpha\mu\theta_i}{1-\alpha\mu} v_{t+1} + \frac{\alpha\mu(1+\theta_i)}{1-\alpha\mu}\varphi_{it+1} - \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\left(\frac{\theta_i}{\theta}\right)\epsilon_t - \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\varepsilon_{it}.$$

By substituting for the shock v_{t+1} from the aggregate market return equation we obtain:

$$r_{i,t+1} - r_i = \frac{v_i + \alpha\mu\theta_i}{1+\alpha\mu\theta}(r_{t+1} - r) - \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\left(\frac{\theta_i - v_i\theta}{\theta + \alpha\mu\theta^2}\right)\epsilon_t + \frac{\alpha\mu(1+\theta_i)}{1-\alpha\mu}\varphi_{it+1}$$
$$- \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\varepsilon_{it}$$

which is equation (11) with the addition of the two idiosyncratic fundamental and belief distortion components.

Firm Level Required Return and the CAPM

Consider the perceived distribution of the stock return on firm i by the investor. From previous analysis, the realized return is equal to:

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

Investors' believed distribution of a generic $g_{i,t+s}$ at time t is rational up to the shift $\mu^{s-1}\epsilon_{i,t}$. Assume that investors are unsophisticated about their belief distortion. Specifically, they think that the belief distortion $\mu^{s-1}\epsilon_{i,t}$ they hold at time t with respect to $g_{i,t+s}$ will be unchanged at $t+1$. As a result, investors' believed distribution of the forecast revision $\widetilde{\mathbb{E}}_{t+1}(g_{i,t+1+s}) - \widetilde{\mathbb{E}}_t(g_{i,t+1+s})$ is equal to the rational distribution. Unsophistication also implies that investors believe that next period dividend growth $g_{i,t+1}$ will be upward shifted by $\epsilon_{i,t}$. As a result, they perceive their forecast error to be distributed according to the rational distribution.

Formally, investors' time t perception of the realized return at $t+1$ is equal to:

$$r_{i,t+1} = r_i + \frac{v_{i,t+1}}{1-\alpha\mu}.$$

Given the dependency of the firm level shock on the aggregate shock we have (assuming $\varphi_{it} = 0$):

$$r_{i,t+1} = r_i + \frac{v_i}{1-\alpha\mu} v_t,$$

which can be compared to the aggregate market return:

$$r_{i,t+1} = r + \frac{1}{1-\alpha\mu} v_t.$$

If investors have mean variance preferences, and the risk free rate is r_f, the security market line is:

$$r_i - r_f = \frac{cov(r_{i,t+1}, r_{t+1})}{var(r_{t+1})}(r - r_f),$$

which implies $r_i - r_f = v_i(r - r_f)$.

ONLINE APPENDIX

Appendix B. Further Results on Return Predictability and Aggregate Overreaction

In this Appendix, we collect several results that complement the analysis of expectations-driven return predictability on the aggregate market (Sections 2 and 3.2), specifically assessing the robustness of results in Tables 3, 4, 5, and 6.

B.1 Predictability of aggregate returns. We first reproduce Table 1, which examines the predictability of returns on the basis of expectations, using excess (as opposed to raw) returns. As in Table 1, LTG predicts returns but short term expectations do not.

Table B.1

This table examines the association between excess returns and forecasts for growth in earnings at different horizons. In panels A through D, the dependent variables is the (log) one-year excess return in column [1] and the discounted value of the cumulative 3- and 5-year excess return in columns [2] and [3], respectively. Excess returns are defined as the difference between (log) returns and the 90-day T-bill rate and are computed using monthly data starting with period $t + 1/12$. In Panels E through G, the dependent variables is the one-year excess raw return in column [1] and the discounted value of the cumulative 3- and 5-year raw return in columns [2] and [3], respectively. Excess raw returns $xRet_{t+j}$are the compounded raw returns during the period $t + 1/12$ through $t + j$ minus the compounded return on the risk free asset over the same period. The independent variables are: (a) the forecast for earnings growth in the long run LTG_t, (b) the difference between LTG_t and the forecast for CPI inflation in year $t+1$ by the Survey of Professional Forecasters, $\mathbb{E}_t^O$ [π_{t+1}], (c) the forecast for one-year growth in aggregate earnings between year t+1 and $t + 2$, $\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$, and (d) the difference between $\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$ and, $\mathbb{E}_t^O$ [π_{t+1}]. Variables are normalized to have zero mean and standard deviation of 1. Intercepts are not shown. The sample period is 1981:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (number of lags ranges from 12 in the first column to 60 in the last one). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

	(1)	(2)	(3)
	$xret_{t+1}$	$\sum_{j=1}^{3} \alpha^{j-1} xret_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} xret_{t+j}$
Panel A: Returns and LTG			
LTG_t	-0.3476[a]	-0.5624[a]	-0.5973[a]
	(0.1074)	(0.0916)	(0.1110)
Observations	409	409	409
Adjusted R^2	13%	31%	35%

Panel B: Returns and LTG net of expected inflation			
$LTG_t - \mathbb{E}_t^O[\pi_{t+1}]$	-0.2975[b]	-0.4534[a]	-0.4851[a]
	(0.1318)	(0.1732)	(0.1690)
Observations	409	409	409
Adjusted R^2	9%	20%	23%
Panel C: Returns and growth forecast for year 2			
$\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$	-0.0710	0.0230	0.1979
	(0.1186)	(0.2019)	(0.2176)
Observations	404	404	404
Adjusted R^2	0%	0%	4%
Panel D: Returns and growth forecast for year 2 net of expected inflation			
$\mathbb{E}_t^O[e_{t+2} - e_{t+1} - \pi_{t+1}]$	-0.0755	0.0244	0.1869
	(0.1233)	(0.1901)	(0.2135)
Observations	404	404	404
Adjusted R^2	0%	0%	3%

	Ret_{t+1}	$xRet_{t+3}$	$xRet_{t+5}$
Panel E: Raw returns and LTG			
LTG_t	-0.2591[a]	-0.4143[a]	-0.4648[a]
	(0.0860)	(0.0821)	(0.1092)
Observations	409	409	409
Adj R^2	11%	22%	22%
Panel F: Raw returns and growth forecast for year 1			
$\mathbb{E}_t^O[e_{t+1} - e_t]$	0.0288	0.0937	0.2392[a]
	(0.0950)	(0.0810)	(0.0565)
Observations	404	404	404
Adj R^2	0%	1%	6%
Panel G: Raw returns and growth forecast for year 2			
$\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$	-0.0390	0.0407	0.2524
	(0.0939)	(0.1648)	(0.1758)
Observations	404	404	404
Adj R^2	0%	0%	6%

We next extend the analysis of Table 2 to shorter horizons (1 and 3 years), and include several measures of required returns.

Table B.2
LTG and Return Predictability at Different Horizons

This table examines the association between realized returns and ex-ante proxies for required returns as well as macroeconomic predictors of returns. The dependent variable is the one-year return in Panel A and the discounted value of the cumulative return between year t and $t+3$ in Panel B. Returns are computed using monthly data starting with period $t+1/12$. In addition to the forecast for earnings growth in the long run LTG_t, the independent variables are: (1) the price dividend ratio, (2) the price earnings ratio, (3) forecast for one-year ahead earnings growth, $\mathbb{E}_t^O[e_{t+1}-e_t]$, (4) the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, (5) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, (6) the Martin (2013) expected return on the market, $SVIX_t^2$, (7) the term spread defined as the log difference between the gross yield of 10-year and 1-year US government bonds from the St. Louis Fed, (8) the credit spread defined as the log difference between the gross yield of BAA and AAA bonds from the St. Louis Fed, (9) the Baker et al. (2016) economic policy uncertainty index, (10) the Kelly and Pruitt (2013) optimal forecast of aggregate equity market returns, (11) the forecast for CPI inflation in year $t+1$ by the Survey of Professional Forecasters, $\mathbb{E}_t^O$ [π_{t+1}], and (12) the Nagel and Xu (2021) experienced dividend growth from, $expd_t$. The sample period is 1981:12-2015:12. Data is quarterly in column [2] and monthly elsewhere. All variables are normalized to have zero mean and standard deviation of 1. Intercepts are not shown. The sample period is 1981:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A: Predictability of one-year returns

Dependent Variable: One-year Return, r_{t+1}

Panel A1: LTG and proxies for time-varying returns

	(1)	(2)	(3)	(4)	(5)	(6)
LTG_t	-0.1658	-0.2566[b]	-0.3119[a]	-0.3421[b]	-0.3755[a]	-0.3193[a]
	(0.1113)	(0.1224)	(0.1201)	(0.1369)	(0.1376)	(0.1108)
X_t	-0.2935[a]	-0.2185	-0.0967	0.1065	0.2395[c]	0.1504
	(0.1115)	(0.1428)	(0.1155)	(0.1422)	(0.1390)	(0.1214)
Observations	409	409	404	409	137	193
Adjusted R^2	15%	13%	9%	10%	14%	12%
X_t	pd_t	pe_t	$\mathbb{E}_t^O[e_{t+1}-e_t]$	spc_t	cay_t	$SVIX_t^2$

Panel A2: LTG and other predictors of stock returns

	(1)	(2)	(3)	(4)	(5)	(6)
LTG_t	-0.3743[a]	-0.2891[b]	-0.3228[a]	-0.2952	-0.3020[a]	-0.2522[c]
	(0.1270)	(0.1229)	(0.1175)	(0.2586)	(0.1135)	(0.1503)
X_t	-0.1727	0.0675	0.0176	0.0448	0.2028[c]	-0.2489
	(0.1204)	(0.1295)	(0.0904)	(0.2374)	(0.1068)	(0.1774)
Observations	409	409	372	13400%	409	137
Adjusted R^2	11%	9%	11%	9%	13%	14%
X_t	Term $Spread_t$	Credit $Spread_t$	Uncertainty $Index_t$	Kelly Pruitt MRP_t	1-Yr Fcast Inflation	$expd_t$

Panel B: Predictability of three-year returns

Dependent Variable: Three-year Return $\sum_{j=1}^{3} \alpha^{j-1} r_{t+1}$

	Panel B1: LTG and proxies for time-varying returns					
	(1)	(2)	(3)	(4)	(5)	(6)
LTG_t	-0.3019[a]	-0.4476[a]	-0.4977[a]	-0.5156[a]	-0.5811[a]	-0.4770[a]
	(0.1016)	(0.1043)	(0.1180)	(0.1183)	(0.1449)	(0.1306)
X_t	-0.4104[a]	-0.2147[c]	-0.0273	0.0623	0.3234	0.1662
	(0.1204)	(0.1135)	(0.0880)	(0.0872)	(0.1990)	(0.2509)
Observations	409	409	404	409	13700%	193
Adjusted R^2	37%	28%	24%	24%	33%	28%
X_t	pd_t	pe_t	$E_t[e_{t+1} - e_t]$	spc_t	cay_t	$SVIX_t^2$

	Panel B2: LTG and other predictors of stock returns					
	(1)	(2)	(3)	(4)	(5)	(6)
LTG_t	-0.4356[a]	-0.4826[a]	-0.4862[a]	-0.6561[a]	-0.4923[a]	-0.3676[a]
	(0.1547)	(0.1244)	(0.1134)	(0.1216)	(0.0912)	(0.0989)
X_t	0.1342	0.0502	0.1201	-0.1646	0.2877[b]	-0.5074[a]
	(0.2061)	(0.1728)	(0.2059)	(0.1455)	(0.1133)	(0.1237)
Observations	409	409	372	134	409	137
Adjusted R^2	25%	24%	30%	25%	32%	47%
X_t	Term $Spread_t$	Credit $Spread_t$	Uncertainty $Index_t$	Kelly Pruitt MRP_t	1-Yr Fcast Inflation	$expd_t$

Finally, we confirm that aggregate returns at long horizons are predictable from recent earnings growth, in line with the reduced form analysis of Proposition 2, and consistent with findings by Nagel and Xu (2019).

Table B.3
Return predictability from recent earnings growth

This table examines the association between returns and past growth in earnings. The dependent variables are the (log) one-year return in column [1] and the discounted value of the cumulative return between year t and $t + 3$ in column [2] and between t and $t + 5$ in column [3]. Returns are computed using monthly data starting with period $t + 1/12$. The independent variable is the log of earnings for the S&P500 in year t relative to cyclically-adjusted earnings in year $t - 5$, $e_t - cae_{t-5}$. Variables are normalized to have zero mean and standard deviation of 1. Intercepts are not shown. The sample period is 1981:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (number of lags ranges from 12 in the first column to 60 in the last one). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, [c] significant at the 10% level.

(1) (2) (3)

	r_{t+1}	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
$e_t - cae_{t-5}$	-0.0758	-0.1999	-0.3384[b]
	(0.0830)	(0.1413)	(0.1355)
Observations	409	409	409
Adjusted R^2	1%	5%	13%

B.2 Expectation Revisions, Forecast Errors, and Returns

Table B.4 below extends the results of Table 3, on the determinants of LTG revisions, by considering other measures of required returns. In column 5 we include lagged returns as a proxy for expectations of future returns (see Table 3, Column 2).

Table B.4
Determinants of LTG revisions

This table examines the association between one-year changes in the forecast for growth in the long run and ex-ante proxies for required returns. The dependent variable is the change in the forecast for growth in earnings in the long run LTG_t between year t and $t-1$, ΔLTG_t The independent variables are: (a) the one-year lagged value of LTG_t, (b) log of earnings for the S&P500 in year t relative to cyclically-adjusted earnings in year $t-5$, $e_t - cae_{t-5}$, (c) the term spread defined as the log difference between the gross yield of 10-year and 1-year US government bonds from the St. Louis Fed in column [1], (d) the credit spread defined as the log difference between the gross yield of BAA and AAA bonds from the St. Louis Fed in column [2], (e) the Baker et al. (2016) economic policy uncertainty index in column [3], (f) the Kelly and Pruitt (2013) optimal forecast of aggregate equity market returns in column [4], and (g) the lagged 5-year return on the market in column [5]. The sample period is 1981:12-2015:12. All variables are normalized to have zero mean and standard deviation of 1. Intercepts are not shown. The sample period is 1981:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: ΔLTG_t				
LTG_{t-1}	-0.2903[b]	-0.2746[b]	-0.4187[a]	-0.5720[a]	-0.4716[a]
	(0.1298)	(0.1351)	(0.1474)	(0.0935)	(0.1471)
e_t - cae_{t-5}	0.2400[a]	0.3199[a]	0.3363[a]	0.1544[b]	0.2804[a]
	(0.0505)	(0.0643)	(0.0737)	(0.0608)	(0.0628)
X_t	-0.2642[b]	-0.2079[a]	-0.2413[b]	-0.5268[a]	0.2967[a]
	(0.1138)	(0.0611)	(0.1102)	(0.1004)	(0.0957)
Observations	433	441	432	130	457
Adjusted R^2	36%	35%	36%	66%	38%
X_t	*Term Spread$_t$*	*Credit Spread$_t$*	*Uncertainty Index$_t$*	*Kelly Pruitt MRP$_t$*	*Lagged 5yr return*

We next assess the robustness of Table 4. First, Table B.5 extends it to the 3 year horizon.

Table B.5
Predictability of Forecast Errors and Returns

This table links aggregate forecast errors and market returns. We report regressions using as dependent variable the error in forecasting three-year growth in aggregate earnings in column [1] and the discounted value of the cumulative market return between year t and $t+3$ in columns [2] and [3]. Three-year cumulative market returns (r_{t+3}) are computed using monthly data and run from $t+1/12$ through $t+36/12$. We define the forecast error as the difference between (a) the annual growth in earnings per share between year t and $t+3$, $\Delta_3 e_{t+3}/3$, and (b) the expected long term growth in earnings, LTG_t. The independent variables are the one-year change in LTG_t, ΔLTG_t, the lagged forecast, LTG_{t-1}, and the predicted forecast error, $\Delta_3 e_{t+3}/3 - LTG_t$. We assume that earnings are reported with a with a 3-month lag (i.e. we define e_t as earnings for the calendar period $t-1/4$). Three-year returns are computed using monthly data for the period between $t+1/12$ and $t+36/12$. We report OLS estimates in columns [1] and [2], and second-stage IV results in column [3]. The instrumental variables are ΔLTG_t and LTG_{t-1}. Except for ΔLTG_t, all variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)
	Dependent Variable:			
	$\Delta_3 e_{t+3}$ / 3 - LTG_t	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$
ΔLTG_t	-0.9351[a]	-0.5520[a]		
	(0.1325)	(0.1877)		
LTG_{t-1}	-0.3352[a]	-0.5003[a]		
	(0.1074)	(0.0938)		
$\widehat{\Delta_3 e_{t+3}/3 - LTG_t}$			0.7345[a]	0.3875[c]
			(0.2692)	(0.2089)
pd_t				-0.4800[a]
				(0.1218)
Observations	397	397	397	397
Adjusted R^2	33%	26%	27%	34%
Montiel-Pflueger F-stat			17.6	
Instrument	.	.	ΔLTG_t , LTG_{t-1}	.

Next, Table B.6 instruments LTG revisions used to predict forecast errors.

Table B.6
Predictability of Forecast Errors and Returns

This table links aggregate forecast errors and market returns. We report regressions using as dependent variable the change in the forecast for growth in earnings in the long run LTG_t between year t and $t-1$, ΔLTG_t in column [1], the error in forecasting five-year growth in aggregate earnings in column [2] and the discounted value of the cumulative market return between year t and $t+5$ in column [3]. Five-year returns are computed using monthly data between t+1/12 and t+60/12. We define the forecast error as the difference between (a) the annual growth in earnings per share between year t and $t+5$, $\Delta_5 e_{t+5}/5$, and (b) the expected long term growth in earnings, LTG_t. Five-year cumulative market returns (r_{t+5}) are computed using monthly data for the period between $t+1/12$ and $t+60/12$. The independent variables are: (a) the log of earnings for the S&P500 in year t relative to cyclically-adjusted earnings in year $t-5$, $e_t - cae_{t-5}$, (b) the predicted change in ΔLTG_t from the regression in column [1], $\Delta LTG_t \mid e_t - cae_{t-5}$, (c) the lagged forecast of long term growth in earnings, LTG_{t-1}, and (d) the error in forecasting five-year growth in aggregate earnings. We assume that earnings are reported with a 3-month lag (i.e. we define e_t as earnings for the calendar period $t-1/4$). We report OLS estimates in columns [1] and [2], and second-stage IV results in column [3]. The instrumental variables are $\Delta LTG_t \mid e_t - cae_{t-5}$ and LTG_{t-1}. Except for ΔLTG_t and $\Delta LTG_t \mid e_t - cae_{t-5}$, all variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 12 lags in column [1] and 60 lags in columns [2] and [3]). Superscripts: a significant at the 1% level, b significant at the 5% level, and c significant at the 10% level.

	(1)	(2)	(3)
	Dependent Variable:		
	ΔLTG_t	$\Delta_5 e_{t+5}/5 - LTG_t$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
e_t-cae_{t-5}	0.3283[a]		
	(0.0695)		
$\Delta LTG_t \mid e_t$-cae_{t-5}		-2.1292[a]	
		(0.5155)	
LTG_{t-1}		-0.0122	
		(0.1127)	
$\Delta_5 e_{t+5}/5 - LTG_t$			0.5784[a]
			(0.1299)
Observations	397	397	397
Adjusted R^2	28%	49%	
Montiel-Pflueger F-stat			12.2

Finally, we turn to Table 5, which examines why the price dividend ratio predicts returns. We start by presenting the first stage of Table 5.

Table B.7
Predictability of Returns from Revisions and Forecast Errors

The table reports regressions of realized returns at one, three, and five years horizons ($k = 1,3,5$) on one-year revisions in long term growth in earnings occurring between $t+1$ and $t+k$ (i.e. ΔLTG_{t+1} through ΔLTG_{t+k}) and long-term forecast errors occurring between $t+1$ and time $t+k$ (i.e. $e_{t+1} - LTG_{t-4}$ through $\Delta_5 e_{t+k} - LTG_{t+k-5}$). Returns are computed using monthly data starting with period t+1/12. The dependent variables are the (log) one-year return in column [1] and the discounted value of the cumulative return between year t and $t+3$ in column [2] and between *t* and $t+5$ in column [3]. We define the time-*t* forecast error as the difference between (a) the annual growth in earnings per share between year t and $t+5$, $\Delta_5 e_{t+5}/5$, and (b) the expected long term growth in earnings, LTG_t. Except for ΔLTG_{t+k}, all variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 12 lags in column [1] and 60 lags in columns [2] and [3]). Superscripts: a significant at the 1% level, b significant at the 5% level, and c significant at the 10% level.

	(1)	(2)	(3)
	r_{t+1}	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$

ΔLTG$_{t+1}$	0.9899[a]	0.4496[c]	0.7826[a]
	(0.2821)	(0.2416)	(0.1401)
ΔLTG$_{t+2}$		0.5410[a]	0.8589[a]
		(0.1887)	(0.1969)
ΔLTG$_{t+3}$		0.5026[b]	0.6264[a]
		(0.2270)	(0.1509)
ΔLTG$_{t+4}$			0.6556[a]
			(0.1320)
ΔLTG$_{t+5}$			0.7194[a]
			(0.1523)
$(e_{t+1} - e_{t-4})/5$ - LTG$_{t-4}$	0.0959	-0.1459	-0.3135[a]
	(0.0961)	(0.1007)	(0.0554)
$(e_{t+2} - e_{t-3})/5$ - LTG$_{t-3}$		0.1157	-0.0854
		(0.1252)	(0.0847)
$(e_{t+3} - e_{t-2})/5$ - LTG$_{t-2}$		0.3236[a]	-0.0440
		(0.1127)	(0.0964)
$(e_{t+4} - e_{t-1})/5$ - LTG$_{t-1}$			0.1188
			(0.0894)
$(e_{t+5} - e_{t})/5$ - LTG$_{t}$			0.1526[a]
			(0.0434)
Observations	361	361	361
R^2	32%	64%	81%
Adjusted R^2	31%	63%	80%

Finally, Table B.8 repeats the analysis of Table 5 using only LTG forecast errors.

Table B.8

LTG Explains Why Price Ratios Predict Returns

This table examines why the price-dividend ratio ($p_t - d_t$) predicts stock market returns. As in Table 5, we proceed in two steps. In the first step, we regress realized returns r_{t+k} at the one, three, and five years horizons ($k = 1,3,5$) on the long-term forecast errors occurring between $t+1$ and time $t+k$ (i.e. $e_{t+1} - LTG_{t-4}$ through $\Delta_5 e_{t+k} - LTG_{t+k-5}$), presented in Panel A. We generate return residuals $\tilde{r}_{t+k}$ from these first-stage regression. In the second step, Panel B, we use $p_t - d_t$ to predict r_{t+k} in columns [1], [3], and [5] and $\tilde{r}_{t+k}$ in columns [2], [4], and [6]. Except for ΔLTG_t, all variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 12 lags in columns [1]-[2], 36 lags in column [3]-[4], and 60 lags in columns [5]-[6]). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A. Predictability of Returns from Errors and Revision

	(1)	(2)	(3)
	r_{t+1}	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
$(e_{t+1} - e_{t-4})/5$ - LTG$_{t-4}$	0.3279[a]	-0.0069	0.0526[c]
	(0.1146)	(0.0396)	(0.0279)
$(e_{t+2} - e_{t-3})/5$ - LTG$_{t-3}$		0.2031[a]	0.2348[a]

		(0.0581)	(0.0325)
$(e_{t+3} - e_{t-2})/5 - LTG_{t-2}$		0.5386^{a}	0.2684^{a}
		(0.0448)	(0.0388)
$(e_{t+4} - e_{t-1})/5 - LTG_{t-1}$			0.1924^{b}
			(0.0501)
$(e_{t+5} - e_{t})/5 - LTG_{t}$			0.5587^{a}
			(0.0413)
Observations	361	361	361
R^2	12%	47%	62%

Panel B. Predictability of Returns from the Price Dividend Ratio

	(1)	(2)	(3)	(4)	(5)	(6)
	r_{t+1}		$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$		$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	
	raw	residual	raw	residual	raw	residual
pd_t	-0.3742^{a}	-0.3803^{a}	-0.6219^{a}	-0.3905^{a}	-0.8204^{a}	-0.4120^{a}
	(0.1446)	(0.1177)	(0.2012)	(0.1052)	(0.2115)	(0.1429)
Obs	361	361	361	361	361	361
R^2	10%	12%	28%	20%	48%	32%
Adj R^2	10%	9%	27%	19%	48%	31%

B.3 Robustness of Firm Level Results

We now turn to our firm-level analysis in Table 6. We first extend it to the three year horizon.

Table B.9
Firm-Level Results: 3 year horizon

We present firm-level regressions for all US firms in the IBES sample. We define firm-level forecast errors as the difference between (a) the growth in firm i's earnings per share between year t and $t+3$, $\Delta_3 e_{i,t+3}/3$, and (b) the expected long term growth in firm i's earnings, $LTG_{i,t}$. In column [1] we perform an OLS regression of the error in forecasting the three-year earnings growth on: (a) the one year revision of the forecast for a firm's long-term earnings growth, $\Delta LTG_{i,t}$ and (b) the lagged forecast for long-term growth in earnings $LTG_{i,t-1}$. In column [2] we perform an OLS regression of the discounted cumulative (log) return for firm i between year t and $t+3$, $\sum_{j=1}^{3} \alpha^{j-1} r_{i,t+j}$ on the same two independent variables. In column [3] we perform an IV regression of stock returns $\sum_{j=1}^{3} \alpha^{j-1} r_{i,t+j}$ on the forecast errors fitted in column [1]. In column [4] we perform an OLS regression of stock returns $\sum_{j=1}^{3} \alpha^{j-1} r_{i,t+j}$ on the forecast errors fitted in column [1] and the dividend to price ratio $pd_{i,t}$ (restricting to observations where dividends are paid). Returns are computed using monthly data starting with period *t+1/12* through *t+3*. We assume that earnings are reported with a with a 3-month lag (i.e. we define e_t as earnings for the calendar period $t - 1/4$). Regressions include time- and firm-fixed effects, which we do not report. The sample period is 1982:12-2015:12. We report Driscoll–

Kraay standard errors with autocorrelation of up to 60 lags. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)
	$\Delta_3 e_{i,t+3}/3$ - $LTG_{i,t}$	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$
$\Delta LTG_{i,t}$	-0.2304[a]	-0.1583[a]		
	(0.0147)	(0.0354)		
$LTG_{i,t}$	-0.2661[a]	-0.1871[a]		
	(0.0134)	(0.0362)		
$\widehat{\Delta_3 e_{1.t+3}/3 - LTG_{\iota t}}$			0.6981[a]	0.7108[a]
			(0.1129)	(0.1381)
$pd_{i,t}$				-0.0887[a]
				(0.0101)
Observations	420,127	420,127	420,127	420,063
Adjusted R^2	2%	0%	8%	1%
KP F-stat			197	
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes

We next redo Table 6 using revisions instrumented from fundamentals (similar to the exercise in Table B.5 for aggregate returns) and lagged forecast in long-term growth in earnings LTG_{t-1}.

Table B.10
Firm-Level Results: instrumented revisions

We present firm-level regressions for all US firms in the IBES sample. We define firm-level forecast errors as the difference between (a) the growth in firm i's earnings per share between year t and $t+3$, $\Delta_3 e_{i,t+3}/3$, and (b) the expected long term growth in firm i's earnings, $LTG_{i,t}$. In column [1] we run an OLS regression of the one year revision of the forecast for a firm's long-term earnings growth, $\Delta LTG_{i,t}$ on the log of firm's i earnings growth between year t and $t-5, \Delta_5 e_{i,t}$. In column [2] we perform an OLS regression of the error in forecasting the five-year earnings growth on fitted value of $\Delta LTG_{i,t}$ from column [1] and the lagged forecast $LTG_{i,t-1}$. In column [3] we perform an IV regression of the discounted cumulative (log) return for firm i between year t and $t+5$, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$ on the error in forecasting the five-year earnings growth, instrumented with $LTG_{i,t-1} | \Delta e_{i,t}$ and $LTG_{i,t-1}$. Five year returns are computed using monthly data starting for the $t+1/12$ - $t+5$ period. We assume that earnings are reported with a with a 3-month lag (i.e. we define e_t as earnings for the calendar period $t-1/4$). Regressions include time- and firm-fixed effects, which we do not report. The sample period is 1982:12-2015:12. We report Driscoll–Kraay standard errors with autocorrelation of up to 12 lags in column [1] and 60 lags in columns [2] and [3]. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)
	$\Delta LTG_{i,t}$	$\Delta_5 e_{i,t+5}/5$ - $LTG_{i,t}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
$\Delta_5 e_{i,t}$	-0.0141[b]		
	(0.0055)		

$\Delta LTG_{i,t} \mid \Delta_5 e_{i,t}$		26.7826[a]	
		(1.4496)	
$LTG_{i,t-1}$		-0.0793[a]	
		(0.0120)	
$\Delta_5 e_{i,t+5}/5 - LTG_{i,t}$			0.2712[a]
			(0.0255)
Observations	284,402	284,402	284,402
Adjusted R^2	0%	20%	.
KP F-stat	-0.0141[b]		170.2
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes

Appendix C. Further Results on Return Predictability in the Cross Section

We start by extending Table 7, which examines the determinants of the returns of the LTG portfolio, by controlling each specification by the price dividend ratio.

Table C.1
Determinants of LTG portfolio returns

We predict the return for portfolios formed on the basis of the forecast for long-term growth in earnings for firm i, $LTG_{i,t}$ using expectations about earnings growth for the market. On each month between December 1982 and December 2015, we form decile portfolios based on $LTG_{i,t}$ and report regression results for the five-year cumulative (log) returns on: (a) the lowest decile ($LLTG$) in column [1], (b) the highest decile ($HLTG$) in column [2], and (c) the difference between the two ($PMO = LLTG - HLTG$) in columns [3]-[7]. Five year returns are computed using monthly data starting for the $t + 1/12$ - $t + 5$ period. The independent variables are: the one year forecast revision for long term growth in aggregate earnings, ΔLTG_t, the one-year lagged forecast, LTG_{t-1}, the (log) five-year return of CRSP's value-weighted index between t and $t + 5$, $\ln(Mkt_{t,t+5})$, the price dividend ratio and the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [4], the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [5], the Martin (2013) expected return on the market $SVIX^2$ in column [6], and the price dividend ratio in column [7]. Except for ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. Newey-West standard errors are in parentheses (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: (Log) Five-year Return					
	LLTG	HLTG	PMO	PMO	PMO	PMO
ΔLTG_t	0.3407[b]	-0.8180[a]	1.0545[a]	1.0327[a]	1.0323[a]	1.3574[a]
	(0.1687)	(0.1859)	(0.2098)	(0.1891)	(0.1930)	(0.3106)
LTG_{t-1}	0.4877[a]	-0.4827[a]	0.7299[a]	0.7132[a]	0.7109[a]	0.8516[a]
	(0.1339)	(0.0723)	(0.1071)	(0.1134)	(0.1315)	(0.1976)
$\ln(Mkt_{t,t+5})$	0.6075[a]	0.5698[a]	-0.4176[a]	-0.4015[a]	-0.4286[a]	-0.3943[b]
	(0.1030)	(0.1025)	(0.1001)	(0.0992)	(0.1069)	(0.1665)
pd_t	-0.4865[a]	0.1061	-0.3027[a]	-0.2826[b]	-0.2817[b]	-0.8863[b]
	(0.0965)	(0.0863)	(0.1093)	(0.1183)	(0.1231)	(0.4439)
X_t				0.0485	0.0522	-0.0493
				(0.1215)	(0.1644)	(0.1286)
Obs	397	397	397	397	133	193
Adj R^2	82%	83%	74%	74%	73%	77%
X_t				spc_t	cay_t	$SVIX^2_t$

We next extend the exercise of Tables 7 and 8, which examines the determinants of the returns and of the forecast errors of LTG portfolios, by looking at the shorter, 3 year horizon.

Table C.2
Returns and Forecast Errors of LTG Portfolios

This table predicts returns (Panel A) and forecast errors (Panel B) for portfolios formed on the basis of expected long-term growth in earnings for firm i, $LTG_{i,t}$ using beliefs about aggregate earnings growth. On each month between December 1982 and December 2015, we form decile portfolios based on $LTG_{i,t}$. In Panel A we report regression results for the three-year cumulative (log) returns on: (a) the lowest decile ($LLTG$) in column [1], (b) the highest decile ($HLTG$) in column [2], and (c) the difference between the two ($PMO = LLTG - HLTG$) in columns [3]-[7]. Returns are computed using monthly data starting with period $t + 1/12$. The independent variables are: the one year forecast revision for long term growth in aggregate earnings, ΔLTG_t, the one-year lagged forecast, LTG_{t-1}, and the (log) three-year return of CRSP's value-weighted index between t and $t+3$, $\ln(Mkt_{t,t+5})$. In Panel B, we report regressions for the forecast errors in predicting earnings growth between t and $t+3$ of the following three portfolios: (a) the lowest decile ($LLTG$) in column [1], (b) the highest decile ($HLTG$) in column [2], and (c) the difference between the two ($PMO = LLTG - HLTG$) in column [3]. We define portfolio errors as the mean forecast error of the firms in the relevant *LTG* portfolio, i.e. the time *t* average difference between: (1) the annual growth in firm i's earnings per share between year t and, $\Delta_3 e_{i,t+3}/3$, and (2) the expected long term growth in firm i's earnings, $LTG_{i,t}$. In Panel B, te independent variables are: (a) the one year forecast revision for aggregate earnings, ΔLTG_t, and (b) the lagged one-year forecast, LTG_{t-1}. Except ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 36 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A. Three-Year Returns

	(1) LLTG	(2) HLTG	(3) PMO
ΔLTG_t	0.2222[b]	-0.6565[a]	0.8204[a]
	(0.0886)	(0.2276)	(0.2688)
LTG_{t-1}	0.2150[c]	-0.4067[a]	0.5349[a]
	(0.1274)	(0.0815)	(0.1115)
$\ln(Mkt_{t,t+3})$	0.8445[a]	0.5612[a]	-0.3439[b]
	(0.0697)	(0.1334)	(0.1546)
Observations	397	397	397
Adjusted R^2	80%	78%	63%

Panel B. Three Year Equal Weighted FE

	(1) LLTG	(2) HLTG	(3) PMO
ΔLTG_t	-0.5123[a]	-0.7157[a]	0.5260[a]
	(0.1780)	(0.1452)	(0.1084)
LTG_{t-1}	-0.1778	-0.7922[a]	0.7926[a]
	(0.1263)	(0.0932)	(0.1064)
Observations	397	397	397
Adjusted R^2	9%	60%	56%

Table C.3 extends Table 8 by showing that value-weighted forecast errors of LTG sorted portfolios are predictable at the 5 year horizon. This is remarkable, since we may expect forecast errors to be concentrated in smaller firms, and also that LTG captures only one portion of such errors.

Table C.3
Forecast Errors of LTG Portfolios

This table predicts forecast errors for portfolios formed on the basis of expected long-term growth in earnings for firm i, $LTG_{i,t}$ using beliefs about aggregate earnings growth. On each month between December 1982 and December 2015, we form decile portfolios based on $LTG_{i,t}$ and report regressions for the forecast errors in predicting earnings growth between t and $t+5$ of the following three portfolios: (a) the lowest decile ($LLTG$) in column [1], (b) the highest decile ($HLTG$) in column [2], and (c) the difference between the two ($PMO = LLTG - HLTG$) in column [3]. We define portfolio errors as the mean value-weighted forecast error of the firms in the relevant LTG portfolio, i.e. the time t average value-weighted difference between: (1) the annual growth in firm i's earnings per share between year t and $t+5$, $\Delta_5 e_{i,t+5}/5$, and (2) the expected long term growth in firm i's earnings, $LTG_{i,t}$.The independent variables are: (a) the one year forecast revision for aggregate earnings, ΔLTG_t, and (b) the lagged one-year forecast, LTG_{t-1}. Except ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	(1)	(2)	(3)
	Dependent variable: Five-Year VW Forecast Error		
	LLTG	HLTG	PMO
ΔLTG_t	-0.3595[c]	-0.8597[a]	0.7970[a]
	(0.1905)	(0.1389)	(0.1217)
LTG_{t-1}	0.0911	-0.7493[a]	0.7910[a]
	(0.1687)	(0.0759)	(0.0891)
Observations	397	397	397
Adjusted R^2	6%	52%	55%

Having assessed the robustness of the analysis of LTG sorted portfolios (Tables 7 and 8), we now similarly extend the analysis in Table 9, which examines the predictability of factor returns and their forecast errors, to a shorter, 3 year horizon.

Table C.4
Predictability of factor returns and forecast errors

This table links beliefs about growth in earnings to Fama-French factor returns (Panel A) and forecast errors (Panel B). The dependent variables in Panel A are the compounded (log) return between year t and $t+3$ of the following 4 factors: (a) high-minus-low book-to market (HML) in column [1], (b) robust-minus-weak profitability factor (RMW) in column [2], (c) conservative-minus-aggressive investment (CMA) in column [3], and (d) small-minus-big factor (SMB) in column [4]. Returns are computed using monthly data starting with period $t+1/12$. The dependent variables in Panel B are the forecast errors in predicting the growth in earnings between t and $t+3$ for the: (1) HML, (2) RMW, (3) CMA, and (4) SMB portfolios. In Panel A, the independent variables are: (a) the one-year revision in aggregate earnings growth forecast, ΔLTG_t, (b) the one-year lagged forecast, LTG_{t-1}, (c) the (log) three-year return of CRSP's value-weighted index between t and $t+3$, $\ln(Mkt_{t,t+3})$. In Panel B, the independent variables are ΔLTG_t and LTG_{t-1}. Except ΔLTG_t, variables are standardized. Intercepts are not shown. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 36 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A: Three-year Returns

	(1) HML	(2) RMW	(3) CMA	(4) SMB
ΔLTG_t	0.6430[b]	0.1991	0.6603[b]	-0.0180
	(0.3033)	(0.3302)	(0.2649)	(0.2100)
LTG_{t-1}	0.6816[a]	0.1231	0.6016[a]	0.4614[a]
	(0.1407)	(0.1756)	(0.1547)	(0.1264)
$\ln(Mkt_{t,t+3})$	0.2074	-0.4148[b]	-0.1966	-0.4100[b]
	(0.2156)	(0.2114)	(0.1821)	(0.1597)
Obs	397	397	397	397
Adj R^2	37%	24%	54%	50%

Panel B: Three-Year Forecast Errors

	(1) HML	(2) RMW	(3) CMA	(4) SMB
ΔLTG_t	-0.0200	0.4645[a]	0.2211[c]	-0.3733[a]
	(0.1388)	(0.1678)	(0.1278)	(0.1327)
LTG_{t-1}	0.2450[b]	0.4822[a]	0.1925	-0.4568[a]
	(0.1213)	(0.1020)	(0.1172)	(0.1343)
Obs	397	397	397	397
Adj R^2	6%	31%	6%	24%

Next, in Table C.5, we add *cay* and its interaction with the market return to the benchmark specification in Table 9.

Table C.5
Predictability of factor returns and role of *cay*

This table links beliefs about growth in earnings to Fama-French factor returns. The dependent variables are the compounded (log) return between year t and $t+5$ of the following 4 factors: (a) high-minus-low book-to market (HML) in column [1], (b) robust-minus-weak profitability factor (RMW) in column [2], (c) conservative-minus-aggressive investment (CMA) in column [3], (d) small-minus-big factor (SMB) in column [4], and (e) betting-against-beta (BAB) in column [5]. Returns are computed using monthly data starting with period $t + 1/12$. The independent variables are: (a) the one-year change in the expected long term growth in earnings, ΔLTG_t, (b) the one-year lagged forecast for long term growth in earning, LTG_{t-1}, (c) the (log) five-year return of CRSP's value-weighted index between t and $t+5$, $\ln(Mkt_{t,t+5})$, (d) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, and (e) the interaction between $\ln(Mkt_{t,t+5})$ and cay_t, $\ln(Mkt_{t,t+5}) * cay_t$. Except ΔLTG_t, variables are normalized to have zero mean and standard deviation of 1. Intercepts are not shown. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

(1)	(2)	(3)	(4)	(5)

Dependent Variable: Five-year (log) Return of:

	HML	RMW	CMA	SMB	BAB
ΔLTG_t	0.8462[a]	0.1706	0.6653[a]	0.3516[b]	0.3461[a]
	(0.1317)	(0.2304)	(0.2210)	(0.1595)	(0.0379)
LTG_{t-1}	0.7546[a]	0.0651	0.4404[a]	0.5505[a]	0.3157[a]
	(0.0947)	(0.1402)	(0.0948)	(0.0994)	(0.0357)
$\ln(Mkt_{t,t+5})$	0.5781[a]	-0.5109[b]	0.1971	-0.1076	0.4042[a]
	(0.1758)	(0.2109)	(0.1394)	(0.0797)	(0.0462)
cay_t	1.1376[a]	0.4046	1.4857[a]	0.5593[c]	0.1344
	(0.2730)	(0.3492)	(0.1776)	(0.2949)	(0.1009)
$\ln(Mkt_{t,t+5}) * cay_t$	-0.8570[a]	-0.0441	-1.0620[a]	-0.7292[a]	-0.3207[a]
	(0.2870)	(0.3492)	(0.2001)	(0.1978)	(0.0876)
Observations	133	133	133	133	133
Adjusted R^2	68%	36%	66%	65%	66%

Finally, we present a benchmark for Table 9, namely a univariate regression of factor returns on the market factor. In Table C.4 panel A, we present the regression. Panel B compares the variation explained by the univariate regression to that explained by the bivariate regressions in Table 9.

Table C.6
Predictability of factor returns: benchmark

The dependent variable is the cumulative (log) return between year t and $t+5$ of the low-minus-high LTG portfolio, PMO, in column [1], (b) the high-minus-low book-to market portfolio, HML, in column [2], (c) the robust-minus-weak profitability portfolio, RMW in column [3], (d) the conservative-minus-aggressive investment portfolio, CMA, in column [4], or (e) the small-minus-big factor portfolio, SMB, in column [5]. The independent variable in Panel A is the (log) five-year return of CRSP's value-weighted index between t and $t+5$, $\ln(Mkt_{t,t+5})$. Returns are computed using monthly data starting with period $t+1/12$.

Panel B reports the adjusted R squared from bivariate regressions of the each of the five dependent variables in Panel A on: (a) the one-year change in the expected long term growth in earnings, ΔLTG_t, and (b) the one-year lagged forecast for long term growth in earning, LTG_{t-1}. These can be compared with the adjusted R squared from the univariate regressions in Panel A. Variables are normalized to have zero mean and standard deviation of 1. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags). Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A: Predictability of factor returns from market factor

	(1)	(2)	(3)	(4)	(5)
	PMO	HML	RMW	CMA	SMB
$\ln(\text{Mkt}_{t,t+5})$	-0.5108a	-0.0971	-0.5305a	-0.3903	-0.6512a
	(0.1921)	(0.2855)	(0.1366)	(0.3097)	(0.1671)
Observations	397	397	397	397	397
Adjusted R^2	25%	1%	27%	14%	40%

Panel B: Summary of R^2

	(1)	(2)	(3)	(4)	(5)
	PMO	HML	RMW	CMA	SMB
Adj R^2 Bivariate Regression	46%	50%	20%	52%	49%

Appendix D. Return Predictability, Tangible News, Intangible News

Our key methodological innovation is to use expectations data, and in particular measured expectations of long-term fundamentals, as direct predictors of both returns and forecast errors for the aggregate stock market, for individual firms, and in the cross section of stocks. Throughout, we have been agnostic as to whether the sources of overreaction are tangible or intangible news. Previous work that does not rely on expectations data and tries to measure news directly has found conflicting results. Daniel and Titman (2006) show that past fundamentals do not predict future stock returns, and argue that this is consistent with an outsized role of (unmeasured) intangible news. Nagel and Xu (2019) construct a proxy for five years earnings growth and show that it can predict future aggregate returns, consistent with a key role of tangible news in shaping beliefs (through fading memory).

We revisit this issue. We predict returns using LTG revisions and lagged LTG, but also control for our proxy for five-year fundamental growth $e_t - cae_{t-5}$. In this exercise, the explanatory power of past fundamentals $e_t - cae_{t-5}$ can derive either from their ability to capture fundamentals-driven movements in beliefs not embodied in LTG (e.g., changes in short term expectations or noise in LTG revisions), or from their ability to capture fundamentals-driven discount rate movements. This exercise then allows us to jointly assess the role of tangible news, the informativeness of LTG data, and mechanisms based on fundamentals driven required returns.

Table D.1 reports the results. In panel A we regress the five years ahead return on the LTG revision and lagged LTG, controlling for $e_t - cae_{t-5}$. Column (1) performs the exercise at the aggregate level, using aggregate LTG revision and lagged LTG. Column (2) does the same at the firm level, using firm level LTG revisions and lagged LTG. Here we proxy for past fundamentals using five years earnings growth. Panel B looks at cross sectional return spreads.

Table D.1

This table links beliefs about growth in earnings to the return on the market portfolio (Panel A, column [1]), firm-level returns (Panel A, column [2]), and portfolio returns (Panel B). In Panel A, the dependent variable is the five-year discounted value of the cumulative (log) return for: (a) the market, $\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$, in column [1], and (b) firms on IBES, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$, in column [2]. In column [1] the independent variables are: (a) the one-year change in the expected long term growth in earnings, ΔLTG_t, (b) the one-year lagged forecast for long term growth in earning, LTG_{t-1}, and (d) the log of earnings in year t relative to cyclically-adjusted earnings in year $t-5$, $e_t - cae_{t-5}$. In column [2] the independent variables are: (a) $\Delta LTG_{i,t}$, (b) $LTG_{i,t-1}$, and (c) the growth in firm i's earnings of between t and $t+5$, $\Delta_5 e_{i,t+5}$. In column [2] of Panel A, we also include time- and firm-fixed effects, which we do not report. Returns are computed using monthly data starting with period $t + 1/12$. The dependent variables in Panel B is the cumulative (log) return between year t and $t+5$ of the low-minus-high LTG portfolio (PMO) in column [1], (b) the high-minus-low book-to market portfolio (HML) in column [2], (c) the robust-minus-weak profitability portfolio (RMW) in column [3], (d) the conservative-minus-aggressive investment (CMA) portfolio in column [4], or (e) the small-minus-big factor (SMB) portfolio in column [5]. The independent variables on Panel B are: (a) ΔLTG_t, (b) LTG_{t-1}, (c) (log) five-year return of CRSP's value-weighted index between t and $t+5$, $\ln(Mkt_{t,t+5})$, and (d) $e_t - cae_{t-5}$. We assume that earnings are reported with a with a 3-month lag (i.e. we define e_t as earnings for the calendar period $t - 1/4$). Except ΔLTG_t and $\Delta LTG_{i,t}$, variables are normalized to have zero mean and standard deviation of 1. The sample period is 1982:12-2015:12. We adjust standard errors for serial correlation using the Newey-West correction (with 60 lags), except in column [2] of Panel A where we report Driscoll–Kraay standard errors. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

Panel A: Aggregate and firm-level returns

	(1)	(2)
	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$
ΔLTG_t	-0.4168[b]	-0.1544[a]
	(0.1946)	(0.0360)
LTG_{t-1}	-0.4681[a]	-0.1824[a]
	(0.0975)	(0.0403)
$e_t - cae_{t-5}$	-0.2526	
	(0.1824)	
$\Delta_5 e_{i,t}$		-0.0830[a]
		(0.0116)
Observations	397	284,406
Adjusted R^2	34%	2%

Panel B: Portfolio returns

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: Five-year (log) Return of				
	PMO	HML	RMW	CMA	SMB
ΔLTG_t	0.5230[a]	1.2439[a]	0.2893	1.1471[a]	0.5654[a]
	(0.2003)	(0.1389)	(0.2302)	(0.1993)	(0.1101)

LTG_{t-1}	0.6097[a]	0.9525[a]	0.1864	0.7065[a]	0.5694[a]
	(0.1033)	(0.1270)	(0.1833)	(0.1343)	(0.0811)
$\ln(Mkt_{t,t+5})$	-0.2545	0.3343[a]	-0.3807[b]	-0.0420	-0.4994[a]
	(0.2335)	(0.1265)	(0.1750)	(0.1775)	(0.1206)
$e_t - cae_{t-5}$	-0.1608	-0.3944[a]	0.0643	-0.3373[c]	-0.4115[a]
	(0.2022)	(0.1251)	(0.1931)	(0.1784)	(0.0967)
Observations	397	397	397	397	397
Adjusted R^2	50%	68%	30%	56%	65%

The inclusion of past fundamentals does not change the overall message. The predictive power of past fundamentals is typically economically smaller and statistically less significant than that of measured beliefs. The improvement in R^2 in the aggregate, firm level, and cross-sectional regressions is also small. Cross sectional results are especially striking. A model using only the market return and past fundamentals accounts for 25% of the HML *LTG* spread, 1% of the HML BM spread, 28% of the RMW spread, 14% of the CMA spread, and 43% of the SMB spread. Measured expectations thus appear crucial for accounting for cross sectional anomalies.

Appendix E. Simulated p-values

We follow closely the procedure of Kothari and Shanken (1997). Specifically, we run univariate OLS regressions of one-month returns r_{t+1} on the forecast for long-term growth in earnings, LTG_t:

$$r_{t+1} = \alpha + \beta\, LTG_t + u_{t+1}$$

Next, we estimate an AR(1) process for LTG_t

$$LTG_{t+1} = c + \phi\, LTG_t + \nu_{t+1}$$

We adjust the estimated β and ϕ to account for bias as in Kendall (1954) and set the intercepts so that the average fitted values equal the sample means. Using these adjusted coefficients, we obtain adjusted residuals (u, ν). We then set the initial value of LTG to its historical value and obtain 469 simulated values of r_{t+1} and LTG_{t+1} under the null hypothesis of no predictability by randomly selecting (u, ν) pairs with replacement. We then compute 409 values of (overlapping) 5-year returns based on the bootstrapped monthly observations and run regressions of 5-year returns on LTG. We repeat the process 5,000 times. We record the average slope coefficients in those regressions and report p-values based on the fraction of the 5,000 coefficients that exceed the slope coefficient in Table 1. We follow the same procedure to obtain simulated p-values for the for the forecast of earnings growth between year: (a) t and $t+1$, $\mathbb{E}_t^O[e_{t+1} - e_t]$, and (b) t +1 and $t+2$, $\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$. The results are presented in the table below:

Table E.1

	(1)	(2)	(3)
	r_{t+1}	$\sum_{j=1}^{3} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
	Panel A: Returns and LTG		
LTG_t	-0.2389[b]	-0.4019[a]	-0.4349[a]
	(0.0928)	(0.0944)	(0.0831)
p-value (bootstrap)	3.24%	3.88%	3.48%
Observations	409	409	409

Adj R^2	9%	24%	25%
	Panel B: Returns and growth forecast for year 1		
$E_t[e_{t+1}-e_t]$	-0.0335	0.0467	0.1556[a]
	(0.1027)	(0.0716)	(0.0587)
p-value (bootstrap)	36.90%	49.72%	36.88%
Observations	404	404	404
Adj R^2	0%	0%	3%
	Panel C: Returns and growth forecast for year 2		
$E_t[e_{t+2}-e_{t+1}]$	-0.0527	0.0408	0.2113
	(0.0885)	(0.1556)	(0.1686)
p-value (bootstrap)	48%	43%	15%
Observations	404	404	404
Adj R^2	0%	0%	6%

Stambaugh (1999) worries about a setting where the independent variable is persistent and innovations in returns u_{t+1} are highly correlated with innovations in the independent variable v_{t+1}. By construction, the latter is true when the independent variable is a scaled price, such as the book-to-market ratio. Kothari and Shanken report that the correlation between shocks to book-to-market and to annual returns is -0.80. In our setting, the correlation between shocks one month returns and LTG is only -0.051. Similarly, the correlation between one-month returns and$\mathbb{E}_t^O[e_{t+2} - e_{t+1}]$ is negligible (i.e., 0.02). In contrast, correlation between shocks to one-month returns and $\mathbb{E}_t^O[e_{t+1} - e_t]$ is -0.41.