Belief Overreaction and Stock Market Puzzles

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

April 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 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 earnings growth, for the market as a whole but especially for specific firms. This leads to inflated stock prices and, as beliefs are systematically disappointed, to subsequent low returns in the aggregate market and for specific firms in the cross section. Overreaction of measured 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 European Research Council for Financial Support under the ERC Consolidator 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, and Adi Sunderam for extremely helpful comments. Julien Manili provided outstanding research assistance.

1. Introduction

In the dividend discount model, the price of a stock at time t is given by:

$$P_t = \sum_{s=t+1}^{\infty} \frac{\mathbb{E}_t(D_s)}{(1+r)^{s-t}}$$

where r is the constant required return and $\mathbb{E}_t(D_s)$ is the rational expectation of the dividend per share at time s. Research over the past four decades shows that this model is a poor description of stock market movements. There are two related problems. First, stock prices are much more volatile than the present value of dividends or earnings (Shiller 1981, Leroy and Porter 1981). In the dividend discount model, all price volatility should be due to news about these fundamentals. Second, the current price dividend ratio has a weak positive correlation with future growth in dividends or earnings but a strong negative correlation with future returns (Campbell and Shiller 1988). This is also inconsistent with the model, in which the price dividend ratio reflects rational forecasts of future dividends and required returns are constant.

The conventional explanation of such "excessive" stock price movements is time varying required returns. Returns may vary due to changes in risk preference (e.g., Campbell and Cochrane 1999), or due to long run or disaster risk (Rietz 1988, Bansal and Yaron 2004, Barro 2006). In these theories, expectations are rational. In good times investors require and expect low returns, so stocks are expensive. As conditions revert to normal, required returns rise and stocks become cheaper. This approach has two problems. First, the prediction that investors expect low returns in good times is counterfactual: survey expectations of returns are higher in good than in bad times (Greenwood and Shleifer 2014). Second, changes in risk attitudes, in long run risk, or in disaster risk are hard to measure.

Here we pursue an orthogonal approach: we keep required returns constant and relax rational expectations of fundamentals. We discipline departures from rationality using data on analyst expectations of future earnings growth of listed firms. We show that expectations of long term growth, LTG, display a remarkable ability to forecast returns for the aggregate market and in the cross section, indicating that beliefs about the long term are a key source of anomalies.

In Section 2 we showcase the promise of expectations data for helping solve major stock market puzzles. We present two facts. First, measured expectations of earnings growth are volatile enough that they solve Shiller's excess volatility puzzle, even with constant required returns. Second, high current expected long-term earnings growth predicts lower future aggregate stock returns. The predictive power of LTG is large compared to that of model-based measures of required returns or of prominent macroeconomic predictors.

Section 3 studies the mechanism behind these facts by linking forecast errors and returns. We first show that LTG overreacts: upward LTG revisions predict disappointment of growth forecasts. We connect this fact to price anomalies by showing that predicted disappointment is associated with low returns. This link also holds at the level of individual firms, controlling for multiple aggregate shocks. These findings point to a model in which an overreaction to good news causes excess optimism about aggregate fundamentals, inflating stock prices. In the future, systematically disappointing earnings growth causes a price reversal and hence low returns.

In Section 4 we study whether fluctuations in aggregate optimism can account for return differentials and co-movement in the cross-section. La Porta (1996) showed that stocks with high LTG earn lower average returns than stocks with low LTG. We first show that this return differential widens after times of high aggregate optimism. We next show the connection to overreaction: high aggregate optimism also predicts stronger forecast disappointment for high LTG firms compared to low LTG ones. This evidence is consistent with a mechanism in which high LTG stocks overreact more than low LTG stocks to aggregate good news, perhaps because they 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.

In sum, overreaction in measured expectations of long-term fundamentals can help unify longstanding puzzles in finance, ranging from aggregate stock market volatility (typically explained with time varying required returns) to cross sectional return differentials (typically explained with cross sectional differences in exposure to risk factors (Fama and French 1993)). Our results show the promise of using beliefs data and non-rational belief formation models for finding realistic and parsimonious mechanisms of return predictability.

Several 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 it to the aggregate market. De la O and Myers (2020) show that analysts' forecasts of short-term earnings have strong explanatory power for the aggregate price earnings and price dividend ratios. Nagel and Xu (2019) show that past dividend growth negatively predicts future market returns and positively correlates with expectations of earnings growth. Compared to the last two papers, our innovation is to use beliefs data to jointly predict returns and forecast errors, in both the aggregate and the cross section, highlighting the common mechanism of overreaction.

We offer a new angle on macro-financial 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 et al. (2020), who find overreaction by individual professional forecasters, we find overreaction in consensus expectations and connect it to excess stock market volatility. Our analysis opens new avenues for thinking about macro-financial volatility as the byproduct of belief volatility, in line with recent work in macroeconomics (Bianchi et al. 2021, Bordalo et al. 2021, L'Huillier et al. 2021).

[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).

2. 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." Data coverage starts on 3/1976 for EPS_{it} and 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.[3] 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.[4]

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

---

[3] 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.

[4] 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.

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

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.[5]

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

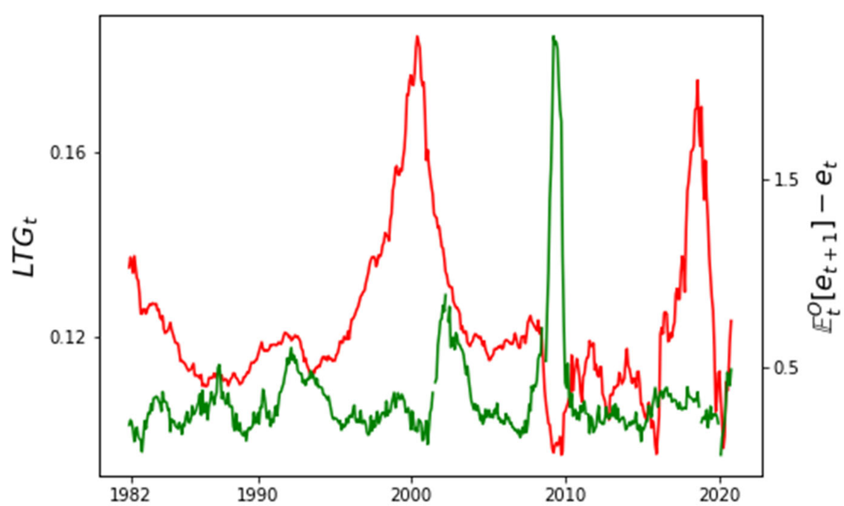

Figure 1. We plot the expected short- and long-term growth in earnings ($\mathbb{E}_t^O[e_{t+1}] - e_t$ in green and LTG_t in red, respectively, where $e_t = \log EPS_t$ and $\mathbb{E}_t^O$ represents measured expectations). The scale for short-term earnings ($\mathbb{E}_t^O[e_{t+1}] - e_t$) is on the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

2.1 Shiller's Excess volatility Puzzle

To assess whether beliefs about fundamentals can account for stock market volatility, we construct a price index $\tilde{p}_t$ using measured expectations of future earnings growth and compare it to the actual stock price p_t. We follow Campbell and Shiller (1987, 1988), who express the log return r_{t+1} obtained by holding the stock market between t and $t+1$ as:

[5] 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.

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

where p_t and p_{t+1} are the log stock prices at t and $t+1$, d_{t+1} is the log dividend at $t+1$, k is a constant, and $\alpha = e^{pd}/(1+e^{pd}) < 1$ depends on the average log price dividend ratio pd. 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 $t+s+1$.

In Equation (2), variation in the 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 explanations of excess stock price volatility introduce time varying required returns. We instead keep required returns and expectations of future returns constant at r, while allowing for non-rational beliefs about future dividend growth, which we denote by $\widetilde{\mathbb{E}}_t(.)$. Taking the expectation of (2) under these assumptions gives:

$$\tilde{p}_t = d_t + \frac{k-r}{1-\alpha} + \sum_{s \geq 0} \alpha^s \widetilde{\mathbb{E}}_t(g_{t+1+s}), \tag{3}$$

where $\tilde{p}_t$ denotes the price under possibly non-rational beliefs $\widetilde{\mathbb{E}}_t(.)$. In Equation (3), excess price volatility arises from possibly non-rational changes in $\widetilde{\mathbb{E}}_t(g_{t+1+s})$.

We build an empirical counterpart of $\tilde{p}_t$ by rewriting Equation (3) in terms of earnings per share and expected earnings growth, and by inserting measured expectations into it:

$$\tilde{p}_t = e_t + \frac{\tilde{k}-r}{1-\alpha} + \ln\left(\frac{\mathbb{E}_t^O EPS_{t+1}}{EPS_t}\right) + \alpha \ln\left(\frac{\mathbb{E}_t^O EPS_{t+2}}{\mathbb{E}_t^O EPS_{t+1}}\right) + \sum_{s=2}^{10} \alpha^s \, LTG_t + \frac{\alpha^{10}}{1-\alpha} \, g, \tag{4}$$

where as in Figure 1 $\mathbb{E}_t^O$ denotes measured beliefs, which proxy for market beliefs $\widetilde{\mathbb{E}}_t$, and where we set the conventional values $\alpha = (1+e^{-pd})^{-1} \sim 0.9774$, $r = 8.48\%$ (the sample mean) and $\tilde{k} = k + (1-\alpha)\, de = 0.0927$, where de is the average log payout ratio.

We measure expected growth between t and $t+2$ using forecasted earnings, and between $t+3$ and $t+10$ using LTG_t.[6] We employ LTG_t up to 10 years ahead because this is the average duration of a business cycle in our data. We have no expectations data for longer horizons. Beyond $t+11$ we set the constant growth $g = 6.35\%$ such that the level of our price index p_t^O is consistent with the average stock price.[7] We use nominal earnings growth values, but Appendix B (Figure B.1) shows that our results are robust when we account for inflation.

Figure 2 plots the market price p_t (green line) against our price index $\tilde{p}_t$ (red line). It also plots the rational price p_t^{RE} (blue line), computed following Shiller's (2014) methodology.[8]

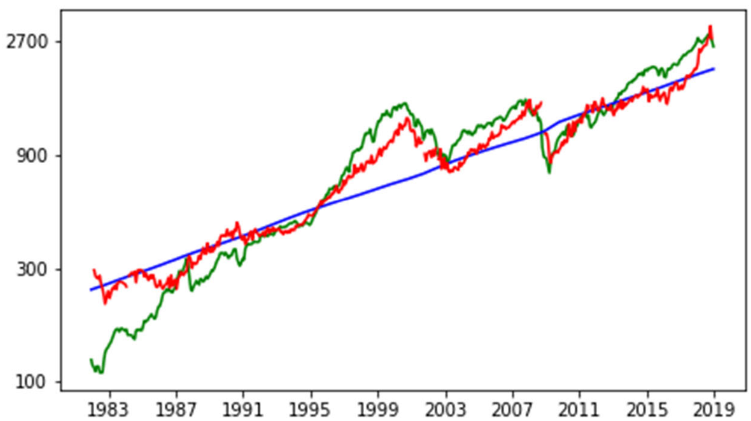

Figure 2. We plot in logscale the levels of the S&P500 index (green line), the rational benchmark index (p_t^{RE}, blue line, footnote 10), and the price index based on earnings forecasts ($\tilde{p}_t$, red line, Equation 4).

Expectation-based prices $\tilde{p}_t$ are remarkably well aligned with the actual price p_t , especially at low frequencies. When the actual price is above the rational benchmark p_t^{RE}, so is

[6] 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.

[7] That is, g is the average of the growth rate g_t obtained by solving, at each time t, the equation $p_t = e_t + \frac{\tilde{k}-r}{1-\alpha} + \alpha \ln\left(\frac{\mathbb{E}_t^O EPS_{t+2}}{\mathbb{E}_t^O EPS_{t+1}}\right) + \sum_{s=2}^{10} \alpha^s \, LTG_t + \frac{\alpha^{10}}{1-\alpha} \, g_t$. Results are virtually identical if we apply to *LTG* the estimated decay of observed cyclically adjusted earnings (see Appendix B, Table B.2).

[8] Starting from the terminal price $p_T^* = \ln\left(\frac{D_T}{r-g}\right)$ at $T = 2020$, the index p_t^{RE} is computed backwards, using the actual dividends over time, and setting $g = 5.81\%$ and $r = 8.48\%$ to reflect sample averages. That is:

$$p_t^{RE} = d_t + \sum_{s=t}^{T} a^{s-t}(d_{s+1} - d_s) + \alpha^{T-t} * (p_{2020}^* - d_{2020}) + \sum_{s=t}^{T-1} a^{s-t}(k-r).$$

$\tilde{p}_t$; and conversely when the actual price is below the rational benchmark. Table 1 shows that the standard deviation of one year changes in our index $\tilde{p}_t$ is quantitatively very close to that of the one-year change in the actual stock price, and much higher than that obtained using the rational benchmark. In Appendix B, Table B.2 we show that very similar results obtain if we construct a price index using expectations of dividends, which are available only after 2002.[9] Measured expectations solve Shiller's excess volatility puzzle.

Table 1
Volatility of log price changes

This table reports the standard deviation and 95th confidence interval of one-year change in: (a) the log of the price of the S&P500 index, Δp, (b) the rational benchmark index, Δp^{RE} (footnote 8), and (c) the price index based on earnings forecasts (Equation 4), $\Delta\tilde{p}$. The sample period is 12/1982 to 12/2020.

	Earnings-based index		
	Δp	Δp^{RE}	$\Delta\tilde{p}$
σ	14.8%	0.7%	14.6%
95th Confidence Interval	13.9%-15.9%	0.6%-0.7%	13.7%-15.6%

Appendix B, Table B.1 further shows that the price earnings ratio $\tilde{p}_t - e_t$ constructed using our index has strong explanatory power for time variation in the actual ratio $p_t - e_t$. De la O and Myers (2021, 2022) show that expectations of short term earnings growth explain most of the variation in the price earnings ratio, consistent with the idea such variation mostly comes from the dynamics of earnings as opposed to prices (see also Adam and Nagel 2022). To construct $\tilde{p}_t$ we use in addition expectations of long term earnings growth, LTG_t. Fluctuations in LTG_t are critical to account for prices and hence for return predictability, as we show next.[10]

2.2 Return Predictability

[9] Following Campbell and Shiller (1987), Appendix B, Table B.3 shows that the cointegrated series $\tilde{P}_t - \frac{D_t}{r}$, where $\tilde{P}_t = e^{\tilde{p}_t}$ is the price level from Equation (4), captures most of the excessive volatility of its empirical counterpart.

[10] 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.

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 2 reports the results (results are similar if we use excess returns, see Appendix B, Table B.4).

Table 2
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. 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
Adjusted R^2	9%	24%	25%
Panel B: Returns and growth forecast for year 1			
$\mathbb{E}_t^O[e_{t+1} - e_t]$	-0.0335	0.0467	0.1556[a]
	(0.1027)	(0.0716)	(0.0587)
Observations	404	404	404
Adjusted 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
Adjusted R^2	0%	0%	6%

High current expectations of long term earnings growth strongly predict low future returns. LTG accounts for 25% of variation in realized returns over the following five years.[11] In contrast,

[11] 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 2. We find that, under the null that the LTG_t coefficient is zero, the predicted values of the LTG_t coefficients are: -0.0128 in column 1, -0.0197 in column 2, and -0.0163 in column 3. The p-values for the LTG coefficients in Table 2 under the null of no predictability are: 2.11% in column 1, 3.30% in column 2, and 0.71% in

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 long term earnings growth expectations.[12]

Where does the predictive power of measured expectations come from? One possibility is that LTG_t spuriously reflects time varying required returns. This could happen if analysts estimate LTG_t by fitting in Equation (2) the growth rate of earnings that justifies the current stock price while erroneously assuming constant required returns. We assess this in two ways.

As a first test, we run the predictive regression of Table 2 controlling 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 indicator proxies for fluctuations of marginal utility in habit formation models, the second for time varying 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 3 reports the results from this exercise in columns (1), (2) and (3). In columns (4)-(7) we assess the robustness of LTG_t to 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). These additional measures are not constructed to proxy for required returns but allow us to assess how LTG_t fares compared to them in predicting returns.

Table 3

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 $t+5$. The independent variables are: (a) the forecast for earnings growth in the long run LTG_t, (b) the Campbell and Cochrane (1999) surplus consumption ratio spc_t in column [1], (c) the Lettau and Ludvigson (2001) consumption-wealth ratio cay_t in column [2], (d) the Martin (2013) expected return on

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 future returns r_{t+1} and future LTG_{t+1} on LTG_t is below 0.05.

[12] Nagel and Xu (2021) show that past earnings growth predicts returns and LTG forecasts, but they do not show either that LTG predicts returns or that overreaction in LTG accounts for return predictability. De la O and Myers (2022) show that short term earnings expectations help predict returns at a very long (10 year) horizon, consistent with our result in Table 2, panel B. However, this relationship disappears once we control for LTG.

the market $SVIX_t^2$ in column [3] (at one year horizon), (e) 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 [4], (f) the credit spread defined as the log difference between the gross yield of BAA and AAA bonds from the St. Louis Fed in column [5], (g) the Baker et al. (2016) economic policy uncertainty index in column [6], and (h) the Kelly and Pruitt (2013) optimal forecast of aggregate equity market returns in column [7]. The sample period is 1981:12-2015:12. Data is quarterly in column [2], 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.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Dependent Variable: Five-year Return $\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$				
LTG_t	-0.4522[a]	-0.5569[a]	-0.3946[a]	-0.4761[a]	-0.4345[a]	-0.4682[a]	-0.7542[a]
	(0.1033)	(0.1179)	(0.1016)	(0.1198)	(0.1031)	(0.1217)	(0.2648)
X_t	-0.1387	0.1894	0.3852[b]	0.2297	0.1672	0.1945	-0.2800
	(0.1035)	(0.1766)	(0.1782)	(0.1875)	(0.1365)	(0.1994)	(0.2708)
Observations	409	137	193	372	409	409	134
Adjusted R^2	27%	28%	47%	37%	27%	29%	27%
X_t	spc_t	cay_t	$SVIX_t^2$	*Term Spread*$_t$	*Credit Spread*$_t$	*Uncertainty Index*$_t$	*Kelly Pruitt MRP*$_t$

The explanatory power of LTG is extremely robust: its coefficient is fairly stable between -0.4 and -0.5 of a standard deviation in magnitude and highly statistically significant across specifications. Across proxies for time varying returns, the spc and cay measures are themselves insignificant and do not add explanatory power. $SVIX^2$ adds substantial explanatory power, but in a way essentially orthogonal to LTG_t: the R^2 of $SVIX^2$ alone is 19%.

Thus, LTG_t does not proxy for existing measures of time varying returns, validating it as a measure of expectations. Columns (4)-(7) show that expectations data have strong explanatory power compared to standard macro predictors of stock returns as well. None of these predictors is statistically significant once we control for LTG_t and the gain in R^2 compared to Table 2 is modest. Expectations of future fundamentals are a promising driver of stock returns.[13]

As a second test for the ability of LTG_t to capture changes in beliefs about earnings growth as opposed to changes in required returns, we analyse LTG_t revisions. If analysts mechanically fit

[13] These results are robust to considering shorter horizons (1 and 3 years) and also to the inclusion of other measures of required returns (Appendix C, Table C.1).

LTG from prices, producing a spurious correlation with required returns, then changes in LTG_t should be primarily explained by past or expected stock returns. If instead analysts revise LTG by adjusting their previous forecast based on fundamental news, measures of past fundamentals should be a more important driver of changes in LTG, even after controlling for past or expected stock returns. Because LTG_t may also be updated based on intangible news, such as the arrival of new technologies, we should not expect past fundamentals to account for 100% of its revisions.

Table 4 reports, in column (1), the regression of the one year revision ΔLTG_t on LTG_{t-1} and on earnings surprises relative to cyclically adjusted earnings, $e_t - cae_{t-5}$. The latter proxy of past fundamentals captures periods of sustained earnings growth, which are more relevant to assess long term fundamentals than temporary growth episodes. In column (2) we add stock returns in the past year and one-year ahead expected return from the CFO Survey[14]. In columns (3), (4), and (5) we add the proxies for discount rates we already used in Table 3.

Table 4
Determinants of LTG revisions

We study 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 (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 Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [3], (f) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [4] and (g) the Martin (2013) expected return on the market SVIX^2 in column [5]. Data is monthly (quarterly) in columns [1], [3], and [5] ([2] and [4]). 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)
	Dependent Variable: ΔLTG_t				
LTG_{t-1}	-0.4349[a]	-0.4624[a]	-0.4393[a]	-0.3232[a]	-0.3338[b]
	(0.1616)	(0.1090)	(0.1429)	(0.1187)	(0.1510)
$e_t - cae_{t-5}$	0.3938[a]	0.3006[a]	0.3274[a]	0.3883[a]	0.4663[a]
	(0.0827)	(0.0561)	(0.0770)	(0.0889)	(0.1173)
r_{t-1}		0.0572			
		(0.1023)			
$\mathbb{E}_t^O[r_{t+1}]$		0.1222			
		(0.1367)			

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

X_t			0.2291[a]	-0.0928	0.1459
			(0.0655)	(0.1214)	(0.1754)
Observations	457	76	457	148	193
Adjusted R^2	31%	38%	36%	32%	52%
X_t			spc_t	cay_t	$SVIX_t^2$

In column (1), the negative and highly significant coefficient on LTG_{t-1} shows that LTG_t, while persistent, tends to mean revert. It is sharply revised upward after periods of sustained earnings growth, as evident from the positive and highly significant coefficient on past earnings. These two forces alone account for roughly one third of the variation in LTG revisions.

None of these conclusions changes materially when we control for past and expected returns: in column (2) both coefficients are insignificant and only marginally improve explanatory power. The same is true when in columns (3), (4) and (5) we control for the discount rate proxies spc, cay, and $SVIX^2$. In Appendix C, Table C.2, we also control for the macroeconomic predictors, and the explanatory power of LTG_{t-1} and $e_t - cae_{t-5}$ is confirmed.

Overall, then, LTG_t appears to be a reliable proxy for expectations about long term earnings growth and a strong predictor of future stock market returns. This is consistent with the possibility that LTG_t captures the role of non-rational market expectations of long term fundamentals.[15] The next section characterizes the departures of LTG_t from rationality and directly links them to return predictability. Because in Table 2 only LTG_t reliably predicts future returns, in the rest of the paper we focus on this measure of expectations.

3. Expectations and Stock Returns

To organize the analysis, we lay out a reduced form model of beliefs that nests the two 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

[15] The notion that LTG_t proxies for long term aggregate fundamentals is consistent with work showing that analyst earnings expectations are a stronger predictor of aggregate investment than stock market based measures of firm-level q (Cummins, Hasset, Oliner 2006).

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.

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

$$g_{t+1} = \mu g_t + v_{t+1}, \tag{6}$$

where is v_{t+1} is an i.i.d. Gaussian shock with mean zero and variance σ_v^2 and $\mu \in [0,1]$. We make this assumption for tractability. In a previous version of the paper we showed that our key results hold under a general covariance stationary process (BGLS 2020).

The shock v_{t+1} captures tangible news arriving at $t+1$ such as earnings surprises, proxied for instance 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. In the latter case the shock to earnings growth is $v_{t+1} = \tau_{t+1} + \eta_t$, where τ_{t+1} is tangible news, η_t intangible news, and the variance of v_{t+1} reflects the two components $\sigma_v^2 = \sigma_\tau^2 + \sigma_\eta^2$. Daniel and Titman (2006) stress the role of intangible news by showing that returns cannot be predicted using past fundamentals. Nagel and Xu (2019) instead show that a measure of five years dividend growth helps predict future returns. We are agnostic about the source of news. By using expectations data we capture both tangible and intangible news. In Appendix E, we show that expectations data have considerable explanatory power even controlling for past fundamental growth.

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. Formally, at time t expectations about dividend growth at $t+s$ are given by:

$$\widetilde{\mathbb{E}}_t(g_{t+s}) = \mathbb{E}_t(g_{t+s}) + \mu^{s-1}\epsilon_t, \tag{7}$$

where $s > 1$ is the forecast horizon, $\mathbb{E}_t(g_{t+s}) = \mu^{s-1}(\mu g_t + \eta_t)$ is the rational forecast based on (6). The believed distribution of g_{t+s} is rational up to the time varying shift $\mu^{s-1}\epsilon_t$.

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 4, namely the negative correlation between LTG_t revisions and lagged forecast LTG_{t-1}, i.e., $cov[\widetilde{\mathbb{E}}_{t+1}(g_{t+s}) - \widetilde{\mathbb{E}}_t(g_{t+s}), \widetilde{\mathbb{E}}_t(g_{t+s})] < 0$. 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 compared to the rational case. 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 (7) is a special case of the diagnostic expectations model (Bordalo et al. 2018) with a slow-moving benchmark distribution in which overreaction to past news decays exponentially at rate (ρ/μ).[16]

We use Equation (7) to develop our empirical strategy: a test detecting departures from rationality, in particular discriminating between under and overreaction of expectations to news ($\theta \lesseqgtr 0$), and a second test linking such departures from rationality to return predictability.

3.1 Model Predictions

Departures from rationality are usually detected by showing that forecast errors are predictable using information available at the time the forecast was made. To discriminate between over and underreaction to news, Coibion and Gorodnichenko (2015) propose to predict the forecast error, defined as the realization minus the forecast, using the current forecast revision. If beliefs overreact, a positive forecast revision (which captures good news) should predict future

[16] Bianchi and Ilut (2022) study the implications of "slow moving" diagnostic expectations in a New Keynesian macro model. The assumption that expectations shocks depend on fundamental shocks rules out the non-fundamental belief distortions of noise trader models (Black 1986, DeLong et al. 1990a). Such distortions can be easily introduced in the analysis and capture an extreme form of overreaction, in which beliefs react to wholly irrelevant factors.

disappointment (negative forecast error). If instead beliefs under-react, a positive forecast revision (which again captures good news) should predict a future positive surprise (positive forecast error). One advantage of this test is that the forecast revision captures the agent's reaction to any news, so it does not require the econometrician to measure the agent's information set directly.

We develop a forecast error predictability test based on this logic. Because in Equation (7) belief distortions are persistent, in our model systematic forecast errors do not just reflect the reaction to current news, but also that to past news. As a result, forecast errors can be predicted based both on the current forecast revision $\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s})$, as in the Coibion and Gorodnichenko test, and on the lagged forecast $\widetilde{\mathbb{E}}_{t-1}(g_{t+s})$. The lagged forecast captures precisely the distorted reaction to past news. We then obtain the following result:

Proposition 1. *Under Equation (7), the error predictability regression:*

$$g_{t+s} - \widetilde{\mathbb{E}}_t(g_{t+s}) = \beta_0 + \beta_1\left[\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s})\right] + \beta_2\widetilde{\mathbb{E}}_{t-1}(g_{t+s}) + z_{t+s}, \tag{8}$$

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

In our model the Coibion and Gorodnichenko test retains its validity, as evident from the fact that $\beta_1 < 0$ is equivalent to overreaction, $\theta > 0$. In addition, due to the persistence of beliefs, overreaction implies negative predictability of forecast errors from the lagged forecast.

Using Equations (1), (3) and (7) we obtain Proposition 2, which links systematic forecast errors in earnings growth to return predictability.

Proposition 2 *Define the following linear combination of tangible and intangible news $\omega_{t+1} = \left(\frac{1+\alpha\theta}{1-\alpha\mu}\right)\tau_{t+1} + \alpha\left(\frac{1+\theta}{1-\alpha\mu}\right)\eta_{t+1}$. Then, the realized return at t+1 is given by:*

$$r_{t+1} = r - \left(\frac{1-\alpha\rho}{1-\alpha\mu}\right)\epsilon_t + \omega_{t+1}. \tag{9}$$

The realized stock return depends on news ω_{t+1}, which affect both the rational revision of expectations and belief distortions, but also on the past belief distortion ϵ_t. This creates a link with

Proposition 1, because ϵ_t also determines the systematic forecast error for earnings growth at $t + 1$, i.e., $\mathbb{E}_t[g_{t+1} - \widetilde{\mathbb{E}}_t(g_{t+1})] = -\epsilon_t$. Thus, current excess optimism about future earnings growth, $\epsilon_t > 0$, predicts both negative forecast errors and low returns. Intuitively, excess optimism causes the current stock price to be inflated. At $t + 1$ beliefs are systematically disappointed, causing a downward price correction and hence a low realized return.

Propositions 1 and 2 outline a strategy to jointly predict forecast errors and future returns using the current forecast revision $\widetilde{\mathbb{E}}_t(g_{t+s}) - \widetilde{\mathbb{E}}_{t-1}(g_{t+s})$ and the lagged forecast $\widetilde{\mathbb{E}}_{t-1}(g_{t+s})$.

Prediction 1 *If LTG_t overreacts, $\theta > 0$, forecast errors in earnings growth and future stock returns should be negatively predicted by the current LTG_t revision and by LTG_{t-1}. If LTG underreacts, $\theta < 0$, the LTG_t revision positively predicts forecast errors and returns.*

3.2 Predictability of Aggregate Stock Returns

Table 5 tests Prediction 1. Column (1) estimates Equation (8): 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: 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.

Table 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 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]. 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 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)
	Dependent Variable:		
	$\Delta_5 e_t/5 - LTG_t$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$

ΔLTG_t	-0.8407[a]	-0.6403[a]	
	(0.1533)	(0.0764)	
LTG_{t-1}	-0.2157	-0.5252[a]	
	(0.1369)	(0.0864)	
$\widehat{\Delta_5 e_{t+5}/5} - LTG_t$			0.8460[a]
			(0.2501)
Observations	397	397	397
Adjusted R^2	25%	31%	
Montiel-Pflueger F-stat			10.9

Column 1 shows that beliefs overreact, $\theta > 0$. Upward revisions of LTG_t predict future disappointment, suggesting that beliefs become too optimistic when good tangible or intangible news arrives. This confirms, at the level of the S&P 500 index, the overreaction of portfolio-level LTG forecasts documented by BGLS (2019) for individual stocks. Here overreaction occurs at the consensus level, since we are using the median LTG forecast. This is a strong result: as shown by Woodford (2003), Coibion and Gorodnichenko (2015), and Bordalo et al. (2020), information frictions are a powerful force toward detecting consensus underreaction even when individual forecasters overreact. Column 1 also shows that a higher lagged forecast LTG_{t-1} is associated with a lower forecast error. In this specification, the association is not significant at conventional levels, but all other results show statistical significance for LTG_{t-1}.[17]

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 is consistent with the mechanism of the model. Overreaction to news received during the current year causes excessive upward LTG_t revisions, high ϵ_t, and hence an excessive stock market boom at t. This predicts future 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) shows that the link between predictable forecast errors and future returns holds empirically: periods of excess pessimism in which future forecast errors in LTG_t are predictably high (reality is above expectations) are systematically followed by high stock returns. Conversely,

[17] Appendix C, Table C.3 shows that LTG_{t-1} significantly predicts 3- and 7-year forecast errors, as well as returns.

periods of excess optimism in which future forecast errors are predictably low (reality is below expectations) are systematically followed by low returns.

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 between 19 and 25 months' worth of returns over five years.

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, improving over the specification of Table 1, in which only the current forecast LTG_t is used as a predictor. The explanatory power of expectations is also much higher than that of past fundamentals. Our proxy for fundamentals in the past five years, $e_t - cae_{t-5}$, predicts lower five year returns but the R^2 of this regression is 13%, less than half the explanatory power of expectations data.

In sum, overreaction of long term expectations, measured by systematic forecast errors in earnings growth, can help account for excess stock market volatility and return predictability.

3.3 Predictability of Firm Level Stock Returns

The results in Table 5 might be influenced by a few outlier episodes, such as the internet bubble. One way to address this concern is to consider the link between belief overreaction and return predictability at the firm level. Indeed, Prediction 1 should also hold for an individual firm with fundamentals that follow an AR(1) process as in (6) and with expectations about them that satisfy (7). Empirically, at the firm level we can control for time dummies, purging the effects of common shocks (including potential common shocks to discount rates). We can also include firm fixed effects, which control for constant differences in average returns across firms.

Table 6 estimates the specifications of Table 5 at the firm level, 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]

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 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 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]. Except $\Delta LTG_{i,t}$, all variables are standardized. 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)
	Dependent Variable:		
	$\Delta_5 e_{i,t}/5 - LTG_{i,t}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$
$\Delta LTG_{i,t}$	-0.3286[a]	-0.1773[a]	
	(0.0248)	(0.0409)	
$LTG_{i,t-1}$	-0.3626[a]	-0.2163[a]	
	(0.0256)	(0.0446)	
$\widehat{\Delta_5 e_{i,t}/5 - LTG_{i,t}}$			0.5768[a]
			(0.0919)
Observations	371,571	371,571	371,571
Adjusted R^2	4%	1%	
Kleibergen-Paap F-stat	.	.	101.8
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Instrument	.	.	$LTG_{i,t-1}, \Delta LTG_{i,t}$

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 $LTG_{i,t-1}$, both in line with the aggregate results. Column (2) confirms, at the firm 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.

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).

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 plays a role at the level of both the aggregate market and individual firms, underscoring the generality of the belief overreaction mechanism.

4. Return Predictability in the Cross Section

Return predictability is not only an aggregate phenomenon. Decades of asset pricing research have unveiled differences in average returns across stocks grouped based on observed characteristics such as the book to market ratio, profitability, etc. Some view these return differences as reflecting departures from market efficiency, others as reflecting differences in the extent to which stocks are exposed to different sources of fundamental risk, or "factors".

Critically, much cross-sectional return 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). Cochrane (2011) argues that such systematic comovement supports the fundamental risk explanations. The argument is that, when a certain risk materializes, all stocks highly exposed that risk are affected. This argument assumes that behavioral factors are unsystematic, so they cannot explain comovement, in contrast to work emphasizing the importance of systematic psychological factors (DeLong et al 1990, Shleifer and Summers 1990, Kozak Nagel Santosh 2018).

Expectations data allow us to empirically assess this debate. We just showed that there is a systematic source of over-optimism, LTG_t, that predicts both disappointment in earnings growth forecasts and aggregate stock returns. Can LTG_t also shed light on the comovement of returns in the cross section and link it to forecast errors? 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.

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 quantitatively account for this finding. In their analysis, however, there is no systematic driver of these average return differences, and it is not obvious that such a driver should exist. The average return differential could be entirely due to idiosyncratic overpricing of high LTG stocks compared to low LTG ones.

Consider, then, whether the future return differential between high and low LTG firms can be predicted using current optimism about aggregate long term earnings growth. We regress the return of portfolios of stocks sorted based on LTG on our proxies for aggregate over-optimism, 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.

In Table 7, column (1) reports the regression results for the five year log return of 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 return on 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. Here we adopt 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), (5) and (6) 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 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]-[6]. 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 $t+5$, $\ln(Mkt_{t,t+5})$, (d) the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [4], (e) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [5], and (f) the Martin (2013) expected return on the market $SVIX^2$ in column [6]. 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	0.2335[c]	-0.7946[a]	0.9878[a]	0.9207[a]	0.9119[a]	1.0709[a]
	(0.1414)	(0.1779)	(0.1991)	(0.1895)	(0.2045)	(0.2928)
LTG_{t-1}	0.3445[c]	-0.4515[a]	0.6408[a]	0.5990[a]	0.5767[a]	0.5150[a]
	(0.1831)	(0.0851)	(0.1462)	(0.1356)	(0.1403)	(0.1157)
$\ln(Mkt_{t,t+5})$	0.8461[a]	0.5177[a]	-0.2691[a]	-0.2451[b]	-0.3241[a]	-0.2396
	(0.1037)	(0.0958)	(0.0973)	(0.0958)	(0.1228)	(0.1962)
X_t				0.1875	0.1937	0.0370
				(0.1371)	(0.1294)	(0.1623)
Observations	397	397	397	397	133	193
Adjusted R^2	69%	83%	70%	71%	71%	73%
X_t				spc_t	cay_t	$SVIX_t^2$

There is a strong systematic variation in the LTG return spread. Part of it is accounted for by the contemporaneous market return, as in CAPM. When the market does well, both high and low LTG firms do better contemporaneously, as captured by the positive coefficients on the market return in columns (1) and (2). However, the return of the $HLTG$ portfolio is more procyclical than that of the $LLTG$ portfolio, so the PMO return loads negatively on the market.[19]

Crucially, a large chunk of variation in the return of the LTG portfolios is captured by a new source of systematic variation: past belief optimism. In column (1), good news about long term earnings growth, reflected in high beginning of period LTG_t revisions, is followed by higher returns for $LLTG$ stocks. The same holds when the lagged forecast LTG_{t-1} is high. In contrast, times of high aggregate optimism are followed by sharply lower returns of the $HLTG$ portfolio, as shown in column (2). These co-movements account, in column (3) for the dynamics of the PMO portfolio: times of aggregate optimism are followed by sharply higher portfolio returns. These are 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. The explanatory power of the LTG measures is high. The model in column (3) accounts for 70% of the time variation in the LTG return spread, compared to only 25% the contemporaneous market return alone. Measures of discount rates play no role in explaining the data (see columns (4), (5) and (6)).

In sum, measured expectations of aggregate long-term earnings growth predict systematic co-movement among $HLTG$ stocks and among $LLTG$ stocks. This fact is consistent with the idea that during times of aggregate optimism $LLTG$ stocks are undervalued compared to $HLTG$ stocks. It is not obvious, however, where the predictive power of LTG_t comes from. As we saw in Section 3, past aggregate optimism is already reflected in the realized market return, for which we control in Table 7. Where does the additional predictive power of the LTG proxies come from?

To address this issue, we extend our model to allow for firm heterogeneity. For simplicity, we abstract from intangible news by setting $\eta_t = 0$, but this is not critical (see footnote 22). Suppose there are many firms i, each of which exhibits AR(1) dividend growth:

$$g_{i,t+1} = \mu g_{i,t} + v_{i,t}, \tag{10}$$

[19] 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*.

As in Equation (7), expected growth at horizon $s \geq 1$ for firm i is believed to be:

$$\widetilde{\mathbb{E}}_t(g_{i,t+s}) = \mathbb{E}_t(g_{i,t+s}) + \mu^{s-1}\epsilon_{i,t}. \tag{11}$$

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 aggregate 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. We can 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 can create differential exposure of firms to aggregate waves of optimism and pessimism.[20] One implication of this assumption is that the firm level belief distortion is 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). For instance, if the market is optimistic due to the rapid growth of some high-tech firms, then such optimism may contaminate other high-tech firms. Such excess optimism may in part reflect higher fundamentals (high v_i), and in part mere similarity, which increases θ_i for given v_i. The distinction between these two effects is key for understanding returns.

Using Equation (1) the realized stock return for firm i at $t+1$ is given by:

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

By plugging in firm specific price dividend ratios as in Equation (3) we obtain:

$$r_{i,t+1} = r_i + \left(\frac{v_i + \alpha\theta_i}{1-\alpha\mu}\right)v_{t+1} - \left(\frac{1-\alpha\rho}{1-\alpha}\right)\frac{\theta_i}{\theta}\epsilon_t, \tag{12}$$

[20] The simplifying assumption that firms perfectly comove with the market neglects idiosyncratic fundamental and expectations shocks but makes the model tractable. As we show in Appendix A, this assumption is not necessary.

where θ is the extent of market overreaction and r_i is the firm-specific required return.[21] The firm loads on aggregate news v_{t+1} but crucially it also loads on past aggregate optimism ϵ_t with firm specific coefficients. To see the implications for the return regressions in Table 7, we can use Equation (9) to rewrite the realized return for firm i in (12) as follows:

$$r_{i,t+1} - r_i = \left(\frac{v_i + \alpha\theta_i}{1 + \alpha\theta}\right)(r_{t+1} - r) - \left(\frac{1 - \alpha\rho}{1 - \alpha}\right)\left(\frac{\theta_i - \theta v_i}{\theta + \alpha\theta^2}\right)\epsilon_t, \tag{13}$$

The firm's realized return depends on the realized market return $(r_{t+1} - r)$ and on past excess optimism ϵ_t according to firm specific coefficients. There are two cases to consider.

First, if a firm's exposure to aggregate optimism is pinned down purely by its exposure to market fundamentals, $\theta_i = \theta v_i$, the model boils down to the CAPM. That is, the return of firm i loads with coefficient v_i on the market return, which is then the only source of systematic return movements. In this case, even if expectations are non-rational, the market return fully incorporates both the news and the unwinding of past over-optimism. If $\theta_i = \theta v_i$, then, non rational beliefs do not account for the additional predictive power of the LTG proxies in Table 7. Even though the aggregate market displays excess volatility and return predictability, the cross section is correctly priced in terms of market exposure.

If instead firms overreact more or less than warranted by their exposure to fundamentals, $\theta_i \neq \theta v_i$, then 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 comparative overreaction to past shocks. This has two consequences. First, the returns of firms that disproportionally overreact, $\theta_i - \theta v_i > 0$, exhibit a stronger contemporaneous comovement with the market, due to stronger overreaction to contemporaneous news v_{t+1}. Second, the same firms are disproportionally inflated during periods in which aggregate excess optimism ϵ_t is high. As a result, they exhibit a stronger reversal in the future, which implies that their future returns

[21] 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: $r_i = r_f + v_i(r - r_f)$.

are predictably lower during these times. The reverse is the case for firms that exhibit a disproportionally small overreaction $\theta_i - \theta v_i < 0$.[22]

We can now go back to the return regressions in Table 7. 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. Our model accounts for Table 7 as long as high LTG firms exhibit disproportionate overreaction compared to low LTG firms, formally if $(\theta_H - \theta_L) > \theta(v_H - v_L)$. In this case, good aggregate news cause excess optimism for high LTG firms compared to low LTG firms. As a conseqence, during times of high aggregate optimism high LTG firms are overvalued compared to low LTG firms. As a result, high LTG firms subsequently experiences low realized returns compared to low LTG firms, causing the future PMO spread to be high.[23]

To probe whether this mechanism is at play, we can look at the behavior of forecast errors. Recall that in our model the belief distortion for firm i (which is inversely related to the forecast error) is given by:

$$\epsilon_{i,t} = \left(\frac{\theta_i}{\theta}\right)\epsilon_t. \tag{14}$$

Excess optimism 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. This has the following implication.

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 forecast revision ΔLTG_t and lagged forecast LTG_{t-1}. If the two portfolios are similarly exposed to fundamental risk, $v_L \approx v_H$, the same LTG proxies predict a higher PMO spread.*

---

[22] The presence of intangible news simply adds to Equation (13) a third factor capturing contemporaneous aggregate intangible news η_{t+1}. For simplicity, we omit this factor.

[23] In particular, the positive loading of the $LLTG$ return on the expectations proxies is consistent with the stronger case in which low LTG firms are undervalued relative to the market, $\theta_L - \theta v_L < 0$, while the positive loading of the $HLTG$ return on the expectations proxies suggests that high LTG firms are overvalued relative to the market.

If Table 7 offered evidence related to the second part of this prediction, namely the PMO spread widens after periods of high aggregate optimism, the behavior of forecast errors allows us to assess the first part of the prediction. This is done in Table 8. 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 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 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. As a result of these patterns, the PMO LTG portfolio exhibits systematically positive earnings growth surprises after periods of aggregate excess 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, $HLTG$.

We can more precisely connect returns in Tables 7 and errors in Table 8 using our model. The positive predictability of PMO forecast errors in Table 8 points to excess pessimism about $LLTG$ firms compared to $HLTG$ ones in good times, $(\theta_H - \theta_L) > 0$. The positive predictability of PMO LTG returns in Table 7 suggests, as seen before, 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 $HLTG$ firms overreact more than $LLTG$ firms do, compared to their differential fundamental risk. In fact, the two conditions are identical if these firms are similarly exposed to fundamentals, $(v_H - v_L) \approx 0$. 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 (13) 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.

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 7 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. Intuitively, overreactions to current and past news move the spread in opposite directions. Compared to $LLTG$ firms, contemporaneous overreaction drives up the return of $HLTG$ firms, while disappointment of past overreaction drives it down.

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 the explanatory power for cross sectional return spreads is greatly improved if one adds to the classic market return several other return factors constructed using specific 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 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 (Lakonishok, Shleifer, and Vishny 1994). Our previous analysis, and the logic of Equation (13), suggests one way to assess this possibility: to use expectations data and analyze the differential exposure of portfolio firms to changes in aggregate optimism.

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. 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, consistent with the possibility that the cross-sectional return differentials reflect weaker overreaction, and hence relative undervaluation, of the long arm of the portfolio during good times (compared to the short arm). Comparing Table 9 to Appendix D, Table D.1, where we regress factor returns on the market factor alone, reveals that the explanatory power of the LTG proxies is high: 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. Aggregate excess optimism helps explain cross sectional return co-movement.[24]

[24] The predictive power of LTG_t in Panel A is robust 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 (Appendix D, Table D.2). 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.

As in the case of Prediction 2 for the PMO spread, overreaction makes joint predictions for forecast error predictability, reported in Panel B. Consider forecast errors for 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 belief overreaction for high BM stocks compared to low BM ones, $\theta_{HBM} < \theta_{LBM}$. This is also consistent with the return regression in Panel A, which suggests comparative undervaluation of high BM stocks during periods of aggregate excess optimism. 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 firms that are highly profitable and invest conservatively exhibit less disappointment than firms that are less profitable and invest aggressively, respectively. This is also consistent with the relative undervaluation of firms that are profitable or invest conservatively during times of aggregate excess optimism, as captured by columns (2) and (3) in Panel A.

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.

Proposition 2 and Equation (13) can most naturally account for the results about $HML\ BM$, RMW, and CMA. In these cases, the weaker belief overreaction of the long arm of each portfolio measured in Panel B and its comparative undervaluation measured in 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.[25] In the case of the SMB factor, Equation (13) is consistent with the results in Table 9 if small firms are sufficiently more exposed to fundamental risk compared to large firms $v_S > v_L$. In this case, it is possible that small firms are undervalued

[25] According to (13), 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.

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 they disappoint after good time because they display larger absolute overreaction $\theta_S > \theta_L$.[26] The SMB factor is not as intuitively accommodated by our mechanism.

Overall, measured beliefs about future fundamentals shed new light on some leading cross sectional return anomalies via the following mechanism: beliefs about low book to market, low profitability, and aggressively investing firms overreact more to aggregate conditions than those of high book to market, profitable and conservatively investing ones. This leads the latter firms to be aggressively undervalued during good times, leading to higher future returns.[27]

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.

[26] Specifically, one needs that $0 < (\theta_S - \theta_B) < \max[\theta(v_S - v_B), (v_B - v_S)/\alpha]$.

[27] 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. Our results offer an alternative view: long term expectations overreact, so high *LTG* firms are overvalued and yield poor returns going forward. This view entails a key new prediction: returns on the factors are (partially) linked to errors in long term growth forecasts, and these are in turn predictable from fundamental aggregate shocks. In particular, this may help 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.

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, beliefs about long term growth have remarkable predictive power for aggregate returns even when we control for leading proxies for required returns. 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 to 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, *SSRN* wp 3045658.

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*, forthcoming.

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

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

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

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 in which $\gamma_n = \gamma^{n-1}$, in which $\gamma \in (0,1)$ parameterizes the speed of the decay of overreaction to past news. In this case we can write:

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

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

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

We now show that this equation is equivalent to (7) for a suitable choice of the decay parameter γ.

By iterating backward the belief distortion ϵ_t, we can write Equation (7) as:

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

By plugging the expression for the expectations shock we obtain:

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

In the AR(1) process of Equation (6), 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 (6) 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 (6) 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 (7) 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 (8) and to characterize their sign as a function of θ. Using Equation (7), 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:

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

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

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

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

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

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

From this, 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. At this effect, 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, starting from the fact that $G\left(-(1-\rho^2)\right) > 0$, that $G(\theta) > 0$ for all $\theta \geq -(1-\rho^2)$.

To demonstrate 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 verify 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, express $G(\theta)$ as follows:

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

Note, then, that $G(\theta)$ has the following 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:

$$\begin{aligned} f'(\theta) - h'(\theta) &= \frac{\rho(\mu-\rho)}{1-\rho^2}\cdot\left(\frac{2\theta}{1-\rho^2}+\frac{2}{1-\mu\rho}\right)-\frac{\rho(\mu-\rho)}{1-\rho^2}\cdot\left(1+\frac{1}{1-\mu\rho}+\frac{2\theta}{1-\rho^2}\right) \\ &= \frac{\rho(\mu-\rho)}{1-\rho^2}\cdot\left(\frac{1}{1-\mu\rho}-1\right) \geq 0. \end{aligned}$$

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 $h\left(-(1-\rho^2)\right) \geq 0$, 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:

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

By 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:

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

Using Equation (7), we obtain that the forecast revision is equal to:

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

Which is in turn equal to:

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

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.

Stock Returns in the Cross Sections (also with firm level idiosyncratic shocks).

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:

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

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:

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

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 (13) with the addition of the two idiosyncratic fundamental and belief distortion components.

Firm Level Required Return and the CAP M

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

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

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 the latter component, 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 stay unchanged at $t+1$. As a result, investors' believed distribution of the forecast revision $\widetilde{\mathbb{E}}_{t+1}\left(g_{i,t+1+s}\right) - \widetilde{\mathbb{E}}_t\left(g_{i,t+1+s}\right)$ 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. Specifically, 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 preferences are mean variance, and the risk free rate is r_f, the security market line is:

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

which implies:

$$r_i - r_f = v_i\left(r - r_f\right).$$

Appendix B. Robustness and Further Results on Price Indices

In this Appendix, we collect several results that complement the analysis of the synthetic price indices in Section 2.

Due to data availability, our analysis focused on expectations of earnings growth. We start by extending our analysis to expectations of dividend growth. We gather monthly data on stock market analyst forecasts for S&P500 firms from the IBES Unadjusted US Summary Statistics file, focusing on (median) annual forecasts of dividends per share (DPS). Coverage starts on 1/2002 for DPS. We aggregate DPS following the same procedure as for EPS forecasts described in the text, and build synthetic price dividend ratios following:

$$\tilde{p}_t^D - d_t = \frac{k-r}{1-\alpha} + \ln\left(\frac{\mathbb{E}_t^O DPS_{t+1}}{DPS_t}\right) + \sum_{j=1}^{10} \alpha^{j-1}\, \mathbb{E}_t^O \Delta \mathrm{d}_{t+j+1} + \frac{\alpha^{10}}{1-\alpha}\, g \qquad (B.1)$$

where we assume that expectations of long run dividend growth are also described by LTG. The table below shows that this index is highly correlated with the measured price dividend ratio.

Table B.1

We present partial correlations between: (a) the (log) price-dividend ratio, pd_t, (b) the difference between the (log) price index based on dividend forecasts $\tilde{p}_t^D$ (Equation B.1) and log dividends, d_t, (c) the (log) price-earnings ratio, pe_t, (d) the difference between the rational benchmark index based on earnings ($p_t^{RE,E}$, footnote 8) and log earnings e_t, and (e) the difference between the price index based on earnings forecasts, $\tilde{p}_t$, (Equation 4) and log earnings, e_t. The sample period is 1982:12-2020:12. Superscripts: [a] significant at the 1% level, [b] significant at the 5% level, and [c] significant at the 10% level.

	pd_t	$\tilde{p}_t^D - d_t$	pe_t	$p_t^{RE,E} - e_t$
$\tilde{p}_t^D - d_t$	0.5293[a]			
pe_t	0.5829[a]	-0.5558[a]		
$p_t^{RE,E} - e_t$	-0.2337[a]	-0.6670[a]	0.6143[a]	
$\tilde{p}_t - e_t$	0.0580	-0.4578[a]	0.7740[a]	0.8687[a]

Alternative definitions and excess volatility. Here we consider an alternative definition of price $\tilde{p}$ where expectations at time t of growth beyond year $t+5$ is inferred by applying the observed decay of observed cyclically adjusted earnings to LTG_t. Regressing $caeps_t - caeps_{t-t}$ on

$caeps_{t-5} - caeps_{t-10}$ yields a slope coefficient of roughly 0.4. Thus, for a ten-year forecasting horizon we set:

$$\tilde{p}_t^{10} = e_t + \frac{\tilde{k} - r}{1 - \alpha} + \ln\left(\frac{E_t^O\ EPS_{t+1}}{EPS_t}\right) + \sum_{j=1}^{5} \alpha^{j-1}\ E_t^O \Delta \mathrm{e}_{t+j+1}$$

$$+ \sum_{j=6}^{10} \alpha^{j-1} \ln(1 + 0.4 * E_t^O \Delta \mathrm{e}_{t+5}) + \frac{\alpha^{10}}{1 - \alpha}\ g_{10}. \qquad (B.2)$$

and similarly for a 15 and 20-year forecasting horizon, as well as for an alternative dividend based index $p_t^{D,10}$ (where long term growth is assumed to be described by LTG). Table B.2 shows the results.

Table B.2

Panel A reports the standard deviation of one-year change in: (a) the log of the price of the S&P500 index, Δp_t, (b) the one-year change in the index based on dividend forecasts, $\Delta \tilde{p}_t^D$ (analogue of equation 4 for dividends), and (c) the one-year change in the alternative index based on dividend forecasts over 10, 15, and 20-year horizons ($\Delta \tilde{p}_t^{D,10}$, $\Delta \tilde{p}_t^{D,15}$, $\Delta \tilde{p}_t^{D,20}$, see Equation B.2). The sample period ranges from 11/2006 to 12/2020. Panel B reports the standard deviation of one-year change in: (a) the log of the price of the S&P500 index, Δp_t, (b) the one-year change in the index based on earnings forecasts, $\Delta \tilde{p}_t$ (equation 4), and (c) the one-year change in the alternative index based on earnings forecasts over 10, 15, and 20-year horizons ($\Delta \tilde{p}_t^{10}$, $\Delta \tilde{p}_t^{15}$, $\Delta \tilde{p}_t^{20}$see Equation B.2). The sample period ranges from 12/1982 to 12/2020.

Panel A: Dividend based synthetic price

	Δp_t	$\Delta \tilde{p}_t^D$	$\Delta \tilde{p}_t^{D,10}$	$\Delta \tilde{p}_t^{D,15}$	$\Delta \tilde{p}_t^{D,20}$
σ	14.8%	19.0%	15.8%	16.7%	17.0%

Panel B: Earnings based synthetic price

	Δp_t	$\Delta \tilde{p}_t$	$\Delta \tilde{p}_t^{10}$	$\Delta \tilde{p}_t^{15}$	$\Delta \tilde{p}_t^{20}$
σ	14.8%	14.6%	12.4%	13.0%	13.2%

Volatility of cointegrated series. Finally, following Campbell and Shiller (1987) we assess the volatility of the cointegrated series $P_t - \frac{D_t}{r}$ for different measures of prices. The first column in Table B.3 reproduces the volatility of changes in log prices from Table 1. The remaining columns assess the volatility of the co-integrated series $P_t - \frac{D_t}{r}$ for different measures of prices.

Table B.3

Panel A reports the standard deviation of the difference between: (a) the S&P index (P_t) and the present value of dividends D_t/r in column [1], (b) the rational benchmark index based on dividends ($P_t^{RE,D}$, see footnote 10 for the log earnings version) and D_t/r in column [2], and (c) the price index based on dividend forecasts, $\tilde{P}_t^D$ (see equation 4 for the log earnings version) and D_t/r in column [3]. Panel B repeats the calculations in Panel A but price indices are based on earnings rather than dividends. Panel C repeats the calculation in Panel B but differences are based on the present value of earnings, E_t/r, rather than dividends, D_t/r . The sample period ranges from 11/2005 to 12/2020 in Panel A and from 12/1981 to 12/2020 in Panels B and C. We use $r = 8.48\%$, the average return of the S&P during the period 1981:12-2020:12.

Panel A: Dividend-based indices relative to the present value of dividends

	(1)	(2)	(3)
	$P_t - \frac{D_t}{r}$	$P_t^{RE,D} - \frac{D_t}{r}$	$\tilde{P}_t^D - \frac{D_t}{r}$
σ	563	175	664

Panel B: Earnings-based indices relative to the present value of dividends

	(1)	(2)	(3)
	$P_t - \frac{D_t}{r}$	$P_t^{RE} - \frac{D_t}{r}$	$\tilde{P}_t - \frac{D_t}{r}$
σ	658	273	482

Panel C: Earnings-based indices relative to the present value of earnings

	(1)	(2)	(3)
	$P_t - \frac{E}{r}$	$P_t^{RE} - \frac{E_t}{r}$	$\tilde{P}_t - \frac{E}{r}$
σ	476	156	294

We next reproduce Table 2, which examines the predictability of returns on the basis of expectations, using excess (as opposed to raw) returns.

Table B.4

This table examines the association between excess returns and forecasts for growth in earnings at different horizons. 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. 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)
	$xret_{t+1}$	$\sum_{j=1}^{3} \alpha^{j-1} xret$	$\sum_{j=1}^{5} \alpha^{j-1} xret$
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%

We next reproduce Figure 1, which examines the expectation based price index (Equation 5), adjusting for inflation.

Figure B.1
Prices adjusted for inflation

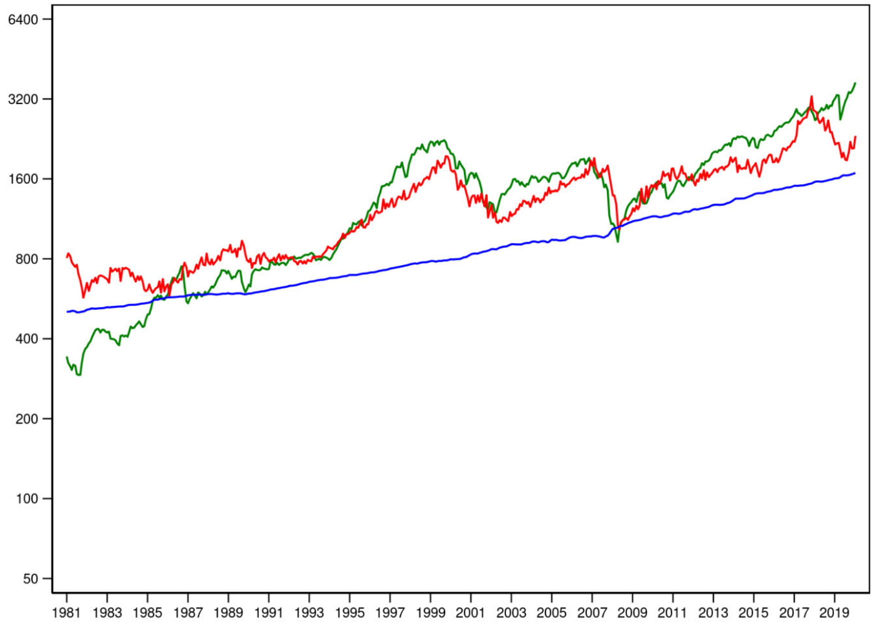

Figure B.1. We plot the S&P500 index (green line), the rational benchmark index (p_t^{RE}, see footnote 10, blue line) and our benchmark price index based on earnings expectations ($\tilde{p}_t$ from equation [4], red line). All values are adjusted for inflation using the CPI index.

Appendix C. Robustness and other results on aggregate overreaction

In this Appendix, we collect robustness checks regarding the analysis of Sections 2 and 3. We start by extending the analysis of Table 2, on the predictability of returns from LTG, to shorter horizons (1 and 3 years) and by including other measures of required returns.

Table C.1
Return predictability and proxies for required returns 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 The independent variables are: (a) the forecast for earnings growth in the long run LTG_t, (b) the Campbell and Cochrane (1999) surplus consumption ratio, spc_t, in column [1], (c) the Lettau and Ludvigson (2001) consumption-wealth ratio, cay_t, in column [2], (d) the Martin (2013) expected return on the market, $SVIX_t^2$, in column [3], (e) 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 [4], (f) the credit spread defined as the log difference between the gross yield of BAA and AAA bonds from the St. Louis Fed in column [5], the Baker et al. (2016) economic policy uncertainty index in column [6], and (g) the Kelly and Pruitt (2013) optimal forecast of aggregate equity market returns in column [7]. 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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: One-year Return r_{t+1}						
LTG_t	-0.3421[b]	-0.3755[a]	-0.3193[a]	-0.3743[a]	-0.2891[a]	-0.3228[a]	-0.2952[c]
	(0.1334)	(0.1272)	(0.1066)	(0.1120)	(0.1109)	(0.0954)	(0.1544)
X_t	0.1065	0.2395	0.1504	-0.1727	0.0675	0.0176	0.0448
	(0.1352)	(0.1561)	(0.1104)	(0.1348)	(0.1544)	(0.1245)	(0.1587)
Observations	409	137	193	409	409	372	134
Adjusted R^2	10%	14%	12%	11%	9%	11%	9%
X_t	spc_t	cay_t	$SVIX_t^2$	*Term Spread*$_t$	*Credit Spread*$_t$	*Uncertainty Index*$_t$	*Kelly Pruitt MRP*$_t$

Panel B: Predictability of three-year returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Three- year Return $\sum_{t=1}^{3} \rho^t\, r_{t+j}$						
LTG_t	-0.5156[a]	-0.5811[a]	-0.4770[a]	-0.4356[a]	-0.4826[a]	-0.4862[a]	-0.6561[a]
	(0.1063)	(0.1378)	(0.1391)	(0.1626)	(0.1244)	(0.1161)	(0.0900)
X_t	0.0623	0.3234[c]	0.1662	0.1342	0.0502	0.1201	-0.1646
	(0.0694)	(0.1946)	(0.2621)	(0.2115)	(0.1885)	(0.2241)	(0.1112)

Observations	409	137	193	409	409	372	134
Adjusted R^2	24%	33%	28%	25%	24%	30%	25%
X_t	spc_t	cay_t	$SVIX_t^2$	*Term Spread*$_t$	*Credit Spread*$_t$	*Uncertainty Index*$_t$	*Kelly Pruitt MRP*$_t$

We next extent the results of Table 4, on the determinants of LTG revisions, by considering other measures of required returns.

Table C.2
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], and (f) the Kelly and Pruitt (2013) optimal forecast of aggregate equity market returns in column [4]. 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)
	Dependent Variable: ΔLTG_t			
LTG_{t-1}	-0.2903[b]	-0.2746[b]	-0.4187[a]	-0.5720[a]
	(0.1281)	(0.1337)	(0.1455)	(0.0923)
e_t - cae_{t-5}	0.2400[a]	0.3199[a]	0.3363[a]	0.1544[a]
	(0.0505)	(0.0632)	(0.0727)	(0.0597)
X_t	-0.2642[b]	-0.2079[a]	-0.2413[b]	-0.5268[a]
	(0.1110)	(0.0602)	(0.1094)	(0.0982)
Observations	433	441	432	130
Adjusted R^2	36%	35%	36%	66%
X_t	*Term Spread*$_t$	*Credit spread*$_t$	*Uncertainty Index*$_t$	*Kelly Pruitt MRP*$_t$

We next generalize Table 5 on the predictability of returns from predicted forecast errors, to horizons of 1 and 3 years.

Table C.3
Return predictability and predictable forecast errors

We present results for firm-level IV regressions for firms in the IBES sample. The dependent variable is: (a) the (log) one-year return, $r_{i,t+1}$, in column [1], (b) the discounted value of cumulative (log) return for firm *i* between year *t* and *t+3*, $\sum_{j=1}^{3} \alpha^{j-1} r_{i,t+j}$, in column [2], or (c) the discounted value of cumulative (log) return for firm *i* between year *t* and *t+5*, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$, in column [3]. The independent variable is the forecast error, defined as the difference between (a) the annual 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, $LTG_{i,t}$. We instrument forecast errors using: (a) the one year change in the forecast for firm's growth in long-term growth in earnings, $\Delta LTG_{i,t}$ and (b) the forecast for growth in earnings in the long term made in year *t-1*, $LTG_{i,t-1}$. Except for $\Delta LTG_{i,t}$, variables are normalized to have zero mean and standard deviation equal to 1. 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)
	Dependent Variable:		
	r_{t+1}	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$	$\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$
$\Delta_5 e_{i,t+5}$ / 5 - $LTG_{i.t}$	0.3815[a]	0.5164[a]	0.5768[a]
	(0.0598)	(0.0865)	(0.0919)
Obs	371,571	371,571	371,571
Adj R^2	-1%	8%	20%
KP F-stat	275.8	124.6	101.8
Adj R^2 Reduced Form	1%	2%.	2%
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Instrument	$LTG_{i,t-1}$, $\Delta LTG_{i,t}$	$LTG_{i,t-1}$, $\Delta LTG_{i,t}$	$LTG_{i,t-1}$, $\Delta LTG_{i,t}$

We next confirm that aggregate returns are predictable from recent earnings growth, in line with the reduce form analysis of Proposition 2, and consistent with findings by Nagel and Xu (2019).

Table C.4
Return predictability from recent earnings growth

This table examines the association between returns and past growth in earnings. The dependent variables is 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]. 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%

Finally, we examine the result in Table 5, Column 1, and confirm that one year revisions and lagged LTG predicts forecast errors at 3 and 7 year horizons.

Table C.5
Predictability of forecast errors at different horizons

	(1)	(2)	(3)
	$\Delta_3 e_t/3 - LTG_t$	$\Delta_5 e_t/5 - LTG_t$	$\Delta_7 e_t/7 - LTG_t$
ΔLTG_t	-0.8740[a]	-0.7667[a]	-0.5305[c]
	(0.1500)	(0.1828)	(0.2753)
LTG_{t-1}	-0.3480[a]	-0.2199	-0.4566[a]
	(0.1061)	(0.1350)	(0.1208)
Observations	409	409	388
Adjusted R^2	30%	22%	22%

Appendix D. Robustness and other results on cross sectional returns

We present a benchmark for Table 9, namely a univariate regression of factor returns on the market factor. In Table D.1 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 D.1
Predictability of factor returns and forecast errors

The dependent variables 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 $t+5$, $\ln(Mkt_{t,t+5})$. Panel B reports the adjusted R squared from bivariate regressions using as independent variables: (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}. For comparison purpose, we also report 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
$\ln(Mkt_{t,t+5})$	-0.5108[a]	-0.0971	-0.5305[a]	-0.3903	-0.6512[a]
	(0.1921)	(0.2855)	(0.1366)	(0.3097)	(0.1671)
Obs	397	397	397	397	397
Adj R2	25%	1%	27%	14%	40%

Panel B: Summary of R^2

	(1)	(2)	(3)	(4)	(5)
	PMO	HML	RMW	CMA	SMB
Adjusted R^2 Univariate Regression	25%	1%	27%	14%	40%
Adjusted R^2 Bivariate Regression	46%	50%	20%	52%	49%

Finally, in Table D.2, we add *cay* and its interaction with the market return to the benchmark specification in Table 9.

Table D.2
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], and (d) small-minus-big factor (SMB) in column [4]. 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)
	Dependent Variable: Five-year (log) Return of			
	HML	RMW	CMA	SMB
ΔLTG_t	0.8462[a]	0.1706	0.6653[a]	0.3516[b]
	(0.1317)	(0.2304)	(0.2210)	(0.1595)
LTG_{t-1}	0.7546[a]	0.0651	0.4404[a]	0.5505[a]
	(0.0947)	(0.1402)	(0.0948)	(0.0994)
$\ln(Mkt_{t,t+5})$	0.5781[a]	-0.5109[b]	0.1971	-0.1076
	(0.1758)	(0.2109)	(0.1394)	(0.0797)
cay_t	1.1376[a]	0.4046	1.4857[a]	0.5593[c]
	(0.2730)	(0.3492)	(0.1776)	(0.2949)
$\ln(Mkt_{t,t+5}) * cay_t$	-0.8570[a]	-0.0441	-1.0620[a]	-0.7292[a]
	(0.2870)	(0.3492)	(0.2001)	(0.1978)
Observations	133	133	133	133
Adjusted R^2	68%	36%	66%	65%

Appendix E. 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 10 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 E.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 the market, $\sum_{j=1}^{5} \alpha^{j-1} r_{t+j}$, and for all firms on IBES, $\sum_{j=1}^{5} \alpha^{j-1} r_{i,t+j}$, in columns [1] and [2], respectively. 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. 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) $e_t - cae_{t-5}$. 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.