Investor (Mis)Reaction, Biased Beliefs, and the Mispricing Cycle*

Azi Ben-Rephael† Steffen Hitzemann‡ Yuanyuan Xiao§

November 2022

Abstract

We construct a new measure that captures the disparity between the market reaction to earnings information and the earnings surprise ("Return-Earnings Gap", "*REG*"). High *REG* positively predicts analyst forecast errors and firm mispricing (overvaluation). Analyst forecast errors are particularly increased when *REG* provides confirming information. In turn, *REG* is positively predicted by analyst forecast errors and higher mispricing, leading to a continuation of firm overvaluation over several quarters. Overall, our results reveal how the market's (mis)reaction feeds back into the belief formation of analysts and investors, leading to a slow correction of firm mispricing. A simple structural model explains the predictive power of *REG* for analyst forecast errors and the build-up of mispricing by incorporating the dynamic expectation formation between different agents.

JEL Classification: G00, G12, G14, G40, G41

Keywords: investor beliefs, analysts, analysts' expectations, mispricing, misreaction, anomalies

*We thank Justin Birru, Martijn Boons, John Campbell, Yixin Chen, Anthony Cookson (discussant), Zhi Da, Alexander Hillert, Heiko Jacobs (discussant), Peter Kelly (discussant), Toomas Laarits (discussant), Sophia Zhengzi Li, Yueran Ma, Sean Myers, Simon Rottke, Andrea Tamoni, Marliese Uhrig-Homburg, as well as participants of the 2022 China International Conference in Finance, the 2022 Midwest Finance Association Annual Meeting, the 2022 Financial Management Association Annual Meeting, the 2022 Research in Behavioral Finance Conference, the 2022 TBEAR Network Asset Pricing Workshop, and seminar participants at Rutgers University and Texas A&M University for helpful comments and suggestions.

†Azi Ben-Rephael, Rutgers Business School. Email: abenrephael@business.rutgers.edu.

‡Steffen Hitzemann, University of Houston. Email: hitzemann@bauer.uh.edu.

§Yuanyuan Xiao, Rutgers Business School. Email: yuanyuan.xiao@rutgers.edu.

1 Introduction

The traditional view of financial markets states that the market aggregates dispersed information efficiently and incorporates new information correctly (Hayek, 1945; Grossman, 1976; Holthausen and Verrecchia, 1988). A more modern view, however, recognizes that investors' expectations can be biased,[1] leading to under- or overreaction to information (Barberis et al., 1998).[2] Biases in market participants' expectations also take an important part in the longstanding debate on risk vs. mispricing (e.g., Kozak et al., 2018) and are relevant for explaining anomaly returns. In particular, recent evidence indicates that analyst forecasts are systematically over-optimistic (pessimistic) for anomaly short (long) positions (La Porta, 1996; La Porta et al., 1997; Engelberg et al., 2018; Bordalo et al., 2019; Engelberg et al., 2020).

While analyst errors line up inversely with anomaly returns, a natural question to ask is why such biases and observed mispricing persist over a long time. Specifically, why does the market follow analysts' expectations that are systematically biased? And, the other way round, why do analysts not detect mispricing and adjust their forecasts accordingly? Is it possible that, alternatively, analysts and other investors influence each other in a way that amplifies the formation of their biased beliefs? While a plethora of studies explore the existence of a bias and link it to overall mispricing, there is still little focus on the *interplay* between the expectations formation of different market participants and how it contributes

[1]The formation of biased expectations can stem from various mechanisms and include extrapolative and diagnostic beliefs (Greenwood and Shleifer, 2014; Barberis et al., 2015, 2018; Cassella and Gulen, 2018; Bordalo et al., 2019; Da et al., 2021), confirmation bias (Nickerson, 1998; Pouget et al., 2017; Cookson et al., 2021; Hirshleifer et al., 2021), sticky belief dynamics (Bordalo et al., 2019), and catch-all sentiment (De Long et al., 1990; Barberis et al., 1998; Baker and Wurgler, 2006).

[2]Examples include overreaction at the market level (Greenwood and Shleifer, 2014; Cassella and Gulen, 2018), overreaction in the cross-section of stocks (Bordalo et al., 2019; Da et al., 2021), and underreaction to analysts forecasts and earnings information (e.g., Bernard and Thomas, 1989, 1990; Chan et al., 1996; Barberis et al., 1998; Cready and Gurun, 2010; Hartzmark and Shue, 2018).

to the correction or amplifications of such biases.[3]

A key challenge towards answering these questions is the difficulty of quantifying the extent of market biases relative to the available fundamental information at a given point in time. To do so, we take a non-parametric approach and measure market participants' reaction to earnings information on earnings announcement days. Specifically, we compare the ranking of the return response (i.e., market participants' beliefs) to the ranking of unexpected earnings (i.e., the fundamental information). We call this measure the *R*eturn-*E*arnings *G*ap (*REG*). Specifically, a higher (lower) *REG* would be an indication of a more positive (less positive) response by market participants for a given earnings surprise.

Using *REG*, we provide new empirical evidence on the dynamic interaction between two prominent groups of agents: analysts, who publish forecasts and provide valuable information to market participants, and investors, who trade and reflect their beliefs into stock prices. Ex-ante, *REG* could reflect a rational market reaction to "soft" information released together with the earnings announcement and be unrelated to future analyst forecast errors and mispricing. On the other hand, a higher (lower) *REG* could reflect excessive optimism (pessimism) towards the firm prospects, which in turn also affects analyst forecasts and firm mispricing. Empirically, we find support in favor of the latter. Specifically, we show that the market's initial response to earnings announcements feeds back into analysts' expectations in a way that it increases their earnings forecast errors in the same direction over the subsequent quarters, leading to persistently biased expectations. These dynamics are accompanied by a build-up in mispricing that takes several quarters to correct as can be seen in Figure 1, which

[3]Recent papers explore how agents form their beliefs based on their private and publicly available signals. Gerken and Painter (2019) focus on retail firms and show that analysts incorporate local information into their forecasts, and that analysts rely more on local signals when less firm-wide information is available. Ludvigson et al. (2021) find that that survey participants including professional forecasters over-weight their own beliefs relative to publicly available information. Importantly, these studies do not explore the *interplay* between analysts and investors and its effect on beliefs formation and firm mispricing.

plots what we define as "the mispricing cycle." We also find that analysts' biased expectations and firm mispricing conditions feed back into future market responses to earnings, pointing to a dynamic amplification effect. Overall, using our *REG* measure, we are able to identify new mispricing build-up dynamics that were not explored in previous research.

Our results indicate that analysts fail to disentangle the noise from information contained in the market response on earning announcements days, such that the market's (mis)reaction results in a positive increase in analyst forecast errors (*AFE*s). Indeed, we find that *REG* positively predicts analyst forecast errors over the subsequent quarters, controlling for current *AFE* and a number of additional stock-specific control variables. This predictability suggests that market participants' reaction feeds back into analyst expectations in a way that distorts their expectation formation. The effect is economically meaningful. While Stambaugh et al. (2012) mispricing scores (*MISP*), for example, also have a positive effect on *AFE*, the effect of *REG* is twice as large in magnitude, highlighting the importance of *REG*.

The behavioral literature suggests a host of explanations for predictable analyst biases. One explanation is "confirmation bias" (Pouget et al., 2017; Cookson et al., 2021; Hirshleifer et al., 2021), according to which analysts interpret the market response in favor of their initial expectations. Other explanations include extrapolative or diagnostic expectations (Bordalo et al., 2019), where analysts put more weight on the market's relative overreaction response signal. While it is not the main focus of our study to quantitatively disentangle these potential behavioral explanations, we find evidence suggesting that a confirmation channel is present in the relation between *REG* and *AFE*. Specifically, when considering cases where *AFE* and *REG* are both positive or both negative (that is, *REG* "confirms" *AFE*) and cases where the two are in the opposite direction, we show that the bias in analysts' expectations becomes more pronounced when both signals are in the same direction.

These results apply to average *AFE*s, and there is ample variation in the cross-section of different analysts with respect to their response to *REG*. Such heterogeneity can exist, for example, due to factors that affect the quality of analysts' private signals, which also determines their reliance on public signals. Potential factors include ability, experience, specialization, as well as the attention and effort that analysts pay to a specific stock in their portfolio. Exploring the cross-section of analysts in detail, we find that analysts exhibit a lower (higher) sensitivity to *REG* when they are more (less) industry-concentrated. This result is consistent with more specialized analysts having better private information and being less reliant on (potentially biased) public signals, in line with Kacperczyk et al.'s (2005) findings for mutual funds. Using analyst stock-level forecast accuracy as a "catch-all" proxy for their information quality confirms this intuition, as we find that analysts with a lower (higher) degree of forecast accuracy exhibit a higher (lower) sensitivity to *REG*.

We complete our investigation of *REG*'s effect on *AFE*s by testing and confirming that the observed effects are equally strong on the positive side and on the negative side. Finally, we extend our analysis to analysts' price targets as well as analyst recommendation changes (as explored by Brav and Lehavy, 2003; Jegadeesh et al., 2004; Da and Schaumburg, 2011; Engelberg et al., 2020), and find results consistent with our main findings.

After pinning down how *REG* captures and drives the expectation formation of investors and analysts, we analyze the relation to firm mispricing scores (*MISP*). Consistent with biased beliefs, we find a significantly positive predictive relation between *REG* and *MISP* over the subsequent quarters. Remarkably, the predicted effect of *REG* on *MISP* is more pronounced two quarters ahead than one quarter ahead, and even larger three quarters ahead, resulting in the mispricing cycle documented in Figure 1. *REG* is thus able to capture an early build-up phase in firm mispricing, while previous research mostly identifies the correction phase of mispricing. As an example, analyst forecast errors (*AFE*) also positively predict *MISP*; however, the predictability is quantitatively concentrated one quarter ahead and declines after that. Furthermore, the predicted effect of *REG* on the next quarter *MISP* is over three times larger than that of *AFE*.

We investigate the effect of *REG* on firm mispricing in detail by exploring, first, the cross-section of anomalies with respect to which mispricing scores are computed, and second, variation in *REG*'s effect on mispricing in the cross-section of firms. As part of the first set of tests, we take advantage of van Binsbergen et al.'s (2021) classification of anomalies into *Build-Up* and *Resolution* anomalies.[4] Using van Binsbergen et al. (2021)'s 57 anomalies set and their classification, we construct two new *MISP* variables, one for *Build-Up* and one for *Resolution*. We find that the mispricing cycle captured by *REG* — that is, the positive predictability of mispricing over several quarters, with increasing magnitude — appears and is very pronounced in the set of the *Build-Up* anomalies, in line with *REG* predicting the exacerbation of firm mispricing. In particular, it takes up to 2 years to reach the peak in mispricing for this subset of anomalies. In stark contrast, the predictive relation of *REG* to mispricing scores for *Resolution* anomalies is negative and of much smaller economic significance. Note that while van Binsbergen et al. (2021) use fundamental cash flow information to classify *Build-Up* and *Resolution* anomalies, we are able to rediscover the stark difference between both types of anomalies simply based on *REG*'s predictive relation to them. In the second set of tests, we find that the response of *MISP* to *REG* is stronger for firms with lower analyst coverage, firms with smaller market cap, firms with lower institutional base, and firms with higher analyst dispersion. These findings reconfirm

[4] van Binsbergen et al. (2021) show that anomalies can be classified as "build-up" or "resolution" anomalies, where build-up anomalies capture the creation of mispricing, while resolution anomalies capture the correction of mispricing.

once more that the amplification and reinforcement of biases captured by *REG* is more pronounced when agents' private signals are not very strong and they are susceptible to taking over biased public signals.

We connect and compare our findings to studies such as La Porta (1996) and Engelberg et al. (2018), who show that public information arrival, especially on earnings announcement days, attenuates existing mispricing. That is, anomaly returns are generally negatively related to mispricing, and this relation is particularly pronounced on earnings days as new information arrives and investors correct their biased beliefs. While this research identifies the correction phase of mispricing, *REG* is able to identify an early build-up phase in biased expectations and mispricing. Consistent with that, we find that the anomaly correction (i.e., *MISP* return predictability) attenuates when *REG* is in the same direction of *MISP*. For example, overvalued stocks, captured by high *MISP* scores, exhibit positive returns in the earnings month and subsequent months when *REG* is positive, and this effect dominates the generally negative relation between *MISP* and subsequent returns.

In a last set of tests, we explore what affects *REG*. We find that *REG* is positively predicted by *AFE* and *MISP*, which is consistent with previous research (e.g., Hughes et al., 2008; Frankel and Lee, 1998; So, 2013) and points to an amplification effect between *REG*, *AFE*, and *MISP*. To better understand this dynamic relation, we analyze impulse response functions of *REG*, *AFE*, and *MISP* based on vector autoregressions (VAR). The VAR analysis highlights the importance of *REG* for predicting *AFE* and *MISP* even after controlling for contemporaneous and lagged effects of these variables.

We round off our paper by presenting a simple structural model of beliefs that explains the predictive power of *REG* for analyst forecast errors and the build-up of mispricing that we observe in the data. In the model, investors and analysts dynamically update their expectations of the firms' earnings growth rate. They observe current earnings as well as a private signal, and each type of agent tries to infer the other agent's private signal from public information. In particular, analysts learn about the investors' signal from the market response to earnings, and investors learn about the analysts' signal from published analyst forecasts. When all agents' expectation formation is unbiased, analyst forecast errors and firm mispricing revert quickly and expectations align with firms' actual earnings growth. However, when the investors' belief formation is biased but both types of agents are unaware of this bias, then the bias is propagated and reinforced between the two groups of agents. The reason is that analysts falsely interpret an abnormal market reaction caused by the investors' bias as an informative signal and thus incorporate this bias into their own expectations. Investors, in turn, observe the analysts' revised forecasts and also assume that these were updated due to the analysts' private information; as a result, the investors' own original bias is reinforced when they update their expectations. These dynamics lead to positive predictability of AFEs and to a build-up in mispricing, which is only resolved once the effect of fundamental information from current earnings overweighs the propagation of the bias. The model demonstrates that the dynamic expectation formation between different types of agents is critical for understanding pricing patterns and the persistence of biases in financial markets.

Our paper contributes to two broad strands of literature. The first is the literature on risk, mispricing, and anomaly returns. Previous papers link analysts' biases with anomaly returns, and also show that information arrival pushes anomaly returns in the right direction, consistent with the correction of biased expectations. Our unique and novel contribution to this literature is that we explore the interaction between investors' and analysts' expectations. We show that analysts' and investors' biased expectations positively predict each other, pointing to an amplification effect. In contrast to previous findings, which capture the peak of firm mispricing and its subsequent resolution, we are able to identify an early build-up phase of mispricing. Relatedly, we find that institutional investors trade heavily in the direction of *REG*, consistent with Edelen et al. (2016) who find that institutional investors often trade in the "wrong" (i.e., mispricing-enhancing) direction of anomaly returns.

Of equal importance, our paper contributes to the fast-growing literature on market participants' beliefs, in particular on how agents form their expectations and process the arrival of fundamental news. The existing literature mostly focuses on how analysts revise their forecasts to earnings information, and more recently how they weigh their private and publicly available signals (e.g., Gerken and Painter, 2019; Ludvigson et al., 2021). Our paper shows that, in addition, analysts also take into account how the market reacts to the arrival of new information. Specifically, we show that analysts fail to fully disentangle the noise from information in public signals, leading to larger biases. In addition, we provide evidence concordant with analysts consuming biased confirmatory information.[5]

2 Measures Construction and Data

*2.1 The Return-Earnings Gap (*REG*) Measure*

Earnings announcements are one of the most important sources of firm-specific information, where firms convey valuable cash flow information to market participants. In response, investors reflect their updated beliefs on the firm's valuation into stock prices.

[5]In addition to biased expectations on earnings forecasts, biased beliefs have been documented about credit spreads (Bordalo et al. 2018), interest rates (Cieslak 2018), cash flow growth (De la O and Myers 2021), and macroeconomic quantities such as inflation and GDP growth and inflation (e.g., Bordalo et al., 2020; Ludvigson et al., 2021). None of these papers analyzes how the market reaction to new information feeds into the belief formation of different agents.

While the relation between earnings announcements and stock returns has extensively been explored in previous studies, we propose a new measure that is designed to capture the *relative* market reaction to firm earnings information. Specifically, we use a non-parametric approach to measure the relative rankings of adjusted standardized earnings surprises ($AdjSUE$) and characteristic-adjusted abnormal returns ($DGTW$, according to Daniel et al. 1997) of each earnings announcement. This allows us to capture the extent to which the stock price response to the cash flow information deviates from the average response. A positive (negative) number indicates that the market reaction is higher (lower) than one would expect, on average.

We first construct the adjusted standardized earnings surprise component ($AdjSUE$). For each earnings announcement, we obtain the actual earnings per share (EPS), the median of analysts' EPS forecasts, and the standard deviation of their EPS forecasts. Following Mendenhall (2004) and Hirshleifer et al. (2021), we estimate the standardized unexpected earnings (SUE) as follows:

$$SUE_{i,t} = \frac{EPS_{i,t}^{Actual} - \text{Med}(EPS_{i.t}^{Estimate})}{\text{SD}(EPS_{i.t}^{Estimate})} \tag{1}$$

Here, $EPS_{i,t}^{Actual}$ is the firm's actual EPS reported on the earnings announcement day, where after-market-close announcements are shifted to the next trading day. $\text{Med}(EPS_{i.t}^{Estimate})$ and $\text{SD}(EPS_{i.t}^{Estimate})$ are the last available median and standard deviation of analysts' EPS forecast consensus reported in I/B/E/S prior to the earnings announcement day. We use I/B/E/S unadjusted information and adjust the actual EPS, the median and the standard deviation of analyst forecasts for dividends and splits using the cumulative adjustment factor from the Center for Research in Security Prices (CRSP) database.

Small or value firms may have different properties than large or growth firms. In addition, announcements of different days or different months may result in systematically different magnitudes of earnings surprises. To make SUE comparable across stocks, we keep the residual, which we denote as $AdjSUE$, from the following regression:

$$SUE_{i,t} \;=\; \beta_0 + \beta_1 LnSIZE_{i,t} + \beta_2 LnBM_{i,t} + \sum_{d=Mon}^{Sat} D_d + \sum_{m=Jan}^{Nov} D_m + \epsilon_{i,t}, \tag{2}$$

where $LnSIZE_{i,t}$ and $LnBM_{i,t}$ are the natural log of the size and book-to-market ratio of stock i as of day t, respectively. D_d and D_m are day-of-week and month-of-year dummies, which control for periodical variations in unexpected earnings. The regression residual $\epsilon_{i,t}$ is our $AdjSUE_{i,t}$ component for each stock i and earnings day t. Finally, to prevent a look-ahead bias, we use information from a one-year backward rolling window up to day t for each earnings day t.

Next, to construct the second component of our REG measure, the stock price adjustment, we compute daily characteristics-adjusted abnormal returns following the approach of Daniel et al. (1997), which accounts for differences in expected returns that are associated with firm size, book-to-market ratio, and momentum. We denote the daily abnormal return of stock i on day t as $DGTW_{i,t}$.

With both components at hand, we turn to construct our REG measure. For each earnings announcement of firm i on day t, we independently sort all earnings announcements over the past year (including day t) by their $DGTW$ and $AdjSUE$ components into 1,000 bins. We denote the relative rankings of its $DGTW_{i,t}$ and $AdjSUE_{i,t}$ as $Rank_{i,t}^{DGTW}$ and $Rank_{i,t}^{AdjSUE}$, respectively. We then define REG as follows:

$$REG_{i,t} \;=\; \frac{Rank_{i,t}^{DGTW} - Rank_{i,t}^{AdjSUE}}{(1,000-1)+(1,000-1)}, \tag{3}$$

where for ease of interpretation, we scale the difference in relative rankings between $DGTW$ and $AdjSUE$ by the number of bins minus one. This makes the REG measure range from -0.5 to 0.5. Thus, a one-unit change in REG from -0.5 to 0.5 implies a flip from the most negative market reaction to the most positive market reaction, relative to the earnings surprise.

The value of $REG_{i,t}$ is determined based on available information up to day t. Specifically, we use a ranking procedure based on a one-year rolling window that expands backward from day t (inclusive). Besides preventing a look-ahead bias, the use of one-year information allows us to take into account time-varying changes that can affect the relative ranking.[6] Finally, in Appendix A.1 we show that constructing REG based on the relative rankings of (i) unadjusted earnings surprises ($SUE_{i,t}$) and raw returns ($RET_{i,t}$) or (ii) adjusted earnings surprises ($AdjSUE_{i,t}$) and long-horizon abnormal returns ($DGTW_{i,t:t+20}$) yields qualitatively similar results. We focus on the one-day return response in our baseline analysis since it is highly visible, captures the attention of the media (and analysts), and is directly tied to earnings, while longer horizon returns are confounded by other events that may occur.

2.2 Analyst Expectations and Firm Mispricing Measures

To explore the interplay between the market response to earnings information and the belief formation of analysts, we infer analyst expectations from three different sets of information reported to I/B/E/S: earnings forecasts, price targets, and stock recommendations. While our baseline analysis focuses on analyst earnings forecast errors (AFE) in line with the majority of the literature, we also provide results using price target forecast errors (i.e.,

[6] We start the one-year backward rolling window rankings from 1985. In addition, we also construct REG using the relative rankings of $DGTW$ and $AdjSUE$ within a five-year backward rolling window, which produce similar results to those with the one-year backward rolling window.

the implied return forecast errors, *RetForeErr*) and buy-and-sell recommendations changes (*RecChng*). Earnings forecasts reveal analysts' perception of a firm's future prospect, while recommendations and price targets provide direct information that investors can act on. Thus, we are able to explore the effect of *REG* on different dimensions of analyst output.

We obtain information about analysts' quarterly EPS forecasts, 12-month price target estimates, and buy-and-sell recommendations from the I/B/E/S database. The analyst forecast error (*AFE*) is the difference between the median of analysts' EPS forecast and the actual EPS, scaled by the standard deviation of analysts' EPS forecasts. Notice that, by construction, the value of *AFE* is exactly opposite to that of *SUE* for each stock i on earnings announcement day t.

Unlike earnings forecasts, future price target estimates and recommendations by analysts offer more straightforward and actionable guidance for investors. Future price targets scaled by the current price provide an estimate of an analyst's return forecast of the stock. Using the 12-month price target estimates from I/B/E/S, we first obtain all the price target estimates issued by analysts over the subsequent 60 trading days following each earnings announcement. Then we estimate the analyst 12-month return forecast by scaling the future price target by the current stock price and subtracting one from the ratio. Next, we compute the actual 12-month return using the actually realized future price and the current price. Finally, we calculate the average return forecast error (*RetForeErr*) as the average of the difference between the forecast return minus the realized return across all analysts.

Recommendations from analysts offer investors explicit trading advice: strong buy, buy, hold, underperform, or sell. Each of them is assigned a numerical number in the I/B/E/S database, from 1 (strong buy) to 5 (sell). We construct analyst recommendation changes (*RecChng*) as the average numerical change of recommendations issued by analysts during the following few weeks after the earnings announcement day. We multiply the change by -1 such that a positive (negative) change is associated with increased optimism (pessimism).

To measure firm mispricing, we adopt the mispricing score from Stambaugh et al. (2015). The mispricing score (*MISP*) of a stock, ranging from 0 to 100, is an arithmetic average of a stock's ranking with respect to 11 well-established anomalies. For each anomaly, a stock is ranked higher if the degree of over-pricing according to that anomaly is greater. Thus, the higher the value of a mispricing score, the greater the degree of over-pricing. In extension of our baseline analysis, we also consider mispricing scores using performance (*PERF*) and management (*MGMT*) anomaly groups as classified by Stambaugh and Yuan (2017), and *Build-Up* and *Resolution* anomaly the classification by based on van Binsbergen et al.'s (2021). Details on the definition of mispricing scores and the different anomalies are provided in Appendix B.

2.3 Firm Controls

We construct the set of firm control variables used in our analysis using information from both the CRSP and I/B/E/S databases following the standard literature. *LnSIZE* is the natural logarithm of the firm's size, defined as the market capitalization of the stock in millions of dollars. *LnBM* is the natural logarithm of the stock's book-to-market ratio. *RET5* and *RET21* are a stock's cumulative past returns from day $t-5$ to day $t-1$ and from day $t-21$ to day $t-1$, respectively. *MOM* is the momentum of a stock, which is the average of daily returns over the period from day $t-251$ to day $t-21$. *RET5*, *RET21*, and *MOM* are all expressed in percentages. *RVOL* is the realized volatility of a stock, defined as the square root of the annualized realized variance, which is 252 times the average squared daily return from day $t-21$ to day $t-1$. *ILLIQ* is the Amihud (2002) illiquidity measure, which is the average ratio of absolute daily return by daily total dollar trading volume of stock from day $t-21$ to day $t-1$. *DISP* is the dispersion of analyst earnings forecasts, which is the standard deviation of analyst earnings forecasts scaled by the stock price. *NUMEST* captures the logarithmized number of analysts that issue earnings forecasts for a firm, defined as the natural logarithm of one plus the number of analysts issuing forecasts.

Note that one of the key variables of interest in our paper, *MISP*, is observed on a monthly basis. Therefore, in the analysis of *MISP*, the daily firm control variables are recorded at the end of each month instead of end-of-day. Accordingly, we also construct the following additional control variables. *MRET* is the monthly return in percentage. *MMOM* is the monthly momentum, which is the monthly return in percentage over the past 11 months. *MRVOL* denotes the monthly realized volatility, which is defined as the standard deviation of monthly returns over the 12 months ending in each June. If at least 9 monthly returns are available, then the *MRVOL* is applied to the following 12 months, i.e., July of the same year to June of the next year. *MILLIQ* is the illiquidity of the month, which is the average daily Amihud (2002) illiquidity ratio over all trading days of the month.

2.4 Sample and Descriptive Statistics

Our sample period runs from 1985 to 2018, where 1985 is the first year for which we construct *REG* using one-year historical earnings data from the I/B/E/S database. We match the I/B/E/S tickers to CRSP using the ICLINK table.[7] As common in the literature, we shift an earnings announcement that occurs after market close to the next trading day. Our final sample includes 228,266 earnings announcements for 8,434 distinct stocks on 6,531 trading days between January 1985 to December 2018. On average, we have 35 distinct stocks with

[7]The ICLINK table provides the mapping between I/B/E/S TICKER and CRSP PERMNO and is created following the WRDS Macro for ICLINK.

earnings announcements reported on an ordinary day in our sample.

Table 2 provides descriptive statistics of our main variables. Panel A presents the time-series average of the cross-sectional mean, standard deviation, and different quantiles for each variable. By construction, *REG* is zero on average. The average daily *DGTW* abnormal return is centered around zero as well, while *SUE* (*AFE*) exhibits a positive (negative) average of 0.193 (−0.193), consistent with Mendenhall (2004). The average mispricing score (*MISP*) is around 50, implying that an average stock in our sample is fairly valued based on Stambaugh et al.'s (2012) metric.

[Table 2]

Panel B of Table 2 reports the time-series average of the cross-sectional correlation of these main variables in our sample. Not surprisingly, *REG* is positively correlated with *DGTW* and negatively correlated with *SUE*, but the correlations of 0.514 and −0.436 with these variables also clearly indicate that *REG*'s relative ranking contains relevant information beyond the mere values of *DGTW* and *SUE*. At the same time, *REG* is positively correlated with the mispricing score, *MISP*, suggesting that stocks with a greater (lower) degree of over-pricing are more likely to experience investor overreaction (underreaction). Finally, the positive correlation between *SUE* and *DGTW* further confirms that investors respond in the same direction as the sign of the earnings surprise in general.

3 *REG*, Expectation Formation, and Mispricing

The *Return-Earnings Gap* captures the extent to which stock prices respond disproportionately to earnings surprises, relative to the average response as captured by the rank-based measurement. We first establish that this response is permanent to a large extent, suggesting that it is driven by investors' beliefs regarding a firm's future fundamentals and not caused by temporary noise. Additional evidence from institutional trading data shows that these beliefs are also revealed by actions (trading), as institutional investors are net buyers on and after earnings announcement days when REG is high. We proceed to investigate the relation of REG to analyst expectations and find that REG positively predicts analyst forecast errors one to several quarters ahead, indicating on the one hand that investors' beliefs captured by REG translate to analysts, and on the other hand that the predictive power of REG is a reflection of biased expectations. Analyzing firm mispricing strongly supports this hypothesis: Indeed, REG predicts cross-sectional mispricing scores several quarters ahead and, remarkably, allows us to capture the build-up phase of mispricing in the cross-section of firms.

3.1 REG*, Abnormal Stock Returns, and Institutional Trading*

To find out whether disproportionate stock price responses to earnings surprises as captured by REG are temporary or rather permanent in nature, we begin our investigation by analyzing the relation between REG and abnormal stock returns on day t (i.e., the earnings announcement day) and the subsequent 21 trading days. If REG captures investor beliefs, we should expect the market reaction on earnings announcement days to persist over a longer period. We indeed find that this is the case.

To examine the relationship between REG and abnormal stock returns, we form portfolios sorted on REG and explore a long-short strategy to illustrate the return pattern associated with REG and its economic significance. We sort all stocks with earnings announcements reported on a given earnings announcement day t into deciles based on their level of REG. We then compute portfolio returns by equally weighting the $DGTW$ abnormal return on day t ($DGTW_t$) of the stocks in each decile portfolio, and we repeat the same for the cumulative

$DGTW$ abnormal return from day $t+1$ to day $t+20$ ($DGTW_{t+1:t+20}$) and the cumulative $DGTW$ abnormal return from day t to day $t+20$ ($DGTW_{t:t+20}$). The REG high-minus-low (H-L) return is then computed based on a portfolio that goes long stocks in the top REG decile and shorts stocks in the bottom REG decile.

[Table 3]

Table 3 reports the results. The $DGTW$ abnormal return grows monotonically from -5.26% in the bottom decile to 5.14% in the top decile as the value of REG increases (a high-low spread of 10.40%). This result is not surprising due to the construction of REG; it reflects the fact that a larger (smaller) value of REG is associated with a relatively higher (lower) market response. What we are mainly interested in is whether the observed return is followed by a major reversal, or if it is rather permanent in nature. The analysis of the subsequent 20 trading days reveals that there is only a small reversal of -0.99% for the long-short portfolio, such that the cumulative return from day t up to $t+20$ still amounts to 9.41%.[8] Overall, this evidence from abnormal returns provides support to the significance and persistence of the market reaction on the day of the earnings announcement. Although a return reversal is observed, it is relatively small (in line with Da et al., 2014), supporting the idea that the gap between the market reaction and a firm's earnings information captures investors' belief about future firm prospects that are not captured by SUE.

We provide additional support to the idea that REG captures investors' belief based on institutional trading data. Investors reflect their beliefs into prices via trading; therefore, if REG reflects investor beliefs, we expect high REG to be associated with a greater amount of

[8]We obtain similar results using a cross-sectional regression analysis, which controls for a wide set of firm characteristics. In particular, the coefficients on REG on day t, $t+1:t+20$, and $t:t+20$, are 18.615, -2.139, and 16.547, respectively. In addition, extending the window to 60 trading days shows that the reversal does not become significantly larger in magnitude.

net buying by institutional investors.

We obtain institutional trading data from ANcerno. The data overlaps with our baseline sample from February 2002 to December 2015. Institutional directional trading (*InstDirTrd*) is defined as the institutional shares bought minus their shares sold normalized by daily total share volume. We employ Fama and MacBeth (1973) regressions and explore the contemporaneous and predictive relation between *REG* and *InstDirTrd*. The contemporaneous regression specification takes the form

$$\begin{aligned} InstDirTrd_{i,t} &= \gamma_{0,t} + \gamma_{reg,t} REG_{i,t} + \gamma_{sue,t} SUE_{i,t} + \gamma_{dgtw,t} DGTW_{i,t} + \\ &\sum_{k=1}^{K} \gamma_{k,t} CONTROLS_{k,i,t} + \epsilon_{i,t}, \end{aligned} \tag{4}$$

where $InstDirTrd_{i,t}$ is the directional trading by institutional investors on earnings day t. For the predictive regressions, the dependent variable is $InstDirTrd_{i,t+1:t+d}$, the cumulative directional trading by institutional investors over the period from day $t+1$ to $t+d$ ($d = 5, 10, 15$), while the right-hand side is identical to (4). $REG_{i,t}$, $SUE_{i,t}$, and $DGTW_{i,t}$ are the return-earnings gap, earnings surprise, and the DGTW-adjusted daily abnormal return of stock i on earnings announcement day t in quarter q. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, and *ILLIQ* as introduced in Section 2.3. Finally, given that the number of firms reporting their earnings is very scarce on some days, we report value-weighted averages in the second stage of the Fama-MacBeth procedure based on the daily number of cross-sectional observations.

[Table 4]

Table 4 reports the regression results. The *REG* coefficient estimate on day t according to column (1) indicates that institutional investors' net buying is positively and significantly related to *REG* on the day of the earnings announcement. Note that this relation holds after controlling for *SUE* and *DGTW*, reflecting the partial association between *REG* and net institutional trading. Columns (2)–(4) investigate the trading behavior of institutional investors during the subsequent d trading days ($d = 5, 10, 15$). The results indicate that institutional investors continue to be net buyers of stocks with positive *REG*, further supporting the argument that *REG* captures a rather permanent component of investor beliefs. Economically, a change in *REG* from its 25^{th} percentile to its 75^{th} percentile results in an increase in institutional net buying of shares by 2.152% ($= (0.114 - (-0.113)) \times 9.481$) of the trading volume over the next five days, which is approximately 0.430% per day, equivalent to 9.45% ($= 0.430/(2.670 - (-1.880))$) of the difference between the 25^{th} percentile and the 75^{th} percentile of overall *InstDirTrd*. Finally, cumulative net buying attenuates in the third week, but it still positively predicted by *REG*.

Taken together, the results reported in this section demonstrate that the gap between a firm's earnings fundamentals and the market reaction captured by *REG* is consistent with the reflection of investors' beliefs into prices and not driven by temporary (e.g., liquidity-related) effects.

3.2 REG *and Analyst Forecast Errors*

Our results on cross-sectional returns and institutional trading strongly suggest that *REG* captures and predicts investors' expectations. Motivated by these findings, we are interested in whether *REG* also translates to the expectations of analysts, and in particular to future analyst forecast errors. Analyst expectations are often used as a proxy for general investor expectations in the literature, but we also emphasize in this paper that analysts and general investors are two different types of agents whose expectation formation processes dynamically interact with each other. If *REG* stems from investors' rational expectations of future fundamentals, or if analysts can perfectly disentangle noise and information from market participants' earnings reactions, then we should find no or a negative predictive relation between *REG* and analyst earnings forecast errors. On the other hand, a positive predictive relation between *REG* and analyst forecast errors would suggest that *REG* stems to some extent from biased investor beliefs, and analysts carry over this bias as they incorporate the observed market reaction.

We explore the effect of *REG* on analyst forecast errors (*AFE*) using Fama and MacBeth (1973) cross-sectional regressions. We use *REG* in quarter q to predict *AFE* over the subsequent quarters up to $q + 12$. The regression specification takes the following form:

$$\begin{aligned} AFE_{i,q+n} &= \gamma_{0,t} + \gamma_{reg,t} REG_{i,t(q)} + \gamma_{afe,t} AFE_{i,t(q)} + \gamma_{dgtw,t} DGTW_{i,t(q)} + \\ &\qquad \gamma_{misp,t} MISP_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} CONTROLS_{k,i,t} + \epsilon_{i,t}. \end{aligned} \tag{5}$$

Here, $AFE_{i,q+n}$ is the analyst earnings forecast error of stock i for the earnings announcement q quarters ahead ($n = 1, \ldots, 12$). $REG_{i,t(q)}$, $AFE_{i,t(q)}$, and $DGTW_{i,t(q)}$ are the return-earnings gap, analyst earnings forecast error, and the DGTW-adjusted daily abnormal return of stock i on earnings announcement day t in quarter q. $MISP_{i,t(q)}$ is the Stambaugh et al. (2015) monthly mispricing score of the month of the earnings announcement. We control for *AFE* and *DGTW* to make sure that the measured effect of *REG* is not due to the persistence in analyst forecast errors or the impact of past returns. We also control for *MISP* to account for the relation between firm mispricing and analyst forecast errors documented in previous studies. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. As in the previous section, we report value-weighted averages based on the daily number of cross-sectional observations in the second stage of the Fama-MacBeth procedure.

[Table 5]

The results are presented in Table 5. In all regressions, the coefficients on *REG* are positive and significant, implying a positive impact of *REG* on future analyst earnings forecast errors. The predictive relation of *REG* decays from quarter 1 to quarter 12, which is expected given that new information enters into the calculation of the analysts' forecasts as time proceeds. Specifically, the coefficient on *REG* ranges from 2.464 (t-statistics of 11.93) in the prediction of quarter $q+1$ to 0.979 (t-statistics of 4.10) in the prediction of quarter $q+12$. Note, again, that this effect is beyond the persistence in *AFE*, which is reflected in the positive relation between *AFE* in quarter q and *AFE* over the subsequent 12 quarters.

Next, we assess the economic significance of *REG*'s impact on *AFE*s. Given the relation between firm mispricing and analyst earnings forecast errors known from the literature, we are particularly interested in comparing the effects of *REG* and *MISP* on *AFE* in subsequent quarters. Take, for example, the *AFE* one quarter ahead (column (1) of Table 5). A change in *REG* from its 25^{th} percentile to its 75^{th} percentile leads to an increase in the next quarter's *AFE* by 0.559 ($= (0.114 - (-0.113)) \times 2.464$), which is around 21.10% ($= 0.559/(0.829 - (-1.820))$) of the difference between the 25^{th} percentile to the 75^{th} percentile in *AFE*. As a comparison, the coefficient on *MISP* is 0.016, which means that a change in *MISP* from its 25^{th} percentile to its 75^{th} percentile results in a rise in the next quarter's *AFE* by 0.278 ($= (58.779 - 41.404) \times 0.016$), which is equivalent to 10.49% ($= 0.278/(0.829 - (-1.820))$) of the change from the 25^{th} percentile to the 75^{th} percentile of *AFE*. This shows that the economic significance of the impact of *REG* on next quarter's analyst earnings forecast errors is nearly twice as large as that of the mispricing score.

Overall, these findings suggest that when forming expectations, analysts take into account how the market reacts to fundamental news. In particular, they suggest that analysts cannot fully disentangle the noise from information, and market participants' reaction to earnings information feeds back into and distorts the analysts' expectation formation.

3.3 REG *and Firm Mispricing*

If *REG* captures and predicts market participants' biased expectations, as our results indicate, it appears likely that this results in a mispricing of firms' stocks. In fact, evidence on the link between anomaly returns and analyst forecast errors (AFE) has been growing rapidly since the early findings by La Porta (1996), and analysts have been found to be overly optimistic (pessimistic) for anomaly shorts (longs). Consequently, we examine the predictive relation between investors' reaction to earnings information (REG) and subsequent firm mispricing scores in this subsection.

As in previous tests, we employ Fama and MacBeth (1973) regressions for predicting *MISP* in the quarters following each earnings announcement:

$$\begin{aligned} MISP_{i,q+n} &= \gamma_{0,t} + \gamma_{emr,t} REG_{i,t(q)} + \gamma_{afe,t} AFE_{i,t(q)} + \gamma_{dgtw,t} DGTW_{i,t(q)} + \\ & \gamma_{misp,t} MISP_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} CONTROLS_{k,i,t} + \epsilon_{i,t} \end{aligned} \tag{6}$$

Here, $MISP_{i,q+n}$ is the Stambaugh et al. (2015) monthly mispricing score observed n quarters ahead at the end of the quarter. $REG_{i,t(q)}$, $AFE_{i,t(q)}$, and $DGTW_{i,t(q)}$ are the return-earnings gap, analyst earnings forecast error, and DGTW-adjusted abnormal return of earnings announcement day t in quarter q. $MISP_{i,t(q)}$ denotes the Stambaugh et al. (2015) monthly mispricing score of the month of earnings announcement day t in quarter q. Firm control variables include *LnSIZE*, *LnBM*, *MRET*, *MMOM*, *MRVOL*, and *MILLIQ* as introduced in Section 2.3. All firm controls are recorded as of the end of the month of day t. As before, we compute the observation-weighted time-series average of each slope coefficient.

[Table 6]

We predict *MISP* for 1, 2, 3, 4, 8, and 12 quarters ahead and report the regression coefficients in Table 6. The collective results clearly indicate that *REG* has a significant and positive influence on firm mispricing. Starting with *MISP* in quarter $q+1$, a change in *REG* from its 25^{th} percentile to its 75^{th} percentile results in a rise in *MISP* of 0.523 ($= (0.114 - (-0.113)) \times 2.304$). For comparison, the increase in *MISP* caused by a change from the 25^{th} percentile to the 75^{th} percentile in *AFE* is 0.156 ($= (0.829 - (-1.820)) \times 0.059$). Thus, the effect of *REG* on firm mispricing is over three times that of *AFE*. Given the large amount of evidence in the literature on the relation between *AFE* and firm mispricing, this comparison establishes that the effect of *REG* on *MISP* warrants attention.

Another remarkable observation that emerges from our results is that effect of *REG* becomes larger in magnitude for up to 4 quarters ahead, while the effect of *AFE* is most pronounced 1 quarter ahead and declines afterwards. This comparison suggests that on a timeline, *REG* is able to better predict the mispricing build-up stage, while *AFE* captures a later stage, where mispricing is at its peak and starts to resolve. For example, comparing the economic effect of *AFE* and *REG* after 3 quarters (column (3) of Table 6), an increase in *REG* from its 25^{th} percentile to its 75^{th} percentile would lead to an increase in *MISP* of 0.681 ($= (0.114 - (-0.113)) \times 2.999$). In contrast, the effect of *AFE* is only 0.106 ($= (0.829 - (-1.820)) \times 0.040$).

[Figure 1]

In order to have a better understanding of the mispricing cycle captured by *REG*, we plot the effect of *REG* on *MISP* over three years in Figure 1. Specifically, the figure illustrates the estimated coefficients on *REG* and the corresponding 95% confidence intervals in the *MISP* predictive regressions. As depicted in the figure, the investors' disproportionate reaction to a stock's earnings surprise positively affects its degree of mispricing in the subsequent quarters, controlling for current mispricing that may already be present. The effect gradually escalates and peaks in the third quarter following the earnings announcement. After that, it resolves rather sharply and is only marginally significant after 12 quarters.

4 Understanding *REG*'s Effect on Analyst Beliefs

Our results show that *REG* predicts average analyst forecast errors, indicating that an overly positive (or negative) market response to earnings announcements leads to an increased bias in analyst expectations. In this section, we investigate the formation of analyst expectations in response to *REG* in detail. First, we analyze the effect of *REG* on the whole cross-section of analysts and examine whether there are cross-sectional differences in *AFE* predictability, for example between generalist and specialist analysts. An increased predictability for generalists (which we find) further supports the idea that a biased market reaction translates to analyst biases particularly for those analysts who do not have very strong private information. Second, we show that the bias in analyst expectations has features of a confirmation bias as it is stronger when *REG* and *AFE* go in the same direction. Finally, we confirm that the predictability of *AFE*s by *REG* is similarly strong for positive and for negative *REG*, and it also translates to analyst *return* forecast errors (based on price targets) as well as to analyst recommendation changes.

4.1 Analyst Heterogeneity and Biased Expectations

In the previous section, we document the average effect of REG on AFE. It is important to note that analysts can differ in their response to REG; this can be driven by ability, experience, and other factors that affect the quality of analysts' private signals such as attention and effort paid to a specific stock in an analyst's portfolio. To explore such cross-analyst heterogeneity, we focus on two dimensions: (i) the degree of analyst industry concentration (in analogy to, e.g., Kacperczyk et al., 2005, for mutual fund managers), and (ii) an analyst's past stock-level forecast accuracy. To capture industry concentration, we construct $Rank(NumInd)$ as the decile ranking based on the number of industries covered by an analyst in the given quarter. To capture accuracy, we use Clement's (1999) $PMAFE$ (Proportionate Mean Absolute Forecast Error) measure, and construct $Rank(PMAFE)$ as the decile ranking, in a given quarter, of an analyst stock-level $PMAFE$ over the past four quarters.

We perform our analysis at the stock-analyst-quarter level using panel regressions. For each analyst, stock, and quarter, we keep the most recent forecast before the upcoming earnings announcement, where we make sure to control for the number of days in between analyst's earnings forecast and firm's earnings announcement. We capture the heterogeneity via the interaction between REG and the analyst characteristics rankings mentioned above. The panel regressions take the following form:

$$\begin{aligned} AFE_{j,i,q+n} &= \gamma_{0,t} + \gamma_{reg,t} REG_{i,t(q)} + \gamma_{rank,t} Rank(Char)_{j,i,t(q)} + \\ &\quad \gamma_{regrank,t} REG_{i,t(q)} \times Rank(Char)_{j,i,t(q)} + \gamma_{afe,t} AFE_{j,i,t(q)} + \\ &\quad \gamma_{dgtw,t} DGTW_{i,t(q)} + \gamma_{misp,t} MISP_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} CONTROLS_{k,i,t} + \epsilon_{j,i,t} \end{aligned} \tag{7}$$

where $AFE_{j,i,q+n}$ is analyst j's earnings forecast error (AFE) for stock i for the earnings announcement n quarters ahead ($n = 1, \ldots, 4$). $REG_{i,t(q)}$ and $DGTW_{i,t(q)}$ are the return-earnings gap and DGTW-adjusted daily abnormal return of stock i on earnings announcement day t in quarter q. $Rank(Char)$ is $Rank(NumInd)$ or $Rank(PMAFE)$. $MISP_{i,t(q)}$ is the Stambaugh et al. (2015) monthly mispricing score of the month of the earnings announcement. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. All regressions include analyst and quarter fixed effects and standard errors are clustered by analyst and quarter.

[Table 7]

Table 7 reports the results. We find that greater industry concentration results in a smaller response of *AFE* to *REG*. That is, analysts that focus on a smaller number of industries are less sensitive to the market's reaction. This is consistent with the idea that analysts that are industry-concentrated generate higher-quality private signals. As such, they are less influenced by the market response when updating their beliefs. For example, a change from a rank of 1 to 10 in industry concentration results in an additional response of $0.101 \times 9 = 0.909$, which is 50% larger than the baseline result.

The second set of regressions indicate that analysts with lower past stock-level forecast accuracy are more affected by *REG*. The effect is economically significant and about half of the effect of industry concentration. Importantly, past forecast accuracy can be viewed as a "catch-all" proxy for analyst ability, experience, or the attention paid by the analyst to the stock. It does not govern — ex ante — the direction of the response to *REG*, as lack of accuracy can be driven by either a positive or a negative bias. Our results show that lower accuracy corresponds to a higher reliance on the market's response to earnings.

Overall, our collective tests reveal that analysts who are less focused on a specific industry and analysts that demonstrate lower past forecast accuracy are more prone to rely on the market reaction to earnings when updating their beliefs. As a result, a biased market response translates more strongly to future forecast errors for these sets of analysts.

4.2 Confirmation Bias and Analyst Expectations

A natural question to ask is why analysts do not better disentangle the bias and noise presented in *REG* from fundamental information to adjust their subsequent forecasts. One explanation suggested in the literature is "confirmation bias" (Pouget et al., 2017; Cookson et al., 2021; Hirshleifer et al., 2021). In our setting, this would suggest that analysts tend to interpret the market response in favor of their own (biased) expectations. To analyze this possibility, we investigate the effect of *REG* on *AFE* over subsequent quarters separately for cases where *REG* has the same sign as the contemporaneous *AFE* and thus confirms the analysts' expectations, and for cases where *REG* disconfirms contemporaneous *AFE* (opposite sign).

We broadly follow Pouget et al. (2017) and employ a linear probability model. Our dependent variable, D(AFE_q & AFE_{q+n} Same Sign) is a dummy variable that equals one when AFE_q is in the same direction of AFE_{q+n}. The main explanatory variable of interest is D(AFE_q & REG_q Same Sign), a dummy variable that is equal to one if *REG* in a given quarter q (observed after *AFE*) is in the same direction as *AFE*. Our hypothesis is that *AFE* in the next quarter will have a higher likelihood of being in the same direction as current *AFE* when current *AFE* and *REG* are in the same direction. We include our standard set of controls as well as another dummy variable D(AFE_q & AFE_{q-1} Same Sign) that accounts for the natural persistence in *AFE*. This dummy is equal to one when AFE_q is in the same direction as AFE_{q-1}. We employ a Fama and MacBeth (1973) regression using the following specification:

$$
\begin{aligned}
D(AFE_q \text{ \& } AFE_{q+n} \text{ Same Sign}) = {} & \gamma_{0,q} + \gamma_{1,q} D(AFE_q \text{ \& } REG_q \text{ Same Sign}) + \\
& \gamma_{2,q} D(AFE_q \text{ \& } AFE_{q-1} \text{ Same Sign}) + \\
& \gamma_{afe,q} AFE_{i,t(q)} + \gamma_{dgtw,t(q)} DGTW_{i,t(q)} + \\
& \gamma_{misp,t(q)} MISP_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} CONTROLS_{k,i,t} + \epsilon_{i,t}.
\end{aligned} \tag{8}
$$

Table 8 reports the results. Consistent with our conjecture, the results across all horizons indicate that when the disproportionate market reaction (REG) is in the same direction as the analysts' initial bias (AFE), analysts will view this is a confirmation of their signal. As a result, we observe a higher likelihood of a continuation of the bias in these cases. The coefficient estimate in column (1) is 0.127, which indicates that for the subsequent quarter analysts will have a 12.7% higher probability of issuing a forecast that is biased in the same direction as their current forecast if REG provides confirming information.

[Table 8]

In addition, the coefficient estimate that captures the persistence in AFE is about 20.8%. Comparing the two indicates that the confirmation effect is 61% of the magnitude of own persistence, indicating its economic significance. The effect remains significant across all horizons (up to $q + 12$), both statistically and economically in comparison to the effect of AFE persistence.

4.3 Alternative Measures of REG *and Analyst Expectations*

We confirm that our results do not critically hinge on particular details of how we specify REG or of how analyst expectations are measured. In Appendix A.1, we show that constructing REG based on the relative rankings of (i) unadjusted earnings surprises ($SUE_{i,t}$) and raw returns ($RET_{i,t}$) or (ii) adjusted earnings surprises ($AdjSUE_{i,t}$) and long-horizon abnormal returns ($DGTW_{i,t:t+20}$) yields qualitatively similar results. In Appendix A.2, we additionally investigate whether the predictive effect of REG on AFE is more pronounced for positive or negative REG, and find that the effect is of similar magnitude on both sides.

Next, instead of using analyst earnings forecasts to infer analyst expectations, we consider analyst price targets and analyst recommendation changes in Appendix A.3 and A.4 and confirm our main results based on these data. Further robustness of our results is provided in Appendix A.5 and A.6, where we consider subsamples for different time periods within our sample as well as panel regressions instead of the Fama and MacBeth (1973) approach.

5 Understanding *REG*'s Effect on Firm Mispricing

Our results in Section 3.3 demonstrate that the biases in expectation formation captured by REG translate to the mispricing cycle. Precisely, REG predicts an increase in mispricing scores in the cross-section of firms for several quarters, before it starts to decline. In this section, we investigate this finding and its underlying economics in further detail. While our baseline results build on mispricing scores according to Stambaugh et al. (2015), we analyze the mispricing cycle using anomaly dissections from the recent literature in Section 5.1. Remarkably, we find that REG predicts a pronounced mispricing cycle for *Build-Up* anomalies, while this is not the case for *Resolution* anomalies. In Section 5.2, we analyze the mispricing cycle for different subsets of firms and show that *REG* predicts a more pronounced mispricing cycle for firms with lower analyst coverage, for small firms, and for firms with high analyst disagreement. These results indicate that *REG* predicts an increase and perpetuation of mispricing especially for firms for which market participants do not have very strong private information.

5.1 Buildup vs. Resolution Anomalies

In a recent paper, van Binsbergen et al. (2021) classify asset pricing anomalies into those that exacerbate mispricing (*Build-Up*) and those that resolve mispricing (*Resolution*). We investigate the predictive relation of *REG* to mispricing scores with respect to these two sets of anomalies. Since we have established that *REG* is associated with biased expectations and increased mispricing, we conjecture that our findings are more pronounced in the *Build-Up* anomalies group. To explore this conjecture, we obtain the set of anomalies used in van Binsbergen et al. (2021) and construct two mispricing measures with respect to build-up and resolution anomalies: $MISP_{BUILD}$ and $MISP_{RES}$.[9]

[Table 9]

We closely follow the average anomaly ranking approach of Stambaugh et al. (2015) and average the stock's rankings according to the anomalies within each cluster. We then re-estimate the regressions in Eq. (6) by replacing *MISP* with the new mispricing scores $MISP_{BUILD}$ and $MISP_{RES}$. The coefficients on *REG* for predicting the mispricing scores 1, 2, 3, 4, 8, and 12 quarters ahead are presented in Table 9. We find a stark difference between

[9] The stock-level anomaly data is obtained from Chen and Zimmermann's (2022) Open Source Asset Pricing dataset (`https://www.openassetpricing.com/data/`). Details about the matching of anomalies to Chen and Zimmermann's (2022) stock-level characteristics as well as about the ranking procedure are provided in Appendix B.1.

the results for the *Build-Up* and *Resolution* anomaly classifications. In the predictions of future $MISP_{BUILD}$, the results from all regressions indicate a positive and significant effect of REG on stock mispricing. Moreover, we find that it takes up to 2 years for the mispricing cycle to reach its peak. These results align with the finding by van Binsbergen et al. (2021) that build-up anomalies result from a continuation of over-pricing and a slow correction and therefore drive stock prices further away from firms' fundamentals. When predicting $MISP_{RES}$, the coefficients on REG are of much smaller absolute magnitude than for $MISP_{BUILD}$ and they are negative, in line with the notion that resolution anomalies resolve existing mispricing and alleviate price dislocation. Note that while van Binsbergen et al. (2021) use fundamental information on future cash flows in order to classify *Build-Up* and *Resolution* anomalies, we do not use such information for our analysis and simply investigate the predictive relation of REG to the two groups of anomalies. It is remarkable that this predictive relation reflects the underlying structural difference between the two anomaly types. Altogether, the comparison between the effect of REG on $MISP_{BUILD}$ and $MISP_{RES}$ implies that REG is an important signal in predicting the exacerbation of mispricing for build-up anomalies and the onset of the correction of mispricing for resolution anomalies.

To have a better understanding of REG's impact on mispricing over time, we plot the coefficients on REG for predicting each mispricing score 1, 2, 3, 4, 8, and 12 quarters ahead in Figure 2. The figure illustrates the coefficients on REG and the corresponding 95% confidence intervals when predicting $MISP_{BUILD}$ and $MISP_{RES}$. The differences between $MISP_{BUILD}$ and $MISP_{RES}$ are very eye-catching, with the impact of REG showing a pronounced mispricing cycle for *Build-Up* anomalies, which can clearly not be observed for *Resolution* anomalies.

[Figure 2]

5.2 Cross-Sectional Firm Heterogeneity

In extension of the mispricing cycle that *REG* predicts on aggregate, it is reasonable to assume that some firms are more sensitive to *REG* than others due to their characteristics. For example, firms can be more prone to mispricing due to their clientele base, media and analyst coverage, and dispersion of opinions. To explore this idea, we revisit the relation between *REG* and *MISP* using four cross-sectional subsamples based on the monthly stock-level medians of (i) analyst coverage, (ii) firm market cap, (iii) institutional ownership, and (iv) analyst disagreement. It is important to note that the variation of *REG* itself is almost identical across these subsamples.[10]

[Table 10]

The first set of results indicates that analyst coverage is relevant for the cross-sectional relation between *REG* and *MISP*. The effect is both statistically and economically significant. For example, in quarter $q+1$, firms with below-median analyst coverage exhibit a 47% (= 0.825/1.769) higher sensitivity of *MISP* to *REG*. The effect is also stronger for small firms, consistent with lower analyst and media coverage, but the effect is weaker (a 31% (= 0.651/2.092) increase in sensitivity in quarter $q+1$). The third set of results indicates that the institutional clientele base is also an important determinant. Not surprisingly, firms with a lower institutional base are more prone to mispricing, which results in a higher sensitivity of *MISP* to *REG*. The effect reaches its peak after four quarters, where firms with below-median institutional ownership present a 50% (= 1.050/2.110) higher sensitivity of *MISP* to *REG*. Finally, the fourth set of tests reveals that firms with higher analyst dispersion present a 104%

[10] The standard deviations of *REG* for low and high medians based on analyst coverage, stock market cap, institutional ownership, and analyst disagreement are: 0.168 and 0.174, 0.164 and 0.177, 0.169 and 0.173, and 0.173 and 0.170, respectively.

higher sensitivity to *REG* in quarter $q+1$. If analyst dispersion also reflects the difference of opinions across market participation, this finding is in line with Miller's (1977) argument.

In sum, the collective set of results shows that *REG* contributes to the build-up and perpetuation of mispricing most strongly for firms for which market participants, and analysts in particular, do not have very strong private information.

5.3 *Effect of* REG *on Anomaly Returns Conditional on Prior Mispricing*

Our results show that *REG* predicts a build-up in future mispricing that is significant both statistically and with respect to its economic magnitude. We can gain another perspective on the economic importance of *REG*'s effect by investigating its interaction with the general relation between mispricing scores and subsequent returns.

As *MISP* measures overvaluation, there is generally an unconditional negative relation between *MISP* in month $m-1$ and the cross-section of stock returns in month m.[11] The literature also documents that anomaly returns converge faster on earnings announcement days (La Porta, 1996; La Porta et al., 1997; Engelberg et al., 2018; Bordalo et al., 2019). Connecting these result with our findings, we are interested to learn if the effect of *MISP* on subsequent stock returns is modulated based on the sign of *REG*. For example, do overvalued (undervalued) stocks exhibit overall positive (negative) returns during the subsequent month or quarter when *REG* is positive (negative)? We acknowledge that the earnings day return, which is part of the cumulative month- or quarter-ahead return, is positively correlated with *REG*. Nevertheless, it is interesting to see whether the effect of *REG* outweighs the general negative relation between *MISP* and returns. We argue that *REG* is very likely indicative of

[11] We verify that this holds in our sample using cross-sectional Fama-MacBeth regressions, and confirm that a one-standard-deviation increase in *MISP* (a value of 12.582) results in a DGTW abnormal return of -0.303% over the next month ($= -0.02412 \times 12.582$, where -0.02412 is the second-stage regression coefficient on *MISP*).

biased beliefs if it goes in the same direction as a stock's prior mispricing; on the other hand, *REG* with an opposite direction of prior *MISP* can accelerate the correction of mispricing.

[Table 11]

We define m as the month in which *REG* is observed and sort firms first by their mispricing scores in month $m-1$ and then by *REG*. Based on this sorting, we consider average returns for the subsamples of *MISP* quintiles and for *REG* greater and smaller than 0. We track these returns over a quarter, from month m to m+2. Table 11 reports the results. As expected, higher *MISP* is associated with negative subsequent returns, with an average long-short return spread of -0.53% (-1.40%) over the subsequent month (quarter). We now consider two long-short portfolios constructed based on our sorting on *REG* and prior *MISP*. The first portfolio goes long stocks in the highest prior-*MISP* quintile with positive *REG*, and shorts stocks in the lowest *MISP* quintile with negative *REG* — as *REG* goes in the same direction as *MISP*, we call it the "With" portfolio. The second portfolio goes long stocks in the highest prior-*MISP* quintile with negative *REG*, and shorts stocks in the lowest *MISP* quintile with positive *REG* — the "Against" portfolio. Positive returns for these portfolios indicate that the related stocks behave contrary to the baseline negative relation between mispricing and subsequent returns, while negative returns are in line with that relation.

We can immediately observe that the "Against" portfolio shows a clear negative return, suggesting that *MISP* converges in the 'right' direction, where the magnitude is large ranging from -3.75% in the first month to -4.37% after a quarter. In contract, the "With" portfolio, which captures investors biased beliefs, presents a positive return that reflects the exuberance of mispricing. Interestingly, the effect of 2.70% in month m attenuates to 1.57% over the quarter, consistent with the the fact that the market realizes at some point that the stocks in the "With" portfolio are mispriced.

In sum, while high *MISP* in month $m-1$ results in negative returns in month m on average, the effect is clearly dominated by *REG*. As a result, an overvalued (high *MISP*) stock with positive *REG* exhibits significant positive returns even when accumulated over the whole quarter, leading to an increase in overvaluation (i.e., higher *MISP*) instead of correction.

6 Determinants of *REG* and Dynamic Interrelations

6.1 Determinants of REG

So far, we have focused on the relation between *REG* and subsequent *AFE* and *MISP*. In this section, we turn our investigation to the factors that possibly contribute to the observed gap between investors' reactions and earnings fundamentals. In particular, we assess the predictive relation of the bias in analyst expectations (*AFE*) and firm mispricing (*MISP*) on *REG*.

We employ Fama and MacBeth (1973) regressions to predict *REG* in the next quarter, REG_{q+1}, for all stock-earnings-announcement observations in our sample. The regression specification takes the following form:

$$\begin{aligned} REG_{i,q+1} \; &= \gamma_{0,t} + \gamma_{afe,t} AFE_{i,t(q)} + \gamma_{misp,t} MISP_{i,t(q)} + \gamma_{emr,t} REG_{i,t(q)} + \\ &\quad \gamma_{dgtw,t} DGTW_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} CONTROLS_{k,i,t} + \epsilon_{i,t}, \end{aligned} \tag{9}$$

where $REG_{i,q+1}$ denotes the the return-earnings gap in the next earnings announcement quarter. $AFE_{i,t(q)}$, $REG_{i,t(q)}$, and $DGTW_{i,t(q)}$ are the analyst forecast error, return-earnings gap, and DGTW-adjusted daily abnormal return of stock i observed at the end of the current earnings announcement day t in quarter q. $MISP_{i,t(q)}$ is the Stambaugh et al. (2015) monthly mispricing score of stock i for the month of the current earnings announcement day t. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. As done in previous analyses, we obtain the slope coefficients from the cross-sectional regression and compute the time-series observation-weighted average of each slope coefficient. Table 12 presents the results.

[Table 12]

The positive and significant coefficients on AFE and *MISP* suggest that stocks with a positive analyst bias and high mispricing scores are more likely to experience a positive overreaction by investors in the next quarter. The results in column (4) show that the positive effect of AFE and *MISP* on *REG* in the next quarter remains intact after controlling for the *REG* and the daily abnormal return on the current earnings announcement day t. The coefficient on AFE is 0.003 with a t-statistics of 13.26, implying that a change in AFE from its 25^{th} percentile to its 75^{th} percentile would result in an increase of 0.0079 $(= (0.829 - (-1.820)) \times 0.003)$ in *REG*, which is 3.48% $(= 0.0079/(0.114 - (-0.113)))$ of the difference between the 25^{th} percentile and the 75^{th} percentile of *REG*. In addition, the coefficient of 0.00045 $(= 0.045/100)$ on *MISP* with a t-statistics of 10.05 indicates that the mispricing score has additional influence on future *REG*. In terms of the economic magnitude, it shows that when *MISP* rises from its 25^{th} percentile to its 75^{th} percentile, it will lead to an increase of 0.0078 $(= (58.779 - 41.404) \times 0.00045)$ in *REG*, which is about 3.44% $(= 0.0078/(0.114 - (-0.113)))$ of the gap between the 25^{th} percentile and the 75^{th} percentile of *REG*.

Interestingly, we also find that the past month's returns (*RET21*) and the earnings announcement day return ($DGTW$) negatively predict *REG*. This is consistent with the correction phase documented in Engelberg et al. (2018), and further highlights the difference between "raw" returns and the relative ranking captured by *REG*.

The evidence documented above suggests that both biased analyst expectations and stock's mispricing contribute to investors' future (mis)reaction to earnings information. In other words, a stock with greater (smaller) analyst earnings forecast errors and a higher (lower) degree of over-pricing would be exposed to a more pronounced investor overreaction (underreaction) to earnings surprises. Altogether, our findings indicate a dynamic amplification effect where higher *REG* leads to greater *AFE* and *MISP*, which in turn lead to higher future *REG*.

6.2 Impulse Responses

To further examine the dynamic relation between *REG*, *AFE*, and *MISP*, we estimate a quarterly vector autoregression (VAR) system of these variables and analyze the corresponding impulse response functions. We consider four lags of each variable. The regressions include the full set of firm control variables together with firm fixed effects and quarter fixed effects. Each graph in Figure 3 plots the response of *AFE*, *MISP*, and *REG* to shocks in the other two variables in the subsequent $0, 1, 2, \ldots, 12$ quarters, respectively.

[Figure 3]

The first plot in Figure 3 depicts the cumulative response of *AFE* to a one-standard-deviation shock in *REG* and *MISP*. As shown in the plot, both *REG* and *MISP* positively affect *AFE* in the following quarters. The impulse responses also confirm the result from our regression analysis that the effect of *REG* on *AFE* is much larger in magnitude compared to the effect of *MISP*. Precisely, a one-standard-deviation shock to *REG* leads to a nearly five times larger response of *AFE* than a one-standard-deviation shock in *MISP*. Next, the response of *MISP* to shocks in *REG* and *AFE* is shown in the second graph. Again, the impulse responses clearly confirm that while *MISP* reacts positively to a one-standard-deviation shock in both *REG* and *AFE*, the impact of *REG* is much larger than that of *MISP*. These results provide further supporting evidence for the economic importance of *REG* for future analyst forecast errors and mispricing.

The last graph shows the response of *REG* to shocks in *AFE* and *MISP*. Consistent with the findings in Section 6.1, a one-standard-deviation shock in both *AFE* and *MISP* leads to a positive response in *REG* in the following quarters, indicating that a stock with greater *AFE* and *MISP* is exposed to more pronounced *REG* in the future.

7 A Simple Model of *REG* and the Mispricing Cycle

We present a simple model that explains the predictive power of *REG* for analyst forecast errors and the build-up of mispricing that we observe in the data. The model explicitly accounts for the dynamic updating of beliefs between investors and analysts and demonstrates how biases are propagated and reinforced between the two groups of agents. Our approach thus extends the literature on belief updating, which often implicitly or explicitly equalizes the expectations of analysts and other investors.[12] We show the precise implications of our model for the case of diagnostic expectations in line with Bordalo et al. (2019) and for confirmatory bias (Rabin and Schrag, 1999).

[12]Notable exceptions are Malmendier and Shanthikumar (2007) and Ke et al. (2020).

Setup We consider a cross-section of firms that is indexed by i. The earnings per share $x_{i,t}$ of firm i evolve as

$$x_{i,t} = bx_{i,t-1} + f_{i,t} + \varepsilon_{i,t}, \tag{10}$$

with mean-reversion parameter b, earnings growth trend $f_{i,t}$, and temporary earnings shocks $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2)$. The dynamics of the earnings growth trend are given by

$$f_{i,t} = af_{i,t-1} + \eta_{i,t} \tag{11}$$

with mean-reversion parameter a and growth trend shocks $\eta_{i,t} \sim N(0, \sigma_\eta^2)$.

Investors and Analysts There are two types of agents in the model, investors and analysts. The agents form expectations about the unobservable fundamental $f_{i,t}$ based on observations of the firm's earnings per share $x_{i,t}$, a private signal, as well as an inferred private signal of the other type of agent. In particular, analysts infer the investors' private signal from the market response on the earnings announcement day, while investors infer the analysts' private signal from the previously published analyst forecasts.

Formally, investors observe a private signal $s_{i,t} = f_{i,t} + \chi_{i,t}$, with $\chi_{i,t} \sim N(0, \sigma_\chi^2)$, and analysts observe a private signal $c_{i,t} = f_{i,t} + \phi_{i,t}$, with $\phi_{i,t} \sim N(0, \sigma_\phi^2)$. Investors infer the private signal of the analysts from time $t-1$ by observing published analyst forecasts, assuming that $\tilde{c}_{i,t-1} = f_{i,t-1} + \phi_{i,t-1}$ holds for the inferred signal $\tilde{c}_{i,t-1}$. They update their expectations about $f_{i,t}$ according to

$$\tilde{f}^I_{i,t} = a\tilde{f}^I_{i,t-1} + K_1\theta^I_1(x_{i,t} - bx_{i,t-1} - a\tilde{f}^I_{i,t-1}) + K_2\theta^I_2(s_{i,t} - a\tilde{f}^I_{i,t-1}) + K_3\theta^I_3(\tilde{c}^I_{i,t-1} - \tilde{f}^I_{i,t-1}). \tag{12}$$

The vector $\theta^I = (\theta_1^I, \theta_2^I, \theta_3^I)$ represents biases in the investors' expectation formation, and the case $\theta^I = (1, 1, 1)$ stands for unbiased beliefs, for which (12) corresponds to the standard Bayesian updating rule.[13]

Analysts infer the time-t private signal of the investors by observing market prices, assuming that $\tilde{s}_{i,t} = f_{i,t} + \chi_{i,t}$. Their expectations about $f_{i,t}$ follow the process

$$\tilde{f}_{i,t}^A = a\tilde{f}_{i,t-1}^A + K_1\theta_1^A(x_{i,t} - bx_{i,t-1} - a\tilde{f}_{i,t-1}^A) + K_2\theta_2^A(\tilde{s}_{i,t}^A - a\tilde{f}_{i,t-1}^A) + K_3\theta_3^A(c_{i,t} - a\tilde{f}_{i,t-1}^A), \quad (13)$$

with biases captured by $\theta^A = (\theta_1^A, \theta_2^A, \theta_3^A)$.

Inferred Signals and Biases As a baseline, consider our model for the case that no biases are present, $\theta^I = \theta^A = (1, 1, 1)$. By observing the market reaction to an earnings announcement, analysts can, together with the other observed variables, perfectly infer the investors' private signal, $\tilde{s}_{i,t}^A = s_{i,t}$. The other way round, investors observe analyst forecasts $bx_{i,t-1} + a\tilde{f}_{i,t-1}^A$ and can thus perfectly back out the analysts' private signal, $\tilde{c}_{i,t-1}^I = c_{i,t-1}$. As a result, there are no notable interactions between the expectations of the two types of agents, and temporary deviations revert quickly.

Now we introduce a bias into the model in the following way: We assume that $\theta^I \neq (1, 1, 1)$, but analysts and investors are unaware of this bias and assume that $\theta^I = (1, 1, 1)$. As a consequence of their unawareness, they also believe that the other group of agents extracts the private signal correctly.[14]

With these biases, the analysts' belief of the investors' fundamental growth expectation,

[13]See Liptser and Shiryaev (2001). $K = (K_1, K_2, K_3)$ is the vector of Kalman gains, which weight the different signals according to their precision.

[14]The mechanism works analogously for $\theta^A \neq (1, 1, 1)$. We first introduce the bias on the investors' side, employing diagnostic expectations in line with Bordalo et al. (2019). Afterwards, we consider a model variant which additionally features a confirmation bias on the analyst side, as motivated by our results in Section 4.2.

$\mathcal{A}(\tilde{f}^I_{i,t})$, differs from the actual fundamental growth expectation of the investors $\tilde{f}^I_{i,t}$ as given by (12):

$$\mathcal{A}(\tilde{f}^I_{i,t}) = a\tilde{f}^I_{i,t-1} + K_1(x_{i,t} - bx_{i,t-1} - a\tilde{f}^I_{i,t-1}) + K_2(\tilde{s}^A_{i,t} - a\tilde{f}^I_{i,t-1}) + K_3(c_{i,t-1} - \tilde{f}^I_{i,t-1}). \quad (14)$$

In particular, analysts observe $\tilde{f}^I_{i,t}$, which follows the equation (12), but they believe that the investors' expectation is formed according to (14). Equalizing (12) and (14), the analysts infer the investors' private signal as

$$\tilde{s}^A_{i,t} = a\tilde{f}^I_{i,t-1} + \frac{1}{K_2}(\tilde{f}^I_{i,t} - a\tilde{f}^I_{i,t-1} + K_1(x_{i,t} - bx_{i,t-1} - a\tilde{f}^I_{i,t-1}) - K_3(c_{i,t-1} - \tilde{f}^I_{i,t-1})). \quad (15)$$

Similarly, the investors' belief of the analysts' fundamental growth expectation, $\mathcal{I}(\tilde{f}^A_{i,t})$, differs from the actual fundamental growth expectation of the analysts $\tilde{f}^A_{i,t}$ as given by (13):

$$\mathcal{I}(\tilde{f}^A_{i,t}) = a\tilde{f}^A_{i,t-1} + K_1(x_{i,t} - bx_{i,t-1} - a\tilde{f}^A_{i,t-1}) + K_2(s_{i,t} - a\tilde{f}^A_{i,t-1}) + K_3(\tilde{c}^I_{i,t} - a\tilde{f}^A_{i,t-1}). \quad (16)$$

Investors observe $\tilde{f}^A_{i,t}$ but believe that the analysts' expectation is formed according to (16), which leads to a biased inference of their private signal $c_{i,t-1}$.

***REG*, Expectation Formation, and Mispricing** We demonstrate that through the dynamic belief updating between the two groups of agents, an initial bias is reinforced and translates to the dynamics of analyst forecast errors and firm mispricing as observed in the data. We first introduce a bias into our model that follows diagnostic expectations in line with Bordalo et al. (2019). Precisely, we consider the case $\theta^I = (1+\beta, 1, 1)$ with $\beta = 0.5$. We simulate the expectation formation of investors and analysts based on our model for a cross-section of 25,000 observations.

[Figure 4]

Figure 4 illustrates the results. In the upper left panel, we plot the cross-sectional average of actual earnings growth as well as the investors' and analysts' expectations for the whole cross-section and confirm that all of them average out to zero. In contrast, we compute averages only of those observations with a large positive REG at $t = 1$ in the other panels of the figure. These observations are, by selection, those ones for which the actual earnings growth at $t = 1$ is positive and increased relative to the previous observation, but for which the investors' expected earnings growth is even larger and more strongly increasing. We investigate for our model with diagnostic investor expectations (lower left panel) how the positive REG for these observations translates to analyst expectations, and directly observe that they clearly exceed the actual earnings growth at $t = 1$. This positive bias of analyst expectations is a direct result of the overly positive investor reaction to the earnings announcement, based on which analysts update their beliefs. At $t = 2$, the upward-biased analyst expectations feed back into the belief formation of investors again, such that their expectations further increase and depart from the actual earnings growth. As a result, the expectations produced by the model directly correspond to the build-up in mispricing — the deviation of market valuations from fundamental valuations — that we document empirically. The model also shows that after $t = 2$, the investor and analyst expectations converge towards actual earnings growth only at a slow rate, such that mispricing and positive AFEs persist for a while.

These dynamics are in stark contrast to the case where we consider the identical model but without biases, as the upper right panel of Figure 4 illustrates. In that case, both analyst and investor expectations very much converge towards the actual rate of earnings growth by

$t = 2$, even though they start far away from it as a result of the large initial *REG*. Finally, we introduce an analyst confirmation bias into our model, as motivated by our empirical results in Section 4.2. We do so by setting the analysts' bias θ_2^A on the inferred investor signal to 1.5 if the current analyst forecast error $a\tilde{f}_{i,t-1}^A + bx_{i,t-1} - x_{i,t}$ and the inferred investor signal $\tilde{s}_{i,t}^A - a\tilde{f}_{i,t-1}^A$ (relative to prior expectations) have the same sign, and to 1 otherwise. As a result, analysts put a 50% greater weight on the signal observed from the market if it provides confirming information. The lower right panel of Figure 4 shows that the confirmation channel reinforces analyst and investor biases and leads to larger and more slowly converging deviations of expectations from the firms' actual earnings growth.

8 Conclusion

How investors form their expectations and how their expectations drive asset prices has been in the interest of academic research over the last several decades. Recent research highlights cognitive and other constraints that lead to biased expectation formation.

In this paper, we provide new empirical evidence, which adds to this growing line of research. We show that market participants are likely to take into account the (biased) actions of other agents when forming their expectations. Consequently, expectation formation across market participants is a dynamic process featuring feedback effects that can result in an amplification of agents' initial bias.

Based on a new measure that captures market participants' response to earnings information, we explore the interactions between investors who trade and reflect their beliefs when earnings information is released and analysts who provide their expectations about future firm earnings. We uncover a positive dynamic relation between market participants' reaction to earnings, analyst earnings forecast errors, and the degree of firm mispricing. In particular, we show that future analyst forecast errors are predicted by the disproportionate market response on earnings announcement days, and the predictability is more pronounced when the market reaction to earnings confirms the analysts' prior views. We also show that the market's initial reaction to earnings predicts the early build-up stage in firm mispricing.

To formalize our interpretation of these results, we present a simple structural model in which investors and analysts dynamically update their expectations of the firms' earnings growth rate. Each type of agent tries to infer the other agent's private signal from public information. When all agents' expectation formation is unbiased, analyst forecast errors and firm mispricing scores are zero on average. However, when the investors' belief formation is biased but both types of agents are unaware of this bias, then the bias is propagated and reinforced between the two groups of agents.

Overall, the dynamics that we document in this paper suggest that different market participants are affected by the biased belief formation of other agents. Our findings contribute to the understanding of investors' belief formation and their effect on asset prices. In particular, they demonstrate the potential spillover effects in investors' expectation formation, which result in amplification effects. They also add to the ongoing debate on the source of anomaly returns. Future research should consider these dynamics and further assess their impact on trading and asset prices.

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645–1680.

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: An extrapolative capital asset pricing model. Journal of Financial Economics 115, 1–24.

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2018. Extrapolation and bubbles. Journal of Financial Economics 129, 203–227.

Barberis, N., Shleifer, A., Vishny, R. W., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Bernard, V. L., Thomas, J. K., 1989. Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting Research 27, 1–36.

Bernard, V. L., Thomas, J. K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305–340.

Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2019. Diagnostic expectations and stock returns. Journal of Finance 74, 2839–2874.

Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2020. Overreaction in macroeconomic expectations. Journal of Finance 110, 2748–2782.

Bordalo, P., Gennaioli, N., Shleifer, A., 2018. Diagnostic expectations and credit cycles. Journal of Finance 73, 199–227.

Brav, A., Lehavy, R., 2003. An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics. Journal of Finance 58, 1933–1967.

Cassella, S., Gulen, H., 2018. Extrapolation bias and the predictability of stock returns by price-scaled variables. Review of Financial Studies 31, 4345–4397.

Chan, L. K., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. Journal of Finance 51, 1681–1713.

Chen, A. Y., Zimmermann, T., 2022. Open source cross sectional asset pricing. Critical Finance Review 11, 207–264.

Cieslak, A., 2018. Short-rate expectations and unexpected returns in treasury bonds. Review of Financial Studies 31, 3265–3306.

Clement, M. B., 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? Journal of Financial Economics 27, 285–303.

Cookson, J. A., Engelberg, J., Mullins, W., 2021. Echo chambers. Working Paper .

Cready, W. M., Gurun, U. G., 2010. Aggregate market reaction to earnings announcements. Journal of Accounting Research 48, 289–334.

Da, Z., Huang, X., Jin, L. J., 2021. Extrapolative beliefs in the cross-section: What can we learn from the crowds? Journal of Financial Economics 140, 175–196.

Da, Z., Liu, Q., Schaumburg, E., 2014. A closer look at the short-term return reversal. Management Science 60, 658–674.

Da, Z., Schaumburg, E., 2011. Relative valuation and analyst target price forecasts. Journal of Financial Markets 14, 161–192.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

De la O, R., Myers, S., 2021. Subjective cash flow and discount rate expectations. Journal of Finance 76, 1339–1387.

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Positive feedback investment strategies and destabilizing rational speculation. Journal of Finance 45, 379–395.

Edelen, R. M., Ince, O. S., Kadlec, G. B., 2016. Institutional investors and stock return anomalies. Journal of Financial Economics 119, 472–488.

Engelberg, J., McLean, R. D., Pontiff, J., 2018. Anomalies and news. Journal of Finance 73, 1971–2001.

Engelberg, J., McLean, R. D., Pontiff, J., 2020. Analysts and anomalies. Journal of Accounting and Economics 69, 101249.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.

Frankel, R., Lee, C. M., 1998. Accounting valuation, market expectation, and cross-sectional stock returns. Journal of Accounting and Economics 25, 283–319.

Gerken, W. C., Painter, M., 2019. The value of differing points of view: Evidence from financial analysts' geographic diversity. Review of Financial Studies, forthcoming .

Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. Review of Financial Studies 27, 714–746.

Grossman, S., 1976. On the efficiency of competitive stock markets where trades have diverse information. Journal of Finance 31, 573–585.

Hartzmark, M. S., Shue, K., 2018. A tough act to follow: Contrast effects in financial markets. Journal of Finance 73, 1567–1613.

Hayek, F. A., 1945. The use of knowledge in society. American Economic Review 35, 519–530.

Hirshleifer, D., Lourie, B., Ruchti, T. G., Truong, P., 2021. First impression bias: Evidence from analyst forecasts. Review of Finance 25, 325–364.

Holthausen, R. W., Verrecchia, R. E., 1988. The effect of sequential information releases on the variance of price changes in an intertemporal multi-asset market. Journal of Accounting Research 26, 82–106.

Hughes, J., Liu, J., Su, W., 2008. On the relation between predictable market returns and predictable analyst forecast errors. Review of Accounting Studies 13, 266–291.

Jegadeesh, N., Kim, J., Krische, S. D., Lee, C. M., 2004. Analyzing the analysts: When do recommendations add value? Journal of Finance 59, 1083–1124.

Kacperczyk, M., Sialm, C., Zheng, L., 2005. On the industry concentration of actively managed equity mutual funds. Journal of Finance 60, 1983–2011.

Ke, Y., Lo, K., Sheng, J., Zhang, J. L., 2020. Do investors affect financial analysts' behaviors? evidence from short sellers. Working Paper .

Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting factor models. Journal of Finance 73, 1183–1223.

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

La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R., 1997. Good news for value stocks: Further evidence on market efficiency. Journal of Finance 52, 859–874.

Liptser, R. S., Shiryaev, A. N., 2001. Statistics of random processes: I. General theory. Springer Science Business Media.

Ludvigson, S. C., Bianchi, F., Ma, S., 2021. Belief distortions and macroeconomic fluctuations. Working Paper .

Malmendier, U., Shanthikumar, D., 2007. Are small investors naive about incentives? Journal of Financial Economics 85, 457–489.

Mendenhall, R. R., 2004. Arbitrage risk and post-earnings-announcement drift. Journal of Business 77, 875–894.

Miller, E. M., 1977. Risk, uncertainty, and divergence of opinion. Journal of Finance 32, 1151–1168.

Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–08.

Nickerson, R. S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology 2, 175–220.

Pouget, S., Sauvagnat, J., Villeneuve, S., 2017. A mind is a terrible thing to change: confirmatory bias in financial markets. Review of Financial Studies 30, 2066–2109.

Rabin, M., Schrag, J. L., 1999. First impressions matter: A model of confirmatory bias. Quarterly Journal of Economics 114, 37–82.

So, E. C., 2013. A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts? Journal of Financial Economics 108, 615–640.

Stambaugh, R. F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288–302.

Stambaugh, R. F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. Journal of Finance 70, 1903–1948.

Stambaugh, R. F., Yuan, Y., 2017. Mispricing factors. Review of Financial Studies 30, 1270–1315.

van Binsbergen, J. H., Boons, M., Opp, C. C., Tamoni, A., 2021. Dynamic asset (mis)pricing: Build-up vs. resolution anomalies. Working Paper .

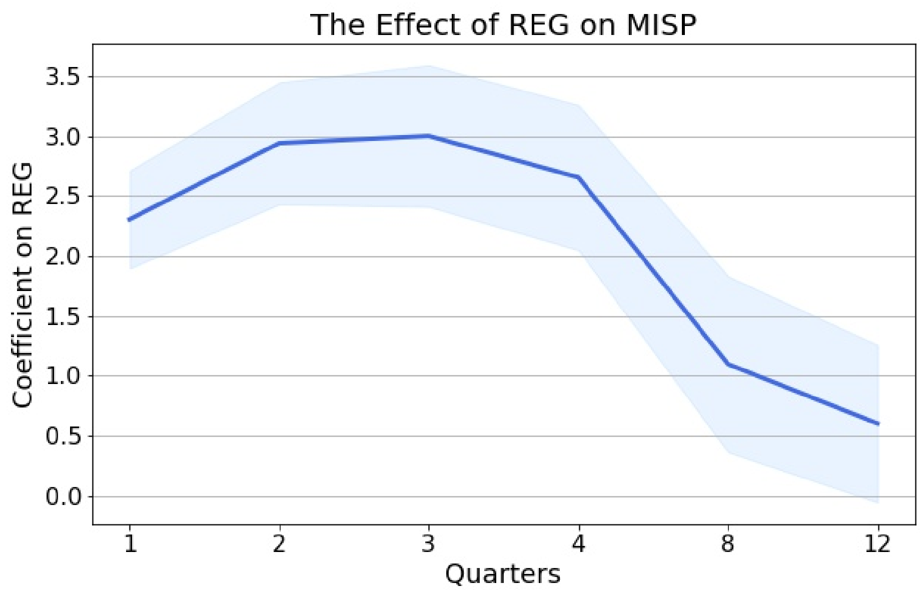


Fig. 1 - The Mispricing Cycle

The figure above show the coefficients and the corresponding 95% confidence intervals on *REG* in the Fama and MacBeth (1973) regressions for predicting *MISP* in 1, 2, 3, 4, 8, and 12 quarters ahead. The sample period is from January 1985 to December 2018. Standard errors are adjusted for serial correlation using Newey and West (1987) correction.

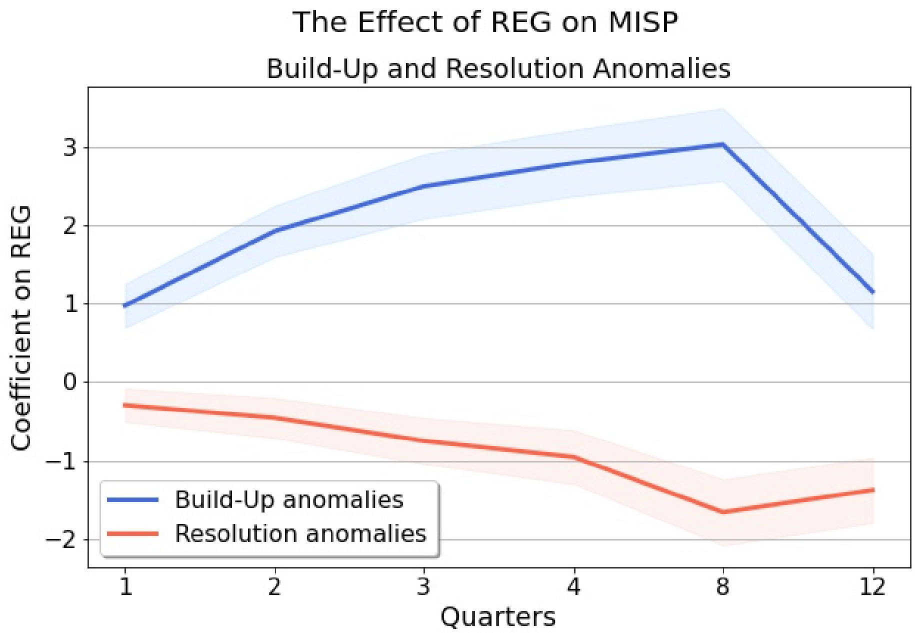
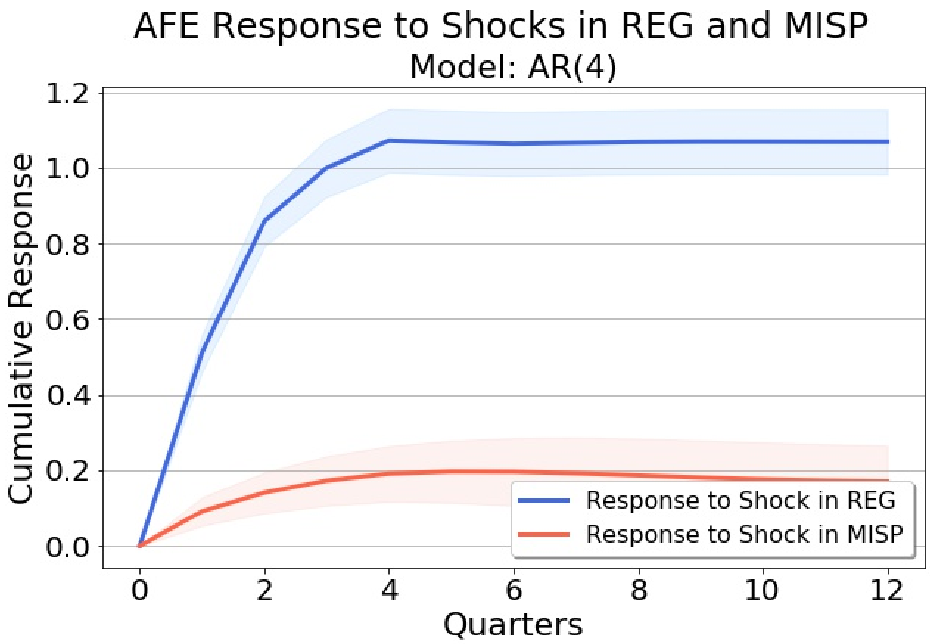


Fig. 2 - Anomaly Dissection

The figure above show the coefficients and the corresponding confidence 95% intervals on REG in the Fama and MacBeth (1973) regressions for predicting the mispricing scores $MISP_{BUILD}$ and $MISP_{RES}$ in 1, 2, 3, 4, 8, and 12 quarters ahead. The sample period is from January 1985 to December 2018. Standard errors are adjusted for serial correlation using Newey and West (1987) correction.



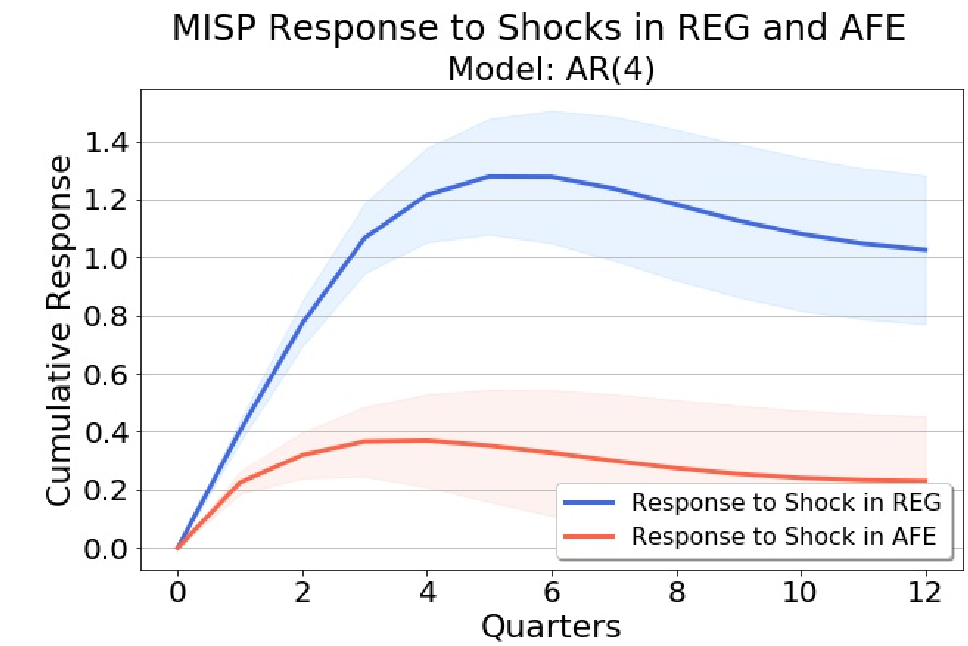
(To be Continued)

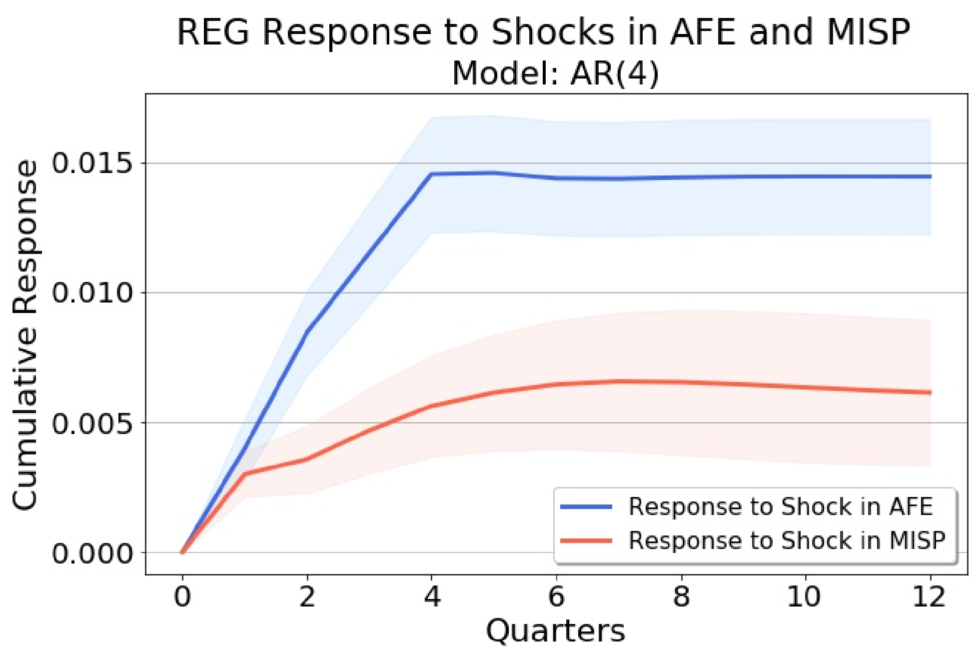
Fig. 3 - Impulse Response

$$
\begin{aligned}
AFE_{j,q} = &\quad \alpha_1 + \sum_{i=1}^{4} \beta_{1,i} \cdot AFE_{j,q-i} + \sum_{i=1}^{4} \gamma_{1,i} \cdot MISP_{j,q-i} + \sum_{i=1}^{4} \theta_{1,i} \cdot REG_{j,q-i} + \delta \cdot X_{j,q-1} + f_j + q_t + \epsilon_{1,j,q};\\
MISP_{j,q} = &\quad \alpha_2 + \sum_{i=1}^{4} \beta_{2,i} \cdot AFE_{j,q-i} + \sum_{i=1}^{4} \gamma_{2,i} \cdot MISP_{j,q-i} + \sum_{i=1}^{4} \theta_{2,i} \cdot REG_{j,q-i} + \delta \cdot X_{j,q-1} + f_j + q_t + \epsilon_{2,j,q};\\
REG_{j,q} = &\quad \alpha_3 + \sum_{i=1}^{4} \beta_{3,i} \cdot AFE_{j,q-i} + \sum_{i=1}^{4} \gamma_{3,i} \cdot MISP_{j,q-i} + \sum_{i=1}^{4} \theta_{3,i} \cdot REG_{j,q-i} + \delta \cdot X_{j,q-1} + f_j + q_t + \epsilon_{3,j,q}.
\end{aligned}
$$

The above and the following figures plot the impulse responses of AFE, $MISP$ and REG to a one-standard-deviation shock to these variables. We estimate a quarterly vector autoregression (VAR) system of AFE, $MISP$ and REG, with four lags of each variable. The regressions include the full set of firm control variables ($X_{j,q-1}$) together with firm fixed effects (f_j) and quarter fixed effects (q_t). The VAR system takes the form as shown in the above equation system. We do not allow quarter-0 shocks to enter the system and affect the variables. Thus, the responses are only based on the lags of the system. Each graph depicts the response in the subsequent 0, 1, 2, 6, 8, 10, and 12 quarters, listed on the x-axis. The solid lines depict the variable responses and the shaded areas depict the 95% confidence intervals. The sample period is from January 1985 to December 2018.

(Continued)


MISP Response to Shocks in REG and AFE
Model: AR(4)
Cumulative Response
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
Quarters
Response to Shock in REG
Response to Shock in AFE


REG Response to Shocks in AFE and MISP
Model: AR(4)
Cumulative Response
0.015
0.010
0.005
0.000
0
2
4
6
8
10
12
Quarters
Response to Shock in AFE
Response to Shock in MISP

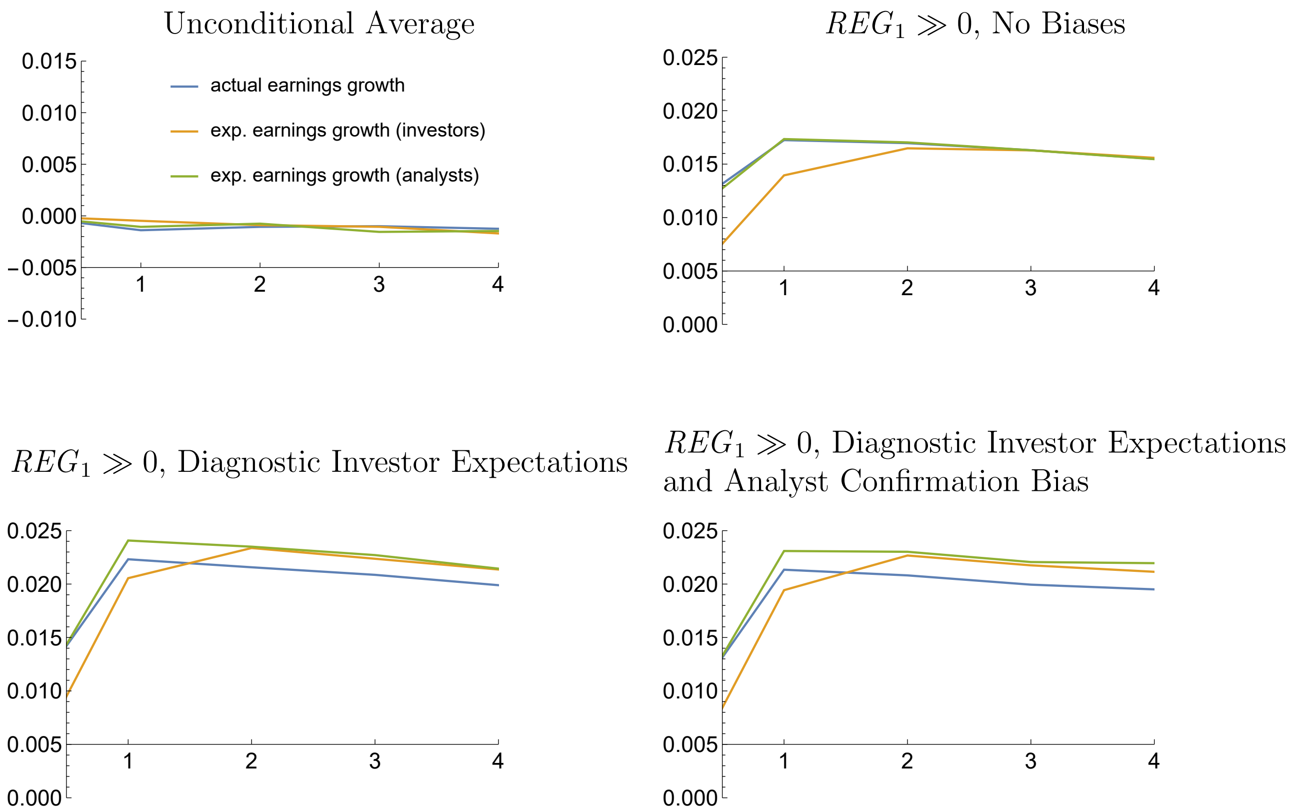


Fig. 4 - Structural Model: Dynamic Expectation Formation

The figure above depicts the average dynamics of actual earnings growth as well as investor and analyst expectations unconditionally (upper left panel) and for observations with a large positive REG at $t = 1$ (lower panels and right panels). We simulate a cross-section of firm-earnings-announcement observations based on our model and plot cross-sectional averages of $f_{i,t}$, $\tilde{f}^I_{i,t}$, and $\tilde{f}^A_{i,t}$ over time, for the whole sample as well as for the subsample where the disparity between market response and earnings surprise at $t = 1$ is in the top quantile. The blue line shows the actual earnings growth, the yellow line the earnings growth as expected by the investors, and the green line the earnings growth as expected by analysts. We consider three model specifications: the model with no biases ($\beta = 0$, upper right panel), the model with diagnostic investor expectations ($\beta = 0.5$, lower left panel), and the model with diagnostic investor expectations and analyst confirmation bias (lower right panel). The other model parameters are calibrated as $b = 0.56$, $a = 0.96$, $\sigma_\varepsilon = 0.08$, and $\sigma_\eta = 0.14$ in line with Bordalo et al. (2019), as well as $\sigma_\chi = 0.09$, $\sigma_\phi = 0.14$.

Table 1 - Variable Definition

This table provides definition for the major variables in our analysis.

Variable	Definition
DGTW	Characteristic-adjusted daily stock return constructed following Daniel et al. (1997), calculated by subtracting the return on a peer portfolio consisting of stocks with similar size, book-to-market ratio and past return momentum.
SUE	The difference between actual EPS and median of analysts' estimated EPS scaled by the standard deviation of analysts' forecasts (adjusted for dividends and stock splits).
AdjSUE	The residual from a regression of *SUE* on *LnSIZE*, *LnBM*, and day-of-week and month-of-year fixed effects.
REG	The difference in the rankings of *DGTW* and *AdjSUE* of the stock on earnings announcement day t.
AFE	Analyst earnings forecast errors. The difference between the median of analysts' estimated EPS and the actual EPS, scaled by the standard deviation of the analysts' forecasts (adjusted for dividends and stock splits).
RetForeErr	Analyst price-target based return forecast error (in %). The average of the return forecast errors across analysts issuing price targets over the subsequent 60 days following an earnings announcement. An analyst return forecast error is defined as ((Future price target - Actual Future Price)/Current price) - 1.
RecChng	The average recommendation changes issued by analysts, multiplied by -1.
MISP	Monthly mispricing score of Stambaugh et al. (2015).
InstDirTrd	Institutional investors' daily shares bought minus shares sold normalized by total daily stock volume (in %).
LnSIZE	The natural log of the firm size.
LnBM	The natural log of the firm Book-to-Market ratio.
RET5	Stock cumulative return over the past 5 trading days (in %).
RET21	Stock cumulative return over the past 21 trading days (in %).
MOM	Momentum. The average of daily returns over the period from t-252 to t-21 (in %).
RVOL	Realized volatility of stock. The square root of the annualized realized variance, which is 252 times the average squared daily returns over the past 21 trading days.
ILLIQ	Amihud (2002) illiquidity measure. The average ratio of absolute daily return by daily total dollar trading volume of stock over the past 21 trading days.
DISP	Dispersion of analyst's earnings forecast. The standard deviation of analysts' earnings forecasts scaled by stock price.
NUMEST	The natural logarithm of one plus the number of analysts issuing earnings forecasts.
MRET	Monthly cumulative return (in %).
MMOM	Monthly momentum. The cumulative monthly return over the past 11 months (in %).
MRVOL	Monthly realized volatility. The standard deviation of monthly returns over the 12 months ending in each June; if at least 9 monthly returns available, then apply the *MRVOL* to the following 12 months, i.e. from July of the same year to June of the next year).
MILLIQ	Monthly illiquidity. The average daily Amihud (2002) illiquidity ratio over all trading days during the month.

Table 2 - Descriptive Statistics

This table reports the descriptive statistics of the main variables in our analysis. Our sample consists of 8,434 distinct companies, which had analyst forecasts on EPS and actual EPS in the I/B/E/S database from January 1985 to December 2018. Panel A reports the observation-weighted time-series average of the cross-sectional mean, standard deviation, and quintiles of each variable. Panel B shows the observation-weighted time-series average of the cross-sectional correlations of some key variables in our regression analysis.

Panel A: Cross-Sectional Summary Statistics

	Mean	SD	P1	P25	Median	P75	P99
REG	0.000	0.172	-0.377	-0.113	0.001	0.114	0.372
SUE	0.193	5.348	-19.404	-0.829	0.421	1.820	12.785
DGTW	0.000	6.122	-17.757	-2.613	0.004	2.688	16.973
AFE	-0.193	5.348	-12.785	-1.820	-0.421	0.829	19.404
RetForeErr	17.719	46.308	119.348	44.444	12.056	-12.336	-64.183
$RecChng_{t+1:t+5}$	0.098	1.405	2.000	1.000	1.000	-1.000	-2.000
$RecChng_{t+6:t+15}$	0.177	1.423	2.000	1.000	1.000	-1.000	-2.000
MISP	50.276	12.582	24.491	41.404	49.832	58.779	78.780
InstDirTrd	0.266	12.858	-42.334	-1.880	0.000	2.670	41.953
LnSIZE	6.822	1.568	3.791	5.697	6.725	7.841	10.608
LnBM	-0.795	0.781	-3.044	-1.228	-0.706	-0.277	0.778
RET5	0.420	5.837	-13.649	-2.583	0.150	3.064	17.831
RET21	0.903	11.463	-26.672	-5.210	0.443	6.335	34.852
MOM	15.556	49.190	-60.433	-12.148	8.568	32.403	196.426
RVOL	0.416	0.234	0.124	0.263	0.362	0.508	1.248
ILLIQ	0.206	1.115	0.000	0.002	0.010	0.051	5.606

Panel B: Selected Cross-Sectional Correlations

	REG	*SUE*	*DGTW*	*AFE*	*MISP*
REG	1.000				
SUE	-0.436	1.000			
DGTW	0.514	0.211	1.000		
AFE	0.436	-1.000	-0.211	1.000	
MISP	0.051	-0.097	-0.017	0.097	1.000

Table 3 - REG and Stock Returns

This table reports the average $DGTW$ abnormal returns on day t, cumulative $DGTW$ abnormal returns from day t+1 to day t+20, and from day t to day t+20 to single-sorted portfolios based on REG of day t. The sample period is from January 1985 to December 2018. The average $DGTW$ abnormal returns on day t, cumulative $DGTW$ abnormal returns from day t+1 to day t+20, and from day t to day t+20 on a high-minus-low (H-L) portfolio that longs stocks in the top decile and shorts stocks in the bottom decile are presented in the last column. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. Standard errors are adjusted for serial correlation using Newey and West (1987) correction. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	Decile Portfolios Sorted By REG_t										
	Low	D2	D3	D4	D5	D6	D7	D8	D9	High	H - L
$DGTW_t$	-5.257***	-3.281***	-2.365***	-1.769***	-0.842***	0.708***	1.791***	2.545***	3.233***	5.138***	10.395***
	(-95.97)	(-71.83)	(-55.51)	(-38.69)	(-14.84)	(14.19)	(41.67)	(56.28)	(66.17)	(95.91)	(115.47)
#Obs	3,439	3,439	3,439	3,439	3,439	3,439	3,439	3,439	3,439	3,439	3,439
$DGTW_{t+1:t+20}$	0.543***	0.296***	0.310***	0.242***	0.315***	0.228***	-0.092	-0.081	-0.406***	-0.445***	-0.988***
	(4.70)	(2.84)	(3.46)	(2.76)	(3.44)	(2.86)	(-1.21)	(-0.88)	(-4.68)	(-3.58)	(-6.68)
#Obs	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419
$DGTW_{t:t+20}$	-4.711***	-2.963***	-2.046***	-1.494***	-0.470***	1.005***	1.752***	2.513***	2.846***	4.697***	9.408***
	(-26.28)	(-21.06)	(-17.9)	(-12.71)	(-4.17)	(10.48)	(16.43)	(18.47)	(20.33)	(24.00)	(30.91)
#Obs	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419	3,419

Table 4 - *REG* and Institutional Trading

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting institutional investors' directional trading. The sample includes 5,005 distinct stocks from February 2002 to December 2015. Each column name signifies the dependent variable. $InstDirTrd_t$ is the normalized directional trading (net buying) by institutional investors on day t. $InstDirTrd_{t+1:t+d}$ indicates the cumulative directional trading by institutional investors over the period from day $t+1$ to $t+d$ ($d = 5, 10, 15$). Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, and *ILLIQ* as introduced in Section 2.3. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. Standard errors are adjusted for serial correlation using Newey and West (1987) correction. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) $InstDirTrd_t$	(2) $InstDirTrd_{t+1:t+5}$	(3) $InstDirTrd_{t+1:t+10}$	(4) $InstDirTrd_{t+1:t+15}$
REG	3.939***	9.481***	8.587***	5.154
	(8.40)	(5.70)	(3.13)	(1.32)
SUE	0.010	0.257***	0.498***	0.693***
	(0.55)	(3.75)	(4.50)	(4.16)
DGTW	0.047***	0.171***	0.404***	0.633***
	(4.20)	(4.07)	(5.66)	(6.00)
LnSIZE	-0.133***	-1.045***	-1.976***	-2.998***
	(-3.61)	(-6.53)	(-6.25)	(-6.25)
LnBM	-0.013	-0.455**	-1.046***	-1.565***
	(-0.21)	(-2.14)	(-2.96)	(-3.38)
RET5	0.189***	0.223***	0.303***	0.355***
	(14.82)	(4.81)	(3.66)	(3.23)
RET21	0.026***	0.040*	0.037	0.011
	(4.20)	(1.79)	(0.89)	(0.21)
MOM	0.002*	0.024***	0.055***	0.077***
	(1.91)	(4.65)	(6.45)	(6.43)
RVOL	-0.291	-1.752	-4.577*	-8.678***
	(-0.87)	(-1.37)	(-1.96)	(-2.70)
ILLIQ	1.250	1.243	-1.913	-7.255
	(0.82)	(0.21)	(-0.20)	(-0.53)
Intercept	1.078***	9.021***	17.554***	28.016***
	(3.20)	(6.31)	(6.15)	(6.52)
Adj. R-squared	1.11%	0.56%	0.67%	0.92%
#Days	1,265	1,265	1,263	1,262
#Obs	100,594	100,534	100,455	100,367

Table 5 - The Effect of *REG* on Analyst Earnings Forecast Errors

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting the *AFE* in the following quarters. The sample period is from January 1985 to December 2018. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) AFE_{q+1}	(2) AFE_{q+2}	(3) AFE_{q+3}	(4) AFE_{q+4}	(5) AFE_{q+8}	(6) AFE_{q+12}
REG	2.464***	1.699***	1.397***	1.541***	1.253***	0.979***
	(11.93)	(7.23)	(5.21)	(5.87)	(4.29)	(4.10)
AFE	0.135***	0.096***	0.077***	0.068***	0.062***	0.047***
	(13.62)	(8.33)	(6.28)	(5.45)	(4.03)	(3.98)
DGTW	-0.076***	-0.048***	-0.048***	-0.043***	-0.035***	-0.015*
	(-9.19)	(-5.89)	(-3.27)	(-5.04)	(-3.53)	(-1.84)
MISP	0.016***	0.020***	0.017***	0.015***	0.018***	0.016***
	(9.74)	(8.77)	(9.01)	(8.47)	(8.00)	(7.99)
LnSIZE	-0.092***	-0.053**	-0.071***	-0.079***	-0.109***	-0.127***
	(-4.63)	(-2.5)	(-3.32)	(-3.50)	(-4.45)	(-4.46)
LnBM	0.158***	0.149***	0.101**	0.129***	0.13***	0.057*
	(4.45)	(4.08)	(2.55)	(3.08)	(3.77)	(1.65)
RET5	-0.008	0.000	-0.007	-0.005	0.008	0.001
	(-1.62)	(0.03)	(-1.23)	(-0.94)	(1.28)	(0.15)
RET21	-0.010***	-0.008***	-0.007**	-0.008**	-0.002	-0.004
	(-3.78)	(-2.94)	(-2.41)	(-2.28)	(-0.59)	(-1.35)
MOM	-0.006***	-0.005***	-0.003***	-0.001	0.002***	0.002***
	(-10.82)	(-6.67)	(-4.24)	(-1.43)	(2.79)	(2.80)
RVOL	-0.027	0.303*	-0.035	0.161	-0.420**	-0.665***
	(-0.20)	(1.94)	(-0.17)	(0.84)	(-2.22)	(-3.24)
ILLIQ	1.763**	1.702*	2.566**	3.384**	2.202	-2.052
	(2.06)	(1.81)	(2.45)	(2.02)	(0.77)	(-1.07)
DISP	28.684***	9.512	23.271***	20.502***	14.924*	40.856***
	(4.30)	(1.62)	(3.15)	(3.33)	(1.81)	(4.97)
NUMEST	-0.103**	-0.189***	-0.140***	-0.063	-0.068	-0.016
	(-2.07)	(-3.73)	(-2.95)	(-1.12)	(-1.23)	(-0.28)
Intercept	-0.060	-0.454**	-0.216	-0.369*	-0.142	-0.079
	(-0.36)	(-2.37)	(-1.04)	(-1.82)	(-0.69)	(-0.34)
Adj. R-sqaured	9.19%	7.62%	6.28%	5.64%	4.78%	3.61%
#Days	2,355	2,330	2,321	2,297	2,203	2,043
#Obs	172,926	168,681	165,079	162,126	150,073	134,978

Table 6 - The Mispricing Cycle

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting for *MISP* in the following quarters. The sample period is from January 1985 to December 2018. In the regressions, All dependent variables except for *REG*, *AFE*, and *DGTW*, are observed at the end of the month of earnings announcement day t. Firm control variables include *LnSIZE*, *LnBM*, *MRET*, *MMOM*, *MRVOL*, and *MILLIQ* as introduced in Section 2.3. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) $MISP_{q+1}$	(2) $MISP_{q+2}$	(3) $MISP_{q+3}$	(4) $MISP_{q+4}$	(5) $MISP_{q+8}$	(6) $MISP_{q+12}$
REG	2.304***	2.939***	2.999***	2.653***	1.097***	0.602*
	(11.07)	(11.34)	(9.97)	(8.60)	(2.93)	(1.80)
AFE	0.059***	0.030***	0.040***	0.030***	0.022**	0.024**
	(9.91)	(4.21)	(4.64)	(2.80)	(2.20)	(2.51)
DGTW	-0.087***	-0.098***	-0.084***	-0.067***	-0.017	0.005
	(-10.84)	(-8.90)	(-6.77)	(-4.87)	(-1.19)	(0.37)
MISP	0.841***	0.769***	0.662***	0.559***	0.463***	0.409***
	(86.00)	(73.43)	(64.64)	(112.68)	(84.01)	(69.60)
LnSIZE	-0.232***	-0.383***	-0.571***	-0.755***	-1.000***	-1.010***
	(-7.74)	(-9.66)	(-11.42)	(-15.20)	(-17.44)	(-15.83)
LnBM	-0.263***	-0.340***	-0.246***	0.007	0.614***	1.096***
	(-5.26)	(-4.98)	(-3.11)	(0.08)	(7.88)	(11.03)
MRET	-0.124***	-0.116***	-0.102***	-0.096***	0.036***	0.017***
	(-32.53)	(-26.42)	(-21.41)	(-17.87)	(6.66)	(3.05)
MMOM	0.008***	0.035***	0.065***	0.091***	0.091***	0.069***
	(6.70)	(24.46)	(36.47)	(42.11)	(37.33)	(29.58)
MRVOL	2.757***	3.677***	4.637***	5.673***	-4.042**	-7.279***
	(2.64)	(2.85)	(3.42)	(4.15)	(-2.34)	(-3.73)
MILLIQ	-0.496***	-0.515**	-0.902***	-1.308***	-1.245***	-0.478
	(-3.53)	(-2.55)	(-3.34)	(-3.20)	(-3.88)	(-1.24)
Intercept	9.082***	13.387***	19.645***	25.697***	33.023***	36.138***
	(14.74)	(19.64)	(27.73)	(59.84)	(71.91)	(69.98)
Adj. R-squared	76.42%	62.77%	47.56%	36.03%	27.06%	22.64%
#Months	203	202	200	197	188	182
#Obs	129,589	125,581	122,006	118,183	106,572	95,984

Table 7 - The Effect of *REG* on Analyst Forecast Errors: Analyst Heterogeneity

This table reports results from panel regressions of individual analyst forecast errors (*AFE*) in quarters q+1 - q+4 on *REG* in quarter q, *REG* interaction with analyst characteristics and other control variables. In particular, we consider two analyst characteristics: 1) the degree of analyst industry concentration, and 2) analyst accuracy. To capture industry generalization we construct $Rank(NumInd)$ as the decile ranking based on the number of industries covered by an analyst in the given quarter. To capture accuracy, we use Clement (1999) $PMAFE$ (Proportionate Mean Absolute Forecast Error) measure, and construct $Rank(PMAFE)$ as the decile ranking of an analyst stock-level $PMAFE$ over the past year. Control variables include *AFE*, *DGTW*, *MISP*, and the number of days between analyst's earnings forecast and the earnings announcement. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. The sample period is from January 1985 to December 2018. All the regressions include analyst and quarter fixed effects and standard errors are clustered on analyst and quarter. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) AFE_{q+1}	(2) AFE_{q+2}	(3) AFE_{q+3}	(4) AFE_{q+4}
	Industry Concentration: Generalist vs. Specialist			
REG	1.809***	1.828***	1.648***	1.359***
	(5.43)	(3.72)	(3.88)	(5.19)
$Rank(NumInd)$	-0.004	0.013	-0.002	-0.001
	(-0.68)	(1.40)	(-0.33)	(-0.09)
$REG \times Rank(NumInd)$	0.101***	0.032	0.040	0.054
	(3.51)	(1.17)	(1.26)	(1.64)
Control Variables	Yes	Yes	Yes	Yes
Fixed Effects	Analyst, Quarter	Analyst, Quarter	Analyst, Quarter	Analyst, Quarter
Adj. R-sqaured	11.93%	3.88%	4.36%	3.28%
#Obs	622,831	572,326	527,047	490,863
	Analyst Accuracy: Accurate vs. Inaccurate			
REG	1.781***	1.134***	1.298***	1.580***
	(3.19)	(2.87)	(6.37)	(6.23)
$Rank(PMAFE)$	-0.020***	-0.015***	-0.015***	-0.009**
	(-7.27)	(-4.94)	(-4.02)	(-2.50)
$REG \times Rank(PMAFE)$	0.052***	0.034**	0.035**	0.033**
	(3.67)	(2.22)	(2.41)	(2.01)
Control Variables	Yes	Yes	Yes	Yes
Fixed Effects	Analyst, Quarter	Analyst, Quarter	Analyst, Quarter	Analyst, Quarter
Adj. R-sqaured	4.64%	4.79%	3.63%	3.53%
#Obs	405,404	376,503	350,896	329,871

Table 8 - *REG* and Confirmation Bias of *AFE*

This table reports the results from regressions detailed in Equation (8). The dependent variables is a dummy that equals one if the *AFE* of a firm in quarter q is of the same sign as the firm's *AFE* in n (n = 1, 2, 3, 4, 8, and 12) quarters ahead. D(AFE_q & $REG_{t(q)}$ Same Sign) is a dummy that equals one if a firm's *AFE* in quarter q is of the same sign as its *REG* (observed later than *AFE*) in the same quarter. D(AFE_q & AFE_{q-1} Same Sign) is a dummy variable that equals one when the *AFE* of a firm in quarter q and quarter $q-1$ are of the same sign. *AFE*, *DGTW* and *MISP* are analyst forecast error, earnings announcement day *DGTW* abnormal returns, and misprcing scores of the firm in quarter q. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. The sample period is from January 1985 to December 2018. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	$q+1$	$q+2$	$q+3$	$q+4$	$q+8$	$q+12$
	(1)	(2)	(3)	(4)	(5)	(6)
D(AFE_q & $REG_{t(q)}$ Same Sign)	0.127***	0.111***	0.104***	0.107***	0.102***	0.098***
	(37.75)	(34.12)	(32.66)	(31.28)	(28.71)	(25.74)
D(AFE_q & AFE_{q-1} Same Sign)	0.208***	0.198***	0.205***	0.176***	0.155***	0.150***
	(64.78)	(58.77)	(60.50)	(50.81)	(42.2)	(38.72)
AFE	-0.013***	-0.013***	-0.013***	-0.014***	-0.016***	-0.018***
	(-25.35)	(-22.63)	(-22.92)	(-24.80)	(-26.50)	(-28.23)
DGTW	0.005***	0.004***	0.004***	0.005***	0.005***	0.004***
	(12.74)	(11.74)	(10.93)	(11.82)	(13.16)	(9.11)
MISP	-0.029**	-0.04***	-0.043***	-0.061***	-0.05***	-0.085***
	(-2.24)	(-2.94)	(-3.16)	(-4.33)	(-3.37)	(-5.27)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-sqaured	10.63%	9.98%	9.67%	8.62%	8.19%	8.38%
#Days	2,197	2,162	2,144	2,119	2,026	1,879
#Obs	159,697	155,970	152,986	150,025	139,139	125,029

Table 9 - Anomaly Dissection and the Mispricing Cycle

This table reports the coefficient on *REG* from the Fama and MacBeth (1973) cross-sectional regressions predicting for mispricing scores, $MISP_{BUILD}$ and $MISP_{RES}$, which are associated with two classes of anomalies: build-up and resolution, respectively. The dependent variables are the stock's average rankings with respect to anomalies within each class. In each month, we rank stocks according to each anomaly. The higher the ranking, the greater the degree of overvaluation. Then for each stock, we compute the equal-weighted average of rankings across all anomalies within the corresponding anomaly class. All dependent variables except for *REG*, *AFE*, and *DGTW*, are observed at the end of the month of earnings announcement day t. Firm control variables include *LnSIZE*, *LnBM*, *MRET*, *MMOM*, *MRVOL*, and *MILLIQ* as introduced in Section 2.3. The sample period is from January 1985 to December 2018. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	$q+1$	$q+2$	$q+3$	$q+4$	$q+8$	$q+12$
	(1)	(2)	(3)	(4)	(5)	(6)
			$MISP_{BUILD}$			
REG	0.974***	1.923***	2.498***	2.795***	3.031***	1.155***
	(6.86)	(11.50)	(11.87)	(12.87)	(12.87)	(4.73)
			$MISP_{RES}$			
REG	-0.297***	-0.458***	-0.753***	-0.957***	-1.662***	-1.380***
	(-2.75)	(-3.52)	(-5.04)	(-5.46)	(-7.75)	(-6.46)

Table 10 - The Mispricing Cycle: Cross-Sectional Heterogeneity

This table reports the coefficient on *REG* from the Fama and MacBeth (1973) cross-sectional regressions of *MISP* over the subsequent quarters (q+1 to q+12) on *REG* in quarter t and other control variables. The difference of the coefficient on *REG* in subsamples and the corresponding t-statistics are also reported. We consider cross-sectional subsamples based on monthly medians of: 1) analyst coverage, 2) firm market cap, 3) institutional ownership, and 4) analyst disagreement. All dependent variables except for *REG*, *AFE*, and *DGTW*, are observed at the end of the month of earnings announcement day t. Firm control variables include *LnSIZE*, *LnBM*, *MRET*, *MMOM*, *MRVOL*, and *MILLIQ* as introduced in Section 2.3. The sample period is from January 1985 to December 2018. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) $MISP_{q+1}$	(2) $MISP_{q+2}$	(3) $MISP_{q+3}$	(4) $MISP_{q+4}$	(5) $MISP_{q+8}$	(6) $MISP_{q+12}$
			Analyst Coverage			
Low Coverage	2.594***	3.198***	3.252***	2.998***	0.509	-0.091
	(9.96)	(9.60)	(8.42)	(7.59)	(1.01)	(-0.17)
High Coverage	1.769***	2.409***	2.538***	2.356***	1.413***	1.093**
	(6.98)	(7.24)	(6.47)	(5.87)	(2.88)	(2.40)
Low - High	0.825**	0.789*	0.714	0.642	-0.904	-1.184*
	(2.27)	(1.68)	(1.30)	(1.14)	(-1.29)	(-1.68)
			Firm Market Cap			
Small	2.680***	3.437***	3.508***	3.022***	0.494	-0.068
	(9.89)	(9.76)	(8.33)	(6.80)	(1.02)	(-0.14)
Large	2.092***	2.705***	2.758***	2.579***	1.425***	1.048**
	(7.74)	(8.02)	(7.12)	(6.06)	(3.01)	(2.15)
Small - Large	0.651**	0.732	0.750	0.443	-0.931	-1.116
	(1.73)	(1.50)	(1.31)	(0.72)	(-1.37)	(-0.23)
			Institutional Ownership			
Low IO	2.509***	3.279***	3.332***	3.160***	1.096**	0.812
	(10.30)	(10.38)	(8.65)	(7.78)	(2.30)	(1.53)
High IO	1.972***	2.461***	2.503***	2.110***	0.922**	0.251
	(7.43)	(7.31)	(6.48)	(5.21)	(1.98)	(0.56)
Low - High	0.537	0.818*	0.829	1.050*	0.174	0.561
	(1.49)	(1.77)	(1.52)	(1.83)	(0.26)	(0.81)
			Analyst Disagreement			
High DIS	2.784***	3.34***	3.204***	2.824***	0.501	-0.558
	(9.73)	(8.71)	(6.83)	(5.72)	(0.92)	(-1.08)
Low DIS	1.365***	2.090***	2.154***	1.723***	0.872*	0.732*
	(5.84)	(6.99)	(5.89)	(4.28)	(1.83)	(1.70)
High - Low	1.419***	1.250**	1.050*	1.101*	-0.371	-1.290*
	(3.84)	(2.57)	(1.77)	(1.73)	(-0.51)	(-1.92)

Table 11 - The Return Predictability of *MISP* Conditioning on *REG*

This table reports the cumulative monthly $DGTW$ abnormal returns from month m (an earning announcement month) to $m+2$ of portfolios formed based on the quintile ranking of *MISP* (an overvaluation score) as of the end of month $m-1$ and the sign of *REG* of the earnings announcement month m. Columns (1) and (2) report the $DGTW$ abnormal returns in month m (the earnings month), $MDGTW_m$. Columns (3) and (4) report the cumulative $DGTW$ abnormal returns in months m and $m+1$, $CumMDGTW_{m:m+1}$. Columns (5) and (6) report the cumulative $DGTW$ abnormal returns from month m to $m+2$, $CumMDGTW_{m:m+2}$. The sample period is from January 1985 to December 2018. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

		$MDGTW_m$		$CumMDGTW_{m:m+1}$		$CumMDGTW_{m:m+2}$	
		$REG > 0$	$REG < 0$	$REG > 0$	$REG < 0$	$REG > 0$	$REG < 0$
$MISP_{m-1}$	Q5	0.0175***	-0.0193***	0.0162***	-0.0194***	0.0097***	-0.0237***
		(9.66)	(-10.00)	(7.21)	(-8.07)	(3.35)	(-8.47)
	Q2-Q4	0.0183***	-0.0119***	0.0189***	-0.0093***	0.0182***	-0.0092***
		(19.23)	(-11.38)	(16.09)	(-7.82)	(12.51)	(-7.32)
	Q1	0.0182***	-0.0095***	0.0209***	-0.0066***	0.02***	-0.006***
		(18.43)	(-7.61)	(16.47)	(-4.53)	(14.19)	(-3.68)
Q5-Q1		-0.07%	-0.98%	-0.47%	-1.28%	-1.03%	-1.77%
Avg. Q5-Q1		-0.53%		-0.88%		-1.40%	
Q5_PosREG-Q1_NegREG		2.70%		2.28%		1.57%	
Q5_NegREG-Q1_PosREG		-3.75%		-4.03%		-4.37%	

Table 12 - *REG* Determinants

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting the *REG* in the next quarter. The sample period is from January 1985 to December 2018. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, *DISP*, and *NUMEST* as introduced in Section 2.3. Coefficients on *MISP*, *RET5*, *RET21*, and *MOM* are multiplied by 100 for readability. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	REG in the Next Quarter			
	(1)	(2)	(3)	(4)
AFE	0.007***	0.007***	0.006***	0.003***
	(48.02)	(42.32)	(33.44)	(13.26)
MISP	0.019***	0.049***	0.047***	0.045***
	(5.37)	(11.25)	(10.69)	(10.05)
REG			0.046***	0.136***
			(13.19)	(25.62)
DGTW				-0.004***
				(-19.80)
LnSIZE		0.016***	0.015***	0.014***
		(31.68)	(29.92)	(26.50)
LnBM		-0.015***	-0.014***	-0.013***
		(-20.12)	(-18.84)	(-16.45)
RET5		-0.040***	-0.028**	-0.031**
		(-2.86)	(-1.99)	(-2.17)
RET21		-0.067***	-0.060***	-0.055***
		(-9.14)	(-8.04)	(-7.39)
MOM		-0.015***	-0.013***	-0.013***
		(-9.21)	(-8.01)	(-7.76)
RVOL		-0.020***	-0.019***	-0.016***
		(-4.75)	(-4.60)	(-3.86)
ILLIQ		0.034	0.038	0.032
		(1.50)	(1.63)	(1.38)
DISP		0.612***	0.593***	0.505***
		(3.22)	(3.16)	(2.71)
NUMEST		-0.006***	-0.006***	-0.006***
		(-5.17)	(-5.17)	(-5.03)
Intercept	-0.002	-0.12***	-0.114***	-0.103***
	(-1.06)	(-27.45)	(-25.82)	(-23.20)
Adj. R-squared	2.84%	7.5%	7.86%	8.8%
#Days	3,197	2,483	2,411	2,353
#Obs	176,128	173,158	172,623	172,623

Appendix A

A.1 Alternative Specifications of REG

In this section, we repeat the main analyses in Section 3.2, 3.3, and 6.1 with two different specifications for *REG*. The first alternative specification constructs the *REG* measure without any adjustment of the earnings surprise or market response. Our unadjusted *REG* specification is based on the relative rankings of the raw return, *RET*, and the *SUE* on earnings announcement day t. The second alternative specification takes into account the total price pattern from day t up to day t+21. This accounts for the fact that analysts may consider the stock price change over a longer window. To this end, we replace the rank of *DGTW* abnormal return on day t with the cumulative *DGTW* abnormal return from day t to $t+20$. [15]

Table A.1 presents the coefficient on key variables of interest from main analyses with the two different alternative specifications of *REG*. [16] Panel A shows the results for predicting *AFE* in the following quarters. With the unadjusted *REG*, the coefficients on *REG* are always positive and significant as the forecasting horizon extends from the next quarter to 12 quarters ahead. To illustrate the economic significance of the impact of *REG* on *AFE*, we take the prediction for next quarter *AFE* for example. Specifically, a change in *REG* from its 25^{th} percentile to 75^{th} percentile leads to an increase in the next quarter's *AFE* by 0.570 $(= (0.111 - (-0.113)) \times 2.545)$, which is around 21.52% $(= 0.570/(0.829 - (-1.820)))$ of the difference between the 25^{th} percentile to the 75^{th} percentile in *AFE*. On the contrary, a rise in

[15] The mean, standard deviation, 25^{th} percentile, median, and 75^{th} percentile for unadjusted *REG* are -0.002, 0.170, -0.113, -0.001, and 0.111, respectively. The mean, standard deviation, 25^{th} percentile, median, and 75^{th} percentile for long-horizon *REG* are 0.000, 0.180, -0.122, 0.001, and 0.121, respectively.

[16] The regression specifications are the same as Eq. (5), Eq. (6), and Eq. (9), except for the variable $DGTW_{i,t}$. When adopting the unadjusted *REG*, we replace $DGTW_{i,t}$ with $RET_{i,t}$. When adopting the long-horizon *REG*, we replace $DGTW_{i,t}$ with $DGTW_{i,t:t+20}$.

MISP from its 25^{th} percentile to 75^{th} percentile would lead the next quarter *AFE* to increase by 0.278 ($= (58.779 - 41.404) \times 0.016$), which is equivalent to 10.49% ($= 0.278/(0.829 - (-1.820))$) of the difference between the 25^{th} percentile to the 75^{th} percentile in *AFE*. Similarly, the positive impact of the long-horizon *REG* on future *AFE* is also statistically and economically significant as evidenced by the positive coefficients on *REG* with t-statistics no less than 2.83 across all regressions. Given that the 25^{th} and the 75^{th} percentiles of the long-horizon *REG* are -0.122 and 0.121, respectively, a change in *REG* from its 25^{th} percentile to 75^{th} percentile would lead to a rise in next quarter *AFE* by 0.702 ($= (0.121 - (-0.122)) \times 2.888$). As a comparison, a change from 25^{th} percentile to 75^{th} percentile in *MISP* would result in a rise in next quarter *AFE* by 0.278 ($= (58.779 - 41.404) \times 0.016$), which is less than half of the change led by the rise in *REG* from its 25^{th} percentile to 75^{th} percentile. In sum, these results imply that the findings documented in Section 3.2 that higher *REG* would lead to greater analyst earnings forecast error is not discovered by chance. Instead, the positive impact of the return earnings gap on future *AFE* would remain intact when the *REG* is measured differently.

The results for predicting *MISP* in 1, 2, 3, 4, 8, and 12 quarters ahead are displayed in panel B of Table A.1. With the unadjusted *REG*, the positive and significant coefficients on *REG* across all regressions indicate that the positive predictive power of *REG* on future *MISP* are persistent. Specifically, the coefficient on *REG* rises from 2.431 (1 quarter ahead) to 3.259 (3 quarters ahead) and then drops gradually to 0.867 (12 quarters ahead), suggesting to the cyclic pattern in the impact of *REG* on stock mispricing. In terms of the economic magnitude, an increase in *AFE* from its 25^{th} percentile to 75^{th} percentile would be followed by a rise in *MISP* by 0.151 ($= (0.829 - (-1.820)) \times 0.057$). In the meantime, an increase in *REG* from its 25^{th} percentile to 75^{th} percentile would be followed by a rise in *MISP* by

0.545 (= (0.111 − (−0.113)) × 2.431), which is more than three times of the change induced by an increase in *MISP* from its 25^{th} percentile to 75^{th} percentile. For the long-horizon *REG*, the positive influence of *REG* on *MISP* also remains significant for *MISP* up to 12 quarters ahead. With respect to the variation of the impact from *REG* over time, the coefficient on *REG* increases from 1.682 (1 quarter ahead) to 2.388 (3 quarters ahead) and declined afterward to 0.800 (12 quarters ahead), implying the pattern of a cycle. The impact of *REG* is also economically significant. In detail, a change in *REG* from its 25^{th} percentile to 75^{th} percentile is leading the next quarter *MISP* to grow by 0.409 (= (0.121 − (−0.122)) × 1.682), which is more than two times of 0.172 (= (0.829 − (−1.820)) × 0.065), the change in *MISP* resulted from an increase in *AFE* from its 25^{th} percentile to 75^{th} percentile. Overall, the observations above demonstrate that when constructed differently, the *REG* measure can still positively predict stock mispricing in a statistically and economically significant way. And the impact of *REG* on *MISP* is always exhibiting a cyclic pattern.

In panel C of Table A.1, we present the results from predicting the next quarter *REG* for each of the two specifications. Align with the findings in Section 6.1, the coefficients on *AFE* and *MISP* are positive and significant, suggesting that stocks with a positive analyst bias and high mispricing scores are more likely to experience a high *REG* in the next quarter. Altogether with the positive impact of *REG* on future *AFE* and *MISP* documented above, it implies the same dynamic amplification inter-reaction among *REG*, *AFE*, and *MISP* as in Section 6.1, where higher *REG* leads to greater *AFE* and *MISP*, which in turn lead to higher *REG*.

A.2 Positive and Negative REG

As evidenced in the extant literature, investor optimism can induce stock misvaluation to a greater extent than pessimism due to the asymmetric ease of buying versus shorting (Stambaugh et al., 2012). Thus, we examine whether the effect of *REG* on various variables is concentrated on one side. While the portfolio analysis reported in Table 3 indicates a balanced effect, in this subsection we repeat our main analysis using positive and negative *REG* splits.

We first generate two dummy variables, one for positive values of *REG* (Dummy(*REG*>0)) indicating an overreaction on the positive side and a negative *REG* dummy (Dummy(*REG*⩽0)) indicating an overreaction on the negative side, respectively. We then repeat the investigation of the effect of *REG* on next quarter's *AFE* and next quarter's *MISP* by replacing *REG* in Eq. (5) and Eq. (6) with *REG**Dummy(*REG*>0), *REG**Dummy(*REG*⩽0), and Dummy(*REG*>0).

The prediction for next quarter's *AFE* and *MISP* are presented in Table A.2. Columns (1) - (4) show the prediction for *AFE* with current *AFE*, *DGTW*, *MISP*, and stock characteristics as controls. Focusing on column (4), the coefficient on the positive *REG* and negative *REG* interaction terms are 2.704 (t-statistics, 8.20) and 2.184 (t-statistics, 7.02), respectively. It suggests that the positive impact of *REG* on next quarter's *AFE* is not dominated by either positive or negative *REG*s. Moreover, the fact that the coefficient on positive *REG* is larger can support the general findings regarding the short-leg of anomalies. However, the difference between the coefficients of 0.52 is not statistically significant (t statistics, 1.33).

In a nutshell, the comparison between positive and negative *REG*s demonstrates that the effect of *REG* on future analyst earnings forecast error and stock's mispricing score is prevalent for both positive and negative *REG*s. The tilt toward the positive side of *REG* is expected and strengthens our findings of the link between market participants' beliefs, analysts' beliefs, and anomaly returns.

Next, columns (5)–(7) display the results for *MISP* prediction. It is clearly shown that the coefficients on both the positive *REG* and negative *REG* interaction terms are positive and statistically significant, which means that the positive influence of *REG* on *MISP* we documented earlier is not triggered entirely by positive *REG*s or negative *REG*s. Interestingly, the difference between *REG* coefficient estimates of 1.281 is statistically significant (t statistics, 3.46) and in line with Stambaugh et al. (2012)'s findings.

A.3 Analyst Return Forecast Errors

Our analysis reveals a robust relation between *REG* and *AFE*. In this Appendix section, we examine whether other outputs provided by analysts are affected by *REG*. In particular, we focus on analysts' price targets and stock recommendations. Both provide explicit and direct information that investors can act on. Overall, we find consistent results with the findings reported using *AFE*, where an increase in *REG* predicts higher price targets (i.e., positive return forecast errors), and positive recommendation changes.

We explore the relation between analysts' price targets and *AFE* using Fama and MacBeth (1973) regression for predicting the analyst implied return forecast error based on their 12-month price targets. In particular, we focus on price targets that occur *after* the quarterly earnings announcement, allowing analysts to be affected by *REG*. Given that analysts may not issue their price targets immediately after the earnings announcement, we track all analysts' price targets over a window of 60 trading days after the earnings announcement, and calculate the average (i.e., the consensus). Our regression takes the following form

$$\begin{aligned} RetForeErr_{i,t+1:t+60} =& \gamma_{0,t} + \gamma_{emr,t} REG_{i,t(q)} + \gamma_{afe,t} AFE_{i,t(q)} + \\ & \gamma_{dgtw,t} DGTW_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} Z_{k,i,t} + \epsilon_{i,t} \end{aligned} \tag{17}$$

where $RetForeErr_{i,t+1:t+60}$ is the average analyst return forecast error of stock i over the subsequent 60 days following each earnings announcement. $REG_{i,t}$, $AFE_{i,t}$, $DGTW_{i,t}$ are market misreaction, analyst earnings forecast error, and DGTW-adjusted daily abnormal return of stock i as of the earnings announcement day t in quarter q. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, and *NUMEST* as introduced in Section 2.3. We compute time-series value-weighted averages of coefficients based on the daily number of cross-sectional observations as done in previous sections.

The regression results are reported in Table A.3. Column (1) shows the result based on all observations: the coefficient on *REG* is 2.841 with a t-statistics of 2.11, implying that analysts are also too optimistic (pessimistic) in terms of their future price target estimations given high (low) values of *REG*.

Columns (2) and (3) repeat the analysis, where we require at least two or three analysts to issue future price targets for the same stock. On one hand, this might reduce the noise induced by a single analyst, on the other hand as observed, this limits the number of observations. For example, requiring at least 2 analysts, the coefficient rises to 3.267 and it is still statistically significant. A change in *REG* from its 25^{th} percentile to 75^{th} percentile would induce an increase in analyst return forecast error of 0.742% (= $(0.114 - (-0.113)) \times 3.267$).

A.4 Analyst Recommendation Changes

Similar to the analysis of analyst price targets, in this Appendix section, we examine how analysts update their recommendations after observing investors' (mis)reaction on earnings announcement days. We run the Fama and MacBeth (1973) regression for average recommendation changes of analysts during the subsequent three weeks after the earnings announcement.

$$\begin{aligned} RecChng_{i,t+b:t+d} =& \gamma_{0,t} + \gamma_{emr,t} REG_{i,t(q)} + \gamma_{afe,t} AFE_{i,t} + \\ & \gamma_{dgtw,t} DGTW_{i,t(q)} + \sum_{k=1}^{K} \gamma_{k,t} Z_{k,i,t} + \epsilon_{i,t} \end{aligned} \tag{18}$$

where $RecChng_{i,t+b:t+d}$ denotes the average of recommendation changes issued by analysts from b day ahead to d day ahead of the earnings announcement day t in quarter q. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, and *NUMEST* as introduced in Section 2.3. In the second stage of the Fama-MacBeth procedure, we compute time-series value-weighted averages of coefficients based on the daily number of cross-sectional observations.

Table A.4 reports the regression results. Similar to the findings documented with *AFE* and *RetForeErr*, *RecChng* also tends to be more positive following a positive *REG*. This provides additional support for the notion that analyst would revise their expectation based on market reaction to earnings information, and market misreaction would lead to distortion in analyst expectation formation.

In sum, the observations with *RetForeErr* and *RecChng* are consistent with our main findings in Section 3.2. It shows that a higher *REG* results not only in an increase in *AFE* but also in greater analyst return forecast errors and upward recommendation changes. This provides further validation for the argument that *REG* is positively impacting the bias in analyst expectations.

A.5 Pre-2001 vs. Post-2002

In this section, we verify that the patterns we document also hold in the latter part of our sample. We split the sample into two periods: pre-2001 and post-2002[17], and repeat our main analysis. We employ the same Fama and MacBeth (1973) regressions for predicting *AFE*, *MISP*, and *REG* in the next quarter on each subsample.

Panel A of Table A.5 displays the results for predicting *AFE* and *MISP* in the next quarter. Columns (1) - (4) show the prediction for next quarter *AFE*, and columns (5) - (8) present the results for forecasting *MISP* in the next quarter. In all regressions, next quarter *AFE* and *MISP* are positively predicted by *REG* in an economically and statistically significant way. Panel B shows the results for regression of next quarter *REG* on current *AFE*, *MISP*, and *REG*. In accordance with the observation in Section 6.1, we find in both pre-2001 and post-2002 periods that greater analyst earnings forecast error and a higher degree of overpricing would lead to a larger value of *REG* in the next quarter.

Overall, the results on subsamples suggest that our main findings are prevalent in both pre-2001 and post-2002 periods. Specifically, the relationships between *REG*, *AFE*, and *MISP* are statistically and economically significant in both the first and the second half of our sample period.

[17] After splitting, *REG*s are generated based on observations available within each subsample to prevent possible information leakage across subsamples. Specifically, *REG*s are available from 1985 in the pre-2001 sample and from 2003 in the post-2002 sample.

A.6 Panel Regressions

In this section, we repeat the forecasting for next quarter *AFE*, *MISP*, and *REG* using panel regression instead of the Fama and MacBeth (1973) cross-sectional regression we employ. In particular, we re-run the prediction regressions with firm and date fixed effects and cluster the t-statistics by firm and date.

Table A.6 reports the results from panel regressions for predicting *AFE*, *MISP*, and *REG* in the next quarter. The positive predictability of *REG* for next quarter *AFE* remains intact in panel regressions. Specifically, the coefficient on *REG* is 2.410 in the full-control specification and it is statistically significant. Economically, it implies that a change in *REG* from the 25^{th} to the 75^{th} percentile results in an increase in *AFE* of 20.57%, relative to its 25^{th} to 75^{th} range $(= (0.114 - (-0.113)) \times 2.401/(0.829 - (-1.820)))$. In the meanwhile, a change in *MISP* from its 25^{th} to 75^{th} percentile leads to a rise of 5.90% in *AFE*, relative to its 25^{th} to 75^{th} range $(= (58.779 - 41.404) \times 0.009/(0.829 - (-1.820)))$. Thus, the impact of *REG* on next quarter *AFE* is more than three times as large as that of *MISP*.

With the panel setting, we still find *REG* positively predicting subsequent quarter *MISP* and this effect is economically and statistically significant. The coefficient on *REG* and *AFE* in the regression with all controls are 2.190 (t-statistics = 10.79) and 0.056 (t-statistics = 9.00). In an economic sense, it implies that a change from the 25^{th} to the 75^{th} percentile in *REG* would result in an increase in *MISP* of 0.4971 $(= (0.114 - (-0.113)) \times 2.190))$. On the other hand, the change in *MISP* triggered by a rise in *AFE* from its 25^{th} to 75^{th} percentile is 0.1483 $(= (0.829 - (-1.820)) \times 0.056))$. Therefore in a horse race between *REG* and *AFE*, the effect of *REG* turns out to be over three times as large.

Along with the findings with Fama and MacBeth (1973) cross-sectional regression, we also identify the positive influence of biased analyst earnings forecast and degree of overpricing on

REG in the subsequent quarter. In all regressions, the positive predictability of AFE and $MISP$ for next quarter REG is statistically and economically significant.

In sum, the dynamic amplification effect among REG, AFE, and $MISP$ still holds and remains economically and statistically significant with the panel setting.

A.7 SYY's MGMT and PERF Anomaly Decomposition

In section 5.1, we explored the relationship between REG and $MISP$ and the build-up and resolution classifications. For completeness, in this section, we follow Stambaugh and Yuan (2017) and decompose SYY's 11 anomalies, which are building blocks of the $MISP$ score, into two clusters: management ($MGMT$) and performance ($PERF$). We closely follow the average anomaly ranking approach of Stambaugh et al. (2015) and average the stock's rankings according to the anomalies within each cluster. Thus, we have two mispricing measures: $MISP_{MGMT}$ and $MISP_{PERF}$.

We repeat the predictions in Eq. (6) by replacing the $MISP$ with the new mispricing scores associated with two classes of anomalies: $MISP_{MGMT}$, $MISP_{PERF}$. The coefficients on REG for predicting the mispricing scores in 1, 2, 3, 4, 8, and 12 quarters ahead are presented in Table A.7. As clearly shown in the table, and also in Figure A.1, REG is positively predicting $MISP_{MGMT}$ up to 4 quarters ahead in a significant way, and it peaks in 1 year ahead. For longer forecasting horizons (8 and 12 quarters ahead), the coefficients on REG drop and become insignificant with t-statistics no greater than 1.29. On the contrary, the coefficient on REG declines gradually when predicting $MISP_{PERF}$ in 1, 2, 3, 4, 8, and 12 quarters ahead. In general, the dissection of $MISP$'s underlying 11 anomalies into $MGMT$ and $PERF$ anomalies demonstrates that the positive impact of REG on stock mispricing still holds for $MISP_{MGMT}$ and $MISP_{PERF}$, which also connects to the two mispricing factors in Stambaugh and Yuan (2017). Besides, the observations about the persistence of REG's coefficient confirm the long (short) nature of the characteristics constituting $MGMT(PERF)$. That is, $MISP_{MGMT}$ takes time to reach its peak, while $MISP_{PERF}$ is reflected quickly in the scores and then decays.

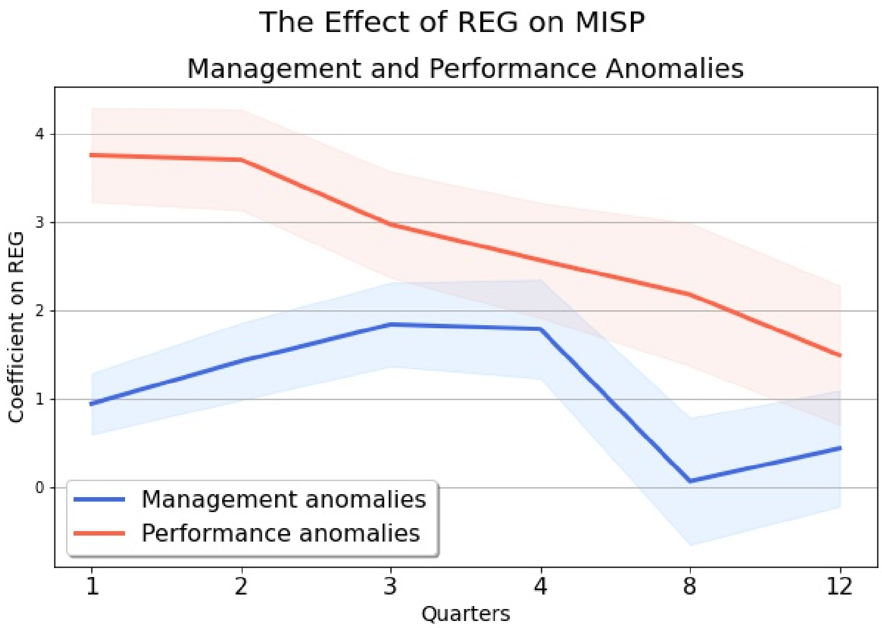


Fig. A.1 - Anomaly Dissection

The figure above show the coefficients and the corresponding confidence 95% intervals on REG in the Fama and MacBeth (1973) regressions for predicting the mispricing scores $MISP_{Mgmt}$ and $MISP_{Perf}$ in 1, 2, 3, 4, 8, and 12 quarters ahead. The sample period is from January 1985 to December 2018.

Table A.1 - Alternative Specifications for *REG*

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting for *AFE*, *MISP*, and *REG*. The sample period is from January 1985 to December 2018. *REG* is constructed based on the differences between the rankings of 1) *SUE* and 1-day *RET*, or 2) *AdjSUE* and 21-day *DGTW*. Panels A and B present the results for predicting *AFE* and *MISP* in the following quarters, respectively. Panel C presents the results for predicting *REG* in the next quarter where the coefficients on *MISP* are multiplied by 100 for readability. Coefficients on lagged dependent variables, and stock control variables are not reported. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. Standard errors are adjusted for serial correlation using Newey and West (1987) correction. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A: Predicting AFE *in the Following Quarters*

	(1) AFE_{q+1}	(2) AFE_{q+2}	(3) AFE_{q+3}	(4) AFE_{q+4}	(5) AFE_{q+8}	(6) AFE_{q+12}
			SUE, 1-day *RET*			
REG	2.545***	1.653***	1.330***	1.431***	1.177***	1.123***
	(12.57)	(7.03)	(5.11)	(5.54)	(4.61)	(4.69)
MISP	0.016***	0.020***	0.017***	0.015***	0.017***	0.016***
	(9.77)	(8.70)	(9.05)	(8.42)	(7.85)	(8.17)
			AdjSUE, 21-day *DGTW*			
REG	2.888***	1.740***	1.714***	1.760***	1.316***	0.730***
	(12.54)	(7.09)	(5.95)	(5.91)	(5.02)	(2.83)
MISP	0.016***	0.020***	0.017***	0.015***	0.018***	0.015***
	(9.75)	(8.90)	(9.34)	(8.75)	(8.14)	(7.72)

Panel B: Predicting MISP *in the Following Quarters*

	(1) $MISP_{q+1}$	(2) $MISP_{q+2}$	(3) $MISP_{q+3}$	(4) $MISP_{q+4}$	(5) $MISP_{q+8}$	(6) $MISP_{q+12}$
			SUE, 1-day *RET*			
REG	2.431***	3.171***	3.259***	2.913***	1.176***	0.867**
	(11.73)	(12.05)	(10.85)	(9.16)	(3.14)	(2.45)
AFE	0.057***	0.026***	0.035***	0.025**	0.021**	0.020**
	(9.81)	(3.57)	(4.13)	(2.24)	(2.04)	(2.05)
			AdjSUE, 21-day *DGTW*			
REG	1.682***	2.352***	2.388***	1.762***	1.082***	0.800***
	(7.37)	(7.88)	(7.20)	(4.85)	(2.61)	(2.16)
AFE	0.065***	0.033***	0.042***	0.038***	0.021**	0.017*
	(10.62)	(4.34)	(4.63)	(3.35)	(2.00)	(1.76)

Panel C: Predicting REG *in the Next Quarter*

	SUE, 1-day *RET*		*AdjSUE*, 21-day *DGTW*	
	(1)	(2)	(3)	(4)
AFE	0.006***	0.003***	0.007***	0.003***
	(33.52)	(12.96)	(36.31)	(9.35)
MISP	0.056***	0.053***	0.044***	0.039***
	(12.50)	(11.88)	(9.40)	(8.44)
1-day RET/21-day DGTW		-0.004***		-0.003***
		(-21.24)		(-28.73)

Table A.2 - Positive and Negative *REG*

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting for *AFE* and *MISP* in the next quarter. For brevity, this table only reports the coefficients of interest. The sample period is from January 1985 to December 2018. Dummy(*REG*>0) is a dummy variable which equals 1 if the *REG* on day t is greater than zero. Otherwise, it is set to be zero. Dummy(*REG*⩽0) takes the value of one when the *REG* is smaller than or equal to zero. Otherwise it is zero. Columns (1) to (4) show the results for predicting *AFE* in the next quarter. Columns (5) to (7) reports the results for predicting *MISP* in the next quarter (i.e., three months ahead). t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	AFE in the Next Quarter				*MISP* in the Next Quarter		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
*REG**Dummy(*REG*>0)	3.313***	3.00***	2.980***	2.704***	1.846***	4.234***	3.154***
	(10.45)	(9.67)	(9.21)	(8.20)	(7.03)	(16.30)	(10.77)
*REG**Dummy(*REG*⩽0)	3.068***	2.668***	2.556***	2.184***	0.600**	2.835***	1.873***
	(12.66)	(10.71)	(8.58)	(7.02)	(2.42)	(10.38)	(6.49)
Dummy(*REG*>0)	-0.143**	-0.094	-0.076	0.018	-0.164***	-0.050	-0.095*
	(-2.23)	(-1.50)	(-1.23)	(0.27)	(-3.24)	(-0.98)	(-1.86)
AFE			0.135***	0.133***	0.089***		0.059***
			(13.45)	(12.96)	(15.28)		(9.88)
DGTW			-0.074***	-0.076***		-0.115***	-0.087***
			(-9.80)	(-9.06)		(-14.48)	(-10.82)
MISP				0.016***	0.842***	0.841***	0.841***
				(9.75)	(86.18)	(85.81)	(85.84)
Control Variables	No	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	1.83%	5.37%	9.21%	9.44%	76.35%	76.39%	76.44%
#Days/#Months	3,377	2,677	2,565	2,250	203	203	201
#Obs	202,079	200,030	200,030	172,926	129,589	129,589	129,589

Table A.3 - *REG* and Analyst Price Target Forecast Errors

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting the analyst implied return forecast error based on their 12-month price targets, averaged over the subsequent 60 trading days (one quarter) after the firm's earnings announcement day. The sample includes 5,733 distinct stocks with valid analysts' price targets (PTG) from January 2000 to December 2018. Column (1) presents the result based on all observations. Columns (2) and (3) show the results on the observations where we require at least two and three analysts to issue future price targets (PTG) for the same stock, respectively. Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, and *NUMEST* as introduced in Section 2.3. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	All Obs	NumPTG⩾2	NumPTG⩾3
	(1)	(2)	(3)
REG	2.841**	3.267**	3.791
	(2.11)	(1.99)	(1.63)
AFE	0.061	0.142*	0.081
	(1.07)	(1.91)	(0.71)
DGTW	-0.448***	-0.411***	-0.421***
	(-12.93)	(-9.41)	(-6.82)
Control Variables	Yes	Yes	Yes
Adj. R-squared	15.79%	17.19%	18.57%
#Days	1,608	1,324	1,055
#Obs	116,568	81,222	53,220

Table A.4 - *REG* and Analyst Recommendation Changes

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions for predicting analyst recommendation changes in the following weeks. The sample period is from January 1985 to December 2018. In the regressions, In columns (1) - (4), the dependent variable is the average recommendation change issued by analysts in the first week after day t (i.e., from day $t+1$ to day $t+5$). The dependent variable in columns (5) - (8) is the average recommendation change issued by analysts in the second and third weeks after day t (i.e., from day $t+6$ to day $t+15$). Firm control variables include *LnSIZE*, *LnBM*, *RET5*, *RET21*, *MOM*, *RVOL*, *ILLIQ*, and *NUMEST* as introduced in Section 2.3. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. Standard errors are adjusted for serial correlation using Newey and West (1987) correction. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	$RecChng_{t+1:t+5}$				$RecChng_{t+6:t+15}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
REG		0.286		-0.334		2.198***		3.132*
		(1.11)		(-0.65)		(2.86)		(1.92)
AFE	-0.053***	-0.058***	-0.040***	-0.044	-0.005	-0.023	0.034	-0.123
	(-4.25)	(-3.66)	(-3.19)	(-1.44)	(-0.10)	(-0.45)	(0.45)	(-1.42)
DGTW			0.008	0.019			0.027	-0.026
			(1.33)	(1.61)			(1.14)	(-0.69)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	3.72%	4.66%	5.33%	5.57%	4.08%	3.94%	9.05%	9.85%
#Days	182	157	157	134	34	22	22	13
#Obs	13,346	13,332	13,332	13,332	7,001	6,996	6,996	6,996

Table A.5 - Pre-2001 vs. Post-2002

This table reports the results of Fama and MacBeth (1973) cross-sectional regressions predicting for *AFE*, *MISP* and *REG* in the next quarter. The sample period is from January 1985 to December 2018. Columns (1) to (4) of Panel A show the results for predicting *AFE* in the next quarter. Columns (5) to (7) of Panel A report the results for predicting *MISP* in the next quarter. Panel B presents the results for predicting *REG* in the next quarter, where the coefficients on *MISP* are multiplied by 100 for readability. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. *t*-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A: Predicting AFE *and* MISP *in the Next Quarter*

	AFE in the Next Quarter				*MISP* in the Next Quarter			
	Pre-2001		Post-2002		Pre-2001		Post-2002	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
REG	2.158***	2.24***	2.806***	2.730***	3.290***	2.374***	3.630***	2.338***
	(6.48)	(6.83)	(10.90)	(10.51)	(14.62)	(7.81)	(14.45)	(7.67)
AFE	0.165***	0.162***	0.118***	0.112***		0.054***		0.064***
	(9.15)	(8.97)	(10.39)	(10.12)		(5.57)		(7.95)
DGTW	-0.099***	-0.102***	-0.062***	-0.061***	-0.149***	-0.121***	-0.086***	-0.059***
	(-5.55)	(-5.62)	(-10.03)	(-9.66)	(-10.57)	(-8.58)	(-10.63)	(-6.65)
MISP		0.008***		0.020***	0.837***	0.838***	0.846***	0.846***
		(3.52)		(9.37)	(59.05)	(59.12)	(61.06)	(61.00)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	10.11%	9.95%	8.42%	8.76%	75.68%	75.75%	77.26%	77.3%
#Days/#Months	1,206	1,142	1,361	1,114	112	111	85	85
#Obs	78,282	74,870	114,360	91,501	61,035	61,035	63,381	63,381

Panel B: Predicting REG *in the Next Quarter*

	REG in the Next Quarter							
	Pre-2001				Post-2002			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AFE	0.008***	0.007***	0.006***	0.004***	0.007***	0.007***	0.006***	0.003***
	(30.67)	(25.30)	(19.57)	(8.76)	(36.93)	(33.81)	(27.28)	(9.71)
MISP	0.013**	0.009	0.011	0.006	0.038***	0.077***	0.075***	0.074***
	(2.31)	(1.31)	(1.46)	(0.84)	(8.51)	(14.24)	(13.70)	(13.29)
REG			0.044***	0.126***			0.051***	0.146***
			(7.78)	(15.32)			(11.03)	(20.18)
DGTW				-0.005***				-0.003***
				(-12.11)				(-17.16)
Control Variables	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Adj. R-squared	2.57%	6.53%	6.82%	7.68%	3.14%	7.07%	7.43%	8.44%
#Days	1,633	1,216	1,169	1,141	1,461	1,185	1,140	1,113
#Obs	76,930	75,299	74,764	74,764	93,851	92,607	91,318	91,318

Table A.6 - Panel Regressions

This table reports the results from panel regressions predicting for future *AFE*, *MISP* and *REG* in the next quarter. Columns (1) and (2) of Panel A show the results for predicting *AFE* in the next quarter. Columns (3) and (4) of Panel A report the results for predicting *MISP* in the next quarter. Panel B presents the results for predicting *REG* in the next quarter, where coefficients on *MISP* are multiplied by 100 for readability. The sample period is from January 1985 to December 2018. All regressions include firm and time fixed effects. Standard errors are clustered on firm and time. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A: Predicting AFE *and* MISP *in the Next Quarter*

	AFE in the Next Quarter		*MISP* in the Next Quarter	
	(1)	(2)	(3)	(4)
REG	2.401***	2.410***	0.736***	2.190***
	(16.40)	(16.83)	(4.47)	(10.79)
AFE	0.023***	0.025***	0.085***	0.056***
	(3.57)	(3.58)	(11.80)	(9.00)
DGTW	-0.060***	-0.066***		-0.072***
	(-14.36)	(-16.79)		(-8.04)
MISP		0.009***	0.719***	0.718***
		(5.76)	(45.01)	(44.92)
Control Variables	Yes	Yes	Yes	Yes
Fixed Effects	Firm, Time	Firm, Time	Firm, Time	Firm, Time
Adj. R-squared	9.82%	9.44%	76.75%	76.80%
#Obs	198,351	171,301	128,878	128,878

Panel B: Predicting REG *in the Next Quarter*

	REG in the Next Quarter			
	(1)	(2)	(3)	(4)
AFE	0.003***	0.003***	0.003***	0.001***
	(25.93)	(24.79)	(22.57)	(9.20)
MISP	0.075***	0.042***	0.041***	0.039***
	(14.84)	(7.97)	(7.88)	(7.51)
REG			0.013***	0.087***
			(3.97)	(20.98)
DGTW				-0.003***
				(-28.30)
Control Variables	No	Yes	Yes	Yes
Fixed Effects	Firm, Time	Firm, Time	Firm, Time	Firm, Time
Adj. R-squared	76.72%	76.78%	76.75%	76.80%
#Obs	130,449	130,449	128,878	128,878

Table A.7 - Anomaly Dissection and the Mispricing Cycle: Performance and Management Anomalies

This table reports the coefficient on REG from the Fama and MacBeth (1973) cross-sectional regressions predicting for mispricing scores, $MISP_{MGMT}$ and $MISP_{PERF}$, which are associated with four classes of anomalies: management and performance, respectively. The dependent variables are the stock's average rankings with respect to anomalies within each class. In each month, we rank stocks according to each anomaly. The higher the ranking, the greater the degree of overvaluation. Then for each stock, we compute the equal-weighted average of rankings across all anomalies within the corresponding anomaly class. Given that the number of firms reporting their earnings is very scarce on some days, we report value weighted averages based on the daily number of cross sectional observations. The sample period is from January 1985 to December 2018. All dependent variables except for REG, AFE, and $DGTW$, are observed at the end of the month of earnings announcement day t. Firm control variables include *LnSIZE*, *LnBM*, *MRET*, *MMOM*, *MRVOL*, and *MILLIQ* as introduced in Section 2.3. Standard errors are adjusted for serial correlation using Newey and West (1987) correction. t-statistics are reported below the coefficient estimates in parentheses. *,**, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

	$q+1$	$q+2$	$q+3$	$q+4$	$q+8$	$q+12$
	(1)	(2)	(3)	(4)	(5)	(6)
			$MISP_{MGMT}$			
REG	0.938***	1.418***	1.835***	1.782***	0.059	0.431
	(5.34)	(6.37)	(7.62)	(6.27)	(0.16)	(1.29)
			$MISP_{PERF}$			
REG	3.750***	3.698***	2.963***	2.560***	2.173***	1.486***
	(13.87)	(12.75)	(9.68)	(7.72)	(5.28)	(3.70)

Appendix B

B.1 Description on Anomaly Dissection

In this section, we describe the construction for misprcing score with respect to four classes of anomalies: management (*MGMT*), performance (*PERF*), build-up (*Build-Up*), and resolution (*Resolution*), as well as the data collecting process for the stock-level characteristics constituting each class of anomalies.

Stambaugh and Yuan (2017) classify the 11 anomalies underlying the *MISP* score into two clusters: management (*MGMT*) and performance (*PERF*). The *MGMT* anomalies include net stock issues, composite equity issues, accruals, net operating assets, asset growth, and investment to assets, all of which are presenting quantities that firm's management can directly impact. The *PERF* anomalies include distress, O-score, momentum, gross profitability, and return on assets, all of which are related more to firm performance and less affected by firm's management. We match each of the 11 anomalies to stock-level characteristics from the Open Source Cross-sectional Asset Pricing dataset by Chen and Zimmermann (2022) according to variable definition, original paper author(s), and publication year. We have successfully matched all 6 *MGMT* and 5 *PERF* anomalies with available characteristics from Chen and Zimmermann (2022). Table B.1 lists the 11 anomalies from *MGMT* and *PERF* classes and their closest matches from Chen and Zimmermann (2022).

van Binsbergen et al. (2021) study 57 asset pricing anomalies and classify them into build-up (*Build-Up*) anomalies that exacerbate stock mispricing and resolution (*Resolution*) anomalies that resolve stock price dislocation. We match each of the 57 anomalies in their Table (C.1) to stock-level characteristics from the Open Source Cross-sectional Asset Pricing dataset by Chen and Zimmermann (2022) according to variable definition, original paper author(s), and publication year. We have successfully matched 16 out of 21 *Build-Up* and 22 out of 36 *Resolution* anomalies with available characteristics from Chen and Zimmermann (2022). Table B.2 lists the 57 anomalies from *Build-Up* and *Resolution* classes and their closest matches from Chen and Zimmermann (2022).

After obtaining the anomalies, we sort stocks in each month according to each anomaly. To be consistent with the Stambaugh et al. (2015) *MISP* score, we rank stocks in each month into 100 bins according to firm's relative degree of overpricing. The greater the degree of overvaluation, the higher the rank with respect to the given anomaly. That is, firms with the highest growth would receive the highest rank in terms of the given anomaly. A stock's mispricing score with respect to each class of anomalies is the average of its rankings in terms of all anomalies within the corresponding anomaly class, and it ranges between 0 and 100, which is the same as the Stambaugh et al. (2015) *MISP* score. By construction, a higher value of the mispricing score associated with an anomaly class implies a greater degree of overpricing with respect to the underlying anomalies.

Table B.1 - Anomaly Dissection: Management and Performance

This table lists the 11 anomalies which the Stambaugh et al. (2015) *MISP* score is constructed based on. According to Stambaugh et al. (2015), the 11 anomalies can be clustered into two classes: Management and Performance. For each anomaly, we present the associated class and name adopted by Stambaugh et al. (2015). The last column indicates the closest match available from Chen and Zimmermann's Open Source Cross-sectional Asset Pricing database.

Classification	Predictor	Closest Match
Management	Accruals	Accruals
Management	Asset Growth	AssetGrowth
Management	Composite Equity Issues	CompEquIss
Management	Investment to Assets	Investment
Management	Net Stock Issues	NetEquityFinance
Management	Net Operating Assets	NOA
Performance	Distress	FailureProbability
Performance	Gross Profitability	GP
Performance	Momentum	Mom12m
Performance	O-score	OScore
Performance	Return on Assets	roaq

Table B.2 - Anomaly Dissection: Build-Up and Resolution

This table lists the 57 anomalies studied by van Binsbergen et al. (2021). van Binsbergen et al. (2021) classifies anomalies into two classes: Build-up and Resolution. For each anomaly, we present the associated class, acronym, and name adopted by van Binsbergen et al. (2021). The last column indicates the closest match available from Chen and Zimmermann's Open Source Cross-sectional Asset Pricing database. "N/A" implies that the corresponding stock-level signal is not available in the Google Drive folder of Chen and Zimmermann's database.

Classification	Predictor	Acronym	Closest Match
Build-up	Bid ask spread	SPREAD	BidAskSpread
Build-up	Cash+Short-term Investments over AT	C2A	Cash
Build-up	Cashflow to Debt	C2D	cashdebt
Build-up	Gross margin - sales (Prc changes)	dGS	GrGMToGrSales
Build-up	Gross profitability over book equity	PROF	N/A
Build-up	Idiosynctatic FF3M volatility	IDIOV	IdioVol3F
Build-up	Income to AT	ROA	roaq
Build-up	Income to lagged BE	ROE	RoE
Build-up	Income to shares outstanding	EPS	N/A
Build-up	Industry adjusted PM	aPM	ChPM
Build-up	Mom12-2	R122	Mom12m
Build-up	Mom12-7	R127	IntMom
Build-up	Mom6-2	R62	Mom6m
Build-up	Operating Inc. after depr. to sales	PM	PM
Build-up	PM scaled by net operating assets	RNA	N/A
Build-up	Pre-tax income over sales	IPM	N/A
Build-up	Return on invested capital	ROIC	roic
Build-up	Sales minus cost of goods	PCM	N/A
Build-up	Stdev of turnover	sdTURN	std_turn
Build-up	Stdev of volume	sdDVOL	VolSD
Build-up	Tangibility	TAN	tang
Resolution	Absolute Operating Accruals	AOA	N/A
Resolution	BEME - IndustryAdjusted	aBEME	N/A
Resolution	Beta	BETAd	Beta
Resolution	Book equity over market equity	BEME	BM
Resolution	Change in PPE and Inventory over AT	dPIA	InvestPPEInv
Resolution	Change in inventories over AT	IVC	ChInv
Resolution	Cost of goods sold+expenses over AT	OL	OPLeverage
Resolution	Debt to Price	D2P	NetDebtPrice
Resolution	Detrended Turnover	DTO	N/A
Resolution	Dividend to Price	DP	DivYield
Resolution	Income to market cap	E2P	EP
Resolution	Industry adjusted SAT	aSAT	N/A
Resolution	Industry adjusted market cap	aSIZE	N/A
Resolution	Log Change in shares outstanding	dSO	N/A
Resolution	Long-term reversal	R3613	N/A

(To be continued)

(Continued)

Classification	Predictor	Acronym	Closest Match
Resolution	Market cap	SIZE	N/A
Resolution	Maximum daily return	MAXRET	MaxRet
Resolution	Net operating assets over AT	NOA	NOA
Resolution	Net sales over operating assets	ATO	AssetTurnover
Resolution	Operating Accruals	OA	PctAcc
Resolution	Percentage growth rate in sales	SG	sgr
Resolution	Prc change in equity book value	dCEQ	DelEqu
Resolution	Prc change in shares outstanding	dSOUT	N/A
Resolution	Prc change in total assets	I2A	AssetGrowth
Resolution	Residual volume	SUV	N/A
Resolution	Return volatility	RETVOL	IdioRisk
Resolution	Sales (sale) to total assets (at).	SAT	N/A
Resolution	Sales to Lagged Total Assets	CAT	N/A
Resolution	Sales to cash	S2C	salecash
Resolution	Sales to price	S2P	SP
Resolution	Short-term reversal	R21	N/A
Resolution	Size + longterm debt - AT to cash	ROC	CashProd
Resolution	Tobins Q	Q	N/A
Resolution	Total assets	AT	N/A
Resolution	Total assets over market cap	A2ME	AM
Resolution	Volume over shares outstanding	TNOVR	ShareVol