Extreme Categories and Overreaction to News

Spencer Y. Kwon and Johnny Tang*

March 12, 2024

Abstract

What characteristics of news generate over-or-underreaction? We study the asset-pricing consequences of diagnostic expectations, a model of belief formation based on the representativeness heuristic, in a setting where news events are drawn from categories with extreme distributions of fundamentals. Our model predicts greater overreaction to news belonging to categories with more extreme outliers, or tail events. We test our theory on a comprehensive database of corporate news that includes news from 24 different categories, including earnings announcements, product launches, M&A, business expansions, and client-related news. We find theory-consistent heterogeneity in investor reaction to news, with more overreaction in the form of greater post-announcement return reversals and trading volume for news categories with more extreme distributions of fundamentals.

*Spencer Yongwook Kwon: Brown University, spencer_kwon@brown.edu. Johnny Tang: Cornell University, johnnytang@cornell.edu. We thank Malcolm Baker, Nick Barberis, Francesca Bastianello, Michael Blank, Pedro Bordalo, John Campbell, Thomas Chaney (editor), Nicola Gennaioli, Mark Egan, Paul Fontanier, Xavier Gabaix, Robin Greenwood, Sam Hanson, Alex Imas, Yueran Ma, Peter Maxted, Josh Schwartzstein, Kelly Shue, Andrei Shleifer, David Solomon (discussant), Evan Soltas, Jeremy Stein, Adi Sunderam, Paul Tetlock, four anonymous referees, and seminar participants at Harvard, Yale, and the 2021 Spring NBER Behavioral Finance meeting for helpful comments and feedback. We would like to thank Jerry He, Purvesh Jain, and Chengyue Zhang for outstanding research assistance.

1 Introduction

The presence of both systematic over-and-underreaction in financial markets remains a major puzzle. On one hand, stock prices can overreact when firms experience high returns, earnings growth (De Bondt and Thaler, 1985; Cutler et al., 1991; Lakonishok et al., 1994; La Porta, 1996; Bordalo et al., 2019), or spikes in media coverage and sentiment (Da et al., 2011; Tetlock, 2007; Antweiler and Frank, 2006). On the other hand, stock prices underreact to other types of information, such as earnings announcements and profitability (Bernard and Thomas, 1989; Bouchaud et al., 2019; Sloan, 1996). The heterogeneity in investor reaction to news raises a key theoretical and empirical question: what characteristics of news predict whether investors underreact or overreact?

We propose and test a novel predictor of investor over-and-underreaction to news. Our approach draws from two key features of investor psychology. First, investors react to news by evaluating it based on similar events in the same category. For example, investors may react to a tech company's product launch by recalling other past product launches. Second, the past events that come to mind tend to be salient outliers: investors are more likely to draw references to the original iPhone launch than any other product launches. For example, Tesla's 2016 launch of its Model 3 vehicle was hailed as its "iPhone moment".[1] These two features of investor reaction to news can reflect cognitive forces, such as associative recall (Kahana, 2012; Bordalo et al., 2020b), or other forces such as biased media coverage (Nimark, 2014; Tetlock, 2014). The selective retrieval of salient past events may distort investor beliefs. These forces imply that whether investors overreact to a news event depends on which category it belongs to: investors are more likely to overreact to news belonging to categories with extreme outliers.

Motivated by these features, we build a formal model of investor psychology with three key components. First, we assume that each news announcement belongs to a news category. Second, to reflect the importance of tail events in shaping investor reaction to news, we model the distribution of fundamentals of each news category as a power-law distribution, or extremal (Gabaix, 2009; Embrechts et al., 2013): while most news

[1] https://www.ft.com/content/28d27254-12da-11e6-839f-2922947098f0

of a given category have modest impact on fundamentals, some tail events have major implications. Third, as we show in the data, news categories differ in their extremeness: while some categories are more extreme (fatter-tailed) – their top 1% news have much greater impact than their median news – others are less so. Differences in the tail will be the key driver of investor under-and-overreaction across news categories.

We combine these assumptions with diagnostic expectations (DE) (Bordalo et al., 2018), a model of belief formation based on Kahneman and Tversky's representativeness heuristic. DE capture the insight that agents overweight in their beliefs states of the world that have become more likely in light of news.[2] When applied to a family of distribution with varying tails, DE exaggerate the degree to which tail events have become objectively more or less likely after a news event. When the news is from a more extreme news category, diagnostic expectations of fundamentals overreact and overshoot the rational benchmark. Conversely, news from a less extreme news category is more representative of non-tail outcomes and generates underreaction. The model predicts differences in investor biases across news categories, not within category differences: within each category, the model predicts a constant amount of over-or-underreaction. We close our model by introducing diagnostic and rational investors into a stylized asset pricing model, where rational arbitrageurs are slow to enter the market and correct prices. Our model predicts greater return reversals and disagreement-driven trading volume for news in more extreme news categories.

In the second part of the paper, we take our theoretical predictions to the data. For news categories, we draw from a comprehensive database of corporate news events in the US from 2011 to 2018, which span a wide range of news categories, including earnings announcements, leadership changes, business expansions, and mergers and acquisitions. Consistent with our model, we document that the distribution of fundamentals for each news category is well-fit by a power-law distribution. Furthermore, we document statistically and economically significant variation in extremeness, i.e., how fat the tails are, across news categories. While news categories such as leadership changes, mergers and

[2]Diagnostic expectations have been used to model exuberance in credit booms, overreaction in macroeconomic forecasts (Bordalo et al., 2020a), and closest to our setting, the overvaluation of firms with high long-term growth prospects (Bordalo et al., 2019).

acquisitions (M&As), and lawsuits have more extreme fundamental distributions, categories such as earnings announcements, guidances, and client announcements tend to be less extreme. Lastly, most news categories have fatter tails than the unconditional distribution of stock returns.

After documenting cross-category differences in extremeness, we test our core prediction that there is greater overreaction to news from more extreme news categories. For each news category, we measure whether announcement-day returns are positively or negatively predictive of the subsequent 90 day returns. We find that while stock prices exhibit significant post-announcement drift following earnings announcements, they also exhibit post-announcement reversals of comparable magnitude for other news categories. Consistent with our core hypothesis, we find a strong link between the extremeness of the news category and post-announcement drifts and reversals. We estimate that news from the most extreme categories exhibit reversals of up to -23% of their announcement-day returns, while the news from the least extreme categories experience drifts of up to 7% of announcement-day returns.

We also test additional predictions of our model regarding trading volume and expectations. We find that holding fixed fundamentals, news from more extreme categories have greater trading volume: conditional on a 10% announcement-day return, we estimate that the daily turnover increases by 32% from the least to the most extreme news categories. Turning to expectations, we measure category-level differences in how investor expectations respond to news using analysts' earnings per share (EPS) forecasts as a proxy for investor beliefs (Bordalo et al., 2020a). We estimate Coibion and Gorodnichenko (2015) regressions of forecast errors on forecast revisions and find suggestive evidence that analyst forecasts react more sensitively to news in more extreme categories.

We conduct a series of robustness exercises to assess the validity of our main findings. We show that our results are robust to using alternative measures of extremeness based on earnings growth or longer-horizon returns, using alternative announcement windows, accounting for potential overlaps in news, computing extremeness only using past data, and using different statistical inference methods. We also find that our results are robust to sample selection and hold consistently in different sets of news categories, as well as excluding both small news and outliers. Lastly, we discuss and test possible alternative explanations for our results, such as the informativeness of news (Bordalo et al., 2023b; Augenblick et al., 2021; Ba et al., 2022), media coverage, familiarity, and other news characteristics like the sign and magnitude of the news.

Our paper contributes to the extensive theoretical (Barberis et al., 1998; Hong and Stein, 1999; Daniel et al., 1998) and empirical (De Bondt and Thaler, 1985; Lakonishok et al., 1994; La Porta, 1996; Daniel and Titman, 2006; Bernard and Thomas, 1989; Bordalo et al., 2019) literature studying investor over-and-underreaction. In particular, our work is part of a growing literature that seeks to find determinants of over-and-underreaction, such as time horizon (Giglio and Kelly, 2018; d'Arienzo, 2020; Wang, 2019; Gormsen and Lazarus, 2023), persistence (Bordalo et al., 2020a; Afrouzi et al., 2023), tangibility (Daniel and Titman, 2006), media sentiment (Tetlock, 2007; Engelberg et al., 2012), and contrast effects (Hartzmark and Shue, 2018). Our focus on informational characteristics brings our paper closer to the recent work by Augenblick et al. (2021) and Ba et al. (2022), which experimentally documents greater overreaction to less informative signals, possibly in complex environments. While these papers focus on characteristics of the individual news, our theory and measure explain over-and-underreaction at the category-level: in particular, we find that properties of the broader news category – how extreme the tail is – shape investor reaction to all news of that category.

Our work also relates to the large empirical literature on how investors react to news (Barber and Odean, 2008; Huberman and Regev, 2001; Tetlock, 2007; Antweiler and Frank, 2006; Engelberg and Parsons, 2011; Da et al., 2011; Neuhierl et al., 2013; Fedyk, 2018), which documents how news events and media coverage can lead to spikes in investor attention and short-term reversals. The literature has highlighted the role of large returns, causal impact of media (Engelberg and Parsons, 2011), and prominence in coverage (Huberman and Regev, 2001; Fedyk, 2018) as possible drivers of the salience of news. We contribute by theoretically investigating what informational characteristics of news make it salient, focusing on an event's association to past significant tail events. We show that this generates systematic differences in over-and-underreaction across news categories, which can be quantitatively captured by measurements of the tail.

Lastly, our model of investor psychology builds on the literature that brings psychological foundations to information processing in financial settings. The fundamental premise of our model – investors react to news by drawing associations with other similar events – resonates with theoretical and empirical work on associative recall (Wachter and Kahana, 2019; Bordalo et al., 2023b; Enke et al., 2020; Charles, 2022). Furthermore, our findings relate to the crucial role played by rare tail events in expectation formation, both in the lab as well as financial and macroeconomic settings (Tversky and Kahneman, 1992; Kozlowski et al., 2020; Malmendier and Nagel, 2011; Bordalo et al., 2022; Barberis, 2013). We contribute by translating these broad psychological insights into a concrete quantitative predictor of investor over-and-underreaction, and systematically testing it on a comprehensive database of corporate news.

The rest of the paper is organized as follows. Section 2 presents the model, and Section 3 describes the data. Section 4 tests the core prediction of our model. We find significant differences in the extremeness of each news categories, and show that short-term return reversals are concentrated in the more extreme categories. Section 5 discusses possible alternative explanations for our findings, and Section 6 concludes.

2 Extreme news categories and reaction to news

In this section, we present a simple model of investor reaction to different categories of corporate news (e.g. product launches vs earnings announcements). The model is based on two core assumptions that we later validate in the data. First, for each news category, the distribution of fundamentals follows a power-law distribution: each news category contains outlier news events. Second, we assume each category differs in how extreme its outliers are. While some categories are fat-tailed – its top 1% events have much greater impact than its median event – others are less so. When combined with diagnostic expectations (DE), our model shows that differences in the tail generate category-level differences in investor over-and-underreaction: the more extreme a news category, the greater the average overreaction to its constituent events.

2.1 Model: set-up

Fundamentals and news categories There is a stock in zero net supply with initial fundamentals $F_0 > 0$. Let v be the future growth of fundamentals, with $F_{final} = \exp(v) \cdot F_0$. At $t = 0$, v is unknown and has a symmetric distribution:

$$\pi^d(v) = \begin{cases} \pi_0(v) & \text{for } |v| < v_{0,d} \\ C \cdot |v|^{-(\zeta_d^{-1}+1)} & \text{for } |v| > v_{0,d}. \end{cases} \tag{1}$$

Relative to standard specifications, the only difference is that we assume that the unconditional distribution of v is power law with index ζ_d^{-1}. While we primarily make this assumption for analytical tractability, this assumption is consistent with the data: the unconditional distribution of stock returns, even outside of news-announcement days, is fat-tailed (Gabaix et al., 2003; Plerou et al., 1999; Oh and Wachter, 2018).[3]

At $t = 1$, a news announcement in news category $\mathcal{C} \in \mathbf{C}$ occurs, where $\mathbf{C}$ is the set of all news categories. Conditional on news in category $\mathcal{C}$, v follows a power-law distribution:

$$\pi_{\mathcal{C}}(v) = \frac{\zeta_{\mathcal{C}}^{-1}}{2} \cdot \frac{v_{0,\mathcal{C}}^{\zeta_{\mathcal{C}}^{-1}}}{|v|^{\zeta_{\mathcal{C}}^{-1}+1}} \text{ for } |v| \geq v_{0,\mathcal{C}}.^{4} \tag{2}$$

The distribution of fundamentals of category $\mathcal{C}$ is specified by two parameters: $v_{0,\mathcal{C}} \geq v_{0,d}$, the scale parameter, and $0 < \zeta_{\mathcal{C}} < 1$, the tail parameter, which governs the extremeness of the distribution. Equation (2) captures the two core assumptions of our model. First, the distribution of fundamentals of a news category is extreme.[5] Second, the degree of extremeness, reflected by the tail parameter $\zeta_{\mathcal{C}}$, varies across each news category $\mathcal{C} \in \mathbf{C}$.

[3]While we focus on the symmetric case for simplicity, it is known that unconditional stock returns have a negative skew (Kelly and Jiang, 2014). Appendix A.5 considers an extension where one allows the reference distribution to also exhibit asymmetric tails (with the left tail potentially being fatter-tailed).

[4]To ensure that equations 1 and 2 are mutually compatible, we assume that the unconditional probability of an announcement in any news category is arbitrarily small. We also assume that the tails of each announcement $\zeta_{\mathcal{C}}$ are known to investors: in particular, we do not model investors learning about the tail from the realization of v, as is done in (Kozlowski et al., 2020).

[5]The extremal distribution literature has a precise way of categorizing non-extreme (such as log-normal, normal, exponential) distributions and extreme distributions (such as power-law, Student-t, Cauchy, etc) distributions. The limit of $\max\{x_1, x_2, ...x_n\}$, suitably normalized, converges to the Gumbel distribution for thin-tailed distributions, and the Frechet distribution for heavy-tailed distributions. For more details, see Embrechts et al. (2013).

The greater the $\zeta_{\mathcal{C}}$, the farther the difference between the tail and modal outcomes of category $\mathcal{C}$: for example, the quantile ratio $q_{1,10}$, the ratio between the top 1% and the top 10% news, increases in $\zeta_{\mathcal{C}}$.[6]

Investor psychology Upon announcement, each investor j learns of the news category $\mathcal{C}$. She also receives an idiosyncratic signal s_j of fundamentals v, which reflect differences in how each investor interprets the news announcement (Kandel and Pearson, 1995). We assume that s_j is drawn from the conjugate distribution $s_j = v \cdot u_j$, $u_j \sim Unif[0,1]$, with $v > 0$: the negative case follows analogously.[7] The rational posterior of v conditional on $(\mathcal{C}, s_j)$ is:

$$\pi_{\mathcal{C}}(v|s_j) = (\zeta_{\mathcal{C}}^{-1} + 1) \cdot \frac{v_{1,\mathcal{C}}^{\zeta_{\mathcal{C}}^{-1}+1}}{v^{\zeta_{\mathcal{C}}^{-1}+2}} \text{ for } v \geq v_{1,\mathcal{C}} = \max\{s_j, v_{0,\mathcal{C}}\}. \tag{3}$$

To generate over-and-underreaction, we depart from rational expectations and assume that investors form diagnostic expectations (DE) of fundamentals given news (Bordalo et al., 2018). DE formalize the psychology of representativeness (Tversky and Kahneman, 1983), where individuals exaggerate states of the world that have become disproportionately likely in light of news.[8] The diagnostic distribution of fundamentals is given by:

$$\pi_{\mathcal{C}}^{\theta}(v|s_j) \propto \pi_{\mathcal{C}}(v|s_j) \cdot \left(\frac{\pi_{\mathcal{C}}(v|s_j)}{\pi^d(v|s_j = 0)} \right)^{\theta}. \tag{4}$$

[6]The quantile ratio is given by $q_{1,10} = 10^{\zeta_{\mathcal{C}}}$. Skewness, another popular measure of the tail, is also monotonically increasing in $\zeta_{\mathcal{C}}$, as long as $\zeta_{\mathcal{C}}$ is sufficiently small so that third moments are defined. While we assume eq. (2) in a reduced-form manner, one can microfound eq. (2) as a stationary outcome of a stochastic dividend growth process, where variation in $\zeta_{\mathcal{C}}$ is endogenized by differences in the expected growth rate and volatility associated with each news category. For details, see Gabaix (2016).

[7]One can generalize to $\pi(s_j|v) = \gamma \cdot s_j^{\gamma-1} v^{-\gamma}$. While the idiosyncratic signal assumption is not necessary for our return predictability result, it is necessary to generate disagreement-driven trading volume, which we also test empirically. Furthermore, while this specification assumes that there is no ambiguity in whether a particular news event is positive or negative, one can easily extend the model to allow for ambiguity in the news, with each news category having potentially different tails in the positive and negative direction. Appendix A.5 discusses the implications of such an extension, and in particular shows that the key comparative statics of our model holds even when allowing for ambiguity.

[8]More broadly, the representativeness heuristic refers to the psychological tendency to overrepresent representative attributes of a class, where "an attribute is representative ... if ... the relative frequency of this attribute is much higher in that class than in a reference class" (Tversky and Kahneman, 1983). In other words, agents overestimate the frequency of trait t that is representative of group G relative to a reference group $-G$. Bordalo et al. (2018) apply the heuristic to reaction to news by setting G as the arrival of new information and $-G$ as the no-news counterfactual, where the realized signal is equal to its expected value.

The likelihood ratio $\frac{\pi_{\mathcal{C}}(v|s_j)}{\pi^d(v|s_j=0)}$ is higher for realizations of v that have become more likely in light of news relative to the no-news benchmark.[9] The parameter θ reflects the degree to which representative outcomes are overweighted, with $\theta = 0$ nesting the rational case. The following assumption is necessary to ensure that the diagnostic distribution has a well-defined mean for all categories in $\mathbb{C}$:

Assumption 1. Let $\zeta_m = \max_{\mathcal{C}\in\mathbb{C}} \zeta_{\mathcal{C}}$. We assume that $\zeta_m > \zeta_d$, and the diagnostic parameter θ is sufficiently small such that $\theta < \frac{\zeta_m^{-1}-1}{\zeta_d^{-1}-\zeta_m^{-1}}$.

Asset markets There are two types of investors: diagnostic and rational. All investors have asset demand that is linear in their subjective expected log returns:

$$D_j^{DE}(s_j, p) = \kappa \cdot \left(E_j^{\theta}[\log(F_{final})] - \log(p)\right), \; D_j^{RE}(s_j, p) = \kappa \cdot \left(E_j[\log(F_{final})] - \log(p)\right).^{10} \quad (5)$$

At $t = 0$, all beliefs are at the prior: $E_j^{\theta}[\log(F_{final})] = E_j^{rat}[\log(F_{final})] = \log(F_0)$, with $p_0 = \log(F_0)$. With the arrival of news at $t = 1$, we assume that initially only a unit mass of diagnostic investors trade the asset, with p_1 adjusting to clear the market: $\int D_j^{DE}(s_j, p_1)di = 0$. Variation in s_j across investors generates disagreement and trading. As is standard, we define the total ($t = 1$) trading volume to be $Vol = \frac{1}{2}\left(\int |D_j(s_j, p)|ds_j\right)$.

At $t = 2$, rational arbitrageurs of mass K enter the market.[11] The late entry of rational arbitrageurs reflects slow-moving arbitrage (Duffie, 2010), where prices initially

[9] Our specification exactly follows that in Bordalo et al. (2018, 2020a), where the no-news benchmark is given by the unconditional distribution π_d with s_j equal to its ex ante expected value. One can also consider an alternative where the no-news benchmark is the prior, with no difference in the qualitative predictions.

[10] In particular, we assume that investors do not learn about v from prices. Furthermore, we also assume that the extremeness of each news category, $\zeta_{\mathcal{C}}$, is known to all investors: investors do not infer about the underlying extremeness of a news category from the signal s_j.

[11] One can easily extend our model to the case where there are also K_1 rational traders at $t = 1$, and $K_2 > K_1$ rational traders at $t = 2$: the slow entry of rational arbitrageurs at $t = 2$ reflects the sluggishness of arbitrage capital. We are also assuming that that diagnostic agents are myopically optimizing their expected returns, not accounting for the future entry of rational investors. One can relax this simplifying assumption by modeling diagnostic investors as short-lived one-period investors who are trying to sell to period 2 investors. In other words, their demand can be re-written as:

$$D_j^{DE}(s_j, p_1) = \kappa \cdot \left(E^{\theta}[\log(p_2)|\mathcal{C}, s_j] - \log(p_1)\right).$$

Given that p_2 settles to rational expectations, the diagnostic expectations of p_2 behaves similarly to the diagnostic expectations of v directly, with no change in the qualitative conclusions.

dislocated by news-driven behavioral demand (Barber and Odean, 2008) are gradually corrected by arbitrageurs. For simplicity, we assume $K \gg 1$, which implies asset prices settle at $t = 2$ to the average rational valuation.

2.2 Model solution: expectations, prices, and volume

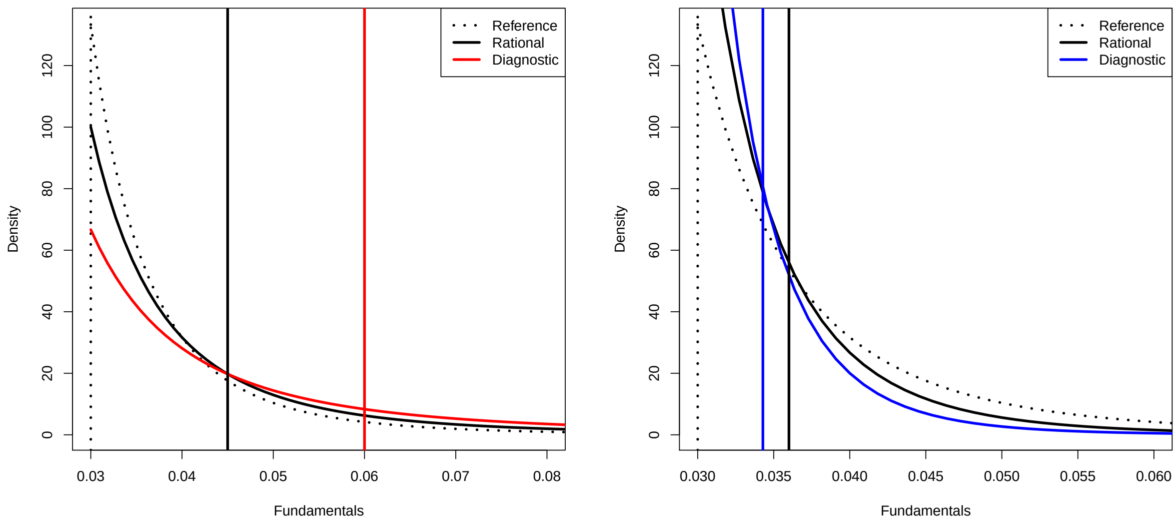
Biased expectations by news category For each news category $\mathcal{C}$, we solve for investor expectations, returns at $t = 1, 2$, and the trading volume at $t = 1$. We begin by characterizing how a diagnostic investor reacts to her signal s_j for an announcement in category $\mathcal{C}$. Combining equations (1), (2), and (4) yields the following.

Proposition 1 (DE with tails)**.** The diagnostic expectations of v is given by:

$$E^{\theta}[v|\mathcal{C}, s_j] = \psi(\zeta_{\mathcal{C}}, \zeta_d, \theta) \cdot E[v|\mathcal{C}, s_j] = \frac{1 + \zeta_{\mathcal{C}} + \theta\left(1 - \frac{\zeta_{\mathcal{C}}}{\zeta_d}\right)}{1 + \zeta_{\mathcal{C}} + (1 + \zeta_{\mathcal{C}}) \cdot \theta\left(1 - \frac{\zeta_{\mathcal{C}}}{\zeta_d}\right)} \cdot E[v|\mathcal{C}, s_j], \tag{6}$$

where $E[v|\mathcal{C}, s_j]$ is the rational expectation. The distortion term $\psi(\zeta_E, \zeta_d, \theta)$ is increasing in $\zeta_{\mathcal{C}}$, with expectations overshooting the rational benchmark ($\psi > 1$) if and only if $\zeta_{\mathcal{C}} > \zeta_d$.

All proofs are relegated to Appendix A. Figure 1 illustrates how DE distorts the expectations of fundamentals. In the left panel, the fundamentals associated with news category $\mathcal{C}_1$ is more extreme than the ex ante distribution of fundamentals ($\zeta_{\mathcal{C}_1} > \zeta_d$). The rational posterior of v, shown in the solid black curve, has a fatter tail than the reference distribution, shown in the dotted curve. In this case, diagnostic expectations, shown in the red curve, exaggerate the prevalence of extreme outcomes, causing the posterior mean to overshoot. This echoes the intuition of Bordalo et al. (2019) – in response to news that increases the right tail of long-term growth prospects, investors exaggerate the probability that the company will become "the next Google." In contrast, the fundamentals of news category $\mathcal{C}_2$ in the right panel are less extreme than the reference distribution. In that case, extreme outcomes become *less* likely in light of news. Diagnostic investors instead reason that the news is instead representative of non-tail outcomes, and underreact. Note that the contrast in Figure 1 is at the *news category* level: our theory produces differences in investor biases across different categories. Within a given category, Propo-



(a) DE, more extreme event-type

(b) DE, less extreme event-type

Figure 1: Diagnostic expectations and the under-overreaction

Note: Figures 1a and 1b show the DE distortions of subjective fundamentals for more and less extreme event-types, respectively. The solid red and blue curves plot the density functions of the distributions of subjective fundamentals under diagnostic expectations. The solid black curves plot the density functions under the rational distributions. The dotted black curves plot the reference distributions. The solid black vertical lines plot the expectation of fundamentals for the rational agents. The solid red and blue vertical lines plot the subjective expectations of fundamentals for the diagnostic agents.

sition 1 implies a constant amount of over-or-underreaction for both large and small news announcements. In other words, the heterogeneity in biases in our model is driven by a news event's association with tail events in the same category, not its mechanical size.

One can translate Proposition 1 into forecast error predictability (Coibion and Gorodnichenko, 2015; Bordalo et al., 2020a), assuming each analyst is also diagnostic and has the same information as an investor. At $t = 1$, the forecaster revises her expectations regarding the growth rate from its ex ante mean 0 to $E^{\theta}_{j,t}[v|\mathcal{C}, s_j]$. Proposition 1 implies that forecaster i's forecast error, $FE_{j,1} = v - E^{\theta}_{j,1}[v|\mathcal{C}, s_j]$, is predictable by her forecast revision, $FR_{j,1} = E^{\theta}_{j,1}[v|\mathcal{C}, s_j]$. The coefficient $\beta^{CG}_{\mathcal{C}}$ from the regression, $FE_{j,1} = \alpha + \beta^{CG}_{\mathcal{C}} FR_{j,1} + \epsilon_j$, is positive if forecasters systematically underreact to announcements in category $\mathcal{C}$, negative if they overreact, and zero under rational expectations.

Corollary 1. *β_C^{CG} decreases in ζ_C, and is negative if and only if $\zeta_C > \zeta_d$.*

Corollary 1 shows that β_C^{CG} decreases in ζ_C: there is greater overreaction, as indicated by a more negative forecast error predictability coefficient, for more extreme news categories. Corollary 1 corresponds directly to the following empirical prediction.

Prediction 1. Forecast errors are more negatively predicted by forecast revisions (greater overreaction) for more extreme news categories.

Return predictability by news category Imposing market clearing at $t = 1$ and 2, we obtain Proposition 2, which relates our results to returns and trading volume.

Proposition 2 (Returns and volume)**.** Denote $\zeta_{\mathcal{C},\theta}^{-1} \equiv \zeta_{\mathcal{C}}^{-1} + \theta\left(\zeta_{\mathcal{C}}^{-1} - \zeta_d^{-1}\right)$, and $\eta_{\mathcal{C}}(v) \equiv \frac{v_{0,\mathcal{C}}^2 + v^2}{2v}$. Period 1 and 2 returns $r_t = \log(p_t) - \log(p_{t-1}), t \in \{1, 2\}$ satisfy

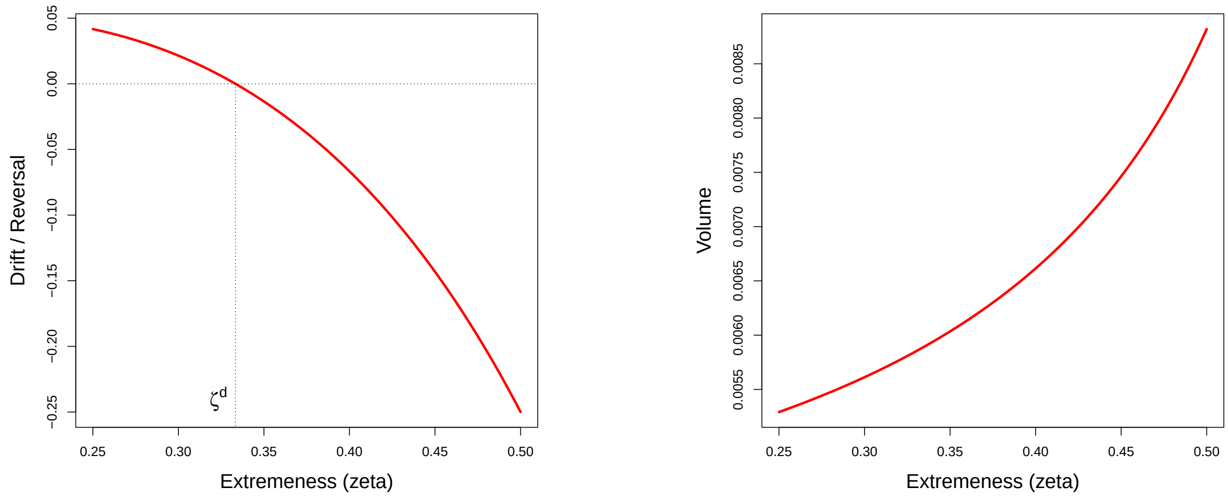
$$r_2 = \beta_{\mathcal{C}}^{ret} \cdot r_1,\ \beta_{\mathcal{C}}^{ret} \equiv \frac{\zeta_{\mathcal{C}} - \zeta_{\mathcal{C},\theta}}{1 + \zeta_{\mathcal{C},\theta}}. \tag{7}$$

The volume at $t = 1$ (announcement-day) is given by $Vol = \frac{1}{2}\kappa \cdot (1 + \zeta_{\mathcal{C},\theta}) \cdot (1 - \eta_{\mathcal{C}}(v))^2$.

$\beta_{\mathcal{C}}^{ret}$ captures the predictive relationship between period 1 returns, the announcement-day returns, and period 2 returns, the post-announcement returns.If $\beta_{\mathcal{C}}^{ret} < 0$, there is overreaction in asset prices to news events of category $\mathcal{C}$: a fraction $|\beta_{\mathcal{C}}^{ret}|$ of initial returns is reversed. Conversely, if $\beta_{\mathcal{C}}^{ret} > 0$, there is underreaction and drift. Corollary 2 summarizes the comparative statics of $\beta_{\mathcal{C}}^{ret}$ and volume with respect to $\zeta_{\mathcal{C}}$.

Corollary 2. *The drift-reversal coefficient $\beta_{\mathcal{C}}^{ret}$ decreases in $\zeta_{\mathcal{C}}$ and the diagnostic parameter θ. News categories whose distribution of fundamentals are more extreme than the reference distribution ($\zeta_{\mathcal{C}} > \zeta_d$) are associated with reversals ($\beta_{\mathcal{C}}^{ret} < 0$), and those that are less extreme ($\zeta_{\mathcal{C}} < \zeta_d$) are associated with drift ($\beta_{\mathcal{C}}^{ret} > 0$). Holding v fixed, trading volume increases in $\zeta_{\mathcal{C}}$.*

The predictions of Corollary 2 are visualized in Figure 2: there is more overreaction, or greater short-term reversals, for more extreme news categories. The post-announcement return r_2 is generated by the gradual entry of rational investors, not the long-run revelation of fundamentals. Furthermore, as depicted by Figure 2b, our theory also has



(a) Drift-Reversals vs Extremeness

(b) Volume vs Extremeness

Figure 2: DE predictions: over- and under-reaction, volume, and extremeness

Note: Figure 2a plots the theoretical relationship between return drift/reversal and the extremeness of the distribution of fundamentals. The dashed vertical line (ζ_d) corresponds to the extremeness of the distribution of fundamentals for the reference distribution. The dashed horizontal line corresponds to a drift/reversal coefficient β_C^{ret} of zero. Figure 2b plots the theoretical relationship between trading volume and the extremeness of the distribution of fundamentals. Volume is defined as half of absolute asset holdings at $t = 1$, holding fixed fundamentals.

implications for trading volume. Holding fixed fundamentals v, as the underlying distribution grows more extreme, diagnostic agents trade more aggressively based on their private signals, leading to greater trading volume. To summarize, Corollary 2 implies the following empirical predictions.

Prediction 2. News categories of more extreme fundamental distribution are associated with greater short-term post-announcement reversals.

Prediction 3. News categories of more extreme fundamental distribution are associated with greater announcement day trading, holding fixed the fundamentals of the news.

Measuring ζ_C Our model makes a final prediction relevant for measuring ζ_C. Directly measuring the long-run impact of news on fundamentals can be challenging. Corollary 3 imply that one can measure ζ_C by instead through the distribution of short-term returns.

Corollary 3. *More extreme categories also have a more extreme distribution of r_1.*

Prediction 4. Across news categories, the extremeness of the distribution of fundamentals is positively correlated with the extremeness of the distribution of short-term returns.

To summarize, our model combines tail events with diagnostic expectations to explain how investors react to different categories of news. Prediction 1 captures the insight that the bias for each news category can be predicted by measuring its tail: the fatter the tail, the greater the overreaction. Predictions 2 and 3 translate these expectational distortions to results on return predictability and trading volume. Lastly, Prediction 4 gives guidance on how to estimate the extremeness of each news category. We now take our model to the data, using a comprehensive database of corporate news announcements.

3 Data

We use two main datasets for news events and stock returns. First, we compile our list of corporate news announcements from the Capital IQ Key Developments dataset. Capital IQ tracks major corporate news events such as earnings announcements, product and client announcements, lawsuits and legal issues, leadership changes, and mergers and acquisitions, but excludes macroeconomic news announcements such as interest rates and unemployment rates that may affect aggregate stock returns. Second, we obtain daily stock returns and trading volume from CRSP. Our sample consists of news announcements made by all US companies listed on a major US stock exchange (NASDAQ, NYSE, and AMEX) between 2011 and 2018. For each news announcement in Capital IQ made by a given firm on a given date, we match the announcement to stock returns and trading volume on the day of the announcement and of the subsequent post-announcement period.[12] To mitigate the effects of market microstructure on our results, we exclude small stocks (less than $2 billion in market capitalization).[13] To measure the intensity of news

[12] We match all news announcements made after trading hours to the next trading day.

[13] We show in robustness exercises that our results hold in small stocks as well. We exclude them because as noted by the market microstructure literature, short-term price reversals can occur due to liquidity concerns: at extremely short time scales, bid ask bounces generate negative return autocorrelation. Even at longer time scales, there may be transient price pressure as market makers demand compensation for liquidity while trading against uninformed flow (Kyle, 1985; Campbell et al., 1993; Nagel, 2012).

coverage, we also use data from RavenPack, a financial news and analytics data provider.

We conduct our analysis on a baseline sample of news categories that directly affect the fundamental value of the firms. We first restrict our sample to news categories that occurred at least 1,000 times across all US companies in our sample, which we list in Table A7. To focus on news categories that directly affect the fundamental values of the firm, we further exclude (1) administrative filings such as announcements of earnings dates or name changes, (2) trading activities such as index exclusion, and (3) debt and equity issuances and repurchases including IPOs and SEOs.[14] To ensure that sample selection choices are not driving our results, we repeat our analyses in Section 4.3 using the full set of news categories, as well alternative selection criteria, such as including small-cap stocks and considering different subsets of news categories.

Summary statistics Table 1 reports the summary statistics of the announcements in our sample. In general, corporate announcement days are characterized by significant price movements and trading behavior. The unconditional means across most categories are largely centered around zero with a small but notable positive mean. Announcement days are also generally associated with large absolute returns: the standard deviation of returns on announcement days for almost all categories exceed 2.1%, the average daily return volatility of stocks in our sample. Announcement days are also characterized by high trading volume, with average daily volume on news days exceeding the average daily volume on no-news days for most news categories. Overall, the data suggest that news announcement days are characterized by higher return volatility and trading volume, consistent with prior work (Solomon, 2012; Neuhierl et al., 2013; Engelberg et al., 2018).

4 Overreaction in extreme news categories

In this section, we present our core empirical findings. We begin with our estimation of the extremeness of each news category. Consistent with the core assumption of our model, we find that the distribution of fundamentals of each news category is well-

[14] We exclude IPOs to avoid conflating IPO announcement-day returns with the IPO premium.

approximated by a power law distribution, with significant variation in extremeness across categories. We then test our core prediction that there is greater overreaction, or short-term reversals, for news in a more extreme category. We support our main finding with several robustness tests to address potential concerns such as announcement timing, overlapping news, and accounting for the magnitude of the news. Finally, we test the additional predictions of our model and find that more extreme news categories are associated with greater trading volume and negative forecast error predictability.

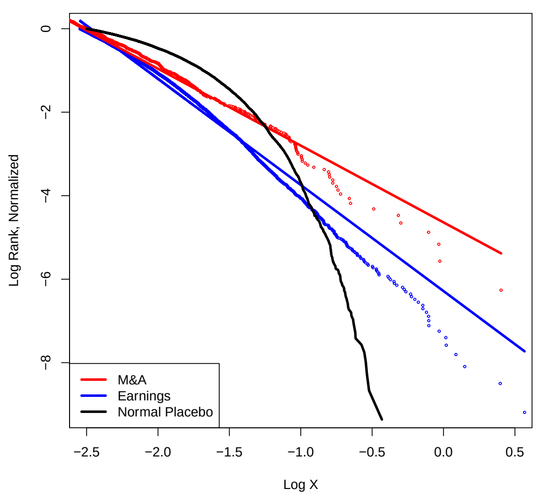
4.1 Extremeness of news categories

Measuring extremeness Our model is based on two core assumptions. First, the distribution of fundamentals of each news category $\mathcal{C}$ is fat-tailed, or extreme. Second, the extremeness of the distribution ($\zeta_{\mathcal{C}}$) differs systematically for each news category. We validate these assumptions by measuring the realized distribution of fundamentals for each news category. For each category $\mathcal{C}$ in our dataset, we collect the set of announcements $\{n_{i,t,\mathcal{C}}\}$, where $n_{i,t,\mathcal{C}}$ refers to an announcement of category $\mathcal{C}$ for firm i at time t. Measuring category $\mathcal{C}$'s extremeness requires two choices. First, for each $n_{i,t,\mathcal{C}}$, we need a measure of the event's impact on firm value, i.e. $\hat{v}(n_{i,t,\mathcal{C}})$. Second, once we construct the set $V_{\mathcal{C}} = \{\hat{v}(n_{i,t,\mathcal{C}})\}$, we need a measure of the extremeness of its empirical distribution, $\hat{\zeta}(V_{\mathcal{C}})$.

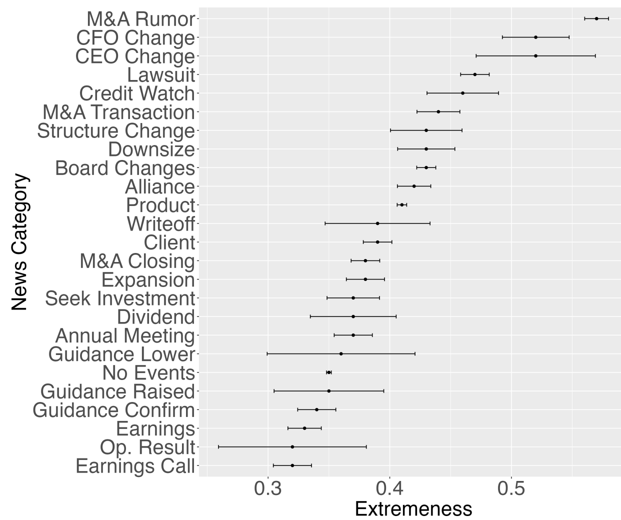
Main specification In our main specification, we use announcement-day returns as a proxy for the news' impact on fundamentals: $\hat{v}(n_{i,t,\mathcal{C}}) = r_{i,t}$. The benefit of this approach is that announcement-day returns can be more reliably attributed to the news than longer term measures. Prediction 4 also implies that $\zeta_{\mathcal{C}}$ can be measured by the tails of $r_{i,t}$ as well as directly from v, which we later validate. For each category $\mathcal{C}$, we fit a power law distribution to the tail of absolute announcement-day returns: taking the top 10% absolute announcement returns, we estimate the log rank-value regression (Gabaix, 2016):

$$\log(Rank_{i,t,\mathcal{C}}) = \xi_{\mathcal{C}} - \hat{\zeta}_{\mathcal{C}}^{-1} \log(|r_{i,t,\mathcal{C}}|), \; |r_{i,t,\mathcal{C}}| > |r_{\mathcal{C},90}|. \tag{8}$$

The relationship is negative by construction. The regression coefficient $\hat{\zeta}_{\mathcal{C}}^{-1}$ captures how much an increase in absolute returns corresponds to a move up in the percentile rank. If the distribution is power-law with $F(x) = 1 - (x/x_{min})^{-k}$, the relationship is exact with $\hat{\zeta}_{\mathcal{C}}^{-1} = k$.



(a) Extremeness estimation: example



(b) Extremeness estimates

Figure 3: Estimating extremeness: $\zeta_{\mathcal{C}}$

Note: Figure 3a plots the estimates corresponding to eq. (8) for M&A events (red), earnings events (blue), and a simulated normally-distributed return distribution (black). The x-axis shows the normalized log value of absolute announcement-day returns, while the y-axis shows the normalized log rank. The solid lines plot the linear best fit corresponding to eq. (8). Figure 3b plots the extremeness $\zeta_{\mathcal{C}}$ estimates for each category corresponding to eq. (8). 95% confidence intervals are computed following Gabaix and Ibragimov (2011).

Figure 3a plots the relationship for two news categories, earnings (in blue) and M&A announcements (in red), and for a simulated normal distribution (in black) with a similar standard deviation. The raw data points are plotted as points and the linear regression estimates following eq. (8) are plotted in solid lines. Figure 3a shows that the tails of the distributions from both news categories are far better fit by power-law distributions than by a normal distribution, whose corresponding curve decays faster than any linear fit. While the plot only shows two categories, the conclusion holds generally: the R^2 associated with the linear fit is close to 1 (above 98%) for all news categories in our sample, compared to an R^2 of 86% for the simulated normal distribution. Overall, the distribution of fundamentals for all news categories are described well by a power-law, with the tail parameters precisely estimated and not driven by a small number of data points. Table 2a reports the estimates of category extremeness (ζ column) and their standard errors (ζ s.e. column): the median standard error across all news categories is 0.01.

Furthermore, there is significant variation in extremeness across news categories. Figure 3b plots the $\zeta_{\mathcal{C}}$ estimates for each category and their 95% confidence intervals, computed following Gabaix and Ibragimov (2011). The coefficient estimates suggest significant variation in the extremeness of fundamentals across news categories, with $\zeta_{\mathcal{C}}$ ranging from 0.32 to 0.57. 20 categories (15 statistically significantly) are more extreme than the no news distribution, i.e., the reference distribution in our model, and 4 categories (2 statistically significantly) are less extreme.[15] To give a sense of the economic magnitudes of these differences, one can translate these results into the magnitude of tail returns. The average announcement-day return greater than 5 percentage points (p.p.) is 8.1 p.p. for earnings calls and 9.5 p.p., or 16.0% greater, for CFO changes. To summarize, we find economically and statistically significant differences in the extremeness of the fundamentals across news categories.

Consistency of tail measures: Prediction 4 We also consider alternative measures of $\zeta_{\mathcal{C}}$. First, we consider alternative measures of fundamentals. Instead of announcement-day returns, we consider longer-term (100-day) returns and earnings growth over k years:

$$\hat{v}(e_{i,t,\mathcal{C}})^{EPS,k} = EPS_{i,t-1+k,\mathcal{C}}/EPS_{i,t-1,\mathcal{C}} - 1, \tag{9}$$

where $1 \le k \le 5$ and $EPS_{i,t}$ is the year t earnings-per-share reported by firm i. We restrict our sample for firms whose earnings per share in year $t-1$ are at least 10 cents. We then compute the power-law coefficients using these alternative measures of fundamentals. Consistent with Prediction 4, we find that all of our measures are highly correlated at

[15] We also compute standard errors using a simple bootstrap. Both approaches yield similar standard errors. We test the significance of the difference in $\zeta_{\mathcal{C}}$ between each news category and the no news distribution assuming the two samples are independent, as they are drawn from distinct days by definition.

the news category level: Table A1 reports the pairwise correlation of our measures of $\zeta_{\mathcal{C}}$. Second, we also consider alternative measures of extremeness, such as skew or quantile ratios. For power law distributions, all of these measures correspond one-for-one with the tail index, although higher-order moments may be unreliably measured for extreme distributions. Consistent with the precision of the power-law estimates, Figure A3 of the Online Appendix shows that all tail measures are highly correlated.

What extremeness rules out On the other hand, extremeness is not captured by other intuitive measures, such as variance or the frequency of large news. What extremeness captures is the relative difference between tail outcomes and typical news within the category, not the average magnitude or unconditional frequency of large news. Earnings announcements provide an illustrative example. While earnings tend to have large announcement-day returns, it is among the least extreme categories, as an outstanding earnings announcement does not result in a much larger impact than other positive earnings announcements. We compare the predictive power of extremeness to these alternative measures in Section 5.2.

4.2 Overreaction to extreme news categories

Drift and reversals Given our measure of category extremeness, we now test our key hypothesis that there is greater overreaction for news in more extreme categories. Our measure of over-and-underreaction in asset prices is given by whether announcement-day returns (r_1) positively or negatively predict post-announcement returns (r_2). Proposition 2 implies the following relationship between the two:

$$r_2 = \beta_{\mathcal{C}}^{ret}(\zeta_{\mathcal{C}}, \theta, \zeta_d) \cdot r_1, \tag{10}$$

where $\frac{\partial \beta_{\mathcal{C}}^{ret}}{\partial \zeta_{\mathcal{C}}} < 0$: more reversals to news belonging in more extreme categories. $\beta_{\mathcal{C}}^{ret}$ corresponds empirically to $\beta_{\mathcal{C}}$ of the following autocorrelation regression:

$$r_{i,t+1,t+k} = \alpha + \beta_{\mathcal{C}} \cdot r_{i,t} + \epsilon_{i,t}, \tag{11}$$

where we pool all occurrences of events in $\mathcal{C}$ that occur on day t for firm i. $r_{i,t}$ is the announcement-day return (corresponding to $t = 1$ in the model) and $r_{i,t+1,t+k}$ is the k-day cumulative post-announcement returns ($t = 2$ in the model).[16] If $\beta_{\mathcal{C}} = 1$, then half of the price movements for news in category $\mathcal{C}$ are realized on the announcement day on average, with a predictable drift of equal proportion over the next k days. If $\beta_{\mathcal{C}} = -0.5$, then half of announcement-day returns would be reversed on average, so the initial price impact would be twice as responsive as the rational benchmark.

Variation in drifts and reversals across news categories Before we test our main hypothesis, we first document the heterogeneity in drift and reversals across our news categories. In our baseline specification, we set $k = 90$ days, similar to the horizon considered by the post-earnings announcement drift (PEAD) literature.[17] To test whether there are significant category-level differences in $\beta_{\mathcal{C}}$, we estimate the following regression using news announcements across all 24 news categories in our sample:

$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \sum_{\mathcal{C} \in \mathbf{C}} \beta_{\mathcal{C}} \cdot 1(News_{\mathcal{C}}) \cdot r_{i,t,\mathcal{C}} + \mu_{\mathcal{C}} + \epsilon_{i,t,\mathcal{C}}, \tag{12}$$

where each observation is a category $\mathcal{C}$ news announcement by firm i on date t. $1(News_{\mathcal{C}})$ is a dummy variable for whether the announcement belongs to news category $\mathcal{C}$, $r_{i,t+1,t+k,\mathcal{C}}$ is the cumulative k-trading days post-announcement returns, and $r_{i,t,\mathcal{C}}$ is the announcement-day return. To ensure our estimates are not driven by outliers, we winsorize announcements for each category at the 1% level. Standard errors are two-way clustered at the firm and day levels.

We wish to test whether there is heterogeneity in post-announcement drifts and reversals across news categories $\mathcal{C}$, $\beta_{\mathcal{C}}$. We conduct two F-tests corresponding to the null hypotheses that all $\beta_{\mathcal{C}}$ are (a) equal to 0, and (b) equal to each other. Table 2b reports the results. We find that $F = 2.55$ for (a) and 1.79 for (b), which rejects both null hypotheses

[16] Concretely, $r_{i,t}$ is the return of firm i from the close of date $t-1$ to the close of date t. $r_{i,t+1,t+k}$ is the return of firm i from the close of date t to the close of date $t+k$.

[17] In Section 4.3, we conduct robustness checks by varying the horizons by setting $k = 30$ and 60, and we also repeat our analysis using both stock returns benchmarked relative to S&P 500 returns and without benchmarking.

with $p < 0.01$, indicating that there is significant heterogeneity in $\beta_{\mathcal{C}}$ across categories. To further illustrate the variation in the data, Table 2a shows the category-level estimates of $\beta_{\mathcal{C}}$ of our news categories. Consistent with the literature on post-earnings announcement drift (Bernard and Thomas, 1989), we find drift for earnings announcements. On the other hand, we find reversals of comparable magnitudes for other news categories, such as leadership changes, mergers and acquisitions, and client-related announcements.[18]

Reversals for extreme categories: testing Prediction 2 We now formally test our core prediction: greater overreaction and reversals for news in more extreme categories. We estimate the following linear regression on our sample of all news announcements:

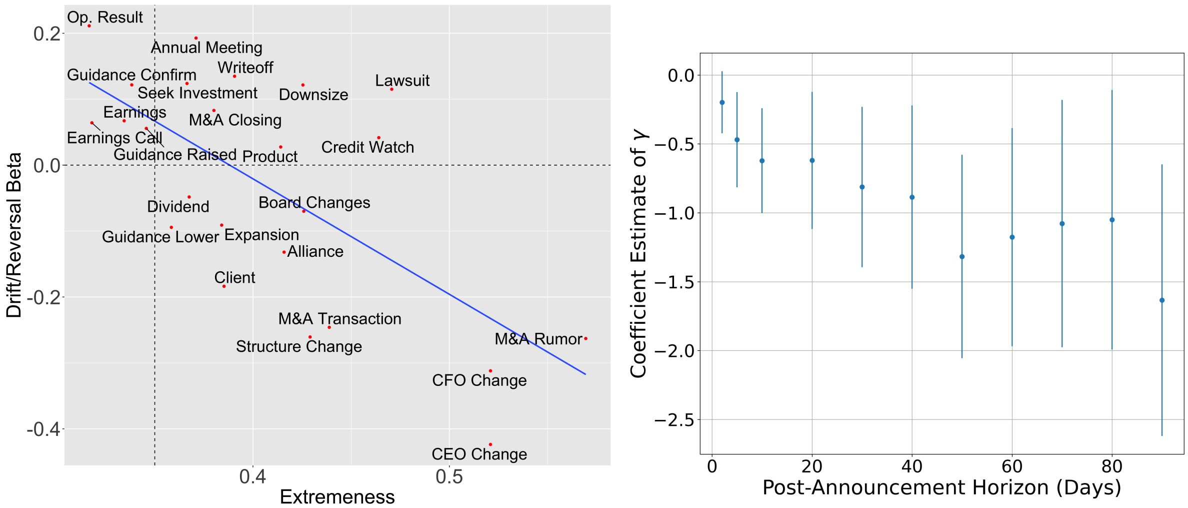
$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \beta \cdot r_{i,t,\mathcal{C}} + \gamma \cdot \zeta_{\mathcal{C},t} \times r_{i,t,\mathcal{C}} + \epsilon_{i,t,\mathcal{C}}, \tag{13}$$

where observations are at the news announcement level. $r_{i,t,\mathcal{C}}$ is the announcement-day return and $r_{i,t+1,t+k,\mathcal{C}}$ is the k-day cumulative post-announcement returns. $\zeta_{\mathcal{C},t}$ is the extremeness of category $\mathcal{C}$ as of time t.[19] Standard errors are two-way clustered at the firm and day level. The coefficient of interest is γ, which captures how post-announcement drifts or reversals vary in category extremeness. A negative γ implies that news from more extreme categories, i.e. a larger $\zeta_{\mathcal{C},t}$, are associated with greater reversals.

Table 3 reports the results corresponding to equation (13), where we set $k = 90$. Column (1) reports our baseline estimate. Column (2) uses returns benchmarked against the S&P 500. Column (3) is a predictive regression that uses only announcements over the past five years to compute extremeness for each announcement. Column (4) uses both S&P 500-benchmarked returns and the past five-year extremeness. For each specification, we estimate a negative and statistically significant γ coefficient, consistent with our

[18]Post-announcement drifts and reversals can alternatively be measured by the returns of a long-short portfolio sorted on announcement-day returns. In Appendix B, we construct these long-short portfolios and confirm that their returns are positively correlated with $\beta_{\mathcal{C}}$'s. We also find that the economic magnitude of reversals and drift across news categories are comparable: a long-short portfolio for news categories with drift gains 61 basis points (bps) over 90 days, while the same strategy for news categories with reversals loses 111 bps.

[19]We estimate two versions of $\zeta_{\mathcal{C},t}$, one over the entire sample period, which uses data after time t, and another using a trailing window of five years. The latter specification ensures that our results are truly predictive and do not use returns from events in the future.



(a) $\beta_{\mathcal{C}}$ vs $\zeta_{\mathcal{C}}$ (b) γ over time

Figure 4: Extreme categories and reversals

Note: Figure 4a plots the relationship between extremeness and post-announcement drift/reversal $\beta_{\mathcal{C}}^{ret}$ for each news category $\mathcal{C}$. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Drift/Reversal Beta is the post-announcement drift or reversal coefficients $\beta_{\mathcal{C}}^{ret}$ estimated following equation (11). The dotted horizontal line indicates where drift/reversal $\beta_{\mathcal{C}}^{ret} = 0$. The dotted vertical line indicates where $\zeta_{\mathcal{C}} = 0.35$, which is the extremeness of the No News distribution. Figure 4b plots the estimated γ over different post-announcement horizons from $k = 2$ to $k = 90$, with the γ for each horizon k being estimated following equation (13). The blue vertical lines plot the 95% confidence intervals for each coefficient estimate.

core prediction that more extreme news categories are more overreacted to. Quantitatively, the variation in extremeness of a news category predicts post-announcement stock price movements ranging from drifts of 7% (95% confidence interval (CI) of [-1%, 15%]) to reversals of -23% (95% CI of [-39%, -6%]).

To visualize our result, Figure 4a plots the drift/reversal coefficients $\beta_{\mathcal{C}}$ against extremeness $\zeta_{\mathcal{C}}$ at the category level. The figure is the empirical analog to the theoretical prediction in Figure 2a. Consistent with the formal regression results, there is a negative relationship between extremeness $\zeta_{\mathcal{C}}$ and the drift-reversal coefficient $\beta_{\mathcal{C}}$: more extreme categories have more reversals, while less extreme categories have drift ($\rho = -0.66$, $p < 0.01$). Figure 4b plots our estimate of γ as we range the horizon from 2 to 90 days. We find that our estimate of γ is robustly negative, with the bulk of the cross-category predictability realized by 40 days.[20]

Underreaction While our estimates imply that our measure predict reversals for the most extreme news categories and drifts for the least extreme, they also imply a small but insignificant degree of drift (3%) for stocks with no events ($\hat{\zeta}_d = 0.35$) (95% confidence interval of [-4%, 11%]). This suggests although there is strong evidence for the cross-category prediction of our theory (more overreaction for more extreme news categories), there may also be slightly more drift in the data than is predicted by our model. This may be due to many forces, such as inattention (DellaVigna and Pollet, 2009), complexity (Engelberg, 2008), capital frictions (Duffie, 2010), or other forces that dampen short-term price reaction to news. While the focus of our paper is to explain the cross-category variation in reaction to news, these forces may modulate the overall level of the bias.

4.3 Robustness

We next test the robustness of our main result. Table 4 summarizes all the robustness exercises and reports the corresponding estimates of our main coefficient of interest γ. We summarize the exercises below and describe them fully in Appendix B.

Alternative measures of ζ_C We first show that our results are robust to how we measure category extremeness. One concern with our main announcement-day returns-based measure is that it may reflect mispricings or fluctuations that may be driven by liquidity or time-varying risk aversion. We address this by constructing alternative extremeness measures based on realized earnings growth, which are not driven by market fluctuations, and longer-horizon returns (from the announcement day to 100 days after). The results are summarized in rows 2 and 3 of Table 4, with our baseline estimate replicated in row 1. Consistent with our earlier findings that our tail measures are highly correlated, we find that our estimates continue to hold for both alternative measures.

[20]The relatively short horizon of return predictability is similar to that of other short-term mispricing; for example, Duffie (2010) document that index deletion effects are reversed also roughly within a comparable period. Given that we do not observe disaggregated trading flows, further work is needed to understand how news-driven mispricings are corrected over time.

Announcement timing Measurement errors in the announcement dates, leakages, or delays may bias our measurement of announcement returns. To account for this, in row 4 in Table 4, we report an alternative specification where we define the announcement window as from 2 days before to 2 days after the announcement date, i.e., the close of date $t-3$ to the close of date $t+2$, and the corresponding post-announcement window as from the close of $t+2$ to the close of $t+90$. Announcement timing can also be strategic. For example, firms may release bad news on Fridays when investors are distracted (DellaVigna and Pollet, 2009). In row 5, we add indicator variables for the hour-by-day-of-the-week that each news announcement was made on (e.g., Friday at 4pm) as both fixed effects and interacted with announcement-day returns to control for variations in the post-announcement drift/reversal patterns across announcement times. In row 6 we repeat our analysis excluding Friday announcements. In all specifications, the coefficient of interest γ remains similar.

Magnitude and sign of the news Our theoretical and empirical focus is on explaining cross-category differences in investor reaction to news. In particular, our theory predicts a uniform degree of over-or-underreaction, $\beta_{\mathcal{C}}$, for large and small news within the same category. In practice, investor biases may also depend on the sign and the relative magnitude of news and announcement returns: Hong et al. (2000) show that negative news is associated with greater drift, while Chan (2003) find that large returns may lead to reversals. These forces may mechanically generate cross-category differences; if there are greater reversals to large returns, news categories with larger announcement-day returns may be associated with reversals. To address this concern, we add functions of announcement-day returns as controls to flexibly capture any unconditional relationship between announcement-day returns and post-announcement returns : in row 7, we consider a a non-parametric decile function $f(r_{i,t,\mathcal{C}}) = \sum_{k=1}^{10} \gamma_k \cdot 1(r_{i,t,\mathcal{C}} \in \Delta_k)$, where Δ_k is the k-th decile of all announcement-day returns, and a cubic polynomial in $r_{i,t,\mathcal{C}}$ in row 8. In both cases, we find similar estimates of γ: extreme news categories are associated with greater overreaction, even after accounting for the magnitude and sign of the news.[21]

[21] Figure A4 plots the estimated $f(r_{i,t})$ for the cubic specification. Consistent with Hong et al. (2000), the estimated function is upward sloping for negative returns (drift for negative news) and downward sloping

We also test whether our results hold in subsamples of announcements of different sizes and valence. Rows 9 and 10 report the results when we exclude announcements in the lowest 25-th and 50-th percentiles absolute announcement-day returns, respectively, to ensure that our findings are not due to small and economically insignificant events. Conversely, to also show that our results are not driven by outliers, rows 11 and 12 exclude outlier events, i.e., those in the top 0.01% and 1% of absolute returns.[22] Lastly, we also split our sample into positive and negative news. Rows 13 and 14 show that our results are most pronounced among events with positive announcement returns, and are smaller in magnitude and statistically insignificant for negative news.

Overlapping news Another potential concern is that the predictability for a given news category can be due to systematic overlap with another category. For example, if there is continued drift for one news category, a news category that systematically follows that category may also be mechanically associated with drift. To account for overlaps, we perform two types of exercises. First, since overlapping announcements are most pronounced for earnings announcements, we exclude news categories whose announcements occur within five days of earnings announcements more than 50% of the time (row 15). More stringently, we remove all announcements from firms that had any other news announcements within the prior 30 days (row 16). Second, instead of excluding announcements, we directly control for the effects of all other news that occurred within the past 90 days on the 90-day post-announcement returns of the current news announcement (row 17). We estimate the following regression specification:

$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \beta_0 \cdot r_{i,t,\mathcal{C}} + \gamma \cdot \zeta_{\mathcal{C},t} \times r_{i,t,\mathcal{C}} + \sum_{\mathcal{C}' \in \mathbf{C}} \sum_{h=1}^{90} \theta_{\mathcal{C}',h} \cdot I(i,\mathcal{C}',t-h) \cdot r_{i,t-h,\mathcal{C}'} + \epsilon_{i,t,\mathcal{C}}, \tag{14}$$

where $I(i,\mathcal{C},t)$ is an indicator for whether firm i has experienced an announcement of category $\mathcal{C}$ on date t. Eq. (14) augments our main specification by also accounting for the

for positive returns (reversals for positive news). Moreover, for positive events, the estimated function is concave, suggesting a larger degree of reversals for large positive announcement-day returns.

[22]While outliers in a category drive investor reaction to news (and implicitly form our predictive measure), the predictability should hold for all news of that category, even excluding said outliers.

component of returns that may be due to past news. $\theta_{\mathcal{C}',h}$ accounts for the component of the return from date $t+1$ to $t+k$ that may be due to reactions to news announcements of category $\mathcal{C}'$ that occurred h days prior to the current announcement. Across all specifications, we find a significant γ of a similar magnitude to our baseline estimates.[23]

Sample selection We also show that our results are robust to various choices of sample selection criteria and the inclusion and exclusion of different news categories. As discussed in Section 3, we select our baseline sample based on two criteria: (1) news categories that have occurred at least 1,000 times in our sample, and (2) excluding news categories that we judged to be not directly pertinent to the fundamentals of the company, including administrative announcements (e.g. news about earnings release date), index inclusions, and capital structure announcements. To show that our results are robust to our selection criteria, rows 18 and 19 repeat our analysis including all news categories that we have excluded, i.e., by undoing both criteria (1) and (2) in row 18 and by using only criteria (1) in row 19.[24] Lastly, row 20 repeats our analysis on small-cap stocks.

Other robustness exercises Table 4 summarizes the remaining robustness exercises. Rows 21 and 22 report the results under different post-announcement return horizons ($k = 30$ and 60). We account for attrition by excluding attrited firms (row 23) and excluding news categories with above-average attrition rates (row 24), alternative standard errors that account for overlapping windows in panel data (Driscoll and Kraay, 1998) (row 25), and compounded estimation errors in using estimated category extremenesses as regressors (Pagan, 1984; Murphy and Topel, 2002) (row 26). Across these additional specifications, we find that our results remain economically and statistically similar.

[23] In Appendix B, we also estimate equation 14 accounting for both news before date t and news after date t, i.e., $h = -90$ to 90.

[24] In Appendix B, we also systematically exclude news categories one-by-one and show that our results are robust to excluding any particular news category in our main sample.

4.4 Testing additional predictions: volume and expectations

Extremeness and volume Prediction 3 implies that more extreme news categories are associated with greater disagreement and trading, holding fixed fundamentals. To test this prediction, we estimate the relationship between volume and extremeness in our sample, holding fixed the absolute announcement returns as a proxy for fundamentals:

$$Turnover_{i,t,\mathcal{C}} = \alpha + \beta \cdot |r_{i,t,\mathcal{C}}| + \delta \cdot |r_{i,t,\mathcal{C}}| \cdot \zeta_{\mathcal{C}} + \mu_t + \mu_{\mathcal{C}} + \epsilon_{i,t,\mathcal{C}}. \tag{15}$$

Observations are at the announcement level. $Turnover_{i,t,\mathcal{C}}$ is the announcement-day volume normalized by the total shares outstanding, $|r_{i,t,\mathcal{C}}|$ is the absolute value of the announcement-day return, and $\zeta_{\mathcal{C}}$ is the extremeness of category $\mathcal{C}$. μ_t and $\mu_{\mathcal{C}}$ are day and category fixed effects. A positive δ implies that extreme categories are associated with more trading, holding fixed the magnitude of announcement returns. Table 5 presents the estimated coefficients. For each specification, we estimate a positive and statistically significant δ: news from more extreme categories generate greater volume holding fixed the magnitude of announcement-day returns. An increase in category extremeness from the least extreme to the most extreme category corresponds to a 32% increase in predicted turnover, from 5.1% to 6.7%.

To visualize our findings, we estimate for each news category the average turnover conditional on a 10% announcement-day return using the following specification:

$$Turnover_{i,t,\mathcal{C}} = \alpha_{\mathcal{C}} + \sum_{\mathcal{C} \in \mathbf{C}} T_{\mathcal{C}} \cdot 1(News_{\mathcal{C}}) \cdot |r_{i,t,\mathcal{C}}| + \epsilon_{i,t,\mathcal{C}}, \tag{16}$$

with $\overline{Turnover}_{\mathcal{C},10} = \alpha_{\mathcal{C}} + \beta_{\mathcal{C}} \cdot 10\%$. Figure 5 plots the relationship between $\zeta_{\mathcal{C}}$ and $\overline{Turnover}_{\mathcal{C},10}$ and is the empirical counterpart to the theoretical prediction in Figure 2b. Consistent with the results in Table 5, we find a strong positive relationship: more extreme news categories have greater turnover adjusted for returns ($\rho = 0.66$, $p < 0.01$).

Forecast error predictability Having tested Predictions 2 and 3 regarding returns and trading volume, we use expectations data to test Prediction 1, that forecast errors are

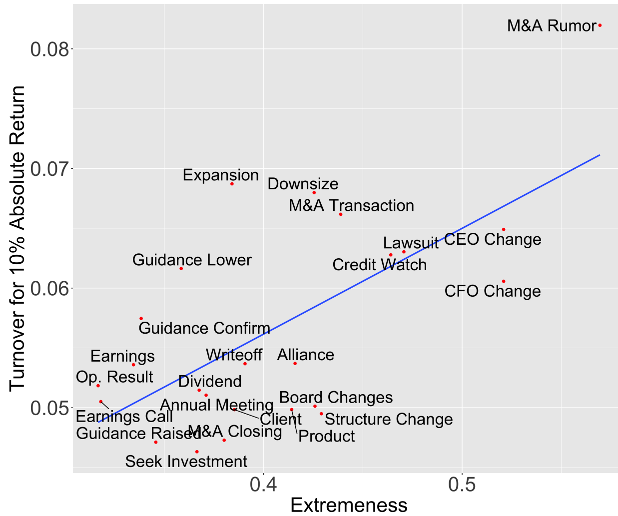


Figure 5: Volume and Extremeness

Note: Figure 5 plots the relationship between extremeness and conditional average turnover for a news announcement with a 10% absolute announcement-day return $\overline{Turnover}_{\mathcal{C},10}$ for each news category $\mathcal{C}$. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Turnover for 10% Absolute Return is the conditional average turnover for a news announcement with a 10% absolute announcement-day return, $\overline{Turnover}_{\mathcal{C},10}$, estimated following eq. (16).

more negatively predictable by revisions for more extreme news categories. We follow a similar procedure to Bouchaud et al. (2019) to compute analyst revisions and forecast errors using I/B/E/S analyst forecasts of earnings.[25] Pooling across all analyst forecasts, we regress forecast errors on forecast revisions (Coibion and Gorodnichenko, 2015), interacted with our category-level extremeness measure $\zeta_{\mathcal{C}}$:

$$ForecastError_{a,i,t,\mathcal{C}} = \alpha + \beta^{CG} \cdot ForecastRevision_{a,i,t,\mathcal{C}} + \gamma^{CG} \cdot \zeta_{\mathcal{C},t} \times ForecastRevision_{a,i,t,\mathcal{C}} + \epsilon_{i,t}. \tag{17}$$

Observations are at the analyst a by announcement level. Equation 17 is the expectations counterpart to the return predictability regression (eq. (13)). $\gamma^{CG} < 0$ implies that forecast revisions more negatively predict forecast errors – revisions are more likely to

[25] First, to remove stale forecasts, we only use forecasts that were issued 90 days or less before the announcement and revised 45 days or less after the announcement. Second, we winsorize the forecasts at the 10% level to remove anomalous forecasts. Third, we focus on forecasts made for two years ahead, although the results are qualitatively similar if we use different forecast horizons. Finally, we normalize each forecast by the price of the stock at the time of the forecasts.

overshoot the rational benchmark – for more extreme categories. Table 6 reports the results. Standard errors are two-way clustered at the analyst and firm levels. Column (1) reports the baseline estimates, and columns (2) and (3) report estimates with analyst and firm fixed effects, respectively. We find suggestive evidence that forecast errors are more negatively correlated with forecast revisions for more extreme news categories. The results become statistically insignificant with the inclusion of firm fixed effects, potentially because a large driver of the variation in the data is differences in news across firms.

5 Discussion

Our core findings of greater overreaction to news in extreme categories relate to a large body of work on investor psychology. In this section, we connect our results to the broader literature and test for alternative explanations for our findings.

5.1 Relation to existing work

Psychology of tail events The key message of our model – that retrieval of past tail events shapes reaction to news – is consistent with a rich literature in psychology and economics. On one hand, people overweight tail outcomes, especially in settings where they are either explicitly described or top of mind. In the lab, participants place greater weight on rare outcomes with explicit probabilities, with potentially salient payoffs (Tversky and Kahneman, 1973; Kahneman, 2011; Bordalo et al., 2022). Barberis and Huang (2008) similarly finds that investors are willing to pay more for stocks with lottery characteristics. In all cases, tail events have a major impact on beliefs. On the other hand, people are also known to neglect tail outcomes in different contexts (Barberis, 2013). Hertwig and Erev (2009) strikingly find that when experimental participants sample draws from a random lottery (instead of being explicitly described their probabilities), they tend to neglect experienced tail outcomes, a phenomenon the authors coin as the experience-description gap. Taken jointly, these findings suggest that whether tail events are overweighted depends on broader contextual forces that influence the availability of these events.[26]

We contribute to this literature by focusing on the role of news as a potential cue for tail events. The key intuition is that news belonging to extreme categories act as stronger cues for tail events, thereby triggering overreaction. Moreover, the exact degree to which a category is representative of tail events can be measured in the data, which generates a quantitative predictor of investor biases. While far from a complete measure of what comes to an investor's mind, our tail measure is able to capture a significant variation of the cross-category differences in reaction to news, as we test in Section 5.2.

Rational learning Our paper also relates to the literature in which agents rationally learn from tail events: when agents are uncertain about the parameters of the data generating process (Hansen, 2007), tail events can have a large impact on beliefs (Kozlowski et al., 2020). Relative to this literature, our focus is on how tail events across in a given category affect agents' reaction to all news, not just tail events, of the same category. Moreover, our work focuses on investor over-and-underreaction, manifested in predictable returns or forecast errors, not just the persistent impact of tail outcomes on beliefs.[27]

Other applications of DE Our paper is part of a growing list of papers that have applied diagnostic expectations to financial and macroeconomic settings. Theoretically, most applications of DE have highlighted overreaction and excess sensitivity of beliefs (Bordalo et al., 2018, 2019; Bianchi et al., 2024; L'Huillier et al., 2023). Our main theoretical innovation is to show that by applying diagnostic expectations to a family of distributions where the tail may vary, one can obtain over-and-underreaction based on the distributional characteristics of the news: if the news is representative of tail outcomes, both the consensus and the individual expectations may overreact. Conversely, if the news

[26]In the field, investors and consumers may neglect tail risk, especially during boom times (Gennaioli et al., 2015). Bordalo et al. (2023b) explicitly model such contextual dependence and find that rare outcomes are oversampled when explicitly cued by the description of the hypothesis.

[27]When one combines learning with convexity, one can also generate biases: for example, if forecast outcomes are convex functions of a given variable, an uncertainty shock over the variable may lead to biased estimates (Orlik and Veldkamp, 2014; Pástor and Veronesi, 2009). We document, however, predictability in not just returns, but also forecasted earnings, which is less subject to convexity concerns. Moreover, to address the concern that our predictability is due to the fact that some news categories are associated with greater uncertainty, we also directly control for the informativeness of the news in Section 5.2.

category is sufficiently thin-tailed, news is representative of non-tail outcomes, and expectations underreact.[28]

5.2 Alternative explanations

We next examine whether alternative explanations could generate our findings. For each alternative explanation we consider, we develop a candidate explanatory variable that proxies for the explanation, and then test the explanatory power of our main measure against these competing variables.

Informativeness One alternative explanation is that category extremeness may instead reflect differences in the informativeness of the news. Tetlock (2014) highlights that attention-grabbing yet "uninformative media content" generates overreaction while "informative content" generates underreaction.[29] If news in extreme categories are typically less informative, this can be an alternative explanation of our findings: for example, an earnings announcement may be much more informative than a CEO firing. We measure the informativeness of a news category from the degree to which prices become more informative of fundamentals post-announcement using the methodology of Dávila and Parlatore (2018) to identify price informativeness. Specifically, for each category $\mathcal{C}$, we pool the announcements to compute $\kappa^p_{\mathcal{C},p10}$ and $\kappa^p_{\mathcal{C},f10}$, which are the price informativeness measures 10 days before and after the announcement. Our measure of the informativeness of $\mathcal{C}$ is given by $\kappa_{\mathcal{C}} \equiv \kappa^p_{\mathcal{C},f10} - \kappa^p_{\mathcal{C},p10}$: if a news category is more informative, it should be reflected in higher price informativeness following the news. We report the details of the methodology in Appendix C.

[28] While Bordalo et al. (2020a) show that DE can be consistent with underreaction in consensus forecasts, it still predicts overreaction in individual forecasts. Consistent with our findings, Bordalo et al. (2019) also show that firms with high long-term-growth expectations have fatter right tails and predict future disappointment, although they do not explicitly model fatter-tailed fundamentals or account for the variation of news in different categories.

[29] Solomon (2012) similarly documents that soft and less informative information can be spun in a positive way, leading to investor overreaction. Griffin and Tversky (1992) and Augenblick et al. (2021) provide experimental evidence that individuals overreact to less informative signals.

Media Investors may react differently to news announcements that are extensively covered by the media. First, media may have a direct causal impact on reaction to news and trading (Engelberg and Parsons, 2011). Alternatively, media coverage of a news announcement can be a measure of its salience (Tetlock, 2014; Bybee et al., 2023). For each news category, we compute $Media_{\mathcal{C}}$ as an alternative explanatory variable, which is the average number of news articles written about news announcements in category $\mathcal{C}$.

Sign and magnitude of news at the category level While we show in Section 4.3 that our main findings hold controlling for the size and sign of the news at the announcement level, we now consider the size of the news as a potential competing measure: on average, some news categories may generate larger returns or absolute returns, which may explain the category-level differences we find. For each $\mathcal{C}$, we compute the mean and standard deviation of announcement returns, $\mu_{\mathcal{C}}$, $SD_{\mathcal{C}}$, and the mean of the five largest announcements, $Largest_{\mathcal{C}}$, as alternative explanatory variables.

Other news characteristics Investors may also react differently to news that they are more accustomed to. We proxy for familiarity by the number of occurrences of category $\mathcal{C}$ divided by 100,000 ($N_{\mathcal{C}}$). Our findings may also reflect the difference between anticipated and unanticipated news: investors may have more time to prepare for the former. We add a variable, $1(Scheduled)_{\mathcal{C}}$, for categories that occur on pre-announced schedules (operating results, earnings, guidances, dividends, earnings calls, and annual meetings).

5.2.1 Testing alternative explanations

We test for whether our results are robust to the inclusion of each of these alternative explanatory variables using the following empirical specification:

$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \beta \cdot r_{i,t,\mathcal{C}} + \gamma \cdot \zeta_{\mathcal{C}} \cdot r_{i,t,\mathcal{C}} + \phi \cdot X_{\mathcal{C}} \cdot r_{i,t,\mathcal{C}} + \epsilon_{i,t}, \tag{18}$$

where observations are at the announcement level and $X_{\mathcal{C}}$ is one of the alternative explanatory variables above. Table 7 reports the corresponding estimates. Across each of the alternative explanations, we find that extremeness remains a statistically and economically significant predictor of post-announcement drifts or reversals.

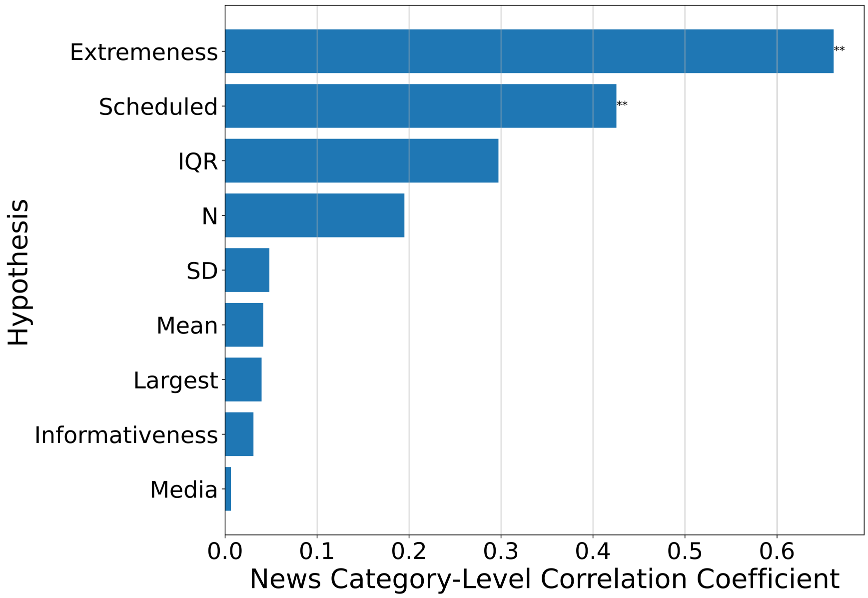


Figure 6: Comparison of explanatory power

Note: Figure 6 plots the category-level correlation coefficients between the category's post-announcement drift or reversal coefficient $\beta_{\mathcal{C}}$ and either extremeness or other alternative explanatory variables (Hypothesis). *** p<0.01, ** p<0.05, * p<0.10.

Lastly, we compare the cross-category explanatory power of our measure relative to these alternative measures. Specifically, for each news category, we correlate the under-overreaction coefficient, $\beta_{\mathcal{C}}$ with each alternative variable $X_{\mathcal{C}}$ from eq. (18). Figure 6 plots the category-level correlation coefficients for extremeness and other alternative measures. Our measure of extremeness most strongly predicts cross-category variation in $\beta_{\mathcal{C}}$ ($\rho = -0.66$), which substantially exceeds that of other alternative measures. Among the list of competing variables, $1(Scheduled)_{\mathcal{C}}$, which has the next highest degree of correlation, may capture the fact that part of the variation comes from earnings announcements having drift while non-earnings announcements tend to have reversals. Other explanatory variables have relatively little correlation with drifts and reversals.

6 Conclusion

Our theory and empirics are motivated by the following intuition: if tail events play a major role in shaping investor beliefs, whether an investor underreacts or overreacts to news depends on how she associates it with past tail events. If investors react to news by drawing references to other news of the same category, our model predicts that the objective distribution of tail events within each category is predictive of over-and-underreaction: the fatter the tails, the greater the overreaction to news of that category. When applied to a comprehensive database of corporate news, our measure predicts the cross-section of over-and-underreaction across different news categories.

We view our current approach as a cautious first step in measuring how investors draw associations between different events. In reality, the category of a news announcement is just one of many features that drive associations between news announcements (Tversky, 1977; Bordalo et al., 2023a). For example, the magnitude and perceived significance of the news, as well as characteristics of the company beyond the announcement itself, such as its past performance, industry, and leadership, surely influence which announcements come to mind for investors. Discovering which features of a news announcement are salient, by leveraging text, surveys, and other richer data, is an important next step in understanding how investors react to news and ultimately how information is incorporated into asset prices.

References

Afrouzi, H., Kwon, S. Y., Landier, A., Ma, Y., and Thesmar, D. (2023). Overreaction in expectations: Evidence and theory. *The Quarterly Journal of Economics*, page qjad009.

Antweiler, W. and Frank, M. Z. (2006). Do us stock markets typically overreact to corporate news stories? *Available at SSRN 878091*.

Augenblick, N., Lazarus, E., and Thaler, M. (2021). Overinference from weak signals and underinference from strong signals. *arXiv preprint arXiv:2109.09871*.

Ba, C., Bohren, J. A., and Imas, A. (2022). Over-and underreaction to information. *Available at SSRN*.

Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies*, 21(2):785–818.

Barberis, N. (2013). The psychology of tail events: progress and challenges. *American Economic Review*, 103(3):611–616.

Barberis, N. and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066–2100.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, 49(3):307–343.

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

Bianchi, F., Ilut, C., and Saijo, H. (2024). Diagnostic business cycles. *Review of Economic Studies*, 91(1):129–162.

Bordalo, P., Conlon, J. J., Gennaioli, N., Kwon, S. Y., and Shleifer, A. (2023a). How people use statistics. Technical report, National Bureau of Economic Research.

Bordalo, P., Conlon, J. J., Gennaioli, N., Kwon, S. Y., and Shleifer, A. (2023b). Memory and probability. *The Quarterly Journal of Economics*, 138(1):265–311.

Bordalo, P., Gennaioli, N., La Porta, R., and Shleifer, A. (2019). Diagnostic expectations and stock returns. *The Journal of Finance*, 74(6):2839–2874.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2020a). Overreaction in macroeconomic expectations. *American Economic Review*.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic expectations and credit cycles. *The Journal of Finance*, 73(1):199–227.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020b). Memory, attention, and choice. *The Quarterly journal of economics*, 135(3):1399–1442.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14:521–544.

Bouchaud, J.-P., Krueger, P., Landier, A., and Thesmar, D. (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance*, 74(2):639–674.

Bybee, L., Kelly, B., and Su, Y. (2023). Narrative asset pricing: Interpretable systematic risk factors from news text. *The Review of Financial Studies*, 36(12):4759–4787.

Campbell, J. Y., Grossman, S. J., and Wang, J. (1993). Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics*, 108(4):905–939.

Chan, W. S. (2003). Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics*, 70(2):223–260.

Charles, C. (2022). Memory moves markets. *Available at SSRN 4019728*.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78.

Cutler, D. M., Poterba, J. M., and Summers, L. H. (1991). Speculative dynamics. *The Review of Economic Studies*, 58(3):529–546.

Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5):1461–1499.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance*, 53(6):1839–1885.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4):1605–1643.

d'Arienzo, D. (2020). Increasing overreaction and excess volatility of long rates. Working paper, Bocconi University.

Dávila, E. and Parlatore, C. (2018). Identifying price informativeness. Technical report, National Bureau of Economic Research.

De Bondt, W. F. and Thaler, R. (1985). Does the stock market overreact? *The Journal of finance*, 40(3):793–805.

DellaVigna, S. and Pollet, J. M. (2009). Investor inattention and friday earnings announcements. *The Journal of Finance*, 64(2):709–749.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of economics and statistics*, 80(4):549–560.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of finance*, 65(4):1237–1267.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.

Engelberg, J. (2008). Costly information processing: Evidence from earnings announcements. In *AFA 2009 San Francisco meetings paper*.

Engelberg, J., McLean, R. D., and Pontiff, J. (2018). Anomalies and news. *The Journal of Finance*, 73(5):1971–2001.

Engelberg, J., Sasseville, C., and Williams, J. (2012). Market madness? the case of mad money. *Management Science*, 58(2):351–364.

Engelberg, J. E. and Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, 66(1):67–97.

Enke, B., Schwerter, F., and Zimmermann, F. (2020). Associative memory and belief formation. Technical report, National Bureau of Economic Research.

Fedyk, A. (2018). Front page news: The effect of news positioning on financial markets. Technical report, Working paper.

Gabaix, X. (2009). Power laws in economics and finance. *Annu. Rev. Econ.*, 1(1):255–294.

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1):185–206.

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937):267–270.

Gabaix, X. and Ibragimov, R. (2011). Rank- 1/2: a simple way to improve the ols estimation of tail exponents. *Journal of Business & Economic Statistics*, 29(1):24–39.

Gennaioli, N., Shleifer, A., and Vishny, R. (2015). Neglected risks: The psychology of financial crises. *American Economic Review*, 105(5):310–314.

Giglio, S. and Kelly, B. (2018). Excess volatility: Beyond discount rates. *The Quarterly Journal of Economics*, 133(1):71–127.

Gormsen, N. J. and Lazarus, E. (2023). Duration-driven returns. *The Journal of Finance*, 78(3):1393–1447.

Griffin, D. and Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive psychology*, 24(3):411–435.

Hansen, L. P. (2007). Beliefs, doubts and learning: Valuing macroeconomic risk. *American Economic Review*, 97(2):1–30.

Hartzmark, S. M. and Shue, K. (2018). A tough act to follow: Contrast effects in financial markets. *The Journal of Finance*, 73(4):1567–1613.

Hertwig, R. and Erev, I. (2009). The description–experience gap in risky choice. *Trends in cognitive sciences*, 13(12):517–523.

Hong, H., Lim, T., and Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of finance*, 55(1):265–295.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance*, 54(6):2143–2184.

Huberman, G. and Regev, T. (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *The Journal of Finance*, 56(1):387–396.

Kahana, M. J. (2012). *Foundations of human memory*. OUP USA.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kandel, E. and Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4):831–872.

Kelly, B. and Jiang, H. (2014). Tail risk and asset prices. *The Review of Financial Studies*, 27(10):2841–2871.

Kozlowski, J., Veldkamp, L., and Venkateswaran, V. (2020). The tail that wags the economy: Beliefs and persistent stagnation. *Journal of Political Economy*, 128(8):2839–2879.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335.

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

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, 49(5):1541–1578.

L'Huillier, J.-P., Singh, S. R., and Yoo, D. (2023). Incorporating Diagnostic Expectations into the New Keynesian Framework. *The Review of Economic Studies*, page rdad101.

Malmendier, U. and Nagel, S. (2011). Depression babies: do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, 126(1):373–416.

Michaely, R., Rubin, A., and Vedrashko, A. (2016). Further evidence on the strategic timing of earnings news: Joint analysis of weekdays and times of day. *Journal of Accounting and Economics*, 62(1):24–45.

Murphy, K. M. and Topel, R. H. (2002). Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics*, 20(1):88–97.

Nagel, S. (2012). Evaporating liquidity. *The Review of Financial Studies*, 25(7):2005–2039.

Neuhierl, A., Scherbina, A., and Schiusene, B. (2013). Market reaction to corporate press releases. *Journal of Financial and Quantitative Analysis*, pages 1207–1240.

Nimark, K. P. (2014). Man-bites-dog business cycles. *American Economic Review*, 104(8):2320–2367.

Oh, S. and Wachter, J. A. (2018). Cross-sectional skewness. Technical report, National Bureau of Economic Research.

Orlik, A. and Veldkamp, L. (2014). Understanding uncertainty shocks and the role of black swans. Technical report, National bureau of economic research.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, pages 221–247.

Pástor, L. and Veronesi, P. (2009). Technological revolutions and stock prices. *American Economic Review*, 99(4):1451–1483.

Plerou, V., Gopikrishnan, P., Amaral, L. A. N., Meyer, M., and Stanley, H. E. (1999). Scaling of the distribution of price fluctuations of individual companies. *Physical review e*, 60(6):6519.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical association*, 89(428):1303–1313.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting review*, pages 289–315.

Solomon, D. H. (2012). Selective publicity and stock prices. *The Journal of Finance*, 67(2):599–638.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, 62(3):1139–1168.

Tetlock, P. C. (2014). Information transmission in finance. *Annu. Rev. Financ. Econ.*, 6(1):365–384.

Tversky, A. (1977). Features of similarity. *Psychological review*, 84(4):327.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, 90(4):293.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5:297–323.

Wachter, J. A. and Kahana, M. J. (2019). A retrieved-context theory of financial decisions. Technical report, National Bureau of Economic Research.

Wang, C. (2019). Under- and over-reaction in yield curve expectations. Working paper, Mendoza College of Business, University of Notre Dame.

7 Tables

Table 1: News Category Summary Statistics

Category	N	Mean	MeanAbs	StDev	P50	Turnover
Alliance	3326	0.15	1.29	2.19	0.09	0.97
Annual Meeting	6259	0.04	1.18	1.76	0.05	1.03
Board Changes	22173	0.06	1.35	2.14	0.08	1.08
CEO Change	1132	0.03	2.12	3.75	0.02	1.88
CFO Change	1429	-0.07	1.66	3.20	0.05	1.44
Client	20990	0.10	1.26	1.98	0.10	0.95
Credit Watch	1323	0.28	2.38	4.00	0.10	2.19
Dividend	1093	0.14	2.37	3.59	0.19	1.85
Downsize	2904	-0.00	1.58	2.76	0.03	1.42
Earnings	23573	0.13	2.59	4.01	0.12	1.97
Earnings Call	15912	0.26	3.00	4.40	0.23	2.20
Expansion	10326	0.06	1.35	2.19	0.06	1.19
Guidance Confirm	19528	0.08	2.66	4.16	0.11	2.12
Guidance Lower	1172	-1.54	3.47	5.64	-0.64	2.53
Guidance Raised	2791	0.90	2.71	4.09	0.56	1.93
Lawsuit	5669	0.03	1.36	2.27	0.04	1.26
M&A Closing	11294	0.10	1.29	1.98	0.08	1.01
M&A Rumor	3622	0.43	1.75	3.21	0.13	1.41
M&A Transaction	5143	0.33	1.71	3.01	0.16	1.32
No Events	2801102	0.07	1.26	2.12	0.07	0.99
Op. Result	1456	0.00	2.37	3.38	0.00	1.95
Product	23500	0.13	1.38	2.56	0.08	1.03
Seek Investment	7900	0.15	2.01	3.22	0.09	1.43
Structure Change	2596	0.10	1.30	2.10	0.14	1.16
Writeoff	3103	0.03	2.49	3.91	0.07	1.90

Note: Table 1 reports the summary statistics of announcement-day stock returns and trading volume for each news category in our dataset. Observations are at the news announcement level from January 1, 2011 to December 31, 2018, inclusive. The sample is all firms listed on the major US stock exchanges with at least $2 billion in market capitalization. N is the number of observations. Mean is the mean, Mean of Abs. is the mean of the absolute value, StDev is the standard deviation, and P50 is the median, of announcement-day returns, respectively. Mean, Mean of Abs., StDev, and P50 are all returns measured in percentage points. Turnover is the announcement-day trading volume defined as the number of shares traded times the share price divided by the total market capitalization times 100.

Table 2: Category-level Heterogeneity

Category	ζ	ζ s.e.	β	β s.e.
Alliance	0.42	0.007	-0.13	0.23
Annual Meeting	0.37	0.008	0.19	0.20
Board Changes	0.43	0.004	-0.07	0.08
CEO Change	0.52	0.025	-0.42	0.19
CFO Change	0.52	0.014	-0.31	0.31
Client	0.39	0.006	-0.18	0.09
Credit Watch	0.46	0.015	0.04	0.20
Dividend	0.37	0.018	-0.05	0.13
Downsize	0.43	0.012	0.12	0.20
Earnings	0.33	0.007	0.07	0.04
Earnings Call	0.32	0.008	0.06	0.04
Expansion	0.38	0.008	-0.09	0.10
Guidance Confirm	0.34	0.008	0.12	0.04
Guidance Lower	0.36	0.031	-0.09	0.12
Guidance Raised	0.35	0.023	0.06	0.09
Lawsuit	0.47	0.006	0.12	0.15
M&A Closing	0.38	0.006	0.08	0.12
M&A Rumor	0.57	0.005	-0.26	0.15
M&A Transaction	0.44	0.009	-0.25	0.11
Op. Result	0.32	0.031	0.21	0.24
Product	0.41	0.002	0.03	0.09
Seek Investment	0.37	0.011	0.12	0.07
Structure Change	0.43	0.015	-0.26	0.27
Writeoff	0.39	0.022	0.13	0.11

(a) Category-Level Extremeness and Drift/Reversal

Hypothesis	(1)	(2)	(3)	(4)
Jointly equal	1.79	3.35	1.77	1.87
p-value	0.0042	<0.0001	0.0053	0.0021
Jointly equal to 0	2.55	1.84	1.72	2.92
p-value	<0.0001	0.0030	0.0072	<0.0001
Industry FEs		X		X
Return Controls			X	X

(b) Testing for heterogeneity

Note: Table 2a reports the extremeness $\zeta_{\mathcal{C}}$ and $\beta_{\mathcal{C}}$, the drift/reversal estimates, for each news category. ζ is the extremeness, i.e., inverse power-law index estimated using eq. (8). ζ s.e. is the standard error for $\zeta_{\mathcal{C}}$ and is computed following Gabaix and Ibragimov (2011). $\beta_{\mathcal{C}}$ is the post-announcement drift/reversal beta estimated using eq. (11). $\beta_{\mathcal{C}}$ s.e. is the standard error for $\beta_{\mathcal{C}}$, two-way clustered at the firm and day levels. Table 2b reports the F-statistics for eq. (12). Jointly equal is the hypothesis that all $\beta_{\mathcal{C}}$'s are equal. Jointly equal to 0 is the hypothesis that all $\beta_{\mathcal{C}}$'s are equal to 0. Industry FEs are industry fixed effects and interaction terms of dummy variables for each industry with the announcement-day return. Return Controls are cubic polynomials of the announcement-day returns.

Table 3: Reversals and Extremeness

	(1)	(2)	(3)	(4)
Announcement-Day Return	0.61^{***}	0.54^{***}	0.77^{***}	0.64^{***}
	(0.18)	(0.17)	(0.25)	(0.23)
Announcement-Day Return × Extremeness	-1.63^{***}	-1.38^{***}	-2.07^{***}	-1.70^{***}
	(0.50)	(0.44)	(0.68)	(0.62)
Constant	0.02^{***}	-0.01^{***}	0.01^{***}	-0.01^{***}
	(0.003)	(0.002)	(0.003)	(0.002)
Time-Varying Tails	No	No	Yes	Yes
Return Benchmark	No	Yes	No	Yes
Observations	197,498	197,498	110,748	110,748

Note: Table 3 reports the estimates corresponding to eq. (13). Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Time-Varying Tails indicates whether Extremeness is computed over a rolling past five-year window (Yes) or over the entire sample (No). Return Benchmark indicates whether the Announcement-Day Return and dependent variable are excess returns relative to the S&P 500 (Yes) or raw returns (No). Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table 4: Summary of Robustness Exercises

	Specification	Coefficient	SE	Observations
1	Baseline	-1.63	0.5	197498
2	Earnings growth extremeness	-1.22	0.43	197498
3	Long-run return extremeness	-1.32	0.62	197498
4	2-day announcement window	-2.71	0.32	197498
5	Hour by day of week controls	-1.62	0.5	197498
6	Exclude Friday announcements	-1.85	0.54	174900
7	Non-parametric return controls (deciles)	-1.68	0.5	197498
8	Non-parametric return controls (polynomial)	-1.46	0.52	197498
9	Exclude smallest 25% of news	-1.61	0.5	148126
10	Exclude smallest 50% of news	-1.68	0.51	98750
11	Exclude top 0.01% of news	-1.39	0.52	197300
12	Exclude top 1% of news	-1.56	0.64	195522
13	Positive news only	-2.16	0.82	104783
14	Negative news only	-0.57	1.21	92715
15	Exclude earnings-overlapping categories	-1.62	0.52	152513
16	Exclude news within 30 days	-3.08	0.9	35879
17	Overlapping news controls	-1.38	0.44	197498
18	All news categories	-1.09	0.46	250852
19	All news categories with 1,000+ occurrences	-1.18	0.48	243966
20	Small-cap stocks	-1.12	0.39	226986
21	Post-announcement horizon k = 30-day	-0.81	0.3	197498
22	Post-announcement horizon k = 60-day	-1.18	0.4	197498
23	Exclude attrited firms	-1.64	0.5	196463
24	Exclude high-attrition news categories	-1.86	0.52	142651
25	Driscoll-Kraay standard errors	-1.63	0.51	197498
26	Block bootstrap	-0.92	0.33	197498

Note: Table 4 summarizes the robustness exercises for eq. (13). Coefficient is the main γ coefficient estimate. SE is the standard error. Observations is the number of observations in the corresponding estimates for each row. Row 1 is the baseline estimate. Rows 2 and 3 use alternative measures of extremeness. Row 4 uses ± 2 days as the announcement window. Row 5 includes hour-by-day-of-week controls. Row 6 excludes Friday announcements. Rows 7 and 8 include decile and cubic polynomial functions of announcement-day returns as controls. Rows 9-12 exclude the smallest 25% and 50%, and largest 0.01% and 1% of news by absolute announcement-day returns, respectively. Rows 13 and 14 report estimates on news with positive and negative announcement-day returns only, respectively. Rows 15 and 16 exclude news categories that overlap with other news. Row 17 includes controls for overlapping news as in eq. (14). Rows 18 and 19 reports the estimates for all news categories and categories with 1,000+ occurrences. Row 20 reports the results on small-cap firms. Rows 21 and 22 set the postannouncement horizons k as 30 and 60 days. Row 23 excludes announcements by firms that attrited during our sample. Row 24 excludes news categories that had above-average attrition rates. Row 25 computes Driscoll and Kraay (1998) standard errors. Row 26 uses a block bootstrap approach (Politis and Romano, 1994) on a full firm-day panel to account for compounded estimation errors.

Table 5: Volume and Extremeness

VARIABLES	(1) Turnover	(2) Turnover	(3) Turnover	(4) Turnover
Abs. Announcement-Day Return	0.25***	0.23***	0.19**	0.19**
	(0.081)	(0.079)	(0.084)	(0.082)
Abs. Announcement-Day Return × Extremeness	0.64***	0.76***	0.90***	0.91***
	(0.22)	(0.22)	(0.23)	(0.23)
Constant	0.0053***	0.0050***	0.0058***	0.0058***
	(0.00028)	(0.00031)	(0.00026)	(0.00028)
Observations	197,498	197,497	197,498	197,497
R-squared	0.371	0.394	0.395	0.412
Trading Day FEs	No	Yes	No	Yes
Return Benchmark	No	No	Yes	Yes

Note: Table 5 reports the estimates corresponding to eq. (15). Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the announcement-day turnover, defined as the volume of shares traded times the share price divided by the market capitalization. Abs. Announcement-Day Return is the absolute value of the announcement-day return and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following eq. (8). Trading Day FEs indicates whether the specification includes trading day fixed effects. Return Benchmark indicates whether the Announcement-Day Return and dependent variable are excess returns relative to the S&P 500 (Yes) or raw returns (No). Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table 6: Expectations and Extremeness

	(1)	(2)	(3)
Forecast Revision	0.83^{***}	0.71^{***}	0.41^{***}
	(0.12)	(0.11)	(0.10)
Forecast Revision × Extremeness	-0.80^{**}	-0.78^{**}	-0.43
	(0.35)	(0.31)	(0.28)
Constant	-0.002^{***}		
	(0.0002)		
Analyst FEs		X	
Firm FEs			X
Observations	949,419	949,419	949,419

Note: Table 6 reports the estimates corresponding to eq. (17). Observations are at the analyst by announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the analyst forecast error, defined as the realized earnings-per-share (EPS) minus the analyst forecast, divided by the stock price at the time of the forecast. Forecast Revision is the change in EPS forecast from before to after the news announcement, divided by the stock price before the announcement. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Analyst FEs and Firm FEs are analyst and firm fixed effect, respectively. Standard errors are two-way clustered at the analyst and firm levels. *** p<0.01, ** p<0.05, * p<0.10.

Table 7: Alternative Explanations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Announcement-Day Return	0.599***	0.846***	0.770***	0.596***	0.759***	0.691***	0.936***	0.595***	0.619**
	(0.180)	(0.278)	(0.221)	(0.182)	(0.218)	(0.211)	(0.287)	(0.180)	(0.241)
Announcement-Day Return x Extremeness	-1.533***	-1.922***	-1.732***	-1.429***	-1.818***	-1.594***	-2.276***	-1.488***	-1.559***
	(0.479)	(0.565)	(0.479)	(0.500)	(0.540)	(0.493)	(0.696)	(0.552)	(0.529)
Announcement-Day Return x Alternative Var.	1.473	-0.032	-2.440	-0.246	-0.095*	-0.167	-0.102	-0.000	-0.880
	(7.270)	(0.032)	(2.416)	(0.168)	(0.049)	(0.162)	(0.065)	(0.000)	(7.943)
Alternative Hypothesis	Mean	IQR	SD	Abs.	N	Largest	Scheduled	Media	Informativeness
Observations	197498	197498	197498	197498	197498	197498	197498	197498	197498

Note: Table 7 reports the estimates corresponding to eq. (18). Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the post-announcement return, i.e., cumulative return from day 1 to day 90 subsequent to the announcement. Announcement-Day Return is the return of the stock on the day of the announcement, expressed in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Alternative Var. is the explanatory regressor for alternative hypotheses, computed at the category level. Mean is the average announcement-day return. IQR is the interquartile range of the announcement-day return. SD is the standard deviation of the announcement-day return. Abs. is the absolute value of the announcement-day return. N is the number of total occurrences of the announcement. Largest is the average of the five largest absolute announcement-day returns for each news category. Scheduled is an indicator variable for news categories that occur on pre-announced schedules (operating results, earnings, guidances, dividends, earnings calls, and annual meetings). Media is the average number of news articles written about the firm on the day of the announcement. Informativeness is the measure of price informativeness computed following Dávila and Parlatore (2018). Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Online Appendix

Extreme Categories and Overreaction to News

Spencer Kwon and Johnny Tang

March 12, 2024

A Proofs and theoretical extensions

A.1 Proof of Proposition 1

We first compute the rational and diagnostic distribution (and the implied expected value) of v. Without loss of generality, let us focus on the case where the signal is positive ($s \geq 0$) (the converse case follows in a symmetric manner). Rational posteriors upon seeing $s \geq 0$ are given by the following expression:

$$\pi_{\mathcal{C}}(v|s_j) \propto v^{-(\zeta_{\mathcal{C},pos}^{-1}+2)} \text{ for } v \geq v_{1,\mathcal{C}} \equiv \max\{s_j, v_{0,\mathcal{C}}\}. \tag{19}$$

Under diagnostic expectations, this is contrasted with the "no news" distribution, which consists of the posterior distribution where a) there is no news (so the correct prior distribution is π^d), and the realization of the signal is equal to its ex ante expected value $s = 0$). That simply consists of:

$$\pi_d(v|s=0) \propto \pi^d(v) \cdot \frac{1}{v} \tag{20}$$

The diagnostic distribution is thus given by:

$$\pi_{\mathcal{C}}^{\theta}(v|s_j) \propto \pi_{\mathcal{C}}(v|s_j) \cdot \left(\frac{\pi_{\mathcal{C}}(v|s_j)}{\pi_d(v|s=0)}\right)^{\theta} \propto v^{-(\zeta_{\mathcal{C}}^{-1}+2)+\theta(\zeta_d^{-1}-\zeta_{\mathcal{C}}^{-1})} \text{ for } v \geq v_{1,\mathcal{C}} \tag{21}$$

One can compute the expected value of the rational and diagnostic posteriors:

$$\begin{aligned} E_{\mathcal{C}}[v|s_j] &= (1+\zeta_{\mathcal{C}})v_{1,\mathcal{C}} \\ E_{\mathcal{C}}^{\theta}[v|s_j] &= (1+\zeta_{\mathcal{C},\theta})v_{1,\mathcal{C}} \end{aligned} \tag{22}$$

where

$$\zeta_{\mathcal{C},\theta}^{-1} = \zeta_{\mathcal{C}}^{-1} + \theta(\zeta_{\mathcal{C}}^{-1} - \zeta_d^{-1}). \tag{23}$$

Simplifying, ψ, the ratio between diagnostic and rational expectations, is given by:

$$\frac{1+\zeta_{\mathcal{C},\theta}}{1+\zeta_{\mathcal{C}}} = \frac{1+(\zeta_{\mathcal{C}}^{-1}+\theta(\zeta_{\mathcal{C}}^{-1}-\zeta_d^{-1}))^{-1}}{1+\zeta_{\mathcal{C}}} = \frac{1+\zeta_{\mathcal{C}}+\theta(1-\frac{\zeta_{\mathcal{C}}}{\zeta_d})}{(1+\zeta_{\mathcal{C}})(1+\theta(1-\frac{\zeta_{\mathcal{C}}}{\zeta_d}))}, \tag{24}$$

as desired. From the expression, it is immediate that $\psi = 1$ when $\zeta_{\mathcal{C}} = \zeta_d$.

To prove the comparative static with respect to $\zeta_{\mathcal{C}}$, note that the derivative of the above with respect to $\zeta_{\mathcal{C}}$ has the same sign as:

$$\begin{aligned}\zeta_d(\zeta_{\mathcal{C}}^2+2\zeta_{\mathcal{C}}-\zeta_d)-\theta(\zeta_{\mathcal{C}}-\zeta_d)^2 &= \zeta_d\zeta_{\mathcal{C}}(1+\zeta_{\mathcal{C}})+\zeta_d(\zeta_{\mathcal{C}}-\zeta_d)-\theta(\zeta_{\mathcal{C}}-\zeta_d)^2 \\ &= \zeta_d\zeta_{\mathcal{C}}\Big(1+\zeta_{\mathcal{C}}+(\zeta_{\mathcal{C}}-\zeta_d)(\zeta_{\mathcal{C}}^{-1}+\theta(\zeta_{\mathcal{C}}^{-1}-\zeta_d^{-1}))\Big).\end{aligned} \tag{25}$$

First, if $\zeta_{\mathcal{C}} > \zeta_d$, the above is clearly positive. Second, if $\zeta_{\mathcal{C}} < \zeta_d$, recall that we have by assumption:

$$\zeta_{\mathcal{C}}^{-1}+\theta(\zeta_{\mathcal{C}}^{-1}-\zeta_d^{-1}) > 1 \tag{26}$$

(for the diagnostic posterior to have a well-defined mean). Thus, the following inequality holds:

$$\zeta_d\zeta_{\mathcal{C}}\Big(1+\zeta_{\mathcal{C}}+(\zeta_{\mathcal{C}}-\zeta_d)(\zeta_{\mathcal{C}}^{-1}+\theta(\zeta_{\mathcal{C}}^{-1}-\zeta_d^{-1}))\Big) > \zeta_d\zeta_{\mathcal{C}}\left(1+\zeta_{\mathcal{C}}+(\zeta_{\mathcal{C}}-\zeta_d)\right) > \zeta_d\zeta_{\mathcal{C}}(2\zeta_{\mathcal{C}}+1-\zeta_d) > 0, \tag{27}$$

where the last inequality is implied by $\zeta_d < 1$.

A.2 Proof of Corollary 1

Proof. This proceeds immediately from Proposition 1 and the following identity:

$$\beta^{CG} = \frac{Cov(FE_{i,1}, FR_{i,1})}{Var(FR_{i,1})} = \frac{Cov(v-E_{i,t}^{\theta}[v|\mathcal{C},s_j], E_{i,t}^{\theta}[v|\mathcal{C},s_j])}{Var[E_{i,t}^{\theta}[v|\mathcal{C},s_j]]} = 1-\psi^{-1}(\zeta_{\mathcal{C}},\zeta_d,\theta), \tag{28}$$

where the last equation uses the identity that rational forecast errors are unpredictable given i's time 1 information set: $Cov(v-E_{i,t}^{rat}[v|\mathcal{C},s_j], E_{i,t}^{rat}[v|\mathcal{C},s_j]) = 0$. □

A.3 Proof of Proposition 2

Deriving the algebraic expressions Market clearing implies that log prices settle to:

$$\begin{aligned} \log(p_1) &= \log(F_0) + \int E^{\theta}_{\mathcal{C}}[v|s_j]di \\ \log(p_2) &= \log(F_0) + \int E_{\mathcal{C}}[v|s_j]di \end{aligned} \tag{29}$$

respectively. By Proposition 1, the average expectations of v integrated across diagnostic investors are given by:

$$\int_{s=0}^{v} E^{\theta}_{\mathcal{C}}[\lambda|s]\cdot f(s|\lambda)ds = (1+\zeta_{\mathcal{C},\theta})\cdot\left(\int_0^{v_{0,\mathcal{C}}} \frac{1}{v}\cdot v_{0,\mathcal{C}}ds + \int_{v_{0,\mathcal{C}}}^{v} \frac{1}{v}s\cdot ds\right) = (1+\zeta_{\mathcal{C},\theta})\cdot\frac{v^2_{0,\mathcal{C}}+v^2}{2v}. \tag{30}$$

By the same token, the analogous expression for rational investors are given by:

$$(1+\zeta_{\mathcal{C}})\cdot\frac{v^2_{0,\mathcal{C}}+v^2}{2v}. \tag{31}$$

Thus, from the above expressions, one immediately obtains:

$$\log(p_1) = \log(F_0) + (1+\zeta_{\theta,\mathcal{C}})\cdot\eta_{\mathcal{C}}(v),\ \ \log(p_2) = \log(F_0) + (1+\zeta_{\mathcal{C}})\cdot\eta_{\mathcal{C}}(v), \tag{32}$$

with period 1 and 2 returns $r_i = \log(p_i) - \log(p_{i-1}), i \in \{1,2\}$ satisfying

$$r_2 = \beta^{ret}_{\mathcal{C}}\cdot r_1,\ \beta^{ret}_{\mathcal{C}} \equiv \frac{\zeta_{\mathcal{C}}-\zeta_{\mathcal{C},\theta}}{1+\zeta_{\mathcal{C},\theta}}. \tag{33}$$

Trading volume at $t=1$ in reaction to the news is given by:

$$Vol = \frac{1}{2}\kappa\cdot(1+\zeta_{\theta,\mathcal{C}})\cdot(1-\eta_{\mathcal{C}}(v))^2. \tag{34}$$

As for volume, note that we can compute the aggregate position of people who are long the asset:

$$Vol = \int_{s^*}^{v} D^{DE}(s, p_1) ds, \tag{35}$$

where s^* is the signal such that the subjective expectation of v equals the average expectations. or:

$$s^* = \frac{\log(p_1) - \log(F_0)}{1 + \zeta_{\mathcal{C},\theta}} \tag{36}$$

Plugging this, we obtain:

$$\begin{aligned} Vol &= \kappa \int_{s^*}^{v} ((1 + \zeta_{\mathcal{C},\theta})s - (\log(p_1) - \log(F_0)))\frac{1}{v} ds \\ &= \frac{1}{2}\kappa(1 + \zeta_{\mathcal{C},\theta})(v - s^*)^2 = \frac{1}{2}\kappa \cdot (1 + \zeta_{\mathcal{C},\theta}) \cdot \left(\frac{v^2 - v_{0,\mathcal{C}}^2}{2v}\right)^2 . \end{aligned} \tag{37}$$

Comparative statics The comparative statics with respect to $\zeta_{\mathcal{C}}$ for volume follows immediately from the above expression. Regarding the drift-reversal coefficient $\beta_{\mathcal{C}}^{ret}$, note:

$$\beta_{\mathcal{C}}^{ret} = \frac{\zeta_{\mathcal{C}} - \zeta_{\mathcal{C},\theta}}{1 + \zeta_{\mathcal{C},\theta}} = \frac{\theta(1 - \frac{\zeta_{\mathcal{C}}}{\zeta_d})}{1 + \zeta_{\mathcal{C}}^{-1} + \theta(\zeta_{\mathcal{C}}^{-1} - \zeta_d^{-1})} = \theta \cdot \frac{\zeta_{\mathcal{C}}(1 - \frac{\zeta_{\mathcal{C}}}{\zeta_d})}{1 + \zeta_{\mathcal{C}} + \theta(1 - \frac{\zeta_{\mathcal{C}}}{\zeta_d})}. \tag{38}$$

From the above expression, $\beta_{\mathcal{C}}^{ret} < 0$ if and only if $\zeta_{\mathcal{C}} > \zeta_d$. To show the comparative static with respect to $\zeta_{\mathcal{C}}$, one can take the derivative of the above expression and find that the above has the same sign as:

$$\begin{aligned} -\Big(\zeta_d(\zeta_{\mathcal{C}}^2 + 2\zeta_{\mathcal{C}} - \zeta_d) - \theta(\zeta_{\mathcal{C}} - \zeta_d)^2\Big) &= -\zeta_d\zeta_{\mathcal{C}}\Big((1 + \zeta_{\mathcal{C}}) + \zeta_{\mathcal{C}}^{-1}(\zeta_{\mathcal{C}} - \zeta_d) + \theta(\zeta_{\mathcal{C}}^{-1} - \zeta_d^{-1})(\zeta_{\mathcal{C}} - \zeta_d)\Big) \\ &= -\zeta_d\zeta_{\mathcal{C}}\Big((1 + \zeta_{\mathcal{C}}) + (\zeta_{\mathcal{C}}^{-1} + \theta(\zeta_{\mathcal{C}}^{-1} - \zeta_d^{-1}))(\zeta_{\mathcal{C}} - \zeta_d)\Big). \end{aligned} \tag{39}$$

The inequality follows in exactly the same way as above: if $\zeta_{\mathcal{C}} > \zeta_d$, the above term is clearly negative. If $\zeta_{\mathcal{C}} < \zeta_d$, the above term is bounded above by:

$$-\zeta_d\zeta_{\mathcal{C}}((1 + \zeta_{\mathcal{C}}) + \zeta_{\mathcal{C}} - \zeta_d) < 0, \tag{40}$$

so $\beta_{\mathcal{C}}^{ret}$ is decreasing in $\zeta_{\mathcal{C}}$, as desired.

A.4 Proof of Corollary 3

Proof. The result follows immediately from Proposition 2, and the fact that power-law distributions are closed under affine transformations: if $X \sim F^{power}(v_0, \zeta^{-1})$, $AX + B \sim F^{power}(Av_0 + B, \zeta^{-1})$. □

A.5 Model extension: asymmetric tails and ambiguous signals

We extend the model to allow for two additional features. First, we allow the default distribution to have asymmetric tails: this is to reflect the fact that the distribution of stock returns (even outside of news events) may have a fatter left-tail. Second, we allow for ambiguity in the interpretation of a news of a given category: an investor may believe for instance that a CEO change may be positive or negative news.

A.5.1 Set-up

Asymmetric tails Thus, we consider the default distribution of the form:

$$\pi^d(v) = \begin{cases} \pi_0(v) & \text{for } |v| < v_{0,d} \\ C_{pos} \cdot |v|^{-(\zeta_{pos}^{-1}+1)} & \text{for } v > v_{0,d} \\ C_{neg} \cdot |v|^{-(\zeta_{neg}^{-1}+1)} & \text{for } v < -v_{0,d}, \end{cases} \tag{41}$$

where C_{pos} and C_{neg} are adjusted such that the mean of v is 0. Similarly, upon the announcement of an event of category $\mathcal{C}$, the distribution of v is power-law with potentially asymmetric tails. In this case, however, we do not impose the expectation of v to be 0: a news of a given category may on average be positive or negative news.

$$\pi_{\mathcal{C}}(v) = \begin{cases} C_{\mathcal{C},pos} \cdot |v|^{-(\zeta_{\mathcal{C},pos}^{-1}+1)} & \text{for } v > v_{0,\mathcal{C}} \\ C_{\mathcal{C},neg} \cdot |v|^{-(\zeta_{\mathcal{C},neg}^{-1}+1)} & \text{for } v < -v_{0,\mathcal{C}} \end{cases} \tag{42}$$

As before, we assume $v_{0,\mathcal{C}} > v_{0,d}$.

Ambiguous signals As before, we assume a conjugate distribution $s_0 \sim Unif[0,v]$ for the distribution of noisy signals, but assume that with probability $q(|s_0|)$, where $q(|s_0|) \leq \frac{1}{2}$, the sign of the signal may be reversed. In other words, we assume:

$$s = s_0 \cdot (-1)^X, \tag{43}$$

where $s_0 \sim Unif[0,v]$ (or $Unif[v,0]$ for $v < 0$) and $X \sim Bern(q(|s_0|))$. In particular, we assume for simplicity that the probability of a sign flip only depends on the magnitude of the realized signal s_0, which nests the case of constant probability. Thus, a positive signal s may arise from a positive realization of v, or potentially a negative realization, with the signs reversed. The higher the q, the greater the ambiguity of the signal.

A.5.2 Rational and diagnostic posteriors

Without loss of generality, let us focus on the case where the signal itself is positive ($s \geq 0$) (the converse case follows in a symmetric manner). Rational posteriors upon seeing $s \geq 0$ are given by the following expression:

$$\pi_{\mathcal{C}}(v|s_j) \propto \begin{cases} (1-q(|s_j|)) \cdot C_{\mathcal{C},pos} \cdot v^{-(\zeta_{\mathcal{C},pos}^{-1}+2)} & \text{for } v \geq v_{1,\mathcal{C}} \equiv \max\{s_j, v_{0,\mathcal{C}}\}. \\ q(|s_j|) \cdot C_{\mathcal{C},neg} \cdot |v|^{-(\zeta_{\mathcal{C},neg}^{-1}+2)} & \text{for } v < -v_{1,\mathcal{C}} \end{cases} \tag{44}$$

Thus, unlike the baseline case, upon seeing a positive signal, the investor may interpret it as good or bad news.

Under diagnostic expectations, this is contrasted with the "no news" distribution, which consists of the posterior distribution where a) there is no news (so the correct prior distribution is π^d), and the realization of the signal is equal to its ex ante expected value $s = 0$). That simply consists of:

$$\pi_d(v|s=0) \propto \pi^d(v) \cdot \frac{1}{v} \tag{45}$$

The diagnostic distribution is thus given by:

$$\pi_{\mathcal{C}}^{\theta}(v|s_j) \propto \pi_{\mathcal{C}}(v|s_j) \cdot \left(\frac{\pi_{\mathcal{C}}(v|s_j)}{\pi_d(v|s=0)}\right)^{\theta} \propto \begin{cases} \frac{((1-q(|s_j|))C_{\mathcal{C},pos})^{\theta+1}}{C_{pos}^{\theta}} v^{-(\zeta_{\mathcal{C},pos}^{-1}+2)+\theta(\zeta_{pos}^{-1}-\zeta_{\mathcal{C},pos}^{-1})} & \text{for } v \geq v_{1,\mathcal{C}} \\ \frac{(q(|s_j|)C_{\mathcal{C},neg})^{\theta+1}}{C_{neg}^{\theta}} v^{-(\zeta_{\mathcal{C},neg}^{-1}+2)+\theta(\zeta_{neg}^{-1}-\zeta_{\mathcal{C},neg}^{-1})} & \text{for } v < -v_{1,\mathcal{C}} \end{cases} \tag{46}$$

One can compute the expected value of the rational and diagnostic posteriors:

$$E_{\mathcal{C}}[v|s_j] = \frac{(1-q(|s_j|))\frac{C_{\mathcal{C},pos}}{\zeta_{\mathcal{C},pos}^{-1}} v_{1,\mathcal{C}}^{-\zeta_{\mathcal{C},pos}^{-1}} - q(|s_j|)\frac{C_{\mathcal{C},neg}}{\zeta_{\mathcal{C},neg}^{-1}} v_{1,\mathcal{C}}^{-\zeta_{\mathcal{C},neg}^{-1}}}{(1-q(|s_j|))\frac{C_{\mathcal{C},pos}}{\zeta_{\mathcal{C},pos}^{-1}+1} v_{1,\mathcal{C}}^{-(\zeta_{\mathcal{C},pos}^{-1}+1)} + q(|s_j|)\frac{C_{\mathcal{C},neg}}{\zeta_{\mathcal{C},neg}^{-1}+1} v_{1,\mathcal{C}}^{-(\zeta_{\mathcal{C},neg}^{-1}+1)}}$$
$$E_{\mathcal{C}}^{\theta}[v|s_j] = \frac{(1-q(|s_j|))\frac{C_{\mathcal{C},pos}}{\zeta_{\mathcal{C},\theta,pos}^{-1}} v_{1,\mathcal{C}}^{-\zeta_{\mathcal{C},\theta,pos}^{-1}} \cdot M_{pos,\theta} - q(|s_j|)\frac{C_{\mathcal{C},neg}}{\zeta_{\mathcal{C},\theta,neg}^{-1}} v_{1,\mathcal{C}}^{-\zeta_{\mathcal{C},\theta,neg}^{-1}} \cdot M_{neg,\theta}}{(1-q(|s_j|))\frac{C_{\mathcal{C},pos}}{\zeta_{\mathcal{C},\theta,pos}^{-1}+1} v_{1,\mathcal{C}}^{-(\zeta_{\mathcal{C},\theta,pos}^{-1}+1)} \cdot M_{pos,\theta} + q(|s_j|)\frac{C_{\mathcal{C},neg}}{\zeta_{\mathcal{C},\theta,neg}^{-1}+1} v_{1,\mathcal{C}}^{-(\zeta_{\mathcal{C},\theta,neg}^{-1}+1)} \cdot M_{neg,\theta}} \tag{47}$$

where

$$M_{pos,\theta} = \left(\frac{(1-q(|s_j|))C_{\mathcal{C},pos}}{C_{pos}}\right)^{\theta}, M_{neg,\theta} = \left(\frac{q(|s_j|)C_{\mathcal{C},neg}}{C_{neg}}\right)^{\theta}$$
$$\zeta_{\mathcal{C},\theta,pos}^{-1} = \zeta_{\mathcal{C},pos}^{-1} + \theta(\zeta_{\mathcal{C},pos}^{-1} - \zeta_{pos}^{-1}), \zeta_{\mathcal{C},\theta,neg}^{-1} = \zeta_{\mathcal{C},neg}^{-1} + \theta(\zeta_{\mathcal{C},neg}^{-1} - \zeta_{neg}^{-1}) \tag{48}$$

A.5.3 Comparative statics: ambiguous signals

Let us first consider the minimum departure from the main model, where the signal itself can be ambiguous: $q(|s_j|) \neq 0$. Otherwise, assume that the distributions are symmetric: both the default distribution and the news category distribution has symmetric tails ($C_{pos} = C_{neg}$, $\zeta_{pos} = \zeta_{neg}$, $C_{\mathcal{C},pos} = C_{\mathcal{C},neg}$, $\zeta_{\mathcal{C},pos} = \zeta_{\mathcal{C},neg}$). In that case, one obtains that the core result in the main theory continues to hold:

Proposition 3. Under the symmetric case, for any specification of signal ambiguity $q(|s_j|)$, the distortion introduced by diagnostic expectations takes the following form:

$$E_{\mathcal{C}}^{\theta}[v|s_j] = \frac{1+\zeta_{\mathcal{C},\theta}}{1+\zeta_{\mathcal{C}}} \cdot \frac{\frac{(1-q(|s_j|))^{\theta+1} - q(|s_j|)^{\theta+1}}{(1-q(|s_j|))^{\theta+1} + q(|s_j|)^{\theta+1}}}{1-2q(|s_j|)} E_{\mathcal{C}}[v|s_j]. \tag{49}$$

As is the unambiguous case, the degree of overreaction is increasing for more extreme distributions ($\zeta_{\mathcal{C}} \uparrow$) and the diagnosticity parameter θ.

Proof. The proof follows immediately from the earlier comparative static proof and the fact that the second multiplicative term of the RHS is independent of $\zeta_{\mathcal{C}}$. $\square$

Relative to the unambiguous case, for $q(s_j) \leq \frac{1}{2}$, the second term is greater than 1: there is a general force for greater overreaction independent of $\zeta_{\mathcal{C}}$. The intuition is simple: under ambiguous signals, a positive realization of s is more diagnostic of positive outcomes of fundamentals than negative outcomes. This implies that in response to a positive signal, the diagnostic investor is overly optimistic, which further boosts the degree of overreaction. Even under this extension, however, our core prediction remains unchanged: greater overreaction for more extreme categories.

A.5.4 Comparative statics: asymmetric tails

Next, consider the other variation: we temporarily shut off signal ambiguity ($q = 0$), and allow for both the reference distribution and the news category tails to be asymmetric. In that case, one obtains the following result:

Proposition 4. The distortion introduced by diagnostic expectations takes the following form:

$$E^{\theta}_{\mathcal{C}}[v|s_j] = \begin{cases} \frac{1+\zeta_{\mathcal{C},\theta,pos}}{1+\zeta_{\mathcal{C},pos}} E_{\mathcal{C}}[v|s_j] & \text{for } s_j > 0 \\ \frac{1+\zeta_{\mathcal{C},\theta,neg}}{1+\zeta_{\mathcal{C},neg}} E_{\mathcal{C}}[v|s_j] & \text{for } s_j < 0. \end{cases} \tag{50}$$

The degree of overreaction for each case is increasing in the relative extremeness $\zeta_{\mathcal{C},pos}/\zeta_{pos}$ and $\zeta_{\mathcal{C},neg}/\zeta_{neg}$.

In particular, the above proposition makes it clear that if the default distribution exhibits a left skew (because of, for instance, broad stock market downturns that are independent of any corporate news events), this implies a greater degree of underreaction for negative news events than positive news events. Intuitively, even if a news event makes a negative tail outcome likely, it is not disproportionately likely relative to what the investor already expects from a default reference distribution.

B Robustness Exercises

B.1 Sorted portfolios

A classic alternative measure of over-or-underreaction to news is given by returns to sorted portfolios. For each news category, one can form a version of a sorted portfolio based on announcement-day returns. Pooling across all announcements in category $\mathcal{C}$, we divide the announcements into deciles by announcement-day returns. We denote:

$$r^{\mathcal{C}}_{sorted,k} \equiv (\bar{r}^{10}_{\mathcal{C},k} - \bar{r}^{1}_{\mathcal{C},k}), \tag{51}$$

where $\bar{r}^{i}_{\mathcal{C},k}$ are average k-day post-announcement returns of announcements in the i-th decile of announcement-day returns pooling all announcements in $\mathcal{C}$.

Similarly, one can pool all announcements in news category $\mathcal{C} \in \mathbb{C}^{+}$, the news categories for which we have obtained a positive drift-reversal coefficient, and announcements in news category $\mathcal{C} \in \mathbb{C}^{-}$, the news categories for which we have obtained a negative drift-reversal coefficient, to form two sorted portfolios:

$$r^{+}_{sorted,k} \equiv (\bar{r}^{10}_{+,k} - \bar{r}^{1}_{+,k}),\ r^{-}_{sorted,k} \equiv (\bar{r}^{10}_{-,k} - \bar{r}^{1}_{-,k}), \tag{52}$$

where $\bar{r}^{i}_{+,k}, \bar{r}^{i}_{-,k}$ are average k-day post-announcement returns of announcements in the i-th decile of announcement-day returns pooling all announcements in $\mathbb{C}^{+}$ and $\mathbb{C}^{-}$.

These returns are not returns to a sorted portfolio in a strict sense, as the news announcements do not occur on the same date. However, it is meant to be an exercise to illustrate the economic magnitude of return predictability across categories. Figure A2 plots both the relationship between $\beta_{\mathcal{C}}$ and $r^{\mathcal{C}}_{sorted,k}$ as well as the cumulative returns for $r^{+}_{sorted,k}$ and $r^{-}_{sorted,k}$. We find that our drift reversal coefficients correspond tightly to the returns of the equivalent sorted portfolio, and that the economic magnitude of drifts and reversals that we find are comparable: a long-short portfolio for news categories with drift gains 61 basis points (bps) over 90 days, while the same strategy for news categories with reversals loses 111 bps in the same period.

B.2 Robustness in measuring extremeness

We show that our measurement of category extremeness, $\zeta_{\mathcal{C}}$, is robust to various specifications. First, and consistent with Prediction 4, Table A1 reports the pairwise correlation between the tail measures, where we use the announcement-day return (our main specification), 100-day cumulative returns (which should be a less biased estimate of fundamentals), and k year realized earnings growth, for $k = 1, 2, 3, 4, 5$. We find a high level of correlation between all of the tail measures at the news category level. In particular, as shown in Figure A3, the extremeness computed using announcement returns and realized earnings growth (at horizon $k = 2$ for the figure) are tightly correlated.

After establishing Prediction 4, we now re-run our main analysis with our alternative measures of category extremeness. Our main specification, using tails estimated from announcement-day returns, is subject to two concerns. First, announcement-day returns can be biased estimates of true fundamentals, and may be affected by other factors, such as time-varying investor risk aversion or market microstructure. Second, if the mispricing driven by investor overreaction itself has fat-tails, one may potentially have a reverse causality concern: categories that are associated with greater overreaction then may have fatter-tailed distribution of announcement-day returns, without any differences in the tails of the underlying fundamentals.

To address both concerns, we repeat our analysis with the category extremeness estimated based on long-run (100-day) returns, as well as realized earnings growth (for firms with EPS greater than \$0.1 at the time of the announcement) for $k = 1, 2, 3, 4, 5$ years. Table A2 shows the results. We estimate a similar negative γ coefficient for each of the alternative measures.

B.3 Accounting for announcement timing

Expanded announcement window We also account for potential issues around the announcement timing. The first issue is whether there may be measurement errors in the true announcement timing. Alternatively, some announcements may have pre-announcement leakages, with some investors being aware of the announcement before the measured announcement timing. To address these issues, we use an expanded announcement window, of ±2 days before and after the news announcement date, as alternative definitions of the announcement returns. Specifically, for a news announcement recorded on date t (either during trading hours on date t or after trading hours on date $t-1$), we consider the announcement window from the close on date $t-3$ to the close on date $t+2$. Concretely, the expanded announcement windows modify our main regression to:

$$r_{i,t+3,t+k,\mathcal{C}} = \alpha + \beta_0 \cdot r_{i,t-2,t+2,\mathcal{C}} + \gamma \cdot \zeta_{\mathcal{C},t} \times r_{i,t-2,t+2,\mathcal{C}} + \epsilon_{i,t,\mathcal{C}}. \tag{53}$$

Observations are at the announcement level. The expanded announcement window is reflected in the announcement return being $r_{i,t-2,t+3,\mathcal{C}}$, i.e., the return from the close of date $t-3$ to the close of date $t+2$. We also modify the dependent variable to be $r_{c,t+3,t+k,\mathcal{C}}$, which is the cumulative post-announcement return from the close on date $t+2$ to the close on date $t+k$. As in the case of main regression in the draft, γ, the coefficient of interest, captures the correlation between the future drift/reversal and the extremeness of each news category. Table A3a reports the results corresponding to eq. (53), also including a version where we use a ±1 day expanded window: our estimates of γ remain negative and economically and statistically significant.

Strategic timing Another potential source of bias is that firms may strategically choose when to release their announcements. To control for strategic announcements, we employ two sets of empirical strategies: (1) we control for the release timing of the announcements, and (2) exclude potentially strategically-released news announcements.

First, we directly control for the release time of the announcement by including indicators for the (1) hour of the day (e.g., 2pm ET), (2) day of week (e.g., Tuesday), and (3) the interaction of hour of the day by day of week (e.g., Tuesday 2pm ET) in our main empirical specification as follows:

$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \sum_{d\in D} \beta_d \cdot r_{i,t,\mathcal{C}} \cdot 1(\text{Time} = \text{d}) + \gamma \cdot \zeta_{\mathcal{C},t} \times r_{i,t,\mathcal{C}} + \mu_d + \epsilon_{i,t,\mathcal{C}}. \tag{54}$$

which is our main specification with two changes. First, we include a set of interaction terms of indicators for whether the announcement time was at a particular hour of day by day of week $1(\text{Time} = \text{d})$ multiplied by the announcement-day return $r_{i,t,\mathcal{C}}$. This accounts for any variation in post-announcement drift or reversals due to different release times, such as if investors pay less attention on Friday afternoons (and hence be associated with greater drift). Second, we include fixed effects for the hour of day by day of week, μ_d, which account for differences in the levels of the returns of the news at different release times, for example if firms strategically release more negative news on Friday afternoons.

Second, the literature on strategic releases of announcement timings finds that strategic releases are primarily concentrated on Friday evenings (Michaely et al., 2016). As such, we also test whether our main result holds among announcements that are less likely to be affected by managers' strategic release decisions. We conduct our analysis by excluding three sets of news: (1) announcements made on Fridays, (2) announcements made on Friday evenings, and (3) announcements made by firms that most frequently (in the top decile) release on Friday evenings.

Tables A3b and A3c report the results. Table A3b shows that our results are largely unchanged excluding after-hours announcements. Columns (1) through (3) of Table A3c report the results controlling for the release time of the news, and Columns (4)-(6) report the results excluding the three sets of announcements likely to be associated with strategic releases. In all specifications, our main coefficient of interest γ remains qualitatively and quantitatively unchanged.

B.4 Accounting for the size and sign of the news

Flexibly controlling for announcement returns In this section, we control for other characteristics of the news that may influence investor reaction. First, investors may generally overreact to large news, or perhaps large movements in prices may just automatically lead to reversals (Chan, 2003). Second, investors may also react differently to positive vs negative news (Hong et al., 2000). To account for any relationship between announcement returns and post-announcement returns unrelated to the news category, we add as controls non-linear functions of announcement-day returns:

$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \beta_0 \cdot r_{i,t,\mathcal{C}} + \gamma \cdot \zeta_{\mathcal{C},t} \times r_{i,t,\mathcal{C}} + f(r_{i,t,\mathcal{C}}) + \epsilon_{i,t,\mathcal{C}}. \tag{55}$$

where $f(r_{i,t,\mathcal{C}})$ is either (a) a cubic polynomial or (b) $f(r_{i,t}) = \sum_{d=1}^{10} \gamma_d \cdot 1(r_{i,t} \in \Delta_d)$, where $1(r_{i,t} \in \Delta_d)$ is an indicator variable for whether $r_{i,t}$ is in the d-th decile of all announcement-day returns. Table A4 reports the results, with the polynomial control in column (1) and the decile control in column (2), showing that the results are again consistent regardless of the exact specification used. Figure A4 plots the estimated function f for the cubic polynomial specification. We find suggestive evidence that consistent with the literature, there tends to be more short-term drift for negative news (and reversals for positive news). Restricted to positive news, the estimated function is also weakly concave, suggesting a greater degree of reversals to large announcement-day return.

Where are our results concentrated? While we have shown that our results are robust controlling for any non-linear relationship between announcement returns and future returns, we also explore where our findings are concentrated in the data. First, one concern may be that our analysis may be dominantly driven by reaction to small, inconsequential news announcements. In Table A5a, we repeat our analysis sequentially removing news announcements with the smallest 25%, 50%, and 75% of absolute announcement-day returns. Our results continue to hold robustly across each subsample, with suggestive evidence that the category extremeness is most strongly predictive for announcements with large announcement-day returns.

Conversely, another concern may be that our results are driven by outliers. To reiterate, even if the main predictor of category extremeness is derived from the distribution of its tail events, our theory and empirics predict biases for *all* news of that given category. To ensure that our results hold for the broad sample, Table A5b repeats our analysis excluding the top 0.01%, 0.05%, 0.1%, 0.5%, and 1% of news events; we still estimate our category extremeness based partially on these events ($\zeta_{\mathcal{C}}$), but exclude them from the sample in our main return predictability regression. We similarly find that our results are robust to excluding outlier events.

Lastly, we also repeat our entire analysis by separating positive and negative news events. Table A5c reports the results. We find that our findings are most pronounced for positive news, and the same direction, but smaller in magnitude and statistically insignificant amongst negative news.

B.5 Accounting for overlapping announcements

In this section, we address the potential concern that the predictability for a given news category can be due to systematic overlap with another category. Suppose that a news category $\mathcal{C}'$ is systematically preceded by an announcement from category $\mathcal{C}$. Then, the return predictability of category $\mathcal{C}'$ can reflect instead return predictability of $\mathcal{C}$. For example, if there is continued drift for one news category, a news category that systematically follows that category may also be mechanically associated with drift.

We address this concern in two ways. First, we exclude announcements for which overlap may be a concern: there may be a closely preceding announcement. In our most stringent criterion, we remove all announcements that had any other news occur to the firm in the 30 days prior to the announcement. This approach is more demanding on the data and results in a smaller sample of 35,000 announcements. Alternatively, given that earnings announcements are the most regular announcements, and hence the most likely to generate systematic overlap with other news categories, we exclude all news categories that co-occur with earnings announcements more than 50% of the time. Table A6a reports the results for both exercises. In both cases, we find that our main result that more extreme categories of news have more post-announcement reversals, holds, as indicated by the negative coefficient estimate on γ.

Second, instead of removing samples for which overlaps can be a concern, we instead directly account for any return predictability driven by the presence of past news announcements. This approach allows us to use all news announcements and the full extent of the data. Formally, we estimate on the full data set the following regression:

$$r_{i,t+1,t+k,\mathcal{C}} = \alpha + \beta_0 \cdot r_{i,t,\mathcal{C}} + \gamma \cdot \zeta_{\mathcal{C},t} \times r_{i,t,\mathcal{C}} + \sum_{\mathcal{C}' \in \mathbf{C}} \sum_{h=1}^{90} \theta_{\mathcal{C}',h} \cdot I(i,\mathcal{C}',t-h) \cdot r_{i,t-h,\mathcal{C}'} + \epsilon_{i,t,\mathcal{C}}, \tag{56}$$

where $I(i,\mathcal{C},t)$ is an indicator for whether firm i has experienced an announcement of category $\mathcal{C}$ on date t. Equation 56 augments our main specification by also accounting for the component of returns that may be due to past news. $\theta_{\mathcal{C}',h}$ accounts for the component of the return predictability due to news announcements of category $\mathcal{C}'$ that occurred h days prior to the current announcement.

While our main specification controls for the impact of past news events, we also consider the case where we estimate equation 56 controlling for future news events: $h = -90, -89, ..., -1, 1, 2, ...90$. This is to address concerns that part of the return predictability may be driven by $r_{i,t,\mathcal{C}}$ anticipating the arrival of future news (for example, a bad earnings announcement can be predictive of future CEO firings). In that case, the regression is no longer predictive: γ instead measures the degree of predictability holding fixed the realization of future news, with a negative γ suggesting greater reversals even after controlling for the arrival of future announcements. Table A6b reports the estimates for our standard specification (only using past events), accounting for future announcements, and both. Across all specifications, we find a significant γ of a similar magnitude to our baseline estimates.

B.6 Sample selection

Table A7 lists all of the news categories that have occurred at least 1000 times in our sample, as well as our criteria for including a given news category in our final sample. We have excluded news categories that we judged to be not directly pertinent to the fundamentals of the company. These include purely administrative announcements (e.g. news about earnings release date), index inclusion announcements, as well as capital structure announcements, such as announcing buybacks transactions and their closing.

There are two types of concerns about our sample selection. First may be that our results are not robust to category exclusion: for example, our results may depend on a particular influential category. Alternatively, one may disagree with a particular news category in our final sample. To ensure that our results are robust to the exclusion of any particular news category, we report the correlation between the category level extremeness $\zeta_{\mathcal{C}}$ and the drift-reversal coefficient $\beta_{\mathcal{C}}$, where we exclude one of each news category in our sample. Table A8 reports the results: the category extremeness and drift-reversal coefficients are tightly correlated across all subsamples.

Second, another concern may be that we have unfairly excluded certain news categories. For example, announcements of dividends, buybacks, IPOs, SEOs, and other capital structure news can be informative about fundamentals. Table A9 repeats our analysis including additional news categories. Our results remain significant as we add back the major captial structure news that we have excluded, as well as including all news categories with more than 1000 occurrences, and even all news categories in our sample. The last result, however, is not particularly informative: the additional announcements form a negligible fraction of the total sample size.

B.7 Other robustness tests

Small cap stocks We have also previously removed small-cap stocks (stocks with lower than \$2B market capitalization) to account for liquidity issues. We show that our results do not depend on this choice: Table A10 shows that our results are unchanged even when including small-cap stocks.

Attrition Another potential concern is that of attrition: certain firms in our sample may drop out after the announcement. This may induce selection bias in our measure of announcement returns, especially for negative categories of news (e.g. a downward guidance event). To ensure that attrition is not confounding our results, we repeat our analysis on the sample of firms that are non-attrited, i.e. present until the end of the sample. Furthermore, in case some news categories are more associated with attrition, we also repeat our analysis dropping all news categories with below-mean rate of attrition among firms experiencing a news in that category. Table A11a shows that our results are largely unchanged for these subsamples. Furthermore, given that our long-run earnings growth measure may also suffer from attrition results, Table A11b also shows that our main results continue to hold when we alternatively use 30, 100, and 250 cumulative returns to estimate category extremeness.

Standard errors Finally, we also use more conservative standard errors. First, given that our announcement windows may overlap in our sample, our errors can be not only correlated at the firm and day level (which our clustered errors account for), but also at the same firm that may have overlapping events. To account for the overlaps in the windows and autocorrelated errors, we use Driscoll-Kraay standard errors (Driscoll and Kraay, 1998). Table A12 reports the results: our estimates remain statistically significant. Another concern is that our results may be subject to compounded estimation error: our standard errors need to account for errors in our estimates of $\zeta_{\mathcal{C}}$ (Pagan, 1984; Murphy and Topel, 2002). To address this, we use a block bootstrap procedure (Politis and Romano, 1994) on a full firm-day panel and estimate the γ coefficient on each bootstrap sample following equation (13). Figure A5 presents the density plot of the γ estimates from the bootstrap samples. The mean is -0.92 and the empirical 95% interval is $[-1.58, -0.29]$, which is broadly in line with the estimated γ coefficients reported in Table 3, with the point estimate being smaller but the statistical inference largely unchanged after accounting for generated regressors.

C Details for measuring price informativeness

Dávila and Parlatore (2018) construct a measure of relative price informativeness based on the difference in the R^2 of the following set of regressions:

$$M_{i,t} = \alpha + \beta_E \cdot E_{i,t} + \epsilon_{i,t},$$
$$M_{i,t} = \alpha' + \beta_{E,1} \cdot E_{i,t} + \beta_{E,2} \cdot E_{i,t+1} + \epsilon'_{i,t},$$

where $M_{i,t}$ is the valuation of firm i at time t, $E_{i,t}$ is realized fundamentals (total earnings) of firm i at time t, and $E_{i,t+1}$ is future fundamentals (total earnings in the subsequent year) for firm i at time t. Denote the R^2 of each regression as R_0^2 and R_1^2 respectively. The measure of relative price informativeness is then given by: $\tau = \frac{R_1^2 - R_0^2}{1 - R_1^2}$. Intuitively, the more informative prices are of future fundamentals, a larger part of current prices should reflect innovations to future earnings, $E_{i,t+1}$, and hence a greater R_1^2. Lastly, given the measure τ, it can be transformed into the Kalman gain measure, $\kappa = \frac{\tau}{1+\tau}$.

For our application, we are interested in measuring the informativeness of news from a given category. For each announcement $n_{i,t_a,\mathcal{C}}$ (an announcement of category $\mathcal{C}$ for firm i that occurred at date t_a), denote $E_{i,t_{pre}}$ as the EPS (earnings-per-share) of firm i announced at date t_{pre}, which is the latest annual earnings announcement for firm i that is at least 180 days before t_a. Similarly, denote $E_{i,t_{post}}$ as the EPS (earnings-per-share) of firm i announced at date t_{post}, which is the earliest annual earnings announcement for firm i that is at least 180 days after t_a. The reason we ensure that there are at least 180 days between t_a and the earnings releases is to have a fair comparison between categories that occur systematically close to earnings announcements and those that do not.

For $M_{i,t}$, the valuation of firm i at time t, we define M_{i,t_a-10} and M_{i,t_a+10} as the price per share of firm i 10 days before and after t_a. With these measures, we run two versions of the above price informativeness regression. First, we run:

$$M_{i,t_a-10} = \alpha + \beta_E^{pre} \cdot E_{i,t_{pre}} + \epsilon_{i,t},$$
$$M_{i,t_a-10} = \alpha' + \beta_{E,1}^{pre} \cdot E_{i,t_{pre}} + \beta_{E,2}^{pre} \cdot E_{i,t_{post}} + \epsilon'_{i,t}.$$

Denote $R^2_{0,pre}$ and $R^2_{1,pre}$ as the R^2 of the two regression specifications. Then, the price informativeness 10-days before the announcement is given by $\tau_{pre} = \frac{R^2_{1,pre} - R^2_{0,pre}}{1 - R^2_{1,pre}}$. To measure the price informativeness 10 days after the announcement, we similarly run:

$$M_{i,t_a+10} = \alpha + \beta_E^{post} \cdot E_{i,t_{pre}} + \epsilon_{i,t},$$
$$M_{i,t_a+10} = \alpha' + \beta_{E,1}^{post} \cdot E_{i,t_{pre}} + \beta_{E,2}^{post} \cdot E_{i,t_{post}} + \epsilon'_{i,t},$$

where we denote $R^2_{0,post}$ and $R^2_{1,post}$ as the R^2 of the two regression specifications. The price informativeness 10-days after the announcement is given by $\tau_{post} = \frac{R^2_{1,post} - R^2_{0,post}}{1 - R^2_{1,post}}$.

For both regressions, we pool all announcements of category $\mathcal{C}$, winsorize at the 10% level, and exclude announcements for which we do not have $E_{i,t_{pre}}$ or $E_{i,t_{post}}$ (i.e. earnings announcements that are at least 180 days away from t_a).

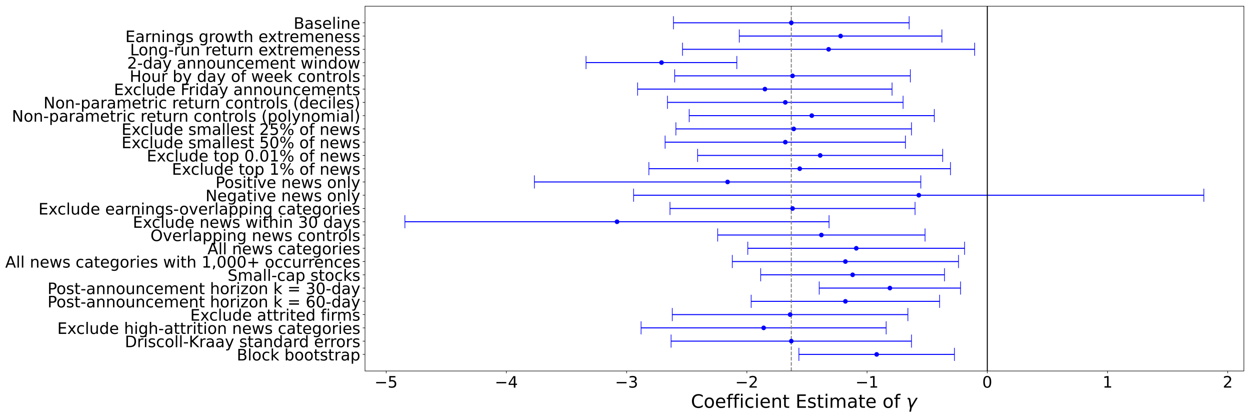
Our measure of the informativeness of news in category $\mathcal{C}$ is given by how much the announcements result in an increase in price informativeness:

$$\kappa_{\mathcal{C}} \equiv \kappa_{\mathcal{C},post} - \kappa_{\mathcal{C},pre},$$

where $\kappa_{\mathcal{C},post} = \frac{\tau_{post}}{1+\tau_{post}}$ and $\kappa_{\mathcal{C},pre} = \frac{\tau_{pre}}{1+\tau_{pre}}$ are the transformations of τ_{post} and τ_{pre} into the equivalent Kalman gain. Intuitively, $\kappa_{\mathcal{C}}$ reflects the increase in the price informativeness after announcements of category $\mathcal{C}$: the more informative a news from a given category, the greater prices should be predictive of future fundamentals after the announcement.

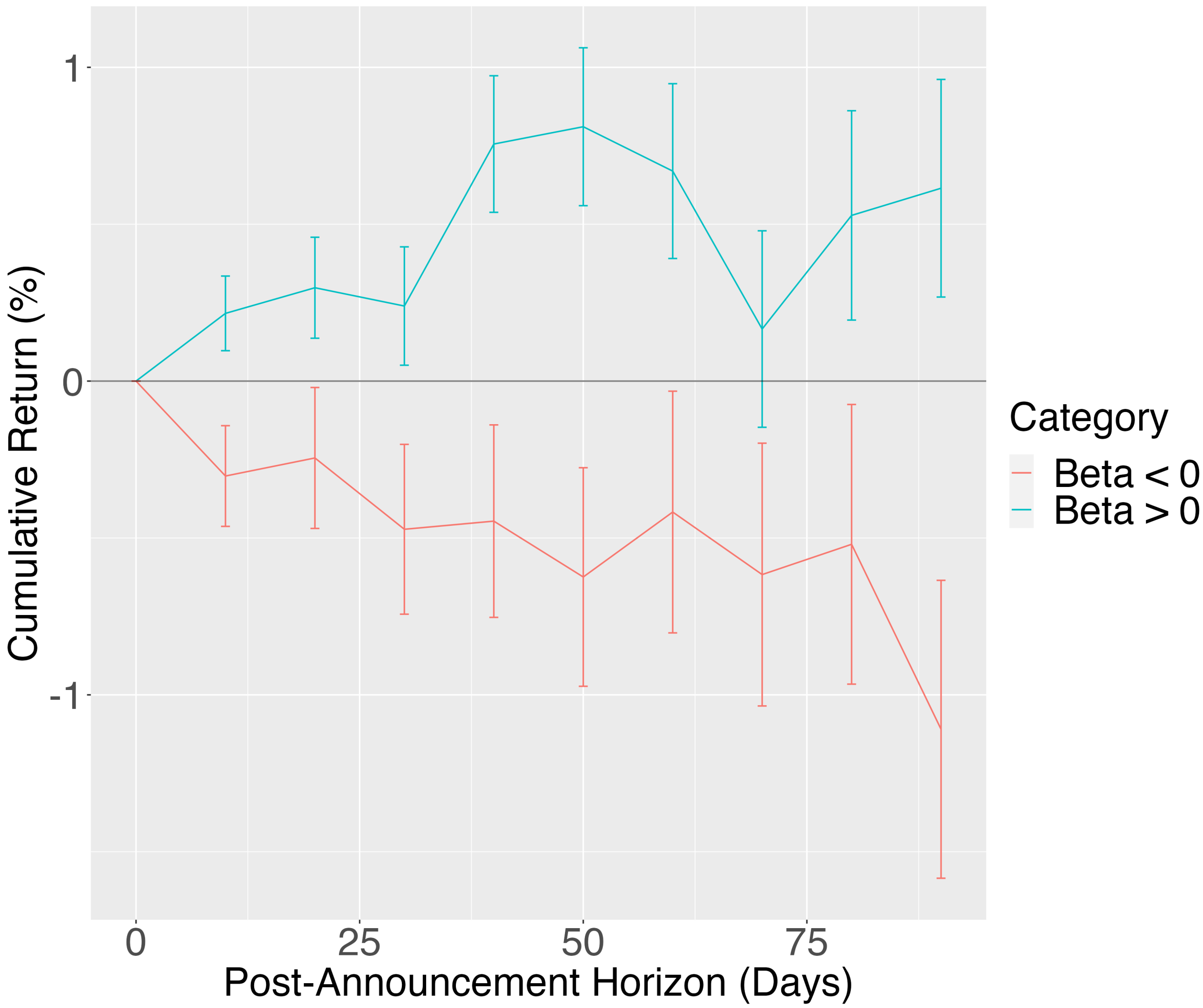
D Additional Tables and Figures

Figure A1: All robustness tests, summary figure



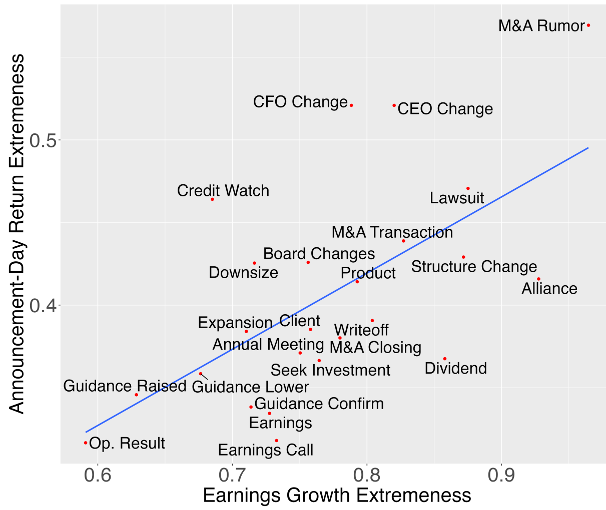
Note: Figure A1 plots the point estimates and 95% confidence intervals of our main coefficient of interest, γ, across each robustness exercise in Table 4. Row 1 is the baseline estimate. Rows 2 and 3 use alternative measures of extremeness. Row 4 uses ± 2 days as the announcement window. Row 5 includes hour-by-day-of-week controls. Row 6 excludes Friday announcements. Rows 7 and 8 include decile and cubic polynomial functions of announcement-day returns as controls. Rows 9-12 exclude the smallest 25% and 50%, and largest 0.01% and 1% of news by absolute announcement-day returns, respectively. Rows 13 and 14 report estimates on news with positive and negative announcement-day returns only, respectively. Rows 15 and 16 exclude news categories that overlap with other news. Row 17 includes controls for overlapping news as in eq. (14). Rows 18 and 19 reports the estimates for all news categories and categories with 1,000+ occurrences. Row 20 reports the results on small-cap firms. Rows 21 and 22 set the post announcement horizons k as 30 and 60 days. Row 23 excludes announcements by firms that attrited during our sample. Row 24 excludes news categories that had above-average attrition rates. Row 25 computes Driscoll and Kraay (1998) standard errors. Row 26 uses a block bootstrap approach (Politis and Romano, 1994) on a full firm-day panel to account for compounded estimation errors.

Figure A2: Magnitude of drift vs reversals: sorted portfolios

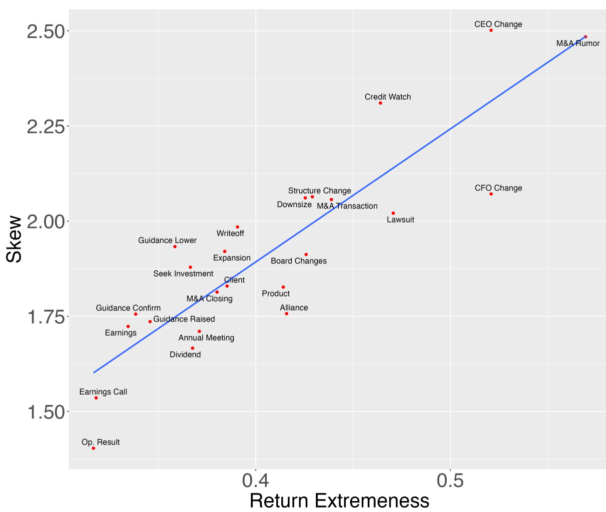


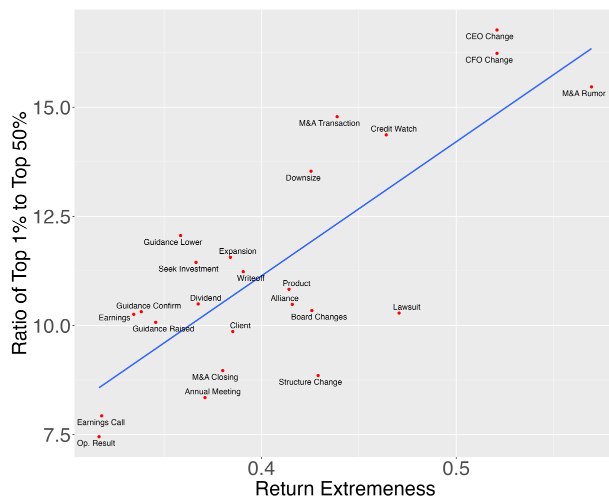
Note: Figure A2 plots the cumulative returns to winner-minus-loser strategies for sorted portfolios created on news categories with drifts (blue), i.e., $\beta_C^{ret} > 0$, and news categories with reversals (red), i.e., $\beta_C^{ret} < 0$. The stocks are sorted based on the announcement-day returns into ten equally-sized portfolios. The winner-minus-loser strategy buys the portfolio with the highest announcement-day returns and shorts the portfolio with the lowest announcement-day returns. 95% confidence intervals are reported in the vertical bars corresponding to each point estimate.

Figure A3: Measuring extremeness in multiple ways



(a) Tails using EPS growth



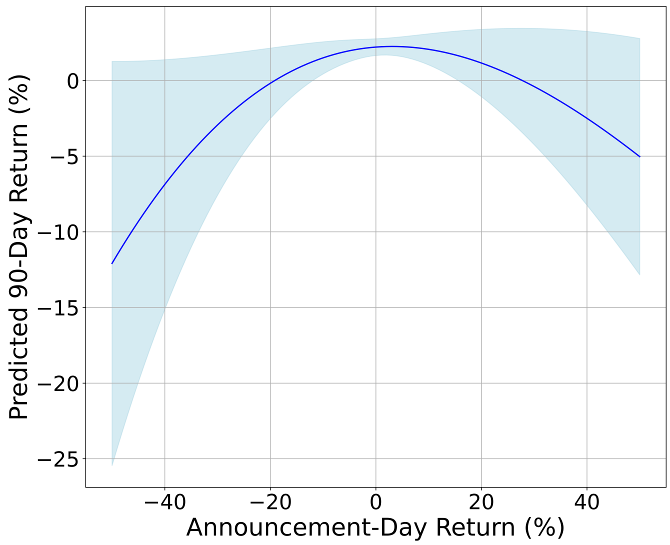


(b) Tails vs skew

(c) Tails vs quantile ratios

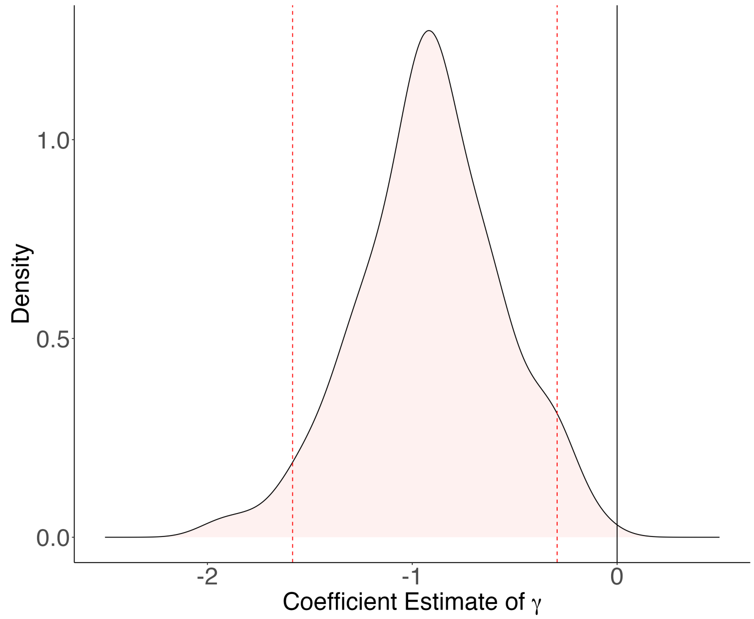
Note: Figure A3 reports the pairwise correlation between our main measure of extremeness based on announcement-day returns and alternative measures: earnings growth in the 2 years after the announcement in panel (a), the skew of the announcement-day return distribution in panel (b), and the quantile ratio of the top 1% of the distribution to the median in panel (c).

Figure A4: Estimate of the unconditional relationship $f(r_{i,t})$



Note: Figure A4 reports the amount of predictable post-announcement drift or reversal due to the announcement-day return based on the cubic specification in Section 4.3 at an extremeness measure of $\zeta_{\mathcal{C}} = 0.35$, corresponding to the No News distribution. Predicted 90-Day Return is The estimated post-announcement drift or reversal for the corresponding Announcement-Day Return, and is expressed as a percent of the Announcement-Day Return. Announcement-Day Return is the return of the stock on the day of the announcement, expressed in percentage points. 95% confidence intervals are reported in the shaded area.

Figure A5: Accounting for compounded estimation errors using block bootstrap



Note: Figure A5 reports the density of estimates of the main coefficient of interest, γ, in eq. (13) using a block bootstrap approach on a full firm-day panel following Politis and Romano (1994). The density is computed over 100 bootstrap samples. The 95% confidence interval is reported by the red dotted lines.

Table A1: Correlation between various tail measures

Announcement-Day Return	1	0.69	0.44	0.62	0.51	0.44	0.49
Long-Run Return	0.69	1	0.25	0.30	0.27	0.21	0.23
Earnings 1yr	0.44	0.25	1	0.60	0.53	0.51	0.63
Earnings 2yr	0.62	0.30	0.60	1	0.87	0.86	0.83
Earnings 3yr	0.51	0.27	0.53	0.87	1	0.98	0.91
Earnings 4yr	0.44	0.21	0.51	0.86	0.98	1	0.93
Earnings 5yr	0.49	0.23	0.63	0.83	0.91	0.93	1

Note: Table A1 reports the pairwise correlation coefficients between different measures of extremeness. Observations are at the news category level. Announcement-Day Return is the extremeness of stock returns of each firm on the day of the announcement and is measured in percentage points for all announcements in each category. Long-Run Return is the extremeness of stock returns of each firm in the 100 days after the announcement, including the announcement day, and is measured in percentage points for all announcements in each category. Earnings 1yr to Earnings 5yr are the extremeness of earnings growths as measured in eq. (9) for all announcements in each category.

Table A2: Robustness to different tail measures

	(1)	(2)	(3)	(4)	(5)	(6)
Announcement-Day Return	0.45***	0.77**	0.92***	0.73**	0.55**	0.47*
	(0.14)	(0.36)	(0.32)	(0.29)	(0.26)	(0.24)
Announcement-Day Return × Extremeness	−1.32***	−0.99**	−1.22***	−0.93**	−0.67**	−0.56*
	(0.41)	(0.47)	(0.43)	(0.36)	(0.31)	(0.29)
Constant	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.0004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Horizon	100 Days	1 Year	2 Year	3 Year	4 Year	5 Year
Measure	Returns	Earnings	Earnings	Earnings	Earnings	Earnings
Observations	197,498	197,498	197,498	197,498	197,498	197,498

Note: Table A2 reports the estimates corresponding to eq. (13), using alternative measures of extremeness. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Horizon is the time range over which the extremeness is measured. Measure is the variable that extremeness is measured on. Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A3: Accounting for announcement timing

	(1)	(2)
Announcement-Day Return	0.817***	0.766***
	(0.130)	(0.120)
Announcement-Day Return x Extremeness	-2.625***	-2.713***
	(0.347)	(0.315)
Event Window (+/-)	1 Day	2 Days
Observations	197498	197498

(a) Expanding announcement window

	(1)
Announcement-Day Return	0.573***
	(0.190)
Announcement-Day Return x Extremeness	-1.538***
	(0.524)
Specification	Exclude After-Hours Announcements
Observations	173256

(b) Excluding after-hours announcements

	(1)	(2)	(3)	(4)	(5)	(6)
Announcement-Day Return	0.585***	0.673***	0.716**	0.695***	0.627***	0.619***
	(0.184)	(0.232)	(0.306)	(0.198)	(0.184)	(0.196)
Announcement-Day Return x Extremeness	-1.525***	-1.651***	-1.619***	-1.851***	-1.695***	-1.662***
	(0.498)	(0.486)	(0.498)	(0.543)	(0.505)	(0.536)
Specification	Hour	Day of Week	Hour by Day of Week	No Fridays	No Fridays Evenings	No Strategic Announcers
Observations	197498	197498	197498	174900	195014	177858

(c) Including date and hour fixed effects

Note: Tables A3a, A3b, and A3c report the estimates corresponding to eq. (13), for different subsamples. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Event Window (+/-) indicates the number of days before and the number of days after the announcement that the announcement window is defined over (e.g., 2 days before to 2 days after). Exclude After-Hours Announcements excludes all announcements made after trading hours. Hour, Day of Week, and Hour by Day of Week report the respective fixed effects and interaction terms included. No Fridays excludes Friday announcements. No Fridays Evenings excludes Friday after-hour announcements. No Strategic Announcers excludes any firm that ever made an announcement on a Friday evening. Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A4: Controlling for the size and sign of the news

	(1)	(2)
Announcement-Day Return	0.539***	0.610***
	(0.191)	(0.193)
Announcement-Day Return^2	-0.431**	
	(0.202)	
Announcement-Day Return^3	0.177	
	(0.166)	
1(Return Decile = 1)		0.001
		(0.004)
1(Return Decile = 2)		0.006
		(0.005)
1(Return Decile = 3)		0.008
		(0.005)
1(Return Decile = 4)		0.009*
		(0.005)
1(Return Decile = 5)		0.009
		(0.006)
1(Return Decile = 6)		0.009
		(0.006)
1(Return Decile = 7)		0.007
		(0.006)
1(Return Decile = 8)		0.005
		(0.006)
1(Return Decile = 9)		0.001
		(0.008)
Announcement-Day Return x Extremeness	-1.463***	-1.682***
	(0.524)	(0.498)
Specification	Cubic Polynomial	Deciles
Observations	197498	197498

Note: Table A4 reports the estimates corresponding to eq. (13), with controls for the size and sign of the news. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points, and Announcement-Day Return^2 and Announcement-Day Return^3 are the same term squared and cubed, respectively. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Cubic Polynomial is a cubic polynomial of Announcement-Day Return. 1(Return Decile = d) is indicator variables for whether the Announcement-Day Return is in the d-th decile. Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A5: Sample selection for small/big news and positive/negative news

	(1)	(2)	(3)
Announcement-Day Return	0.600***	0.625***	0.635***
	(0.184)	(0.185)	(0.185)
Announcement-Day Return x Extremeness	-1.612***	-1.678***	-1.733***
	(0.503)	(0.506)	(0.505)
Absolute Announcement-Day Return Greater Than	25-th Percentile	50-th Percentile	75-th Percentile
Observations	148126	98750	49381

(a) Excluding smallest news

	(1)	(2)	(3)	(4)	(5)
Announcement-Day Return	0.511***	0.544***	0.522***	0.512**	0.601***
	(0.178)	(0.183)	(0.189)	(0.213)	(0.224)
Announcement-Day Return x Extremeness	-1.362***	-1.445***	-1.392***	-1.338**	-1.561**
	(0.491)	(0.509)	(0.525)	(0.604)	(0.636)
Top Percentile Excluded	0.01	0.05	0.1	0.5	1
Observations	197478	197399	197300	196510	195522

(b) Excluding outliers

	(1)	(2)
Announcement-Day Return	0.634**	0.363
	(0.304)	(0.416)
Announcement-Day Return x Extremeness	-2.164***	-0.566
	(0.819)	(1.206)
Subsample	Positive News Only	Negative News Only
Observations	104783	92715

(c) Positive and negative news

Note: Tables A5a, A5b, and A5c report the estimates corresponding to eq. (13), for different subsamples. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index ζ_C estimated following equation (8). Absolute Announcement-Day Return Greater Than refers to the percentile of absolute announcement-day returns below which the announcements were excluded. Top Percentile Excluded refers to the percentile of absolute announcement-day returns above which the announcements were excluded. Positive News Only and Negative News Only refer to announcements that had announcement-day returns greater than 0 and less than 0, respectively. Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A6: Accounting for overlaps

	(1)	(2)
Announcement-Day Return	0.588***	1.125***
	(0.195)	(0.324)
Announcement-Day Return x Extremeness	-1.618***	-3.076***
	(0.519)	(0.899)
Exclude Overlaps	With Earnings	With Any Events 30-Days Prior
Observations	152513	35879

(a) Excluding overlapping announcements

	(1)	(2)	(3)
Announcement-Day Return	0.511***	0.531***	0.453***
	(0.167)	(0.158)	(0.101)
Announcement-Day Return x Extremeness	-1.385***	-1.430***	-1.228***
	(0.438)	(0.419)	(0.253)
Overlap Controls	Pre-Announcement	Post-Announcement	Both
Observations	197498	197498	197498

(b) Fully estimated IRF

Note: Tables A6a and A6b report the estimates corresponding to eq. (14), with controls for overlapping news. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). For Exclude Overlaps, With Earnings refers to the subsample of news categories that do not co-occur with earnings announcements more than 50% of the time, and With Any Events 30-Days Prior refers to the subsample of news announcements that did not have another announcement occur in the 30 days prior. For Overlap Controls, Pre-Announcement refers to $h = -90$ to $h = -1$, Post-Announcement refers to $h = 1$ to $h = 90$, and Both refers to $h = -90$ to $h = 90$ following eq. (14). Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A7: All news categories, exclusion/inclusion criteria

Category	Included	Criteria
Op. Result	Included	
Earnings	Included	
Sales and Trading	Excluded	Trading
Annual Meeting	Included	
Expansion	Included	
Buyback Transaction	Excluded	Capital Structure
Buyback Closing	Excluded	Capital Structure
Structure Change	Included	
Client	Included	
Guidance Lower	Included	
Guidance Confirm	Included	
Guidance Raised	Included	
Credit Watch	Included	
Debt	Excluded	Capital Structure
Downsize	Included	
Dividend	Included	
Earnings Call	Included	
Earnings Release Date	Excluded	Administrative
Board Changes	Included	
CEO Change	Included	
CFO Change	Included	
Fixed Income	Excluded	Capital Structure
Follow-On Equity	Excluded	Capital Structure
Writeoff	Included	
Index Constituents	Excluded	Trading
Lawsuit	Included	
M&A Rumor	Included	
M&A Transaction	Included	
M&A Closing	Included	
Private Placements	Excluded	Capital Structure
Product	Included	
Seek Investment	Included	
Seeking to Sell	Excluded	Trading
Shelf Registration	Excluded	Administrative
Alliance	Included	

Note: Table A7 reports the news categories types in our dataset and the inclusion/exclusion criteria for each news category. Criteria is the selection criteria applied to each news category. Section 3 contains more details on the selection criteria and shows that the results are robust to using different selection criteria.

Table A8: Robustness to category exclusions

Category Excluded	Corr. Coef.	p-value
Op. Result	-0.63	0.001
Earnings	-0.66	0.001
Annual Meeting	-0.66	0.001
Expansion	-0.67	0.000
Structure Change	-0.67	0.000
Client	-0.69	0.000
Guidance Lower	-0.68	0.000
Guidance Confirm	-0.65	0.001
Guidance Raised	-0.66	0.001
Credit Watch	-0.69	0.000
Downsize	-0.69	0.000
Dividend	-0.67	0.000
Earnings Call	-0.66	0.001
Board Changes	-0.66	0.001
CEO Change	-0.59	0.003
CFO Change	-0.61	0.002
Writeoff	-0.67	0.001
Lawsuit	-0.73	0.000
M&A Rumor	-0.63	0.001
M&A Transaction	-0.66	0.001
M&A Closing	-0.66	0.001
Product	-0.67	0.001
Seek Investment	-0.65	0.001
Alliance	-0.66	0.001

Note: Table A8 reports the category-level correlation coefficients between post-announcement drift/reversal and extremeness, excluding each news category one-by-one.

Table A9: Robustness to category inclusions

News Categories	Coefficient	SE	Observations
Baseline	-1.63	0.5	197498
+ Dividends and Buybacks	-1.55	0.5	202191
+ IPOs, SEOs, and Debt	-1.65	0.5	204441
+ Dividends, Buybacks, IPOs, SEOs, and Debt	-1.56	0.5	209134
All News Categories with 1,000+ Occurrences	-1.18	0.48	243966
All News Categories	-1.09	0.46	250852

Note: Table A9 summarizes the robustness exercises for eq. (13) based on different sample selection criteria. Coefficient is the main γ coefficient estimate. SE is the standard error. Observations is the number of observations in the corresponding estimates for each row. Baseline is our main sample of 24 categories. + Dividends and Buybacks adds all dividend and buyback related news announcements. + IPOs, SEOs, and Debt adds all IPOs, SEOs, and debt related news announcements. + Dividends, Buybacks, IPOs, SEOs, and Debt adds both sets. All News Categories with 1,000+ Occurrences refers to the sample of all Capital IQ news categories that occurred at least 1,000 times in our sample. All News Categories refers to the sample of all Capital IQ news categories.

Table A10: Large-cap vs. small-cap firms

	(1)	(2)	(3)
Announcement-Day Return	0.607***	0.427***	0.475***
	(0.184)	(0.153)	(0.133)
Announcement-Day Return x Extremeness	-1.634***	-1.122***	-1.252***
	(0.503)	(0.392)	(0.344)
Sample	Large-Cap (Original)	Small-Cap	All Firms
Observations	197498	226986	424484

Note: Table A10 reports the estimates corresponding to eq. (13), for different subsamples. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Large-Cap refers to firms with market capitalizations of at least \$2 bn at announcement time. Small-Cap refers to firms with market capitalizations of less than \$2 bn at announcement time. All Firms refers to all firms. Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A11: Accounting for attrition

	(1)	(2)
Announcement-Day Return	0.606***	0.687***
	(0.184)	(0.188)
Announcement-Day Return x Extremeness	-1.636***	-1.860***
	(0.502)	(0.521)
Specification	Non-Attrited Firms	Below-Mean Attrition News Categories
Observations	196463	142651

(a) Accounting for non-attrited firms

	(1)	(2)	(3)
Announcement-Day Return	0.601***	0.806***	0.451**
	(0.182)	(0.242)	(0.205)
Announcement-Day Return x Extremeness	-1.783***	-2.220***	-1.321**
	(0.546)	(0.671)	(0.622)
Fundamentals Horizon	30 Days	100 Days	250 Days
Observations	197498	197498	197498

(b) Using alternative horizons

Note: Tables A11a and A11b report the estimates corresponding to eq. (13), for different subsamples. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Non-Attrited Firms refers to the subsample of firms that did not attrite in our sample. Below-Mean Attrition News Categories refers to the subset of news categories that were below the mean attrition rate across all news categories. Fundamentals Horizon refers to the return horizon that extremeness is computed based on. Standard errors are two-way clustered at the firm and day levels. *** p<0.01, ** p<0.05, * p<0.10.

Table A12: Using Driscoll-Kray Standard Errors

	(1)	(2)	(3)	(4)
Announcement-Day Return	0.61***	0.54***	0.77***	0.64***
	(0.20)	(0.18)	(0.27)	(0.24)
Announcement-Day Return × Extremeness	−1.63***	−1.38***	−2.07***	−1.70**
	(0.53)	(0.49)	(0.73)	(0.66)
Constant	0.02***	−0.01***	0.01**	−0.01***
	(0.01)	(0.002)	(0.01)	(0.003)
Time-Varying Tails	No	No	Yes	Yes
Return Benchmark	No	Yes	No	Yes
Observations	197,498	197,498	110,748	110,748

Note: Table A12 report the estimates corresponding to eq. (13) using Driscoll and Kraay (1998) standard errors. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the cumulative post-announcement return from day 1 to day 90 after the announcement. Announcement-Day Return is the stock return of the firm on the day of the announcement and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Time-Varying Tails indicates whether Extremeness is computed over a rolling past five-year window (Yes) or over the entire sample (No). Return Benchmark indicates whether the Announcement-Day Return and dependent variable are excess returns relative to the S&P 500 (Yes) or raw returns (No). Standard errors are computed following Driscoll and Kraay (1998). *** p<0.01, ** p<0.05, * p<0.10.

Table A13: Using Driscoll-Kraay Standard Errors, Volume and Extremeness

VARIABLES	(1) Turnover	(2) Turnover	(3) Turnover	(4) Turnover
Abs. Announcement-Day Return	0.25***	0.23***	0.19**	0.19**
	(0.074)	(0.073)	(0.077)	(0.076)
Abs. Announcement-Day Return × Extremeness	0.64***	0.76***	0.90***	0.91***
	(0.21)	(0.21)	(0.22)	(0.22)
Constant	0.0087***	0.013***	0.0089***	0.014***
	(0.00061)	(0.00062)	(0.00059)	(0.00059)
Observations	197,498	197,498	197,498	197,498
R-squared	0.371	0.394	0.395	0.412
Trading Day FEs	Yes	Yes	Yes	Yes
Return Benchmark	No	No	Yes	Yes

Note: Table A13 reports the estimates corresponding to eq. (15) using Driscoll and Kraay (1998) standard errors. Observations are at the news announcement level from January 1, 2011 to December 31, 2018. The dependent variable is the announcement-day turnover, defined as the volume of shares traded times the share price divided by the market capitalization. Abs. Announcement-Day Return is the absolute value of the announcement-day return and is measured in percentage points. Extremeness is the inverse power-law index $\zeta_{\mathcal{C}}$ estimated following equation (8). Trading Day FEs indicates whether the specification includes trading day fixed effects. Return Benchmark indicates whether the Announcement-Day Return and dependent variable are excess returns relative to the S&P 500 (Yes) or raw returns (No). Standard errors are computed following Driscoll and Kraay (1998). *** p<0.01, ** p<0.05, * p<0.10.

Table A14: Sample headlines for each news category

Event	Headline
Alliance	ChinaNet-Online Holdings, Inc Announces Strategic Partnership with Wuxi Jingtum Network Technology
Annual Meeting	STAAR Surgical Company, Annual General Meeting, Jun 11, 2009
Board Changes	Cellectar Biosciences, Inc. Announces Board Changes
CEO Change	MYOS Corporation Announces Executive Changes
CFO Change	Tetraphase Pharmaceuticals, Inc. Announces Resignation of Kamalam Unninayar as Chief Financial Officer, Effective March 16, 2018
Client	Ocean Power Technologies Enters Into First Commercial PB3 Agreement with Mitsui Engineering and Shipbuilding
Credit Watch	Issuer Credit Rating: BBB/Watch Neg/– From BBB/Negative/–: Local Currency Rating
Dividend	Johnson Controls International plc Approves Quarterly Cash Dividend, Payable on Jan. 6, 2017
Downsize	Pier 1 Imports Inc. Plans to Close 16 Stores
Earnings	Fonar Corp. Reports Unaudited Consolidated Earnings Results for the Third Quarter and Nine Months Ended March 31, 2009
Earnings Call	AtriCure, Inc., Q1 2009 Earnings Call, May-05-2009
Expansion	Aemetis, Inc. Completes Construction of Advanced Biodiesel Pre Treatment Unit Required for BP Supply Agreement
Guidance Confirm	Pareteum Corporation Provides Revenue Guidance for the Second Quarter Ended June 30, 2017
Guidance Lower	Crestwood Revises Earnings Guidance for the Year 2016
Guidance Raised	Hartford Financial Services Group Inc. Revises Earnings Guidance for the Year of 2008
Lawsuit	Hospitality Properties Trust Announces Settlement of Litigation with TravelCenters of America LLC
M&A Closing	Appliance Recycling Centers of America, Inc. (NasdaqCM:ARCI) acquired GeoTraq Inc. for $16 million.
M&A Rumor	PZU Eyes AIG Assets
M&A Transaction	Differential Brands Group Inc. (NasdaqCM:DFBG) entered into a definitive purchase agreement to acquire majority of North American licensing business of GBG USA, Inc. for $1.4 billion.
Op. Result	Delta Air Lines, Inc. Reports Operating Results for the Quarter Ended December 2014
Product	Inovio Biomedical Corporation Influenza Vaccines Demonstrate 100% Protection Against Current Pandemic A/ H1N1 Influenza Viruses in Animal Studies
Seek Investment	Insmed Seeks Acquisitions
Structure Change	Diffusion Pharmaceuticals Inc. Approves Amendment to Certificate of Incorporation
Writeoff	Manitowoc Co. Inc. Announces Impairment Charges for the First Quarter of 2009

Note: Table A14 reports example news headlines for each of the news categories in the dataset.