Do Low Latency Traders Destabilize Prices? Evidence from News Releases

Tarun Chordia, Bin Miao, and Joonki Noh[1]

June 2024

Abstract

This paper provides evidence that the low latency traders (LLTs) hurt market efficiency upon the release of non-earnings news. LLTs behave as speculators and exploit naïve news-chasing investors causing overpriced stocks to become even more overpriced upon the release of high sentiment news. This exacerbation of overpricing, which is subsequently reversed, creates a wedge between prices and the consensus fundamental value and hurts market efficiency. The asymmetric impact of high and low sentiment news on overpriced stocks is driven by more positive news being produced in recent years and investors paying more attention to overpriced stocks.

JEL Classification: G11; G12; G14; G15; G41
Keywords: News releases, Mispricing, High frequency trading, Low latency traders, Market efficiency

Contacts

	Tarun Chordia	Bin Miao	Joonki Noh
Voice:	1-404-727-1620	86-755-2351-8656	1-216-368-3737
Fax:	1-404-727-5238		1-216-368-6249
E-mail:	Tarun.Chordia@emory.edu	miaobin@cuhk.edu.cn	Joonki.Noh@case.edu
Address:	Goizueta Business School Emory University Atlanta, GA 30322, USA	Chinese University of Hong Kong Shenzhen, China	Weatherhead School of Management Case Western Reserve University Cleveland, OH 44106, USA

[1] We are grateful to the conference and seminar participants at the 2022 Global AI finance research conference, 2023 KAFA brownbag seminar, 2023 MFA annual meeting, 2023 APAD annual conference, 2023 CAFM annual meeting, 2024 Eastern FA annual meeting, 2024 French FA annual meetings, 2024 ESADE spring workshop, 2024 CICF annual conference, Case Western Reserve University, Deakin University, Dongbei University of Finance and Economics, KAIST, Texas Tech University, Tianjin University, and Southern University of Science and Technology, as well as Yakov Amihud, Snehal Banerjee, Joey Engelberg, Pedro Garcia-Ares, Clifton Green, Vincent Gregoire, Joel Hasbrouck, Terrence Hendershott, Sahn-Wook Huh, Da-Hea Kim, Juhani Linnainmaa, David Mclean, Ji Yeol Jimmy Oh, Jeffrey Pontiff, Sungjune Pyun, Paul Schmidt-Engelbertz, and Kaushik Vasudevan. In addition, the authors are grateful to Xing Huang for making the SAS code that implements the Barber et al. (2023) algorithm available.

Information is increasingly being released to, interpreted by, and traded on by computers. The impetus for algorithmic trading has come from regulation and advances in technology. Dramatic improvements in technology have allowed computer algorithms to dynamically monitor multiple trading venues and strategically submit orders. These algorithms emphasize speed, and as a result, trade latency has declined to nanoseconds. The literature on low latency trading (LLT) has shown that low latency traders (LLTs) improve market quality through increased liquidity and lower short-term volatility (e.g., see Chordia et al. 2011, Hendershott et al. 2011, Hasbrouck and Saar 2013, Hendershott and Riordan 2013, Boehmer et al. 2018, and Chordia, et al. 2018). Specifically, Brogaard et al. (2014) find evidence that high frequency traders (HFTs)[2] facilitate price discovery by trading in the direction of permanent price changes and against transitory pricing errors. Chordia and Miao (2020) find that LLTs improve the price efficiency around earnings announcements.[3]

While the popular press has argued that HFTs destabilized markets during the Flash Crash on May 6, 2010, Kirilenko et al. (2017) show that the HFTs did not alter their trading strategies as prices fell during the Flash Crash. Also, Brogaard et al. (2018) show that, during periods of extreme price movements in individual stocks, HFTs stabilize prices by absorbing order imbalances. In this paper, we provide evidence that, in certain situations, LLTs behave like short-term speculators and destabilize asset prices. Specifically, we show that in the presence of retail traders who tend to chase news, LLT activity increases

[2] We use LLT to refer to any computer-assisted low-latency trading activity. As such, HFT is a subset of LLT, as specifically defined by the Security and Exchange Commission's (SEC's) concept release on equity market structure (34-61358). We will use the term "HFTs" only for LLTs that fit the SEC definition. Algorithmic trading (AT) refers to the use of computer algorithms to automatically make trading decisions, and thus encompasses both LLT and HFT.

[3] Examples of other papers documenting improvements in market quality due to LLT include the following. Carrion (2013) finds that prices incorporate market-wide return information more quickly on days with high HFT participation. Hagströmer and Nordén (2013) show that HFTs reduce short term volatility. Chaboud et al. (2014) find that LLT improves price efficiency through lower return autocorrelations and fewer arbitrage opportunities. Conrad et al. (2015) find that LLT activity leads prices to more closely resemble a random walk. Hirschey (2020) argues that a subset of HFTs are skilled in predicting future buying and selling price pressure, leading to the improvement in price discovery. Boehmer et al. (2021) provide evidence that in international markets AT improves liquidity and informational efficiency but increases short-term volatility. Chakrabarty et al. (2022) show that HFT improves efficiency around earnings announcements during low-attention periods.

the prices of already overpriced stocks upon the release of positive sentiment news. In this scenario, which is becoming increasingly prevalent, LLTs hurt market efficiency.

To detect the situations in which LLTs might destabilize prices, we examine the return impact of news releases on overpriced and underpriced stocks. This provides a unique setting where investors face two different signals – the degree of mispricing and news releases. Unconditionally, overpriced (underpriced) stocks should have lower (higher) returns. Also, unconditionally, the release of high (low) sentiment news should result in higher (lower) stock returns. However, the *conditional* reactions of mispriced stocks to high and low sentiment news are more complicated since they depend on whether it is the degree of mispricing or the sentiment of the news that prevails in the trading decisions by LLTs. Therefore, examining the conditional reactions allows us to potentially identify the price destabilizing behavior of LLTs, especially when the two signals conflict with each other.

We obtain earnings announcement and news release dates from Compustat and RavenPack, respectively, and we use the 11 mispricing signals in Stambaugh, Yu, and Yuan (2012) to identify relative mispricing. Unconditionally, returns are higher on news days and on earnings announcement days. On days without earnings announcements or other news releases, overpriced stocks earn lower returns. These results are as expected. However, overpriced stocks earn higher returns on news release days and lower returns on earnings announcement days. Thus, for overpriced stocks, mispricing is corrected on earnings announcement days while it is exacerbated on news days. This difference could occur because earnings announcements convey hard quantitative information (signal precision is higher) whereas news releases often convey soft narrative information (signal precision is lower) that is subject to spin.[4]

We measure the information content of earnings announcements by the standardized unexpected

[4] There is evidence that people respond differently to quantitative vs. narrative information. Graeber, Roth, and Zimmermann (2023) use controlled experiments to show that people recall stories more easily than statistics, even if they are unrepresentative of reality. Further, the impact of stories on beliefs fades slowly. Solomon (2012) shows that market reactions can differ across non-earnings news releases vs. earnings announcements as there is more room to spin the non-earnings news releases.

earnings (SUE), and that of news releases by the RavenPack sentiment score (SENT). When SUE is high, underpriced stocks earn higher returns. On days with low SENT, overpriced stocks earn lower returns as the trading signals from the overpricing and from low sentiment news are aligned and both serve to reduce prices. However, when high sentiment news is released, overpriced stocks become even more overpriced. Thus, when the overpricing and high sentiment news conflict with each other, prices deviate further from the consensus fundamental value, thereby, widening the wedge. We also find that a stronger mismatch between the two signals (i.e., higher sentiment news in overpriced stocks) leads to more exacerbation of the overpricing upon news releases. This short-run exacerbation of overpricing hurts price efficiency. On the other hand, low sentiment news helps correct the overpricing and improves efficiency.[5]

We rely on the models of rational speculation[6] in our search for potential underlying economic mechanisms that drive the exacerbation of overpricing upon the release of positive sentiment news. De Long et al. (1990) have theorized that rational speculators can destabilize prices in the presence of positive feedback trading strategies (buying stocks when prices rise and selling stocks when prices fall) pursued by some noise traders. To identify the positive feedback trading strategies followed by naïve news chasers, we isolate retail investor trades by using the Boehmer et al. (2021) algorithm.[7] Indeed, we find that retail traders behave as news chasers and trade in the direction of the news sentiment regardless of the mispricing. Moreover, the exacerbation of overpricing following the release of high sentiment news might be positively related to retail trading activity.

[5] These findings are robust to (i) using signals from 95 anomalies studied by McLean and Pontiff (2016) to determine the degree of mispricing, (ii) using the Stambaugh et al. (2012) method of ranking to categorize underpriced or overpriced stocks, (iii) using only reputable news sources such as the Dow Jones Newswire, and (iv) using three days around each earnings announcement or news release rather than just the day of the announcement or release.

[6] Friedman (1953) and Fama (1965) have argued that rational speculators and sophisticated investors stabilize asset prices. In contrast, Harrison and Kreps (1978) and Allen and Gorton (1993) have argued that investors would be rationally willing to purchase overpriced stocks if they expect to be able to sell these stocks at even higher prices.

[7] Barber and Odean (2008) have argued that retail investors (as opposed to institutional investors) behave as news chasers. Hendershott et al. (2015) have shown that institutions are informed. Our conclusions stay the same when using the Barber et al. (2023) algorithm to identify retail trades.

Next, we examine the impact of LLTs on the news-driven overpricing exacerbation. LLTs have become an important part of financial markets and comprise a large fraction of the trading volume in recent years. We find that the exacerbation of overpricing on positive news days is significantly strengthened by LLTs who can react quickly to news in the direction of its sentiment, i.e., with more intense LLT activity on news release days, positive sentiment news leads to larger positive returns of overpriced stocks. Although negative sentiment news also leads to larger negative returns, the net effect on news release days is the exacerbation of overpricing. While this result is surprising considering that a large literature has shown that LLT improves price discovery and market efficiency, it is akin to short-term speculation. While LLTs are not speculators per se, their algorithms are set up to react fast and trade in the direction that is expected of other naïve and slower market participants. Analogous to speculators, LLTs act quickly to incorporate the naïve traders' return expectations into prices in order to profit from short-term price movements.

More importantly, the impact of LLTs on the news-driven overpricing exacerbation depends crucially on the presence of retail traders, although retail trading by itself does not exacerbate overpricing. Specifically, upon the release of high sentiment news, LLT strengthens the overpricing exacerbation only in the presence of retail traders. In the absence of retail traders, LLT does not exacerbate the overpricing. The results suggest that it is not the naïve news chasers per se, but the speculative behavior by LLTs in the presence of the naïve news chasers that drives the exacerbation of overpricing upon high sentiment news releases. Further, we find that the joint impact of LLTs and retail trading on the overpricing exacerbation depends on the intraday trading patterns of LLTs. Specifically, the significant news-driven overpricing exacerbation obtains only for positive news releases while market is open and is much stronger for morning news releases than afternoon releases. These results are consistent with LLTs having more opportunities for trading and for balancing their positions during trading hours, when news is released in

the morning, since they generally do not carry overnight positions.

One caveat is that even the LLTs might ignore the mispricing signals on news release days because these signals tend to be highly persistent and are thus less timely than news. If the naïve investors and the fast traders focus only on the news, an alternative interpretation of our findings could be that the positive returns upon high sentiment news releases simply reflect the increase in the consensus fundamental values instead of the speculation-driven price changes. To distinguish between the impact of speculation and the updating of the consensus fundamental value, we conduct two additional tests by examining 1) the impact of stale news releases and 2) how the initial return impact changes over the subsequent period after the news releases. The main idea behind the first test is that it shuts down the updating of the consensus fundamental value channel. Upon observing stale news, sophisticated investors are unlikely to update their beliefs about fundamental value, but they still have an incentive to speculate by predicting the reactions of naïve news chasers. We find that even with stale news, the news-driven overpricing exacerbation effect remains significant (although it is weaker), suggesting that short-term speculators rationally purchase overpriced stocks as they expect the news chasers to do the same. Thus, the speculators take advantage of the naïve news chasers who react with a delay possibly due to inattention or due to positive feedback trading strategies. The second test shows that the initial positive returns of overpriced stocks are reversed over the next 30 days after the release of high sentiment news. This return reversal following high sentiment news also points to an overreaction to the news rather than to an updating of the consensus fundamental value. In contrast, there is a continuation in the negative returns of overpriced stocks after the release of low sentiment news. Overall, the reversal and the stale news results are consistent with rational speculation by short-term investors.

The impact of high sentiment news on overpriced stocks is about twice as large as the impact of low sentiment news. These results obtain due to the following reasons. Investors pay more attention (as

measured by the number of unique IP addresses that download 10-K filings) to overpriced stocks than to underpriced stocks, resulting in larger return reactions to both positive and negative sentiment news in overpriced stocks. In addition, the number of news releases with a positive sentiment has been greater, especially in recent years, than that of news releases with a negative sentiment.[8] These factors – more attention paid to overpriced stocks by investors, larger market reactions to news releases for overpriced stocks, and more positive news produced – together with the short-term speculative behavior of LLTs to exploit naïve news chasers are likely to lead to our finding that LLTs can hurt the market efficiency by destabilizing asset prices upon news releases. But why, in recent years, have there been more articles with positive sentiment than with negative sentiment?

Solomon (2012) documents that external investor relations (IR) firms generate more positive sentiment news coverage for their clients by fostering and promoting positive press releases, leading to increased returns around news release days. Using a stock-level monthly proxy for IR activity, we show that overpriced stocks with higher IR activity earn higher returns on days with news releases and on days with positive sentiment news releases. Thus, the creation and fostering of firm-level positive news articles (possibly by internal and/or external IR activities) can lead to overpriced stocks becoming even more overpriced, which results in an increase in the discrepancy between market price and consensus fundamental value and a decrease in price efficiency. However, why don't sophisticated investors account for the IR activity in their investment decisions and prevent this overpricing exacerbation? It is because, in the presence of naïve news chasers, the short-term speculators *rationally* buy overpriced stocks upon the release of high sentiment news to profit from short-term price fluctuations.

We contribute to several strands of the literature. First, we contribute to a burgeoning literature

[8] Hirshleifer (2020) has argued that social transmission bias leads to more positive news being spread. Also, investors may pay more attention to and react more to positive news due to short sale constraints and because stocks are in net positive supply.

that explores the potential impact of LLTs on the market quality by showing that LLTs can destabilize asset prices in certain situations.[9] This is surprising since existing studies have shown that LLT significantly improves market quality including liquidity, price discovery, and market efficiency. Moreover, recent empirical papers have suggested that LLTs' activity leads to faster incorporation of earnings-related hard and quantitative information into prices not only at millisecond frequencies but also at daily frequencies (Chordia and Miao 2020, Bhattacharya et al. 2020, and Chakrabarty et al. 2022). In contrast, we show that the existing mispricing can be exacerbated by LLTs upon releases of non-earnings news whose timing and content are largely unpredictable. Our results suggest that the impact of LLTs on financial markets could be more complicated than previously understood and can be different depending on the underlying information events (earnings announcement vs. other news releases).

Second, we provide empirical support to the idea that rational speculation can destabilize asset prices. There is a large literature on price bubbles where rational behavior by sophisticated investors can lead to price destabilization (e.g., see Abreu and Brunnermeier (2002, 2003) and Brunnermeier and Nagel (2004)). The main economic implication of these studies, which is consistent with our findings, is that rational speculators gain by riding a bubble at the expense of less sophisticated investors. However, price bubbles might occur at much lower frequencies than at the daily frequency studied in this paper.[10]

Third, we contribute to the literature on the impact of news media on asset prices by showing how news releases can lead to a decrease in market efficiency. The literature has focused predominantly on the *unconditional impact* of news releases (without conditioning on mispricing) and the evidence has been mixed. On the one hand, news media has been shown to widely disseminate information and to reduce information asymmetries, thereby lowering expected returns (Fang and Peress, 2009, Tetlock, 2010).

[9] The theoretical model in Jarrow and Protter (2012) predicts that opportunistic LLTs can have a destabilizing price effect when they unknowingly coordinate on a common signal.
[10] Other theories that predict destabilizing speculation in non-bubble settings include Hart and Kreps (1986) and Brunnermeier and Pedersen (2005).

Engelberg et al. (2018) find that news releases help correct mispricing. On the other hand, Solomon (2012) argues that positive press releases by external IR firms can boost stock prices temporarily, which is corrected by subsequent earnings announcements. Hillert et al. (2014) claim that stocks more heavily covered by news media tend to exhibit stronger momentum and subsequent reversals.[11] Our analysis provides novel insights that the joint presence of the naïve news chasers and short-term speculators can play important roles in determining the impact of news media on asset prices.

2. Data

To quantify the stock-level mispricing at the beginning of each month, we rely on the 11 cross-sectional mispricing signals studied in Stambaugh, Yu, and Yuan (2012).[12] These mispricing signals have been shown by prior studies to predict the cross-section of future stock returns. As in Engelberg et al. (2018), we define our stock-level mispricing variable, i.e., *MISP*, at the beginning of each month as the difference in number of long-side and short-side anomaly portfolios that each stock belongs to and we form long-side and short-side portfolios by sorting stocks into quintile portfolios by each mispricing signal.[13] Note that *MISP* is a relative underpricing measure, i.e., stocks with higher value of *MISP* are more underpriced in the cross-section at the beginning of each month. *MISP* remains the same throughout all days in that month.

We obtain earnings announcement dates from the Compustat quarterly file. For a firm that reports earnings after the close of the market, we use the next trading day as the announcement day. We define $EARN_{i,t}$ as an indicator variable for earnings announcement days: $EARN_{i,t} = 1$ when firm i announces earnings on day t and $EARN_{i,t} = 0$ otherwise. We measure the information content of earnings

[11] Other papers in this literature are Tetlock (2007, 2011), Engelberg and Parsons (2011), Gurun and Butler (2011), Dougal et al. (2012), Engelberg et al. (2012), Garcia (2013), Solomon et al. (2014), Hendershott et al. (2015), and Jeon et al. (2022).
[12] We construct the 11 mispricing signals of SYY following their methodology based on the variables available from the CRSP, Compustat, and I/B/E/S databases.
[13] In a robustness check, we employ an alternative definition of *MISP* following Stambaugh, Yu, and Yuan (2012), denoted $MISP^{AR}$, and repeat our tests with $MISP^{AR}$. We find that our test results and conclusion are robust to using $MISP^{AR}$. The results are presented in the internet appendix Table A.6.

announcement by the standardized unexpected earnings (*SUE*). For firms covered by I/B/E/S, $SUE_{i,t}$ is calculated as the actual earnings per share (EPS) minus the median of all outstanding analyst EPS forecasts before the announcement, scaled by the stock price at the end of the fiscal quarter q that relates to the earnings announcement. For firms not covered by I/B/E/S, we measure $SUE_{i,t}$ by the seasonally-adjusted EPS change (= $EPS_q - EPS_{q-4}$) divided by the stock price at the end of the fiscal quarter q. In addition, we winsorize *SUE* at -1 and 1 to mitigate potential effects of outliers and set *SUE* = 0 for non-earnings days, when $EARN_{i,t} = 0$.

We obtain news-related variables from the RavenPack News Analytics database, which provides not only the date and time but also the novelty and relevance scores of each news story. Following Dang, Moshirian, and Zhang (2015) and Chordia, Lin, and Xiang (2021), we include only the news stories that have the novelty and relevance scores of 100 (unless otherwise noted). We define $NEWS_{i,t}$ as an indicator variable for news release days: $NEWS_{i,t} = 1$ if firm i has at least one news article released on day t and $NEWS_{i,t} = 0$ otherwise.[14] To measure the information content of news, we define $SENT_{i,t}$ as the average sentiment score of all news articles about firm i released on day t. Specifically, $SENT_{i,t}$ = (average of ESS – 50)/50, where ESS is the sentiment score for each news article assigned by RavenPack. $SENT_{i,t}$ ranges from -1 to 1 and we assign $SENT_{i,t} = 0$ for non-news days, when $NEWS_{i,t} = 0$. Higher values of $SENT_{i,t}$ thus indicate more positive news articles about firm i released on day t.

We merge the news- and earnings-related variables with daily stock returns and monthly *MISP* variable to investigate whether and how returns change on news release days and earnings announcement days and how the changes are affected by the information content of these events. The sample period is January 2000 through December 2019 and is restricted by the availability of the news-related variables

[14] While there is some trading after the market closes, it has been shown that liquidity is very thin and the LLTs do not actively submit and cancel orders (e.g., see Figure IA.1 in Chordia and Miao, 2020). Therefore, for news releases after the market closes, we use the next trading day as the news release day.

(*NEWS* and *SENT*) from RavenPack. We exclude low-priced stocks with prices less than $5 per share at the beginning of each month. The sample contains over 28.6 million stock-day observations.

2.1. Summary statistics

In our sample, news release days are much more frequent than earnings announcement days, and there is some overlap between the news days and the earnings days. Specifically, Panel A of Table 1 shows that among 5,845,200 news release days in total, 185,119 (3.17%) are also earnings announcement days. These descriptive statistics are similar to those reported in Engelberg et al. (2018) who study the sample period from January 1979 to December 2013.

We also present in Panel B of Table 1 the summary statistics of three mispricing-related variables: the numbers of long-side (i.e., *Long*) and short-side (i.e., *Short*) anomaly portfolios that each stock belongs to and our stock-level mispricing measure (i.e., *MISP*), which is defined as *Long* minus *Short*. Panel C plots the histogram of *MISP*. Overall, these statistics show that many stocks are not significantly mispriced, with the top or bottom quartile of the distribution having only one net buy or sell signal. But the fat tails of the distribution also indicate that a number of stocks seem to be severely mispriced.

3. Returns on information days

In this section, we first examine the return impact of information events on overpriced vs. underpriced stocks, when investors face two different signals: mispricing and news releases. By focusing on the conditional reactions, we search for situations where short-term speculators have an incentive to destabilize asset prices. Then, we examine the trading behavior of naïve news chasers and LLTs to provide a potential economic mechanism that drives our main finding, viz., the model of rational speculation by De Long et al. (1990) who theorize that rational speculators can destabilize prices in the presence of noise traders with positive feedback trading strategies.

3.1. Returns on earnings announcement days and news release days

We first investigate how mispricing relates to returns on two different types of information release days in the following daily panel regression.

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times EARN_{i,t} + \beta_3 MISP_{i,t} \times NEWS_{i,t} + \beta_4 EARN_{i,t} + \beta_5 NEWS_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (1)$$

where $RET_{i,t}$ is the return of stock *i* on day *t*, $MISP_{i,t}$ is the mispricing score of stock *i* measured at the beginning of the calendar month to which day *t* belongs, and $EARN_{i,t}$ and $NEWS_{i,t}$ are, respectively, the indicator variables for earnings announcement day and news release day. $\boldsymbol{X}_{i,t}$ is a (column) vector of control variables that includes daily returns, squared daily returns, and daily share trading volume over the past ten days (from days *t*-10 to *t*-1). Day fixed effects are included, and standard errors are clustered by day and firm to account for the cross-sectional and serial correlations in daily returns.

The results are reported in Panel A of Table 2. Prior research finds that due to short sale constraints, returns are mainly generated by the short leg of trades, i.e., stocks that are overpriced (Stambaugh et al. 2012). So, we estimate the regression model (1) for all stocks as well as for overpriced and underpriced stocks, separately. At the beginning of each month, we sort all stocks by *MISP*, and designate the bottom 30% and top 30% of stocks as overpriced stocks and underpriced stocks, respectively. In fact, we can see from the distribution reported in Panel C of Table 1 that *MISP* is negative in the bottom 30% and positive in the top 30%, suggesting that the designated overpriced stocks have at least one net sell signal (*MISP* = -1) and the designated underpriced stocks have at least one net buy signal (*MISP* = 1). For expositional convenience, we divide $MISP_{i,t}$ by 100 and multiply $RET_{i,t}$ by 100 before estimating the model, implying that the coefficients of *MISP*, *MISP***EARN*, and *MISP***NEWS* are in basis points (bps).

For the sample of all stocks, Column (1) of Panel A shows that *MISP* measured at the beginning of each month positively and significantly predicts the returns on non-information days during that month. While this could be due to positive returns to underpriced stocks or negative returns to overpriced stocks,

Columns (2) and (3) show that, consistent with Stambaugh et al. (2012), it is the negative returns to overpriced stocks that drive the impact of *MISP* on the daily returns in Column (1). The coefficient on *MISP* is 0.996 (t-statistic = 8.04), suggesting that the daily negative return associated with one additional sell signal at the beginning of the month, is about one basis point on "normal" trading days with no information releases. Thus, unconditionally, overpriced stocks earn lower returns. The coefficient on the interaction term *MISP*EARN* is significantly positive at 4.798 (t-statistic = 7.18), and Columns (2) and (3) show that, compared to days without earnings announcements or news releases, overpriced stocks with an additional sell signal have a negative return that is almost nine (=(1.724+13.323)/1.724) times larger in magnitude on earnings announcement days. Therefore, earnings announcements lead to a significant reduction in overpricing and thus a large improvement in price efficiency, which is consistent with Engelberg et al. (2018).

On the other hand, in Column (1), the coefficient on *MISP*NEWS* is significantly negative at -0.967 (t-statistic = -7.44) and Columns (2) and (3) show that this is due to overpriced stocks earning an additional one basis point return and becoming even more overpriced upon the release of news. The negative coefficient on *MISP*NEWS* is surprising and inconsistent with Engelberg et al. (2018). In Internet Appendix B, we provide tests to reconcile these two results. Overall, the adjusted R^2 stays at fairly low levels of less than 1%, which is typical in regressions with daily returns as dependent variable.

In the regression model (1), to the extent that *MISP* captures the deviation between market price and fundamental value, the daily returns could be interpreted as the average amount of mispricing that is corrected by earnings announcements or exacerbated by news releases. Positive (negative) coefficients on the interaction terms *MISP*EARN* (*MISP*NEWS*) indicate the convergence (divergence) between market price and fundamental value, and the magnitude of the coefficients reflects the speed of the convergence (divergence). The results in Panel A suggest that earnings announcements facilitate the convergence

between price and value and improve price efficiency, while other news releases increase the divergence and hurt price efficiency. Note that in Column (1), the absolute coefficient of *MISP*EARN* is about 4.8 times larger than that of *MISP*NEWS,* suggesting that earnings announcements correct the existing mispricing at 4.8 times the rate at which other news releases exacerbate it. This dramatic difference between earnings announcements and non-earnings news releases could be driven by the idea that market participants obtain hard and quantitative information from earnings announcements, while they typically obtain soft and narrative information from news releases, which is subject to spin.

The above results based on *EARN* and *NEWS* dummy variables reflect the impact of earnings announcements and news releases on stock returns. To further understand whether and how these effects vary with the information content of earnings announcements and news releases, we replace *EARN* and *NEWS* with earnings surprises and news sentiment, and estimate the following regression:

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times SUE_{i,t} + \beta_3 MISP_{i,t} \times SENT_{i,t} + \beta_4 SUE_{i,t} + \beta_5 SENT_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t} . \quad (2)$$

SUE is the standardized unexpected earnings (winsorized at -1 and 1), *SENT* is the average sentiment score (transformed to vary from -1 to 1) of all news articles released on each news day, and all other variables are as previously defined. We set *SUE* = 0 for non-earnings announcement days when *EARN* = 0 and *SENT* = 0 for non-news release days when *NEWS* = 0. As shown in Panel B, *SUE* and *SENT* are both strongly and positively correlated with stock returns for all stocks as well as for overpriced and underpriced stocks, suggesting that these two proxies effectively capture the information content of earnings announcements and news releases. The coefficient on *MISP*SUE* is significant only for underpriced stocks in the 10% significance level, suggesting that the returns of underpriced stocks increase with the positive surprise in earnings announcements. In contrast, the coefficient on *MISP*SENT* is strongly negative only for overpriced stocks, implying that, for overpriced stocks, news articles with

positive sentiment are associated with higher returns and more overpricing, and/or news articles with negative sentiment are associated with lower returns and less overpricing. This observation motivates our next set of tests.

To assess the differential impacts of positive and negative sentiment news, we estimate the following daily panel regression, which allows the coefficients of the sentiment variables to vary between positive and negative news days.

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times SUE_{i,t} + \beta_3 MISP_{i,t} \times PSENT_{i,t} + \beta_4 MISP_{i,t} \times NSENT_{i,t} + \beta_5 SUE_{i,t} + \beta_6 PSENT_{i,t} + \beta_7 NSENT_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (3)$$

where *PSENT* is the absolute value of the average sentiment of news articles published on a positive news day (when *SENT* > 0), and 0 otherwise. Similarly, *NSENT* is the absolute value of the average sentiment of news articles published on a negative news day (when *SENT* < 0), and 0 otherwise. Panel C shows that positive and negative sentiment news articles have dramatically different impacts on daily returns. Specifically, Column (1) shows that the coefficient on *MISP*PSENT* is strongly negative at -13.404 (t-statistic = -23.15), while *MISP*NSENT* has a strongly positive coefficient of 13.234 (t-statistic = 16.85). Columns (2) and (3) further show that these results are primarily driven by overpriced stocks. Therefore, overpriced stocks earn lower returns on days with more negative sentiment news articles, but higher returns on days with more positive sentiment news articles, implying that overpricing is exacerbated by positive sentiment news, but it is mitigated by negative sentiment news. The conditional average of *PSENT* on positive news days is 0.239 and that of *NSENT* on negative news days is 0.190 in our sample. In addition, the percentages of positive news days and negative news days (as a fraction of all days) are 10.3% and 6%, respectively in our sample, indicating that there are 72% more positive news days than negative news days. Combining these numbers with the coefficients of *MISP*PSENT* and *MISP*NSENT* suggests that on average, positive sentiment news articles have about 2.2 (=(13.404/13.234)*(0.239/0.190)

*(0.103/0.06)) times larger impact on daily return than negative sentiment news articles. Therefore, the net average effect of the information content of news articles is the exacerbation of existing overpricing driven by positive news content, which provides an explanation for why we obtained the negative and significant coefficient of *MISP*SENT* in Panel B.

Furthermore, the average overall impact of *MISP* is -2.280 (=0.924-13.404*0.239) bps on positive news days, implying that the exacerbation of overpricing on positive news days is stronger than the normal mispricing correction on non-information day. In contrast, the average overall impact of *MISP* is 3.438 (=0.924+13.234*0.190) bps on negative news day, implying that negative sentiment news helps correct the mispricing.

Our results are mainly driven by overpriced stocks. Across all model specifications, the coefficients on *MISP* are significantly positive for overpriced stocks, while they are insignificant for underpriced stocks. For example, Column (2) of Panel A shows that the daily return is about 1.724 bps on non-information days, which is more than 70% larger than that in Column (1). In addition, for overpriced stocks, the coefficients on the interaction terms in Column (2) show that returns significantly decrease on earnings announcements and negative news days but increase on positive news days. These patterns are all consistent with the full sample results in Column (1), but the magnitudes of the coefficient estimates are generally larger, suggesting that the effect of news releases on daily returns is stronger for overpriced stocks. In contrast, for underpriced stocks, Column (3) shows that the coefficients on various interaction terms (except for *MISP*SUE* in the 10% significance level) are insignificant, suggesting that the mispricing signals do not have much impact on the daily returns of underpriced stocks on news release days (as opposed to earnings announcement days).[15]

3.1.1. Impact of retail investors

[15] In the rest of the paper, the earnings-related variables will serve as controls so as to allow us to examine the marginal effects of non-earnings news releases, which is what we focus on in order to discern the impact of short-term speculation on returns.

Since in the theoretical framework of De Long et al. (1990), it is the presence of naïve news-chasing investors that attracts the short-term speculators, in this section we examine the impact of naïve investors on returns following news releases. Retail trading is used as a proxy for trading by naïve investors. The algorithm in Boehmer et al. (2021) is used to classify trades as originating from retail investors.[16] The daily summary measure of retail trading is obtained from the WRDS Intraday Indicator Data over the sample period 2010-2015, which is the same as in Boehmer et al. (2021). For stock *i* on day *t*, we define the retail order imbalance, $RETAIL_OIB_{i,t}$, as the difference between the retail buy and sell share trading volume divided by the sum of the retail buy and sell volume. Panel A of Table 3 presents the results from the following regression with date and firm fixed effects:

$$RETAIL_OIB_{i,t} = \beta_0 + \beta_1 \times SENT_{i,t} + \beta_2 \times MISP_{i,t} + \beta_3 \times RET_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}.$$

The retail order imbalance is positively correlated with returns and news sentiment, indicating that retail investors buy (sell) when high (low) sentiment news is released regardless of mispricing signals and that they are indeed naïve news chasers.

To investigate the impact of naïve news chasers on returns upon news releases, our research design is as follows. The retail trading volume for stock *i* on day *t*, $RETAIL_{i,t}$, is defined as the retail share volume as a fraction of the total share volume. Since the Boehmer et al. (2021) algorithm does not capture all the retail trades, $RETAIL_{i,t}$ is a noisy measure of retail trading volume (see the discussions in Barber et al. (2023) and Battalio et al. (2023)). So instead of $RETAIL_{i,t}$, we use the daily cross-sectional decile rankings of the retail trading volume, *DRT*. We use *DRT* and its interactions with *MISP*, *PSENT*, and *NSENT* in regression (3).[17]

[16] The Boehmer et al. (2021) algorithm relies on the fact that retail marketable orders can receive sub-penny price improvement. When using the Barber et al. (2023) methodology that modifies the Boehmer et al. (2021) algorithm in order to reduce the error rates of signing buying vs. selling retail orders, we find that the conclusions stay essentially unchanged. These results are reported in the internet appendix Table A.1.

[17] For expositional brevity, in the rest of the paper, we focus on the model in equation (3). The test results and conclusions are robust to the models in equations (1) and (2). See the tables in Internet Appendices A and B.

In Panel B of Table 3, the coefficients of *MISP*DRT* are significantly positive in Columns (1) to (3), suggesting that overpriced stocks earn lower returns and underpriced stocks earn higher returns, i.e., existing mispricing is reduced, as retail trading activity increases. In Column (1), the coefficients of *MISP*PSENT* and *MISP*PSENT*DRT* are -5.114 (t-statistic = -3.76) and -22.359 (t-statistic = -5.46), while those of *MISP*NSENT* and *MISP*NSENT *DRT* are 2.311 (t-statistic = 1.39) and 28.124 (t-statistic = 5.91), respectively. We find that the conditional average of *PSENT* on positive news days is 0.214 and that of *NSENT* on negative news days is 0.183 in our retail trading sample. In addition, the percentages of positive news days and negative news days (as a fraction of all days) are 14.2% and 9.4%, respectively, indicating that there are 51% more positive news days than negative news days in our retail trading sample. Combining these numbers with the coefficients of the double and triple interaction terms above implies that for stocks with *DRT* = 1, which are the most popular among retail investors, the average impact of positive sentiment news on daily return is about 1.6 (=(5.114+22.359)/(2.311+28.124)*(0.214/0.183)* (0.142/0.094)) times larger than that of negative sentiment news and thus the net effect on news release days is the exacerbation of existing mispricing by news releases (see Column (1), Panel A, internet appendix Table A.2). Importantly, the magnitude of return impact on news release days (either positive or negative sentiment news days) becomes significantly larger than that in Panel C of Table 2, due to the increased activity of retail investors.

Furthermore, for stocks with *DRT* = 1, the average overall impact of *MISP* on return is –4.463 (=1.416– (5.114+22.359)*0.214) bps on positive news days, implying that the exacerbation of overpricing due to positive sentiment news overwhelms the normal mispricing correction on non-information days. In contrast, the average overall impact of *MISP* on return is 6.563 (=1.416+28.124*0.183) bps on negative news days. In comparison, for stocks with *DRT* = 0, i.e., those not popular among retail investors, the average overall impact of *MISP* on return is –1.094 (= –5.114*0.214) bps on positive news days, while it

is statistically insignificant at 0.423 (=2.311*0.183) bps on negative news day. This implies that the average overall impact of *MISP* on returns increases by more than four times for stocks with *DRT* = 1 relative to those with *DRT* = 0.

For overpriced stocks in Column (2), we find that the coefficients of *MISP*PSENT*DRT* and *MISP*NSENT*DRT* are -113.04 (t-statistic = -8.52) and 88.308 (t-statistic = 5.88), respectively. For underpriced stocks in Column (3), the coefficients are 20.579 (t-statistic = 2.39) and -17.130 (t-statistic = -1.66). These results imply that, as retail trading activity increases, both overpriced and underpriced stocks earn higher (lower) returns on high (low) sentiment news days although overpriced stocks play a much more dominant role than underpriced stocks in driving the results for all stocks in Column (1).

In sum, all the results are consistent with the idea that retail traders are indeed news chasers in the sense that they focus solely on the sentiment of the news while ignoring the mispricing on news release days. The unanswered question in this section is whether it is trading by retail investors that leads to the increase in returns of overpriced stocks upon the release of high sentiment news or whether it is the short-term speculative trading in the presence of naïve news chasers that leads to the exacerbation of overpricing. Next, we focus on the impact of LLTs on news release day returns to search for potential economic explanations for our findings.

3.2. Impact of Speculation

In this section we examine (i) the short-run impact of LLTs on the news-driven exacerbation of overpricing, (ii) the impact of stale news, and (iii) the return pattern over the 30-day period following news releases.

3.2.1. Short-run impact of LLTs

Boehmer, Li, and Saar (2018) show that LLTs follow distinct trading strategies. Some LLTs engage in cross-venue arbitrage, others conduct short-horizon directional arbitrage, and yet others aim to

profit from supplying liquidity. The LLTs who conduct short-horizon directional arbitrage, aim to profit from short-run price movements by trading in the direction of expected price changes. This makes LLT a good setting to study the potential impact of short-term speculation on stock returns on news days.

The dominant view in the literature is that LLTs improve market efficiency, especially around earnings announcements. However, we focus on other unscheduled news releases whose timing and content are largely unpredictable. Following Hasbrouck and Saar (2013) and Chordia and Miao (2020),[18] we define the daily LLT proxy as the number of time-weighted strategic runs as follows:

$$LLT_{i,t} = \frac{1}{2.34\times10^{7}} \sum_{j=1}^{N_{i,t}} T_{i,j,t}, \tag{4}$$

where $T_{i,j,t}$ is the time-in-force in milliseconds for the *j*th strategic run for stock *i* on day *t*, and $N_{i,t}$ is the total number of strategic runs.[19] The total milliseconds in each trading day are 2.34×10^{7}. The sample for the analysis in this section is restricted, by the availability of Nasdaq Historical TotalView-ITCH database used for calculating *LLT,* to stocks traded on Nasdaq from January 2008 to December 2017.

Table 4 presents the results. In Panel A, we regress the log-transformed *LLT* (= *Log(1+LLT)*) on *MISP* and the set of news variables. The coefficient on *MISP* is significantly negative, suggesting that LLTs prefer overpriced stocks, which tend to be characterized by high trading volume.[20] LLT activity increases on news days (as compared to no news days), when positive or negative sentiment news is released, and as the magnitude of news sentiment becomes higher, suggesting that news releases and more extreme sentiment might attract LLTs. To examine the impact of *LLT* on returns on news release days, in Panel B, we interact *DLLT* with *MISP* and the sentiment variables, and then add the three-way interaction

[18] Our choice of the strategic run measure by Hasbrouck and Saar (2013) is further supported by Chakrabarty, Comerton-Forde, and Pascual (2021) who employ proprietary data with known identities of actual HFTs and show that *LLT,* as measured in equation (4), outperforms seven other popular HFT proxies in identifying HFT activities.

[19] For the *j*th strategic run that has at least ten linked messages, we define the time-in-force (= $T_{i,j,t}$) as the time stamp of the last message minus the time stamp of the first message of the run. If a stock has no strategic run for a given day, the values for *T* and *N* are zero, producing *LLT* = 0. For a more detailed discussion about *LLT*, see Chordia and Miao (2020).

[20] We have verified that overpriced stocks have, on average, significantly higher trading volume than underpriced stocks.

terms into equation (3). *DLLT* is a daily dummy variable for stocks in the top decile of *LLT* on each day and identifies stocks with the most intense LLT activity in the cross-section.

Panel B of Table 4 provides strong evidence that LLT significantly increases the prices of overpriced stocks when high sentiment news is released, while it significantly decreases the prices when low sentiment news is released. Also, the magnitude of price changes is larger as the sentiment of the news becomes more extreme. Specifically, in Column (1), the coefficients on *MISP*PSENT*DLLT* and *MISP*PSENT* are -19.767 (t-statistic = -7.34) and -14.062 (t-statistic = -17.16), while those on *MISP*NSENT*DLLT* and *MISP*NSENT* are 22.193 (t-statistic = 6.92) and 14.872 (t-statistic = 14.08), respectively. We find that the conditional average of *PSENT* on positive news days is 0.220 and that of *NSENT* on negative news days is 0.183 in our LLT sample. In addition, the percentages of positive news days and negative news days (as a fraction of all days) are 18.3% and 11.5%, respectively, indicating that there are 59% more positive news days than negative news days in the LLT sample. Then the coefficient estimates of *MISP*PSENT*, *MISP*PSENT*DLLT*, *MISP*NSENT*, and *MISP*NSENT*DLLT* suggest that for stocks with high LLT activity (with *DLLT* = 1), the average impact of positive sentiment news on return is about twice (=(14.062+19.767)/(14.872+22.193)*(0.220/0.183)*(0.183/0.115)) that of negative sentiment news and thus the net effect on news releases days is the exacerbation of existing overpricing driven by positive sentiment news releases. Importantly, the magnitude of the return impact on news release days (regardless of positive vs. negative news releases) become substantially larger than those in Panel C of Table 2, due to the trading activity of LLTs. Comparing the coefficients in Columns (1) to (3) indicates that these results are mainly driven by overpriced stocks.

Furthermore, for stocks with *DLLT* = 1, the average overall impact of *MISP* on daily returns is -6.981 (=0.461-(14.062+19.767)*0.220) bps on positive news days, implying that the exacerbation of overpricing due to positive sentiment news overwhelms the normal mispricing correction on non-

information days. In contrast, the average overall impact of *MISP* on returns is 7.244 (=0.461+(14.872 +22.193)*0.183) bps on negative news days. In comparison, for stocks with *DLLT* = 0, the average overall impact of *MISP* on return is -2.633 (=0.461-14.062*0.220) bps on positive news day, while it is 3.183 (=0.461+14.872*0.183) bps on negative news day, implying that the average overall impact of *MISP* on daily returns are strengthened by two to three times by LLTs.

Next, we examine the supply of liquidity by LLTs on news release days. As in Chordia and Miao (2020), we construct *LLT_SIGN* to capture the difference between the numbers of time-weighted strategic buy runs and sell runs as follows. For stock *i* on day *t*,

$$LLT_SIGN_{i,t} = \frac{1}{2.34\times10^{7}}\left[\left(\sum_{j=1}^{B_{i,t}} T_{i,j,t}^{Bid}\right) - \left(\sum_{k=1}^{A_{i,t}} T_{i,k,t}^{Ask}\right)\right] \times Sign\left(SENT_{i,t}\right), \quad (5)$$

where $T_{i,j,t}^{Bid}$ is the time-in-force in milliseconds for the *j*th strategic buy run at the bid price and $T_{i,k,t}^{Ask}$ is the time-in-force in milliseconds for the *k*th strategic sell run at the ask price, $B_{i,t}$ and $A_{i,t}$ are the numbers of strategic buy runs and sell runs, respectively. $Sign\left(SENT_{i,t}\right)$ has a value of 1 if *SENT* > 0, -1 if *SENT* < 0, and 0 otherwise. Since the difference in the time-in-force between buy and sell runs is signed conditional on the average sentiment of news releases, *LLT_SIGN* can be interpreted as the net quoting activity by LLTs at the bid when *SENT>0* and at the ask when *SENT<0*. Specifically, when *SENT* < 0, a positive *LLT_SIGN* corresponds to posting and cancelling of offer quotes as LLTs aggressively supply liquidity on the ask side to traders looking to buy. To the extent that net order imbalance on news days is mostly in the direction of news sentiment (i.e., more trades at the ask when SENT>0), a large and positive *LLT_SIGN* is, thus, a proxy for the aggressive updating of quotes by LLTs to be able to trade profitably in the direction of the news. In our tests, we employ a daily dummy variable, *DLLTS*, to indicate the top decile of stocks with the highest cross-sectional *LLT_SIGN* and present the results in Panel C of Table 4.

For overpriced stocks in Column (2), the coefficients on *MISP*PSENT*DLLTS* and *MISP*NSENT *DLLTS* are -18.312 (t-statistic = -2.37) and 25.644 (t-statistic = 2.60), respectively, and they are similar

in magnitude to those on *MISP*PSENT* and *MISP*NSENT*. This result implies that returns more significantly increase when high sentiment news is released and decrease when low sentiment news is released for overpriced stocks where LLTs aggressively update their quotes to be able to trade in the direction of the news sentiment. Comparing Columns (1) to (3) indicates that the results for all stocks are mainly driven by overpriced stocks.[21]

We further investigate the joint impact of retail trading and LLT activity on returns on news release days. Panel A of Table 5 first shows that LLT intensity tends to be negatively correlated with retail trading on non-news days (i.e., the coefficient of *DRT* is significantly negative), which is not unexpected because LLT algorithms are designed to mainly interact with institutional order flows. Interestingly, this pattern largely reverses and turns positive on news release days, regardless of the sentiments of the news. The positive and significant coefficients on *DRT*PSENT* and *DRT*NSENT* suggest that the presence of retail traders upon the release of high or low sentiment news is what attracts the LLTs, driving the results in Panel A of Table 4. In contrast, the negative coefficients of *PSENT* and *NSENT* indicate that in the absence of retail traders, LLTs tend to avoid news releases with larger magnitude of high or low sentiment.

We then interact the variables in Panel B of Table 4 with our proxy for retail trading activity, *DRT*, to examine how the joint presence of LLT and retail trading affects stock returns. For brevity, only the three-way and four-way interaction terms are tabulated in Panel B of Table 5.[22] Let's focus on overpriced stocks in Column (2). The positive coefficient on *MISP*PSENT*DLLT* suggests that for stocks in the lowest *DRT* decile, higher LLT activity is also related to a decline in returns, i.e., a decrease in the overpricing. The positive coefficient of *MISP*PSENT*DRT* implies that in the absence of LLTs, retail

[21] Since LLTs are more active in liquid stocks, we repeat the tests in Panels B and C of Table 4 by adding a contemporaneous daily liquidity measure, *LIQ*, proxied by daily share turnover and the interaction terms among *LIQ*, *MISP*, and news variables into the regression models. The internet appendix Table A.3 shows that the results in Panels B and C of Table 4 are robust to these liquidity-related controls.

[22] The internet appendix Table A.7 presents the full estimation results of Panels B and C of Table 5.

trading activity also mitigates the overpricing. In contrast, the coefficient of *MISP*PSENT*DRT*DLLT* is significantly negative suggesting that when high sentiment news is released, there is an increase in prices of overpriced stocks with high retail trading and with high LLT activity. In other words, the short-term speculation by fast traders leads to an exacerbation of overpricing only when the high sentiment news induces trades in the direction of the news by the naïve news-chasing retail investors.

To provide additional support for the LLT driven exacerbation of overpricing, we further decompose *PSENT* based on the intraday timing of news releases. Specifically, we decompose *PSENT* into *PSENT_BMO* (before market open, 0:00:00–9:29:59), *PSENT_AM* (morning trading session, 9:30:00–13:00:00), *PSENT_PM* (afternoon trading session, 13:00:01–16:00:00), and *PSENT_AMC* (after market close, 16:00:01–23:59:59). This decomposition allows us to examine the impact of intraday trading pattern of LLTs on our findings since LLTs are unlikely to carry overnight positions. We expect that the news-driven overpricing exacerbation would be 1) stronger in the morning than in the afternoon since LLTs have more opportunities for trading based on morning news releases before unwinding their intraday positions, and 2) non-existent before market open or after market close. For brevity, we report only the four-way interaction terms based on *PSENT* in Panel C of Table 5. Focusing on overpriced stocks in Column (2), the coefficients of the four-way interaction terms show that at least 69% (=500.62/(500.62 +223.39)) of the high sentiment news-driven overpricing exacerbation is due to morning news releases while the remaining 31% is due to afternoon news releases. In contrast, the four-way interaction terms have insignificant coefficients for news released outside trading hours. These results are supportive of the LLT driven exacerbation of overpricing upon high sentiment news releases.

One takeaway from Table 4 is that LLTs significantly strengthen the exacerbation of overpricing on positive news days and the mitigation of overpricing on negative news days. Thus, the impact of LLTs on overpriced stocks is to accelerate the price movements in the same direction as the underlying price

changes, i.e., prices increase (decrease) for high (low) sentiment news. Table 5 shows that the significant overpricing exacerbation by LLTs becomes prominent only in the presence of the naïve news-chasing retail investors. These results are consistent with O'Hara (2015) who argues that while LLTs trade faster upon the release of news, they do not unearth fundamental information. They are also consistent with (i) Chordia et al. (2018) who show that for macroeconomic announcements, LLTs react within milliseconds in the direction of the surprise, (ii) Chordia and Miao (2020) who have shown that LLTs react to unscheduled firm announcements, and (iii) Biais et al. (2015) who have argued that LLTs have an informational advantage as they trade faster than others. More importantly, given that computers are not subject to behavioral biases, these findings together imply that the algorithms are programmed to react to news and trade faster in the direction that naïve investors are expected to trade in, thereby, significantly widening the wedge between prices and the consensus fundamental value when positive sentiment news is released about overpriced firms. Overall, our results imply that rational speculation drives the short-term returns in the direction of anticipated trading by naïve investors who chase the information contained in news releases.

3.2.2. Impact of stale news

One possible concern about our speculation-based interpretation is that since the anomaly driven mispricing measure (*MISP*) is highly persistent, even the short-term speculators might focus on the news rather than the mispricing as *MISP* may be stale. This raises the possibility that the information contained in news releases might reflect the changes in fundamental values, making both sophisticated and naïve investors update their beliefs. In other words, a high (low) sentiment news release may cause investors to update their expectations of fundamental value positively (negatively) regardless of the degree of mispricing. Our earlier tests cannot separate the impact of speculation from this alternative channel where, upon news releases, investors update their beliefs about the fundamental value.

To determine whether it is the impact of speculation or the updating of the consensus fundamental value that drives the empirical patterns we have found, we shut down the channel wherein investors update the consensus fundamental value by focusing on stale news releases. When stale news is released, sophisticated speculators are unlikely to update their beliefs about fundamental values (as they would have already done so when the news was initially released), but to earn short-term profits, they still have an incentive to trade by predicting the reactions of other naïve news-chasing investors who either follow positive feedback trading strategies or were inattentive to the news.[23]

To operationalize our stale news idea, we apply the following two filters to our sample. First, we construct all news variables based on news articles whose novelty scores are less than 75 (out of 100).[24] Second, in our tests, we exclude firm-day observations that have the releases of genuinely novel news articles whose novelty scores are 100 to avoid a possibility that the variables based on stale news might simply pick up the reactions to novel news on the same day since they tend to cluster and are positively correlated. This filter also inevitably reduces the statistical power of our test since it also eliminates the impact of stale news released on the same day as novel news.

The test results based on the sample of stale news are presented in Table 6. We find that the mispricing exacerbation effect on news release days remains highly significant even though the coefficient estimates are generally smaller than those in Table 2. Specifically, the coefficient of *MISP*PSENT* in Column (1) is negative and highly significant, and its magnitude is 29.2% (=3.910 /13.404) of that in Table 2. Note that this estimate is likely a lower bound due to our conservative choice of filters to select stale news days. The impact of stale news points to the presence of the wedge that is created between prices and the consensus fundamental value upon the arrival of positive sentiment news

[23] Our economic intuition is consistent with Tetlock (2011) who finds that individual investors tend to overreact to stale news, causing temporary price movements, and the reactions to stale news are weaker than to novel news.
[24] RavenPack provides the novelty score (i.e., ENS) for each news article, which ranges from 0 to 100. The higher the novelty score is, the more novel is a given news article.

in overpriced stocks. Thus, the overpricing exacerbation effect on news release days is likely driven by the speculation-induced wedge between prices and the consensus fundamental value.

3.2.3. Reversals

There are at least two reasons to examine the returns of mispriced stocks over the subsequent periods after the news release on day *t*. First, recall that we have documented that news with a positive sentiment leads to more overpricing in already overpriced stocks in the short run. In this section, we examine whether this news-driven overpricing exacerbation is reversed over time. Second, recall that we used stale news in the previous section, to test whether it is the impact of speculation or the updating of the consensus fundamental value that is more likely to drive our findings. By examining the returns over the 30 days following each news release, we can shed further light on the same question. If the exacerbation of overpricing is driven by speculation, then the initial news-day returns should reverse. However, if the news release leads to an updating of the consensus fundamental value, then we should not observe any reversals.

We examine the impact of news release on not only the return on day *t* but also the buy and hold returns over each of the subsequent periods of (*t*+2, *t*+10), (*t*+2, *t*+20), and (*t*+2, *t*+30). Except for the news release on day *t*, we exclude all other future news release days and earnings announcement days when calculating the buy and hold returns. The results are presented in Table 7, where stock-days with non-zero news releases are used to estimate the regression models. While Panel A presents the results for all stocks, Panels B and C present the results for overpriced and underpriced stocks, respectively.

In Panel A, we find that the coefficients of *MISP*PSENT* are, respectively, -11.830 for *RET0*, 2.535 for *RET10*, 4.187 for *RET20*, and 6.549 for *RET30* and all are highly significant, implying that the initial mispricing exacerbation on positive news day is reversed over the next 30 days and that this reversal is stronger for more positive sentiment news. In contrast, the coefficients of *MISP*NSENT* are,

respectively, 14.241 for *RET0*, 4.824 for *RET10*, 6.832 for *RET20*, and 9.572 for *RET30* and all are highly significant, implying that the initial mispricing correction on negative news day continues over the next 30 days and that the continuation is stronger for more negative sentiment news.

Furthermore, recall that the conditional average of *PSENT* on positive news days is 0.239 and that of *NSENT* on negative news days is 0.190. Combining these conditional averages with the coefficients of *MISP* and *MISP*PSENT*, we find that, for positive sentiment news, the average overall impact of *MISP* is -2.246 (=0.581-11.83*0.239) bps for *RET0*, 3.289 (=2.683+2.535*0.239) bps for *RET10*, 6.437 (=5.436+4.187*0.239) bps for *RET20*, and 9.233 (=7.668+6.549*0.239) bps for *RET30*, respectively. This means that after positive news releases on day *t*, the average overall impact of *MISP* on returns also experiences a significant reversal over the next 30 days. In contrast, for negative sentiment news, the average overall impact of *MISP* is 3.287 (=0.581+14.241*0.190) bps for *RET0*, 3.600 (=2.683+4.824*0.190) bps for *RET10*, 6.734 (=5.436+6.832*0.190) bps for *RET20*, and 9.487 (=7.668+9.572*0.190) bps for *RET30*, respectively. Thus, after negative news release on day *t*, the average overall impact of *MISP* on returns is a continuation over the next 30 days. Thus, investors do consider both the level of mispricing and the sentiment of the news releases in their investment decisions.

In Panel B, for overpriced stocks, we find similar patterns with larger magnitudes of coefficients. For instance, the coefficients of *MISP*PSENT* are significant at -30.268 for *RET0* and 6.276 for *RET10*, and those of *MISP*NSENT* are significant at 27.687 for *RET0* and 9.553 for *RET10*. Also, *MISP* continues to have a significant and positive coefficient on the buy and hold returns over the 30 days following the news release. In Panel C, we fail to find evidence that *MISP* has any impact on the returns of underpriced stocks on news day *t* or over any of the subsequent periods.

There are two takeaways from the above results: (i) initial impact of speculation on overpriced stocks is an overreaction that is subsequently reversed, and (ii) the reversal of the initial overreaction following

high sentiment news and the finding that *MISP* continues to predict future returns suggest that investors consider not only the information in the news release but they also consider the level of mispricing when updating their return expectations. Overall, the initial exacerbation of already overpriced stocks on news release day *t* due to high sentiment news is likely due to a speculation-induced overreaction rather than investors' updating their beliefs about fundamental values.

Figure 1 presents the average size and book-to-market adjusted cumulative abnormal returns (CARs) around the releases of positive (in Panel A) vs. negative (in Panel B) sentiment news.[25] For overpriced stocks, the price initially moves in the same direction as news sentiment on day 0 but it experiences a reversal (continuation) over the next 20-day period for positive (negative) news release. For underpriced stocks, the patterns of price movement over the next 20-day period and the interactions between the mispricing signal and news releases are opposite to those for overpriced stocks. Figure 1 also indicates that the magnitude of market reactions to news releases on day 0 for overpriced stocks is about two times larger than that for underpriced stocks, implying that the return impacts of news releases on overpriced stocks dominate those on underpriced stocks. We will provide potential explanations for this asymmetry between overpriced vs. underpriced stocks in the next section.

4. Additional results for the return patterns

In this section, we focus on the following questions: (i) why is the return impact of positive sentiment news on overpriced stocks significantly stronger than that of negative sentiment news, and (ii) why is the return impact of news releases on overpriced stocks substantially larger than that on underpriced stocks?

4.1. Production of positive vs. negative news

Recall from the discussion following equation (3) that the impact of positive sentiment news on daily returns is about twice as large as that of negative sentiment news. We now ask why.

[25] In calculating the average CARs among overpriced vs. underpriced stocks, we exclude the past and future days that have news releases or earnings announcements for each news release day (i.e., day 0) and each stock.

Managers are incentivized to withhold bad news and to engage external investor relations (IR) firms to facilitate the dissemination of good news (see, e.g., Kothari et al. 2009; Solomon 2012). Specifically, Solomon (2012) shows that external IR firms generate more news coverage with positive sentiment for their clients by fostering and promoting positive press releases, leading to increased returns around news release days. Motivated by Solomon (2012), since the IR activity by managers or external IR firms can facilitate the production of a larger number of news articles with positive sentiment and/or news articles with stronger positive sentiment, we hypothesize that this "increased positivity" of news articles might contribute to our finding that overpriced stocks become even more overpriced on news release days.

Figure 2 shows that for an average firm in our sample, positive news articles have consistently and increasingly outnumbered negative news articles during the 20-year period covered by our sample. The enlarging imbalance of positive vs. negative news production is potentially the result of firms boosting their own IR activity or hiring external IR firms to bias media attention in their favor.

To examine its impact on mispricing, we construct a stock-month proxy for IR activity as follows. For firm *i* in month *m*, we first calculate the number of news articles with positive sentiment minus the number of news articles with negative sentiment over the preceding three-month period that ends in month *m-1*, i.e., three-month net positive news coverage. We then run a cross-sectional regression of the three-month net positive news coverage on the averages of firm size (=log of daily market capitalization) and *SENT* over the same three-month estimation period and obtain the residual. The idea is to remove the impact of firm size as larger firms have more news coverage and the impact of the news sentiment as higher average sentiment can be more attention-grabbing, leading to more coverage. This residual variable proxies for the impact of IR activity on the production of more positive news (relative to negative news). Then, at the beginning of month *m*, we sort stocks by the residual into tercile portfolios, and assign $IR_{i,t} =$

1 if the residual value belongs to the top tercile and assign $IR_{i,t} = 0$ otherwise for stock *i* on each day *t* in month *m*. We update this stock-level IR proxy each month by moving the three-month estimation window forward by one month. We interact *IR* with *MISP* and news variables and add the three-way interaction terms to equation (3). We expect the coefficient of the three-way interaction term, *MISP*PSENT*IR*, to be negative and significant if IR activity contributes to how daily returns change on news release days. Table 8 presents the test results.

The coefficient on *MISP*PSENT*IR* is significantly negative in Columns (1) and (2) suggesting that for overpriced firms with high IR activity, a positive sentiment news article exacerbates the mispricing substantially more than for firms with low IR activity. In contrast, the coefficient on *MISP*NSENT*IR* is insignificant. The return spread between positive and negative news days becomes even wider when IR activity fosters and promotes positive sentiment news.

In sum, the results are consistent with the idea that the increase in IR activity produces more positive news articles and contributes to the exacerbation of existing overpricing on positive news days.

4.2. Attention and market reaction

In this section, we investigate whether investor attention contributes to our findings. Specifically, we hypothesize that to the extent that overpriced stocks receive more investor attention, the market reaction to news releases for overpriced stocks will be stronger than that for underpriced stocks. Thus, even if news releases for overpriced stocks and underpriced stocks have similar information content, the stronger market reaction of overpriced stocks suggests that on positive news release days, prices increase by a larger amount for overpriced stocks than for underpriced stocks, thus enlarging the relative mispricing in the cross-section. On the other hand, it also implies that on negative news release days, prices drop by a larger amount for overpriced socks, thus reducing the relative mispricing in the cross-section. The asymmetric market reaction to news releases in Section 3.1, possibly driven by uneven allocation of

investor attention between overpriced and underpriced stocks, could make the increased returns of overpriced stocks on positive news days the main driver of empirical patterns we find.

To test our hypothesis, we first examine whether the overpricing is associated with investors' download activity on the SEC's EDGAR website. Then, we test whether the strength of the market reaction to the content of news releases varies with the degree of mispricing.

4.2.1. Investor attention to mispricing

Filings by firms of 10-K reports are an important source of information for investors. If overpriced stocks receive more investor attention, then there should be a higher download volume of 10-K filings immediately after they are released on SEC's EDGAR website. We examine the relation between stock-level mispricing and investor attention as follows:

$$Log\left(1 + DOWNLOAD_{i,m}\right) = \alpha + \beta_1 \times MISP_{i,m} + \beta_2 \times Log\left(MV_{i,m}\right) + \varepsilon_{i,m}, \quad (6)$$

where $DOWNLOAD_{i,m}$ is the number of unique IP addresses that download firm *i*'s 10-K in month *m*. We count only downloads that are made within 7 days of each 10-K filing day to capture the timely search of information by investors. In addition, to measure investor attention more accurately, we exclude "robot downloads" from our sample, which are defined as IP addresses that access more than 30 firms' filings on a single day, because these downloads are mostly made by computer algorithms that constantly scan the EDGAR server and automatically download newly posted 10-K filings. We control for firm size measured at the end of the previous month *m*-1 and include month and firm fixed effects to control for the impact of time trend or any time-invariant firm characteristics on EDGAR downloads. The standard errors are clustered by month and firm. The sample period for this analysis is restricted to January 2003 through June 2017 as the SEC stopped updating the download data in June 2017. As Panel A of Table 9 shows, the coefficient on *MISP* is significantly negative at -1.035 (t-statistic = -5.82), indicating that the 10-K filings of underpriced stocks are downloaded significantly less often than those of overpriced stocks. This

result is thus consistent with the idea that overpriced stocks tend to receive higher investor attention than underpriced stocks. This higher attention can lead to stronger market reactions of overpriced stocks to news releases, which we now formally test.

4.2.2. Stronger market reaction to news sentiment of overpriced stocks

In this section, we assess the return impact of the sentiment of news releases for overpriced vs. underpriced stocks. We use the following regression model to test whether the market reaction to news content of overpriced stocks is stronger than that of underpriced stocks:

$$RET_{i,t} = \alpha + \beta_1 \times SENT_{i,t} + \beta_2 \times SENT_{i,t} \times OP_{i,t} + \beta_3 \times OP_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (7)$$

where $OP_{i,t}$ is an indicator variable for an overpriced stock, which equals 1 if $MISP_{i,t} < 0$ and 0 if $MISP_{i,t} > 0$, and $MISP_{i,t}$ is the mispricing score of stock *i* measured at the beginning of the month that day *t* belongs to. $\boldsymbol{X}_{i,t}$ includes the same set of control variables as in equation (1). We control for day fixed effects and the standard errors are clustered by day and firm. Equation (7) is estimated using a sample that includes all news release days of mispriced stocks, i.e., stock-day observations with $NEWS = 1$ and $MISP \neq 0$. Panel B of Table 9 reports the results.

The coefficient on *OP* is significantly negative, suggesting that, unconditionally, overpriced stocks earn lower returns. The coefficient on *SENT* is positive at 1.331 (t-statistic = 74.42), indicating that news articles with higher sentiment earn higher returns for underpriced stocks when *OP* = 0. More importantly, the coefficient on *SENT***OP* is significantly positive at 0.669 (t-statistic = 25.87), indicating that the magnitude of market reaction to the sentiment of news increases by about 50% for overpriced stocks. Thus, investors are more attentive to the information releases of overpriced stocks as shown in Panel A, and as a result, the prices of overpriced stocks increase (decrease) by a much larger amount on positive (negative) news days than those of underpriced stocks. This implies that the stronger positive return for overpriced stocks on days that have positive sentiment news articles leads to the overall increase in deviation of

market prices from fundamental values.

5. Robustness

5.1. Three-day event window

When estimating how earnings announcements and news releases affect returns, we assumed that the event dates are accurate and that the market reaction is complete within the same day of information release. In this section, we relax this assumption and use a three-day event window to measure the reactions to earnings announcements and news releases. Specifically, we calculate $SUE_{i,t}$ if firm i makes an earnings announcement over a three-day window [t-1, t+1], and assign $SUE_{i,t} = 0$ otherwise. We calculate $SENT_{i,t}$ as the average sentiment of news articles about firm i released on day t-1, t, or t+1, and assign $SENT_{i,t} = 0$ if no news article is released on any of those three days. *PSENT* and *NSENT* are defined in a similar manner based on the three-day window. The results are reported in the internet appendix Table A.4. While the coefficients on *MISP* are comparable to those in Table 2 in all specifications, the coefficients on the interaction terms between *MISP* and the earnings and news variables are all smaller. This change is expected because while the wider event window is more robust to slight measurement errors in event dates, it also attenuates the estimated market impact of the events by averaging out the effect over three days. But more importantly, the patterns of returns on news release days of Table 2 are all preserved. Positive and negative news releases continue to exhibit the opposite impacts on daily returns.

5.2. News source

Drake et al. (2017) find that information intermediaries are not created equal – the coverage by professional intermediaries, such as Dow Jones Newswires, are associated with positive capital market effects but the coverage by nonprofessional intermediaries, such as nonfinancial websites or blogs, tend to hinder price discovery. In addition, Kogan et al. (2021) argue that falsely promotional articles about firms in social media platforms can temporarily boost firms' stock prices and trading volume by retail

investors around article release days. Motivated by these studies, we examine whether the negative effect of positive news articles on price efficiency might be caused by low quality new articles that may contain misleading information. Specifically, we repeat the tests in equations (2) and (3) with two subsamples of reliable news sources. The first subsample consists of news articles from the high quality news sources with Source Rank = 1 and complete news articles with News Type = Full Article as coded by RavenPack. The second subsample contains RavenPack news articles obtained only from the Dow Jones Newswire.[26]

Results based on the two refined subsamples of news articles are presented in the internet appendix Table A.5. We find that the effect of news days on returns documented earlier with all news articles remains robust for these subsets of high-quality news sources. Specifically, the coefficients of *MISP*SENT* and *MISP*PSENT* are all negative while that on *MISP*NSENT* is positive. The magnitude and statistical significance of these coefficients are comparable to those in Table 2. Thus, the exacerbation of stock mispricing on news release days prevails even among the most credible news sources, and it is unlikely to be driven by false information generated by news sources with low credibility.

5.3. Alternative mispricing measure

In this section, we test whether our results are sensitive to using an alternative measure of stock-level mispricing. Specifically, following Stambaugh et al. (2012), we define $MISP_{i,t}^{AR}$ based on firm *i*'s average percentile ranking of the 11 SYY signals measured at the beginning of the month which day *t* belongs to. For the ease of interpretation, $MISP_{i,t}^{AR}$ is defined to range from -1 to 1 by transforming the raw percentile ranking, which ranges from 0 to 1, by $MISP_{i,t}^{AR} = 2 \times Ranking - 1$. To be consistent with Table 2, we define $MISP^{AR}$ as a measure of relative underpricing, meaning that stocks with higher values of $MISP^{AR}$ are more underpriced. One potential advantage of $MISP^{AR}$ over *MISP* in Table 2 is that $MISP^{AR}$ reflects

[26] In the same spirit as our focusing on the news sources that have high credibility, Kogan et al. (2021) also argue that the news articles from the Wall Street Journal and New York Times are unlikely to be affected by fraud on social media platforms such as SeekingAlpha.com.

the direction as well as the strength of each anomaly signal for a given stock, while *MISP* captures the direction, but not the strength, of the signal. Repeating the tests with this alternative mispricing measure $MISP^{AR}$, we confirm that our main findings remain qualitatively the same. The internet appendix Table A.6 shows that, while the magnitudes of the coefficients are not directly comparable with those in Table 2 because the construction of *MISP* is different, the signs of the coefficients are generally consistent with those in Table 2. In particular, the coefficients of $MISP^{AR}$**SENT* and $MISP^{AR}$**PSENT* are all negative and highly significant, while that of $MISP^{AR}$**NSENT* is positive and highly significant. Thus, once again, positive sentiment news articles tend to exacerbate the overpricing and negative sentiment articles tend to correct it. We thus conclude that the negative effect of positive sentiment news articles about mispriced stocks on price efficiency is not specific to how the degree of stock-level mispricing is measured.

6. Conclusions

There is a large literature showing that LLTs improve market quality and efficiency. In this paper, we provide evidence that in certain circumstances, trading by LLTs can move prices away from fundamentals and hurt efficiency. By conditioning on stock mispricing, we are able to identify situations around news releases when LLT impairs market efficiency. Consistent with the theoretical model of De Long et al. (1990), high sentiment news releases in overpriced stocks, in the presence of naïve news-chasing investors, result in even higher prices as LLTs speculate so as to profit from expected short-term price movements. This creates a speculation-induced wedge between market prices and the consensus fundamental value, hurting price efficiency. The exacerbated overpricing is subsequently reversed. We also find that for overpriced stocks, the impact of high sentiment news on returns is significantly stronger than that of low sentiment news, which could be driven by more positive news being produced in recent years, more attention being paid to overpriced stocks by investors, and the larger return impact of news sentiment on overpriced stocks.

References

Abreu, D. and Brunnermeier, M. 2002. Synchronization risk and delayed arbitrage. *Journal of Financial Economics* 66, 341-360.

Abreu, D. and Brunnermeier, M. 2003. Bubbles and crashes. Econometrica 71, 173-204.

Allen, F. and Gorton, G. 1993. Churning bubbles, *Review of Economic Studies* 60, 813-836.

Barber, B., Huang, X., Jorion, P., Odean, T., and Schwarz, C. 2023. A (Sub)penny For Your Thoughts: Tracking Retail Investor Activity in TAQ. *The Journal of Finance, forthcoming*.

Barber, B., and Odean, T. 2008. All that glitters: The effect of attention and news on the buying behavior of individuals and institutional investors. *Review of Financial Studies* 21, 785-818.

Battalio, R., Jennings, R., Saglam, M., and Wu, J. 2023. Identifying Market Maker Traders as "Retail" from TAQ: No Shortage of False Negatives and False Positives, Working Paper.

Bhattacharya, N., Chakrabarty, B., and Wang, X., 2020. High-frequency traders and price informativeness during earnings announcements. *Review of Accounting Studies* 25, 1156-1199.

Biais, B., Foucault, T., and Moinas, S. 2015. Equilibrium fast trading. *Journal of Financial Economics* 116:292–313.

Boehmer, E., Fong, K., and Wu, J. 2021. Algorithmic Trading and Market Quality: International Evidence. *Journal of Financial and Quantitative Analysis* 56(8), 2659-2688.

Boehmer, E., Li, D., and Saar, G. 2018. The competitive landscape of high-frequency trading firms. *Review of Financial Studies* 31, 2227-2276.

Boehmer, E., Jones, C., Zhang, X., and, Zhang, X. 2021. Tracking retail investor activity. *Journal of Finance* 76, 2249-2305.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. High-Frequency Trading and Price Discovery. *Review of Financial Studies* 27, 2267-2306.

Brogaard, Jonathan, Allen Carrion, Thibaut Moyaert, Ryan Riordan, Andriy Shkilko, and Konstantin Sokolov. 2018. High Frequency Trading and Extreme Price Movements. *Journal of Financial Economics* 128(2): 253-265.

Brunnermeier, M. and Nagel, S. 2005. Hedge Funds and the Technology Bubble. *The Journal of Finance* 59(5), 2013-2040.

Brunnermeier, M. and Pedersen, L.H. 2005. Predatory Trading. *The Journal of Finance* 60(4), 1825-1863.

Carrion, Allen. 2013. "Very Fast Money: High-Frequency Trading on the NASDAQ." *Journal of Financial Markets* 16 (4): 680–711.

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., and Vega, C. 2014. Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance* 69, 2045–84.

Chakrabarty, B., Comerton-Forde, C., and Pascual, R., 2021. Identifying High Frequency Trading activity without proprietary data, Working paper.

Chakrabarty, B., Moulton, P., and Wang, X., 2022. Attention: How high-frequency trading improves price efficiency following earnings announcements, *Journal of Financial Markets* 57, 100690.

Chen, A.Y. and Zimmermann, T., 2022. Open Source Cross-Sectional Asset Pricing. *Critical Finance Review* 11(2), 207-264.

Chordia, T., Green, C., and Kottimukkalur, B. 2018. Rent Seeking by Low Latency Traders: Evidence from Trading on Macroeconomic Announcements. *Review of Financial Studies* 31, 4650-4687.

Chordia, T., Lin, T.C., and Xiang, V. 2021. Risk-neutral Skewness, Informed Trading, and the Cross-section of Stock Returns. *Journal of Financial and Quantitative Analysis* 56, 1713-1737.

Chordia, T. and Miao, B. 2020. Market Efficiency in Real Time: Evidence from Low Latency Activity around Earnings Announcements. *Journal of Accounting and Economics* 70, 101-335.

Chordia, T., Roll, R., and Subrahmanyam, A. 2011, Recent Trends in Trading Activity and Market Quality. *Journal of Financial Economics* 101, 243-263.

Conrad, J., Wahal, S., and Xiang, J. 2015. High Frequency Quoting, Trading, and the Efficiency of Prices. *Journal of Financial Economics* 116, 271-291.

Dang, T., Moshirian, F., and Zhang, B. 2015. Commonality in news around the world. *Journal of Financial Economics*, no. 2, pp. 82-100.

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

Dougal, C., Engelberg, J., Garcia, D., and Parsons, C.A., 2012. Journalists and the stock market. *Review of Financial Studies* 25: 639-679.

Drake. M.S., Rees, L., and Swanson, E.P., 2011. Should Investors Follow the Prophets or the Bears? Evidence on the Use of Public Information by Analysts and Short Sellers. *The Accounting Review* 86(1), 101-130.

Engelberg, J., McLean R.D., and Pontiff, J., 2018. Anomalies and News. *Journal of Finance* 73(5), 1971-2001.

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

Engelberg, J., Sasseville, C., and Williams, J., 2012. Market madness? The case of mad money, *Management Science* 58: 351-364.

Fama, E.G. 1965. The behavior of stock-market prices, *Journal of Business* 38, 34-105.

Fang, L. and Peress, J. 2009. Media coverage and the cross-section of stock returns. *Journal of Finance* 64, 2023-52.

Friedman, Milton. 1953. The case for flexible exchange rates, in Milton Friedman, ed.: Essays in Positive Economics. University of Chicago Press, Chicago, IL.

Garcia, D., 2013. Sentiment during recessions. *Journal of Finance* 68: 1267-1300.

Graeber, T., Roth, C. and Zimmermann, F. 2023. Stories, statistics, and memory. Working paper, University of Cologne.

Gurun, U.G. and Butler, A.W., 2011. Don't believe the hype: Local media slant, local advertising, and firm value. *Journal of Finance* 67: 561-598.

Hagströmer, B. and Nordén, L. 2013. The Diversity of High-Frequency Traders. *Journal of Financial Markets* 16, 741-770.

Harrison, M. and Kreps, D. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics* 92, 323-336.

Hart, O. and Kreps, D. 1986. Price destabilizing speculation, *Journal of Political Economy* 94, 927-952.

Hasbrouck, J. and Saar, G. 2013. Low-Latency Trading. *Journal of Financial Markets* 16, 646–79.

Hendershott, T., Jones, C.M., and Menkveld, A.J. 2011. Does Algorithmic Trading Improve Liquidity? *The Journal of Finance* 66, 1–33.

Hendershott, T., Livdan, D., and Schurhoff, N. 2015. Are institutions informed about news? *Journal of Financial Economics* 117: 249-287.

Hendershott, T. and Riordan, R. 2013. Algorithmic Trading and the Market for Liquidity. *Journal of Financial and Quantitative Analysis* 48, 1001–24.

Hillert, A., Jacobs, H., and Muller, S., 2014. Media Makes Momentum, *Review of Financial Studies* 27(12), 3467-3501.

Hirschey, N. 2021. Do High-Frequency Traders Anticipate Buying and Selling Pressure? *Management Science* 67(6), 3321-3345.

Hirshleifer, D. 2020. Presidential Address: Social Transmission Bias in Economics and Finance. *Journal of Finance* 75(4), 1779-1831.

Jarrow, R.A., and Protter, P., 2012. A dysfunctional role of high frequency trading in electronic markets. *International Journal of Theoretical and Applied Finance* 15, 219-249.

Jeon, Y., McCurdy, T.H., and Zhao, X., 2022. News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. *Journal of Financial Economics* 145(2), 1-17.

Kirilenko, Andrei, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun. 2017." The Flash Crash: The Impact of High Frequency Trading on an Electronic Market." *Journal of Finance* 72 (3): 967-998.

Kogan, S., Moskowitz, T.J., and Niessner, M., 2022. Social Media and Financial News Manipulation. *Review of Finance* 1-40.

Kothari S.P., Shu, S., and Wysocki P.D., 2009. Do Managers Withhold Bad News? *Journal of Accounting Research* 47(1), 241-276.

Loughran T. and McDonald, B., 2011. When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks, *Journal of Finance* 66, 35-65.

McLean, R.D. and Pontiff, J. 2016. Does Academic Research Destroy Stock Return Predictability? *Journal of Finance* 71(1), 5-32.

O'Hara, M., 2015. High Frequency Market Microstructure. *Journal of Financial Economics* 116, 257-270.

Solomon, D., 2012. Selective Publicity and Stock Prices. *Journal of Finance* 67(2), 599-638.

Solomon, D., Soltes, E.F., and Sosyura, D., 2014. Winners in the spotlight: Media coverage of fund holdings as a drive of flows. *Journal of Financial Economics* 113: 53-72.

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

Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62: 1139-1168.

Tetlock, P.C., 2011. All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies* 24: 1481-1512.

Table 1: Summary Statistics

This table presents various summary statistics of the sample, which includes earnings announcement and news release days and all US domestic common stocks with stock prices above $5 from January 2000 to December 2019.

Panel A: Frequencies of Earnings Announcements and News Release Days

Earnings Day	News Day		Total
	No	Yes	
No	22,677,120 (79.11%)	5,660,081 (19.75%)	28,337,201 (98.85%)
Yes	143,571 (0.50%)	185,119 (0.65%)	328,690 (1.15%)
Total	22,820,691 (79.61%)	5,845,200 (20.39%)	28,665,891 (100%)

Panel B: Summary Statistics of *MISP* Variables

	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
Long	1.607	1.639	0	0	1	3	10
Short	1.224	1.516	0	0	1	2	11
MISP	0.382	2.184	-11	-1	0	1	10

Panel C: Distribution of *MISP*

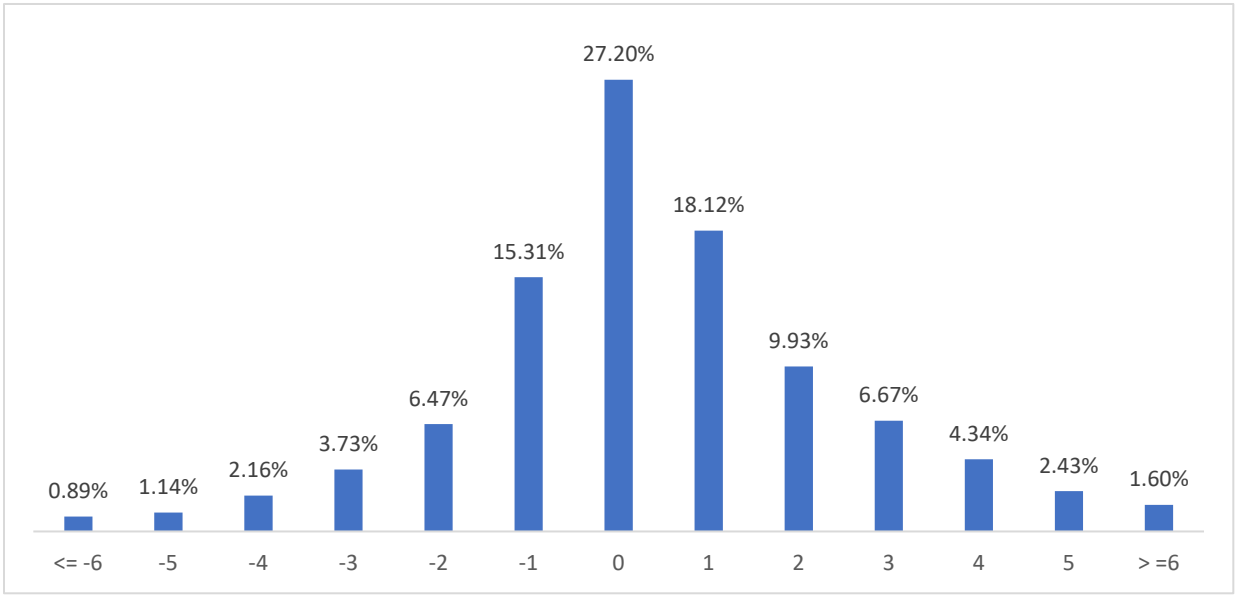


Table 2: Returns on Earnings Days and News Days

This table presents the test results that examine how stock mispricing in the beginning of a month relates to daily returns on earnings announcement days and news release days during that month. We estimate each of the following panel regressions separately:

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times EARN_{i,t} + \beta_3 MISP_{i,t} \times NEWS_{i,t} + \beta_4 EARN_{i,t} + \beta_5 NEWS_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (1)$$

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times SUE_{i,t} + \beta_3 MISP_{i,t} \times SENT_{i,t} + \beta_4 SUE_{i,t} + \beta_5 SENT_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (2)$$

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times SUE_{i,t} + \beta_3 MISP_{i,t} \times PSENT_{i,t} + \beta_4 MISP_{i,t} \times NSENT_{i,t} + \beta_5 SUE_{i,t} + \beta_6 PSENT_{i,t} + \beta_7 NSENT_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}. \quad (3)$$

where $RET_{i,t}$ is the return of stock *i* on day *t*, and $MISP_{i,t}$ is the mispricing score measured at the beginning of the month which day *t* belongs to and it is constructed based on the 11 anomaly signals of Stambaugh, Yu, and Yuan (2012). The higher *MISP* is, the more underpriced the stock is. In equation (1), $EARN_{i,t}$ is an indicator variable for earnings announcement days: *EARN* = 1 when earnings is announced on day *t* and *EARN* = 0 otherwise. $NEWS_{i,t}$ is an indicator variable for news release days: *NEWS* = 1 when a news is released on day *t* and *NEWS* = 0 otherwise. In equation (2), $SUE_{i,t}$ is the standardized unexpected earnings and defined as the actual EPS announced on day *t* minus the consensus forecast of the EPS, scaled by the stock price at the fiscal quarter end of the EPS. As the consensus forecast, we use the median of all earnings forecasts for each firm followed by security analysts or the actual EPS for the same fiscal quarter in the last fiscal year for each firm not followed by security analysts. We winsorize *SUE* at -1 and 1 to mitigate potential effects of outliers, and assign *SUE* = 0 for non-earnings days that have *EARN* = 0. $SENT_{i,t}$ is the average sentiment score of news articles released on day *t* and defined as (average of ESS-50)/50, where ESS is a sentiment variable from RavenPack. Thus, *SENT* ranges from -1 to 1 and we assign *SENT* = 0 for non-news days that have *NEWS* = 0. The higher *SENT* is, the more optimistic the average sentiment is. In equation (3), *PSENT* is the absolute value of the average sentiment of the news articles publicized on positive news days and it is set to zero on negative news days. Similarly, *NSENT* is the absolute value of the average sentiment of the news articles publicized on negative news days and it is set to zero on positive news days. $\boldsymbol{X}_{i,t}$ is a (column) vector that contains three sets of control variables: the values of returns, volatility (return squared), and share trading volume over the past ten days before day *t*. Overpriced stocks are those in the bottom 30% of the *MISP* distribution and underpriced stocks are those in the top 30% of the *MISP* distribution. The sample period of the analysis is January 2000 through December 2019, which is restricted by the availability of RavenPack. Day fixed effects are included in estimating equations (1) to (3). For expositional convenience, we divide *MISP* by 100 and multiply *RET* by 100 before estimating equations (1) to (3). The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Model (1)

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.996***	1.724***	0.209
	(8.04)	(6.14)	(1.44)
MISP x *EARN*	4.798***	13.323***	-1.016
	(7.18)	(6.69)	(-0.59)
MISP x *NEWS*	-0.967***	-1.008***	-0.143
	(-7.44)	(-2.90)	(-0.71)
EARN	0.044**	0.204***	0.179***
	(2.39)	(4.30)	(3.05)
NEWS	0.047***	0.044***	0.034***
	(7.86)	(4.63)	(4.27)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,665,891	8,526,782	7,405,942
Adj. R^2	0.004	0.003	0.003

Panel B: Model (2)

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.891***	2.072***	0.016
	(7.55)	(7.35)	(0.13)
MISP x *SUE*	6.749	-0.290	41.530*
	(1.20)	(-0.02)	(1.92)
MISP x *SENT*	-13.467***	-29.692***	0.339
	(-27.29)	(-19.40)	(0.37)
SUE	2.386***	2.314***	1.581***
	(14.50)	(5.94)	(2.61)
SENT	1.602***	1.299***	1.248***
	(81.52)	(34.63)	(38.52)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,665,891	8,526,782	7,405,942
Adj. R^2	0.008	0.008	0.007

Panel C: Model (3)

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.924***	1.996***	0.040
	(7.76)	(7.10)	(0.30)
MISP x *SUE*	6.839	-0.251	41.606*
	(1.22)	(-0.02)	(1.93)
MISP x *PSENT*	-13.404***	-29.220***	0.350
	(-23.15)	(-16.14)	(0.34)
MISP x *NSENT*	13.234***	29.951***	-0.447
	(16.85)	(13.01)	(-0.29)
SUE	2.384***	2.311***	1.579***
	(14.49)	(5.94)	(2.61)
PSENT	1.512***	1.220***	1.180***
	(65.12)	(27.38)	(31.27)
NSENT	-1.802***	-1.477***	-1.394***
	(-56.16)	(-24.72)	(-25.41)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,665,891	8,526,782	7,405,942
Adj. R^2	0.008	0.008	0.007

Table 3: Impact of the Retail Trading Activity on Returns on News Days

This table examines the retail trading activity on news release days and how it affects the returns. For stock *i* on day *t*, $RETAIL_OIB_{i,t}$ is the order imbalance of retail trading, defined as (retail buy share volume – retail sell share volume) / (retail buy share volume + retail sell share volume). $RETAIL_{i,t}$ is the retail trading volume, defined as retail share volume / total share volume. The retail trades are classified by the algorithm of Boehmer et al. (2021), and the daily summary measures come from the WRDS Intraday Indicator Data. Panel A explores factors that can affect *RETAIL_OIB* on news days. Panel B investigates the impacts of retail trading activity on returns on news release days. *DRT* is the cross-sectional decile ranking of *RETAIL* for each day and ranges from 0 to 1. We interact *DRT* with *MISP* and the news variables, add the three-way interaction terms into equation (3), and re-estimate it. The sample period for this analysis is from January, 2010 to December, 2015 by following Boehmer et al. (2021). The other variables are defined in Table 2. In panel A, day and firm fixed effects are included, and standard errors are clustered by day and firm. In Panel B, day fixed effects are included and standard errors of coefficients are clustered by day and firm. *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Retail Trading Activity on News Release Days

	RETAIL_OIB
SENT	0.015***
	(10.77)
MISP	0.038
	(1.37)
RET	0.050***
	(2.92)
Controls	Yes
Fixed Effects	Day / Firm
Num. of Obs.	1,580,019
Adj. R^2	0.028

Panel B: Return Impact of Retail Trading Activity on News Release Days

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	-0.141	-0.877*	-0.619**
	(-0.91)	(-1.91)	(-2.39)
MISP x *DRT*	1.416***	4.549***	1.263*
	(4.63)	(4.96)	(1.95)
MISP x *SUE*	67.796***	72.698**	306.560*
	(3.26)	(2.11)	(1.90)
MISP x *PSENT*	-5.114***	16.494***	-7.316***
	(-3.76)	(2.99)	(-2.73)
MISP x *PSENT* x *DRT*	-22.359***	-113.040***	20.579**
	(-5.46)	(-8.52)	(2.39)
MISP x *NSENT*	2.311	-11.166*	8.738***
	(1.39)	(-1.71)	(2.81)
MISP x *NSENT* x *DRT*	28.124***	88.308***	-17.130*
	(5.91)	(5.88)	(-1.66)
SUE	7.728***	8.394***	2.171
	(8.83)	(5.48)	(0.51)
PSENT	0.835***	1.361***	0.725***
	(19.98)	(11.36)	(7.76)
PSENT x *DRT*	2.388***	0.450	1.737***
	(19.58)	(1.49)	(6.04)
NSENT	-0.624***	-0.956***	-0.644***
	(-11.29)	(-6.55)	(-5.62)
NSENT x *DRT*	-3.006***	-1.693***	-2.128***
	(-18.26)	(-4.32)	(-5.75)
DRT	-0.031**	0.023	-0.063**
	(-2.38)	(1.02)	(-2.50)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	5,345,188	1,723,192	1,475,992
Adj. R^2	0.011	0.013	0.012

Table 4: Impact of the LLT Activity on Returns on News Day

This table investigates how the low latency trading (LLT) activity affects returns on news release days. Following Hasbrouck and Saar (2013), we measure LLT activity as the number of time-weighted strategic runs for stock *i* on day *t,*

$$LLT_{i,t} = \frac{1}{2.34\times10^{7}}\sum_{j=1}^{N_{i,t}} T_{i,j,t}, \quad (4)$$

where $T_{i,j,t}$ is the time-in-force in milliseconds for the *j*th strategic run, $N_{i,t}$ is the total number of strategic runs, and the total number of milliseconds available for each trading day is 2.34×10^{7}. Following Chordia and Miao (2020), we also measure the aggressiveness of LLT trading activity depending on news sentiment. Specifically, we define *LLT_SIGN* as, for stock *i* on day *t*,

$$LLT_SIGN_{i,t} = \frac{1}{2.34\times10^{7}}\left[\left(\sum_{j=1}^{B_{i,t}} T_{i,j,t}^{Bid}\right) - \left(\sum_{k=1}^{A_{i,t}} T_{i,k,t}^{Ask}\right)\right] \times Sign\left(SENT_{i,t}\right), \quad (5)$$

where $T_{i,j,t}^{Bid}$ is the time-in-force in milliseconds for the *j*th strategic buy run at the bid price and $T_{i,k,t}^{Ask}$ is the time-in-force in milliseconds for the *k*th strategic sell run at the ask price, $B_{i,t}$ and $A_{i,t}$ are the numbers of strategic buy runs and sell runs, respectively, and $Sign\left(SENT_{i,t}\right)$ has a value of 1 if *SENT* > 0, -1 if *SENT* < 0, and 0 otherwise. Panel A examines the intensity of overall LLT activity on news release days, using a natural-log transformed variable *Log(1+LLT)* as the dependent variable. We then define two dummy variables, *DLLT* and *DLLTS*, indicating the top decile of *LLT* and *LLT_SIGN*, respectively, for each day. We interact *DLLT* or *DLLTS* with *MISP* and news variables, add the three-way interaction terms into equation (3), and re-estimate it. Panels B reports the impact of overall LLT activity on returns on news release days, and panel C reports the impact of aggressive LLT activity that competes for liquidity with other market participants. The sample for this analysis contains Nasdaq-listed common stocks only and starts in January 2008 and ends in December 2017, which is restricted by the availability of *LLT* and *LLT_SIGN*. The other variables are defined in Table 2. Day and firm fixed effects are included in the regressions of Panel A, and day fixed effects are included in the regressions of Panels B and C. The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Overall LLT Activity Intensity on News Release Days

	(1)	(2)	(3)
	Log(1+LLT)	*Log(1+LLT)*	*Log(1+LLT)*
MISP	-0.577***	-0.581***	-0.577***
	(-5.68)	(-5.70)	(-5.67)
NEWS	0.036***		
	(27.45)		
SENT		0.012***	
		(4.90)	
PSENT			0.069***
			(17.29)
NSENT			0.093***
			(15.97)
Fixed effects	Day / Firm	Day / Firm	Day / Firm
Num. of Obs.	7,405,791	7,405,791	7,405,791
Adj. R^2	0.002	0.001	0.001

Panel B: Overall LLT Activity

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.461***	1.359***	-0.051
	(2.89)	(3.59)	(-0.32)
MISP x *DLLT*	0.384	2.592*	-0.666*
	(1.06)	(1.84)	(-1.78)
MISP x *SUE*	94.291***	1.130	304.760**
	(5.67)	(0.04)	(2.09)
MISP x *PSENT*	-14.062***	-26.008***	-1.253
	(-17.16)	(-9.25)	(-0.83)
MISP x *PSENT* x *DLLT*	-19.767***	-55.008***	1.859
	(-7.34)	(-3.66)	(0.57)
MISP x *NSENT*	14.872***	24.751***	3.771*
	(14.08)	(6.89)	(1.82)
MISP x *NSENT* x *DLLT*	22.193***	52.684***	3.309
	(6.92)	(3.79)	(0.73)
SUE	7.539***	4.448***	4.099
	(10.05)	(2.73)	(1.03)
PSENT	1.716***	1.505***	1.351***
	(57.95)	(22.19)	(24.13)
PSENT x *DLLT*	0.493***	-0.086	-0.252**
	(5.58)	(-0.28)	(-2.12)
NSENT	-1.772***	-1.609***	-1.452***
	(-46.32)	(-17.76)	(-19.27)
NSENT x *DLLT*	-0.658***	-0.057	-0.069
	(-6.13)	(-0.18)	(-0.40)
DLLT	-0.057***	-0.001	-0.015
	(-3.79)	(-0.02)	(-0.90)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	7,405,791	2,092,590	2,737,735
Adj. R^2	0.013	0.012	0.016

Panel C: Aggressive LLT Activity

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.460***	1.394***	-0.090
	(2.76)	(3.62)	(-0.56)
MISP x *DLLTS*	0.665	3.724**	-0.455
	(1.56)	(2.13)	(-0.73)
MISP x *SUE*	94.708***	1.736	304.770**
	(5.70)	(0.06)	(2.09)
MISP x *PSENT*	-16.376***	-29.422***	-1.259
	(-16.51)	(-8.22)	(-0.82)
MISP x *PSENT* x *DLLTS*	-4.411**	-18.312**	0.513
	(-2.30)	(-2.37)	(0.16)
MISP x *NSENT*	11.142***	19.393***	4.142**
	(9.98)	(4.89)	(2.02)
MISP x *NSENT* x *DLLTS*	18.686***	25.644***	1.962
	(7.71)	(2.60)	(0.46)
SUE	7.524***	4.437***	4.097
	(10.04)	(2.73)	(1.03)
PSENT	1.727***	1.509***	1.282***
	(52.87)	(18.72)	(22.68)
PSENT x *DLLTS*	0.535***	0.229	0.381***
	(9.52)	(1.36)	(3.28)
NSENT	-1.612***	-1.456***	-1.400***
	(-39.86)	(-15.03)	(-18.64)
NSENT x *DLLTS*	-0.289***	-0.223	0.164
	(-3.76)	(-0.96)	(1.03)
DLLTS	-0.151***	-0.087**	-0.111***
	(-11.61)	(-2.25)	(-4.74)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	7,405,791	2,092,590	2,737,735
Adj. R^2	0.013	0.012	0.016

Table 5: Impacts of the Retail Trading and LLT Activity on Returns on News Days

This table investigates how the interaction of retail trading activity with LLT activity affects returns on news release days. Panel A examines how LLT activity varies with retail activity, Panel B presents how this interaction affects returns, and Panel C extends the mispricing return analysis by separately examining news released during different periods of the day: before market open (BMO, 0:00:00–9:29:59), morning session (AM, 9:30:00–13:00:00), afternoon session (PM, 13:00:01–16:00:00), and after market close (AMC, 16:00:01–23:59:59). For each day, we define *DRT* as the daily decile ranking of *RETAIL*, which ranges from 0 to 1, and *DLLT* as the dummy variable indicating the top daily decile of *LLT*, respectively. For Panel B, we add all interaction terms of *DLLT*, *DRT*, *MISP*, *PSENT*, and *NSENT* into equation (3) and re-estimate it. In Panel C, *PSENT* and *NSENT* of each intraday period are separately measured and denoted by _BMO/AM/PM/AMC. For expositional brevity, we report only the three-way and four-way interaction terms in Panel B, and only the four-way interaction terms based on *PSENT* in Panel C. See the internet appendix Table A.7 for the full estimation results. Following Boehmer et al. (2021) for retail trading activity, the sample period is from January, 2010 to December, 2015. For the other aspects of the tests, see the legends in Tables 2 to 4. In all regressions, day fixed effects are included in the estimation and standard errors of coefficients are clustered by day and firm. *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: LLT and Retail Activity on News Days

	(1)	(2)	(3)
	Log(1+LLT)	*Log(1+LLT)*	*Log(1+LLT)*
MISP	-0.539***	-0.546***	-0.541***
	(-4.23)	(-4.28)	(-4.24)
DRT	-0.040***	-0.021***	-0.030***
	(-8.31)	(-4.35)	(-6.24)
DRT x *NEWS*	0.052***		
	(21.03)		
DRT x *SENT*		0.018**	
		(2.17)	
DRT x *PSENT*			0.135***
			(9.46)
DRT x *NSENT*			0.198***
			(10.22)
NEWS	-0.002		
	(-0.71)		
SENT		-0.015***	
		(-3.22)	
PSENT			-0.048***
			(-6.37)
NSENT			-0.047***
			(-4.34)
Fixed effects	Day / Firm	Day / Firm	Day / Firm
Num. of Obs.	3,804,475	3,804,475	3,804,475
Adj. R^2	0.008	0.002	0.001

Panel B: The Impact of LLT and Retail Activity on Returns

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP x *DRT* x *DLLT*	0.655	-0.192	-1.280
	(0.62)	(-0.06)	(-0.68)
MISP x *PSENT* x *DRT*	-11.750***	33.930***	-12.349*
	(-3.12)	(2.99)	(-1.68)
MISP x *PSENT* x *DLLT*	1.889	20.135	-2.698
	(0.69)	(1.33)	(-0.68)
MISP x *PSENT* x *DRT* x *DLLT*	-29.945***	-163.950***	43.884***
	(-4.74)	(-11.78)	(6.54)
MISP x *NSENT* x *DRT*	13.505***	-44.085***	17.359*
	(2.94)	(-2.88)	(1.86)
MISP x *NSENT* x *DLLT*	-3.092	-23.119	2.933
	(-0.89)	(-1.51)	(0.52)
MISP x *NSENT* x *DRT* x *DLLT*	25.982***	137.320***	-44.629***
	(3.55)	(8.81)	(-5.42)
PSENT x *DRT* x *DLLT*	0.960***	0.408	-0.687**
	(3.57)	(0.68)	(-2.44)
NSENT x *DRT* x *DLLT*	-0.928***	-0.169	0.717*
	(-2.72)	(-0.21)	(1.90)
Other interaction terms	Yes	Yes	Yes
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	3,804,475	1,084,497	1,433,084
Adj. R^2	0.014	0.018	0.013

Panel C: The Impact of LLT and Retail Activity on Returns by Intraday Timing of News Release

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
*MISP*PSENT_BMO*DRT*DLLT*	-48.745*	-128.920	-29.061
	(-1.81)	(-0.96)	(-0.61)
*MISP*PSENT_AM*DRT*DLLT*	-131.520***	-500.620***	89.000
	(-3.13)	(-2.75)	(1.41)
*MISP*PSENT_PM*DRT*DLLT*	-52.588*	-223.390*	-40.465
	(-1.67)	(-1.92)	(-0.72)
*MISP*PSENT_AMC*DRT*DLLT*	55.743	114.270	17.031
	(1.47)	(0.53)	(0.29)
Other interaction terms	Yes	Yes	Yes
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	3,804,475	1,084,497	1,433,084
Adj. R^2	0.028	0.034	0.027

Table 6: Returns on Earnings Days and News Days

– Using Stale News

This table presents the test results that examine how the release of stale news affects returns. Specifically, we re-construct the news variables, *PSENT* and *NSENT*, in equation (3) using only news articles with the novelty score less than 75 (out of 100), i.e., ENS < 75, in the RavenPack database. In addition, we estimate equation (3) after excluding firm-day observations that have the releases of news articles whose novelty scores are 100, i.e., ENS = 100. For the other aspects of the tests, see the legend in Table 2. The sample period of our analysis is from January, 2000 to December, 2019. Day fixed effects are included in estimating equation (3). The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	1.039***	1.855***	0.216
	(8.42)	(6.51)	(1.52)
MISP x *SUE*	1.369	11.727	9.617
	(0.25)	(0.91)	(0.46)
MISP x *PSENT*	-3.910***	-11.380***	-0.488
	(-5.21)	(-3.97)	(-0.30)
MISP x *NSENT*	2.231**	6.148*	-1.853
	(2.03)	(1.65)	(-0.93)
SUE	1.598***	2.007***	1.419**
	(11.45)	(5.40)	(2.42)
PSENT	0.213***	0.017	0.154***
	(9.42)	(0.26)	(2.70)
NSENT	-0.298***	-0.226**	-0.145*
	(-8.82)	(-2.32)	(-1.96)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	22,820,691	6,877,813	5,155,542
Adj. R^2	0.005	0.004	0.005

Table 7: Post-news-day Returns

This table examines the returns of mispriced stocks over the subsequent periods after the release of positive or negative sentiment news. Specifically, we extend the tests of equation (3) in Table 2 by calculating the buy and hold return over each of the subsequent periods of (*t*+2, *t*+10), (*t*+2, *t*+20), and (*t*+2, *t*+30) after news release day *t* and by using it as the dependent variable, where stock-day observations with news releases are used to estimate the regression models. We define the following four dependent variables. $RET0_{i,t}$ is the return of stock *i* on news release day *t*, $RET_10_{i,t}$ is the 10-day buy-and-hold return of stock *i* from day *t+2 to t+10*, $RET_20_{i,t}$ is the 20-day buy-and-hold return of stock *i* from day *t+2 to t+20*, and $RET_30_{i,t}$ is the 30-day buy-and-hold return of stock *i* from day *t+2 to t+30*. *RET_10, RET_20,* and *RET_30* are calculated using returns on days with no future news releases or earnings announcement relative to news day *t*. For the other aspects of the tests, see the legend of Table 2. The sample period of our analysis is from January, 2000 to December, 2019. Panel A presents the results for all stocks, while Panels B and C provide those for overpriced stocks and underpriced stocks, respectively. Day fixed effects are included in estimating the regression models. The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: All Stocks

	(1) *RET0*	(2) *RET_10*	(3) *RET_20*	(4) *RET_30*
MISP	0.581***	2.683***	5.436***	7.668***
	(3.53)	(7.86)	(9.88)	(10.34)
MISP x *PSENT*	-11.830***	2.535***	4.187***	6.549***
	(-19.70)	(2.89)	(3.21)	(3.98)
MISP x *NSENT*	14.241***	4.824***	6.832***	9.572***
	(18.16)	(3.80)	(3.63)	(3.94)
MISP x *SUE*	16.299*	1.012	10.716	5.163
	(1.70)	(0.12)	(0.942)	(0.38)
PSENT	1.652***	0.059**	0.074**	0.072*
	(77.86)	(2.56)	(2.12)	(1.66)
NSENT	-1.579***	-0.141***	-0.162***	-0.224***
	(-62.04)	(-4.18)	(-3.19)	(-3.43)
SUE	3.351***	0.784***	1.094***	1.231***
	(11.67)	(4.05)	(4.30)	(3.93)
Control Variables	Yes	Yes	Yes	Yes
Fixed Effects	Day	Day	Day	Day
Num. of Obs.	5,555,092	5,555,092	5,555,092	5,555,092
Adj. R^2	0.015	0.001	0.001	0.001

Panel B: Overpriced Stocks

	(1) *RET0*	(2) *RET_10*	(3) *RET_20*	(4) *RET_30*
MISP	2.666***	7.382***	15.999***	22.474***
	(5.83)	(6.70)	(8.48)	(8.86)
MISP x *PSENT*	-30.268***	6.276**	4.466	7.308
	(-14.97)	(1.97)	(0.92)	(1.19)
MISP x *NSENT*	27.687***	9.553**	4.227	2.527
	(11.34)	(2.19)	(0.64)	(0.30)
MISP x *SUE*	-12.169	14.074	25.938	8.233
	(-0.57)	(0.66)	(0.94)	(0.24)
PSENT	1.296***	0.193**	0.135	0.188
	(27.30)	(2.55)	(1.12)	(1.22)
NSENT	-1.313***	0.030	-0.158	-0.291
	(-21.98)	(0.28)	(-0.96)	(-1.36)
SUE	2.642***	1.162**	1.473**	1.316
	(3.53)	(2.05)	(1.99)	(1.39)
Control Variables	Yes	Yes	Yes	Yes
Fixed Effects	Day	Day	Day	Day
Num. of Obs.	1,600,920	1,600,920	1,600,920	1,600,920
Adj. R^2	0.018	0.001	0.001	0.001

Panel C: Underpriced Stocks

	(1) *RET0*	(2) *RET_10*	(3) *RET_20*	(4) *RET_30*
MISP	-0.311	-0.442	-0.837	-0.853
	(-1.41)	(-0.81)	(-0.90)	(-0.66)
MISP x *PSENT*	1.365	1.851	2.396	4.263
	(1.18)	(0.97)	(0.88)	(1.20)
MISP x *NSENT*	0.152	3.411	2.499	1.623
	(0.09)	(1.32)	(0.64)	(0.32)
MISP x *SUE*	78.226**	-33.450	-10.329	25.518
	(2.20)	(-1.32)	(-0.27)	(0.62)
PSENT	1.318***	0.003	0.040	-0.005
	(31.81)	(0.05)	(0.41)	(-0.04)
NSENT	-1.235***	-0.150	-0.056	-0.024
	(-21.99)	(-1.61)	(-0.39)	(-0.13)
SUE	1.903*	1.962**	1.729	0.679
	(1.85)	(2.00)	(1.31)	(0.51)
Control Variables	Yes	Yes	Yes	Yes
Fixed Effects	Day	Day	Day	Day
Num. of Obs.	2,097,023	2,097,023	2,097,023	2,097,023
Adj. R^2	0.015	0.000	0.000	0.001

Table 8: Returns on Earnings Days and News Days

– Investor Relations Activities

This table examines how the production of positive vs. negative news affects returns on news release days by interacting *MISP* and news variables with our monthly proxy for investor relations (IR) activities. We construct our stock-level monthly IR proxy as follows. For firm *i* in month *m*, we first calculate the number of news articles with positive sentiment minus the number of news articles with negative sentiment over the preceding three-month period that ends at month *m*-1, producing the "three-month net positive news coverage". We then run a cross-stock regression of the three-month net positive news coverage on the averages of firm size (=log of daily market capitalization) and *SENT* over the same three-month period, and obtain the residual. Finally, in the beginning of month *m*, we sort stocks by the residual into tercile portfolios, and assign $IR_{i,t} = 1$ if the residual value belongs to the top tercile and assign $IR_{i,t} = 0$ otherwise for stock *i* on each day *t* in month *m*. We update this IR proxy each month by moving the three-month estimation window forward by one month. We then add the three-way IR interaction terms with *MISP* and new variables to equation (3). The other variables are defined in Table 2. The sample period of our analysis is from January, 2000 to December, 2019. Day fixed effects are included in the estimation. The standard errors of coefficients are clustered by day and firm, and *t*-statistics are in parenthesis.

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.931***	1.907***	-0.014
	(7.44)	(6.42)	(-0.10)
MISP x *IR*	0.005	0.550***	-0.037
	(0.07)	(2.64)	(-0.22)
MISP x *SUE*	6.623	-3.519	48.347**
	(1.18)	(-0.29)	(2.20)
MISP x *PSENT*	-11.846***	-26.244***	0.613
	(-18.61)	(-12.78)	(0.52)
MISP x *PSENT* x *IR*	-3.729***	-8.546***	0.124
	(-4.53)	(-2.76)	(0.07)
MISP x *NSENT*	12.513***	29.446***	-2.580
	(13.89)	(10.24)	(-1.37)
MISP x *NSENT* x *IR*	1.694	-0.006	6.344**
	(1.49)	(-0.00)	(2.53)
SUE	2.373***	2.206***	1.337**
	(14.49)	(5.66)	(2.23)
PSENT	1.500***	1.226***	1.200***
	(61.19)	(24.58)	(27.24)
PSENT x *IR*	-0.011	-0.068	-0.107*
	(-0.50)	(-0.98)	(-1.84)
NSENT	-1.883***	-1.544***	-1.425***
	(-52.91)	(-21.00)	(-20.89)
NSENT x *IR*	0.215***	0.153	0.050
	(7.14)	(1.55)	(0.56)
IR	-0.007***	0.005	-0.008
	(-2.82)	(0.95)	(-1.36)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,077,826	8,401,970	7,273,833
Adj. R^2	0.008	0.008	0.008

Table 9: Investor Attention to Mispricing and Reaction to News Sentiment

In Panel A, we present the test result that examines the relation between stock-level mispricing and the level of investor attention, where investors' attention is proxied by the download volume of 10-K filings from SEC's EDGAR website. Specifically, we estimate the following panel regression:

$$Log\left(1 + DOWNLOAD_{i,m}\right) = \alpha + \beta_1 \times MISP_{i,m} + \beta_2 \times Log\left(MV_{i,m}\right) + \varepsilon_{i,m}, \quad (6)$$

where $DOWNLOAD_{i,m}$ is the number of unique IP addresses that download firm i's 10-K in month m. We count only downloads that are made within seven days of each 10-K filing day to capture the timely search of information by investors. Robot downloads (IP addresses that download more than 30 firms' filings on a single day) are excluded from the count. *MISP* is measured at the beginning of each month m and divided by 100. *MV* is the market value of equity at the end of the previous month m-1. Month and firm fixed effects are included in estimating equation (6). In Panel B, we present the test result that examines whether the market reaction to news contents of overpriced stocks is stronger than that of underpriced stocks. Specifically, we estimate the following panel regression:

$$RET_{i,t} = \alpha + \beta_1 \times SENT_{i,t} + \beta_2 \times SENT_{i,t} \times OP_{i,t} + \beta_3 \times OP_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (7)$$

where $OP_{i,t}$ is an indicator variable for an overpriced stock, which equals to 1 if $MISP_{i,t} < 0$ and 0 if $MISP_{i,t} > 0$. $MISP_{i,t}$ is the mispricing score of stock i measured at the beginning of the month that day t belongs to. $SENT_{i,t}$ is the average sentiment score of news articles released on day t. $\boldsymbol{X}_{i,t}$ includes the same set of control variables as in Table 2. We estimate equation (7) using the sample that includes all news release days of mispriced stocks, i.e., stocks with *NEWS* = 1 and *MISP* ≠ 0 and with day fixed effects included in the model. The sample period is January, 2000 through December, 2019 for Panel B and January, 2003 through June, 2017 for Panel A, which is restricted by the availability of *Download*. The standard errors of coefficients are clustered by month and firm in Panel A, and by day and firm in Panel B, and t-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Investor Attention to Mispricing

	Log(1+DOWNLOAD)
MISP	-1.035***
	(-5.82)
Log(MV)	0.048***
	(3.94)
Fixed effects	Month / Firm
Num. of Obs.	61,667
Adj. R^2	0.898

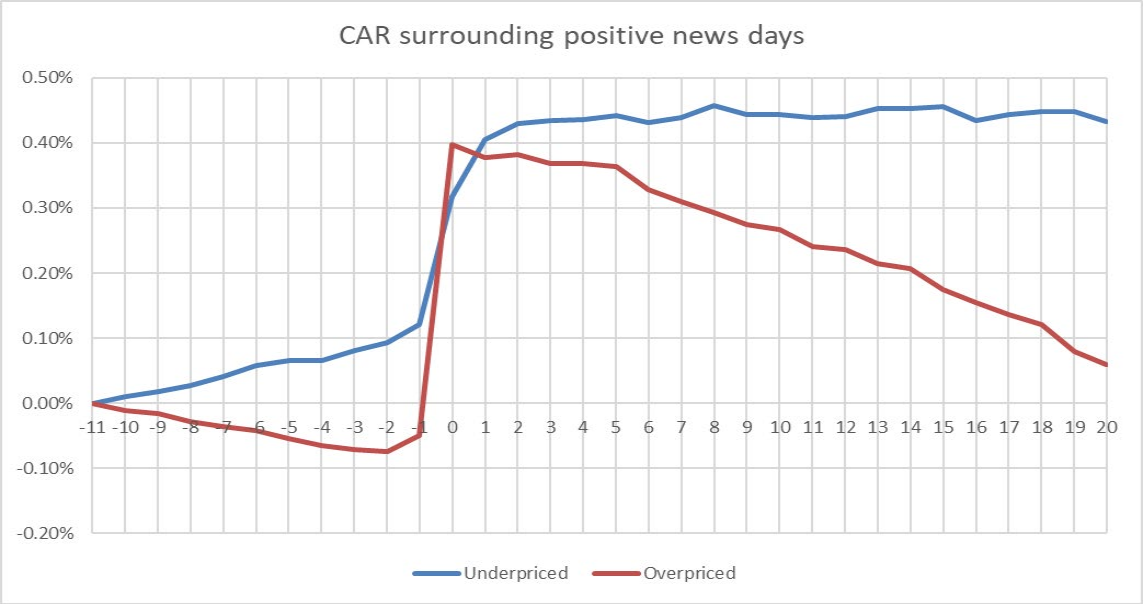
Panel B: Market Reaction to News Sentiment

	RET
SENT	1.331***
	(74.42)
SENT x *OP*	0.669***
	(25.87)
OP	-0.038***
	(-5.07)
Control variables	Yes
Fixed effects	Day
Num. of Obs.	5,845,200
Adj. R^2	0.014

Figure 1: Cumulative Abnormal Returns surrounding News Release Days

This figure presents the plots of the average cumulative abnormal returns (CARs) from days -11 to 20, where day 0 indicates the news release day, for positive sentiment news (in Panel A) and for negative sentiment news (in Panel B). CAR is a size and book-to-market adjusted return, i.e., we subtract the return of the matched value-weighted size and book-to-market sorted 5 by 5 portfolio from the return of each stock for each day. In calculating the average CARs among overpriced vs. underpriced stocks, we exclude the past and future days that have news releases or earnings announcements for each news release day and each stock. Underpriced stocks (in blue solid graphs) and overpriced stocks (in red solid graphs) are, respectively, those within top and bottom 30% of *MISP* at the beginning of each month.

Panel A: Positive News Days (around day 0)



Panel B: Negative News Days (around day 0)

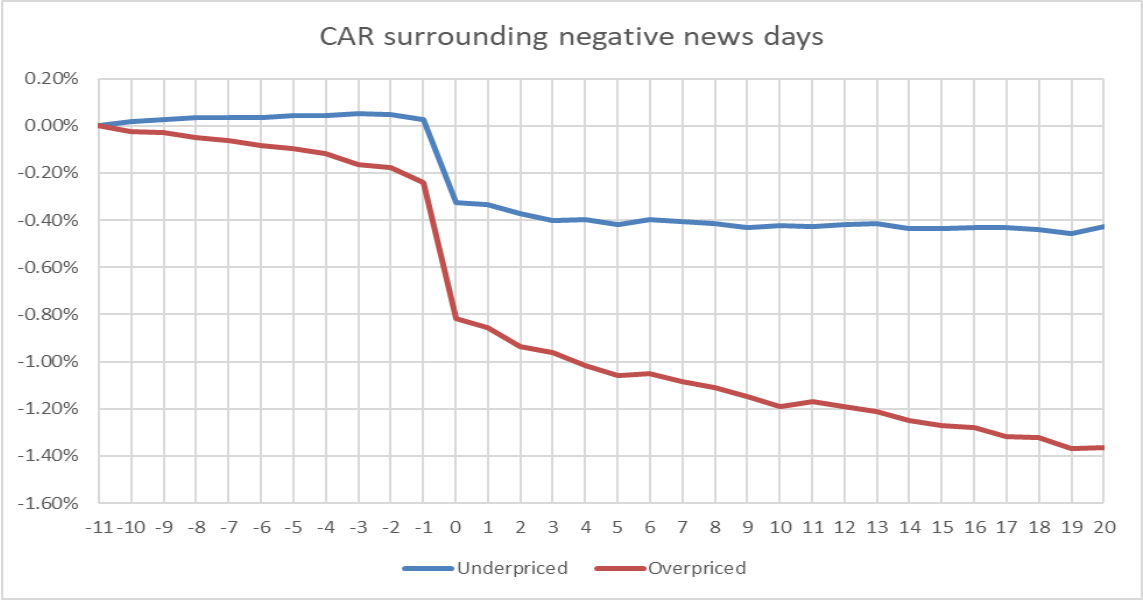
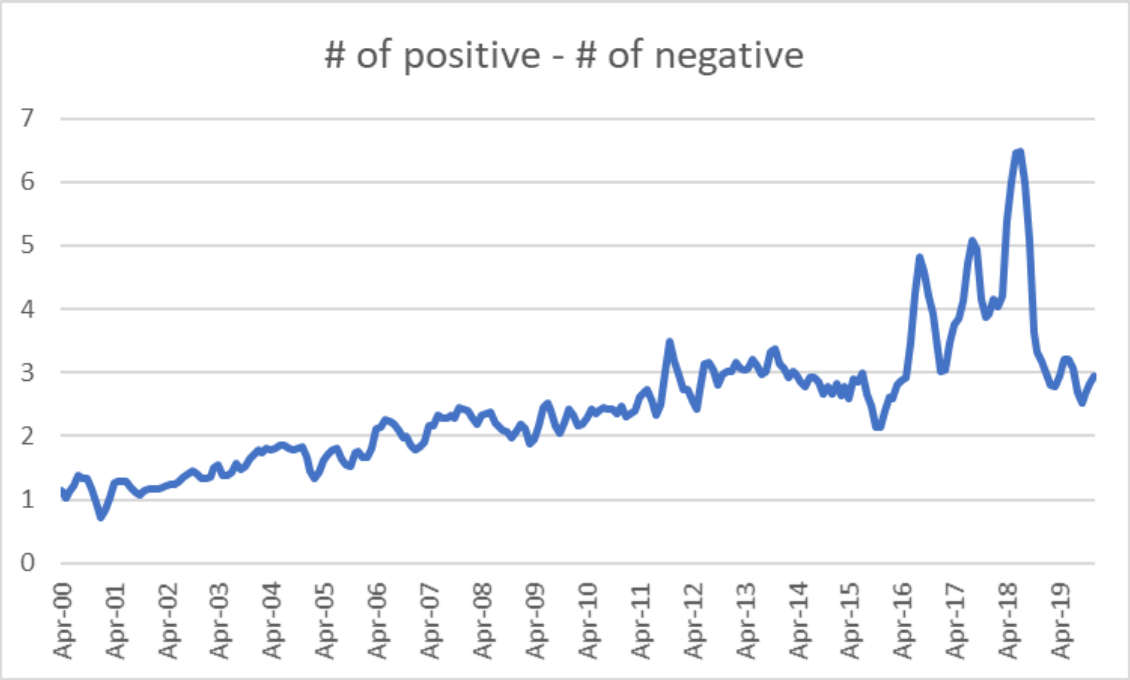


Figure 2: Time-series of One-month Net Positive News Coverage

This figure presents the time-series plot of cross-stock averages of one-month net positive news coverage. For firm *i* in month *m*, to obtain the "one-month net positive news coverage", we first calculate the number of news articles with positive sentiment minus the number of news articles with negative sentiment. We then take the cross-stock average of the one-month net positive news coverage for month *m* and plot its monthly time-series.



Internet Appendix A

Tarun Chordia, Bin Miao, and Joonki Noh

June 2024

This appendix provides additional test results and associated materials for the paper entitled “Do Low Latency Traders Destabilize Prices? Evidence from News Releases” by Chordia, Miao, and Noh (2024).

Table A.1: Impact of the Retail Trading Activity on Returns on News Days – An Alternative Algorithm for Retail Trading Identification

This table repeats the tests in Table 3 by using an alternative algorithm of Barber et al. (2023) to identify retail trading activity. For stock *i* on day *t*, $RETAIL_OIB_{i,t}$ is the order imbalance of retail trading, defined as (retail buy share volume – retail sell share volume) / (retail buy share volume + retail sell share volume). $RETAIL_{i,t}$ is the retail trading volume, defined as retail share volume / total share volume. Panel A explores factors that can affect *RETAIL_OIB* on news days. Panel B investigates the impacts of retail trading activity on returns on news release days. *DRT* is the cross-sectional decile ranking of *RETAIL* for each day and ranges from 0 to 1. We interact *DRT* with *MISP* and the news variables, add the three-way interaction terms into equation (3), and re-estimate it. The sample period for this analysis is from January, 2010 to December, 2015. In panel A, day and firm fixed effects are included, and standard errors are clustered by day and firm. In Panel B, day fixed effects are included and standard errors of coefficients are clustered by day and firm. *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Retail Trading Activity on News Release Days – Alternative Algorithm

	RETAIL_OIB
SENT	0.013***
	(9.83)
MISP	0.050*
	(1.87)
RET	0.390***
	(18.68)
Controls	Yes
Fixed Effects	Day / Firm
Num. of Obs.	1,602,970
Adj. R^2	0.002

Panel B: Return Impact of Retail Trading Activity on News Release Days – Alternative Algorithm

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.052	-0.793*	-0.652**
	(0.36)	(-1.72)	(-2.43)
MISP x *DRT*	1.068***	4.236***	1.193*
	(3.69)	(4.65)	(1.87)
MISP x *SUE*	63.996***	70.614**	320.730**
	(3.08)	(2.06)	(1.96)
MISP x *PSENT*	-4.888***	19.297***	-7.412**
	(-3.17)	(2.93)	(-2.49)
MISP x *PSENT* x *DRT*	-20.800***	-114.030***	20.951**
	(-4.90)	(-7.92)	(2.29)
MISP x *NSENT*	0.688	-15.774**	11.403***
	(0.38)	(-2.25)	(3.37)
MISP x *NSENT* x *DRT*	27.656***	92.658***	-24.304**
	(5.67)	(5.98)	(-2.27)
SUE	7.611***	8.340***	1.397
	(8.74)	(5.48)	(0.32)
PSENT	0.693***	1.291***	0.564***
	(15.69)	(9.58)	(5.52)
PSENT x *DRT*	2.599***	0.559*	2.034***
	(21.58)	(1.76)	(6.78)
NSENT	-0.354***	-0.740***	-0.471***
	(-6.19)	(-4.80)	(-3.85)
NSENT x *DRT*	-3.443***	-2.044***	-2.450***
	(-21.30)	(-5.20)	(-6.52)
DRT	-0.023**	0.031*	-0.067***
	(-2.50)	(1.79)	(-2.85)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	5,591,651	1,796,836	1,510,464
Adj. R^2	0.011	0.013	0.013

Table A.2: Impact of the Retail Trading or LLT Activity on Returns on News Days – Estimating Model (2)

This table repeats the tests in Tables 3 and 4 by estimating the regression model in equation (2). For stock *i* on day *t*, $SUE_{i,t}$ is the standardized unexpected earnings and defined as the actual EPS announced on day *t* minus the consensus forecast of the EPS, scaled by the stock price at the fiscal quarter end of the EPS. As the consensus forecast, we use the median of all earnings forecasts for each firm followed by security analysts or the actual EPS for the same fiscal quarter in the last fiscal year for each firm not followed by security analysts. We winsorize *SUE* at -1 and 1 to mitigate potential effects of outliers, and assign *SUE* = 0 for non-earnings days that have *EARN* = 0. $SENT_{i,t}$ is the average sentiment score of news articles released on day *t* and defined as (average of ESS-50)/50, where ESS is a sentiment variable from RavenPack. Thus, *SENT* ranges from -1 to 1 and we assign *SENT* = 0 for non-news days that have *NEWS* = 0. The higher *SENT* is, the more optimistic the average sentiment is. The sample period for Panel A is from January, 2010 to December, 2015 by following Boehmer et al. (2021). The sample for Panel B contains Nasdaq-listed common stocks only and starts in January, 2008 and ends in December, 2017, which is restricted by the availability of *LLT* and *LLT_SIGN*. For the other aspects of the tests, see the legends in Tables 3 and 4. Day fixed effects are included in estimating equation (2). The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Impact of Retail Trading Activity in Table 3 – Model (2)

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	-0.180	-0.898**	-0.531**
	(-1.15)	(-2.10)	(-2.13)
MISP x *DRT*	1.468***	4.406***	1.270**
	(4.84)	(4.59)	(2.12)
MISP x *SUE*	67.710***	72.805**	306.510*
	(3.25)	(2.12)	(1.90)
MISP x *SENT*	-4.105***	14.863***	-7.862***
	(-3.63)	(3.37)	(-3.76)
MISP x *SENT* x *DRT*	-24.501***	-104.920***	19.490***
	(-7.05)	(-9.76)	(2.78)
SUE	7.731***	8.398***	2.173
	(8.83)	(5.48)	(0.51)
SENT	0.763***	1.225***	0.698***
	(21.06)	(12.23)	(9.41)
SENT x *DRT*	2.595***	0.865***	1.870***
	(24.22)	(3.22)	(7.70)
DRT	-0.039***	0.013	-0.070***
	(-3.08)	(0.54)	(-3.02)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	5,345,188	1,723,192	1,475,992
Adj. R^2	0.011	0.013	0.012

Panel B: Impact of Overall LLT Activity in Table 4 – Model (2)

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.478***	1.331***	0.021
	(2.94)	(3.45)	(0.13)
MISP x *DLLT*	0.526	2.435*	-0.336
	(1.50)	(1.95)	(-1.00)
MISP x *SUE*	94.250***	1.161	304.790**
	(5.67)	(0.04)	(2.09)
MISP x *SENT*	-14.347***	-25.596***	-2.116
	(-20.14)	(-10.97)	(-1.64)
MISP x *SENT* x *DLLT*	-20.789***	-54.065***	-0.092
	(-9.00)	(-4.74)	(-0.03)
SUE	7.541***	4.451***	4.099
	(10.05)	(2.73)	(1.03)
SENT	1.735***	1.540***	1.386***
	(65.56)	(27.27)	(29.15)
SENT x *DLLT*	0.559***	-0.026	-0.132
	(6.98)	(-0.11)	(-1.33)
DLLT	-0.066***	-0.011	-0.034**
	(-4.55)	(-0.38)	(-2.27)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	7,405,791	2,092,590	2,737,735
Adj. R^2	0.013	0.012	0.016

Panel C: Impact of Aggressive LLT Activity in Table 4 – Model (2)

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.355**	1.205***	-0.022
	(2.10)	(3.10)	(-0.14)
MISP x *DLLTS*	1.629***	3.618***	-0.019
	(6.28)	(3.79)	(-0.05)
MISP x *SUE*	94.711***	1.820	304.880**
	(5.70)	(0.06)	(2.09)
MISP x *SENT*	-14.624***	-25.975***	-2.210*
	(-18.02)	(-9.27)	(-1.69)
MISP x *SENT* x *DLLTS*	-9.858***	-20.530***	-0.633
	(-8.00)	(-4.44)	(-0.31)
SUE	7.522***	4.435***	4.094
	(10.04)	(2.73)	(1.03)
SENT	1.688***	1.491***	1.322***
	(59.22)	(22.95)	(27.69)
SENT x *DLLTS*	0.424***	0.223**	0.162**
	(10.22)	(2.08)	(2.13)

DLLTS	-0.118***	-0.082***	-0.066***
	(-13.14)	(-3.61)	(-4.55)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	7,405,791	2,092,590	2,737,735
Adj. R^2	0.013	0.012	0.016

Table A.3: Impact of the LLT Activity on Returns on News Days – Contemporaneous Liquidity Controlled

This table repeats the tests in Panels B and C of Table 4 by adding a contemporaneous daily liquidity measure, i.e. *LIQ*, proxied by daily share turnover and the interaction terms among *LIQ*, *MISP*, and news variables into the regression models. For the other aspects of the tests, see the legends in Table 4. Day fixed effects are included in estimating the regression models. The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Impact of Overall LLT Activity in Panel B of Table 4 with LIQ controlled

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.638**	1.996***	-0.332
	(2.40)	(2.80)	(-1.15)
MISP x *DLLT*	0.415	3.488**	-0.884**
	(1.15)	(2.49)	(-2.27)
MISP x *SUE*	93.713***	0.334	301.620**
	(5.66)	(0.01)	(2.08)
MISP x *PSENT*	-6.075*	-10.184	-6.706**
	(-1.91)	(-1.37)	(-2.36)
MISP x *PSENT* x *DLLT*	-9.729***	-40.689***	3.573
	(-3.52)	(-2.70)	(1.15)
MISP x *NSENT*	8.649***	7.196	7.484*
	(4.75)	(1.24)	(1.84)
MISP x *NSENT* x *DLLT*	11.685***	35.683**	1.652
	(3.68)	(2.52)	(0.37)
SUE	7.520***	4.437***	4.182
	(10.05)	(2.72)	(1.06)
PSENT	1.001***	0.829***	0.760***
	(8.10)	(5.10)	(7.15)
PSENT x *DLLT*	-0.022	-0.736**	-0.519***
	(-0.21)	(-2.38)	(-4.44)
NSENT	-0.978***	-1.033***	-0.665***
	(-12.94)	(-6.00)	(-4.60)
NSENT x *DLLT*	-0.085	0.463	0.264
	(-0.75)	(1.43)	(1.55)
DLLT	-0.045***	0.031	0.001
	(-3.29)	(1.00)	(0.05)
Control variables	Yes	Yes	Yes
LIQ and interactions with *MISP*, *MISP* x *SENT*, *SENT*	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	7,405,791	2,092,590	2,737,735
Adj. R^2	0.014	0.013	0.018

Panel B: Impact of Aggressive LLT Activity in Panel C of Table 4 with LIQ controlled

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	0.640**	2.012***	-0.332
	(2.35)	(2.74)	(-1.15)
MISP x *DLLTS*	0.803*	4.283**	-0.344
	(1.87)	(2.46)	(-0.55)
MISP x *SUE*	94.186***	1.107	301.850**
	(5.68)	(0.04)	(2.08)
MISP x *PSENT*	-6.930**	-9.793	-6.450**
	(-2.19)	(-1.26)	(-2.23)
MISP x *PSENT* x *DLLTS*	-4.447**	-20.682***	-1.289
	(-2.19)	(-2.70)	(-0.41)
MISP x *NSENT*	3.721*	-1.987	7.934*
	(1.88)	(-0.29)	(1.93)
MISP x *NSENT* x *DLLTS*	16.813***	23.240**	0.711
	(7.00)	(2.41)	(0.17)
SUE	7.509***	4.433***	4.180
	(10.04)	(2.72)	(1.06)
PSENT	0.974***	0.833***	0.698***
	(7.82)	(4.97)	(6.50)
PSENT x *DLLTS*	0.443***	0.059	0.321***
	(7.55)	(0.36)	(2.79)
NSENT	-0.760***	-0.863***	-0.589***
	(-9.60)	(-4.55)	(-3.97)
NSENT x *DLLTS*	-0.192**	-0.146	0.249
	(-2.53)	(-0.63)	(1.56)
DLLTS	-0.149***	-0.074*	-0.108***
	(-11.69)	(-1.94)	(-4.67)
Control variables	Yes	Yes	Yes
LIQ and interactions with *MISP*, *MISP* x *SENT*, *SENT*	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	7,405,791	2,092,590	2,737,735
Adj. R^2	0.016	0.016	0.018

Table A.4: Returns on Earnings Days and News Days

– Three-day Event Window

This table repeats the tests in Table 2 with a three-day event window to capture the price reaction to earnings announcements and news releases. We calculate $SUE_{i,t}$ if firm i makes an earnings announcement in a three-day window [t-1, t+1], and assign $SUE_{i,t} = 0$ otherwise. We calculate $SENT_{i,t}$ as the average sentiment of news articles about firm i released on days t-1, t, or t+1, and assign $SENT_{i,t} = 0$ if no news article is released on any of those three days. *PSENT* and *NSENT* are defined in a similar manner based on the three-day window. For the other aspects of the tests, see the legend in Table 2. The sample period of our analysis is from January, 2000 to December, 2019. Day fixed effects are included in estimating equations (2) and (3). The standard errors of coefficients are clustered by day and firm, and t-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1) Model (2)	(2) Model (3)
MISP	0.917***	0.912***
	(7.71)	(7.42)
MISP x *SUE*	1.229	1.280
	(0.56)	(0.59)
MISP x *SENT*	-6.585***	
	(-21.24)	
SUE	1.026***	1.025***
	(15.53)	(15.52)
SENT	0.898***	
	(78.79)	
MISP x *PSENT*		-6.346***
		(-15.96)
MISP x *NSENT*		6.914***
		(11.98)
PSENT		0.849***
		(49.27)
NSENT		-1.010***
		(-41.05)
Control variables	Yes	Yes
Fixed effects	Day	Day
Num. of Obs.	28,665,891	28,665,891
Adj. R^2	0.005	0.006

Table A.5: Returns on Earnings Days and News Days

– News Sources with High Credibility

This table repeats the tests in Table 2 with the two refined subsamples of news sources that have high credibility. For Panel A, we use the most credible news sources by selecting Source Rank = 1 and complete news articles by selecting News Type = Full Article, both of which are coded by RavenPack. For Panel B, we use news articles in RavenPack coming only from Dow Jones Newswire. For the other aspects of the tests, see the legend in Table 2. The sample period of our analysis is from January, 2000 to December, 2019. Day fixed effects are included in estimating equations (2) and (3). The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: High Quality News Sources and Complete News Articles

	(1) Model (2)	(2) Model (3)
MISP	0.896***	0.937***
	(7.63)	(7.98)
MISP x *SUE*	6.509	6.411
	(1.16)	(1.14)
MISP x *SENT*	-20.762***	
	(-20.70)	
SUE	2.382***	2.384***
	(14.48)	(14.49)
SENT	2.153***	
	(61.27)	
MISP x *PSENT*		-23.891***
		(-18.70)
MISP x *NSENT*		16.280***
		(13.17)
PSENT		2.282***
		(52.58)
NSENT		-1.963***
		(-47.12)
Control variables	Yes	Yes
Fixed effects	Day	Day
Num. of Obs.	28,665,891	28,665,891
Adj. R^2	0.006	0.006

Panel B: Dow Jones Newswire

	(1) Model (2)	(2) Model (3)
MISP	0.920***	0.916***
	(7.83)	(7.86)
MISP x *SUE*	6.398	6.407
	(1.14)	(1.14)
MISP x *SENT*	-13.870***	
	(-23.10)	
SUE	2.403***	2.403***
	(14.55)	(14.55)
SENT	1.560***	
	(69.32)	
MISP x *PSENT*		-13.727***
		(-19.33)
MISP x *NSENT*		14.144***
		(15.26)
PSENT		1.552***
		(58.38)
NSENT		-1.576***
		(-42.58)
Control variables	Yes	Yes
Fixed effects	Day	Day
Num. of Obs.	28,665,891	28,665,891
Adj. R^2	0.006	0.006

Table A.6: Returns on Earnings Days and News Days

– An Alternative Definition of Stock-level Mispricing

This table repeats the tests in Table 2 with an alternative definition of stock-level mispricing following the method of Stambaugh, Yu, and Yuan (2012). At the beginning of the month which day *t* belongs to, as a stock *i*'s measure of relative mispricing, we define $MISP_{i,t}^{AR}$ as 2*average percentile ranking-1 (ranging from -1 to 1), where we calculate the average percentile ranking (ranging from 0 to 1) of the 11 mispricing signals of Stambuagh et al. (2012). The higher $MISP^{AR}$ is, the more underpriced the stock is. For the other aspects of the tests, see the legend in Table 2. The sample period of our analysis is from January, 2000 to December, 2019. Day fixed effects are included in estimating equations (2) and (3). The standard errors of coefficients are clustered by day and firm, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1) Model (2)	(2) Model (3)
$MISP^{AR}$	0.062***	0.062***
	(5.19)	(5.17)
$MISP^{AR}$ x *SUE*	0.188	0.198
	(0.36)	(0.38)
$MISP^{AR}$ x *SENT*	-1.444***	
	(-27.53)	
SUE	2.382***	2.380***
	(14.32)	(14.31)
SENT	1.571***	
	(81.40)	
$MISP^{AR}$ x *PSENT*		-1.393***
		(-22.81)
$MISP^{AR}$ x *NSENT*		1.505***
		(17.83)
PSENT		1.483***
		(65.22)
NSENT		-1.763***
		(-56.41)
Control variables	Yes	Yes
Fixed effects	Day	Day
Num. of Obs.	28,050,508	28,050,508
Adj. R^2	0.008	0.008

Table A.7: Impacts of the Retail Trading and LLT Activity on Returns on News Days – Full Estimation Results

This table investigates how the interaction of retail trading activity with LLT activity affects returns on news release days. For each day, we define *DRT* as the daily decile ranking of *RETAIL*, which ranges from 0 to 1, and *DLLT* as the dummy variable indicating the top daily decile of *LLT,* respectively. We then add all interaction terms of *DLLT*, *DRT*, *MISP*, *PSENT* and *NSENT* as follows. In Panel A, for stock *i* on day *t*,

$$RET_{i,t} = \alpha + \beta_1 MISP_{i,t} + \beta_2 MISP_{i,t} \times DRT_{i,t} + \beta_3 MISP_{i,t} \times DLLT_{i,t} + \beta_4 MISP_{i,t} \times DRT_{i,t} \times DLLT_{i,t} + \beta_5 MISP_{i,t} \times SUE_{i,t} + \beta_6 MISP_{i,t} \times PSENT_{i,t} + \beta_7 MISP_{i,t} \times PSENT_{i,t} \times DRT_{i,t} + \beta_8 MISP_{i,t} \times PSENT_{i,t} \times DLLT_{i,t} + \beta_9 MISP_{i,t} \times PSENT_{i,t} \times DRT_{i,t} \times DLLT_{i,t} + \beta_{10} MISP_{i,t} \times NSENT_{i,t} + \beta_{11} MISP_{i,t} \times NSENT_{i,t} \times DRT_{i,t} + \beta_{12} MISP_{i,t} \times NSENT_{i,t} \times DLLT_{i,t} + \beta_{13} MISP_{i,t} \times NSENT_{i,t} \times DRT_{i,t} \times DLLT_{i,t} + \beta_{14} SUE_{i,t} + \beta_{15} PSENT_{i,t} + \beta_{16} PSENT_{i,t} \times DRT_{i,t} + \beta_{17} PSENT_{i,t} \times DLLT_{i,t} + \beta_{18} PSENT_{i,t} \times DRT_{i,t} \times DLLT_{i,t} + \beta_{19} NSENT_{i,t} + \beta_{20} NSENT_{i,t} \times DRT_{i,t} + \beta_{21} NSENT_{i,t} \times DLLT_{i,t} + \beta_{22} NSENT_{i,t} \times DRT_{i,t} \times DLLT_{i,t} + \beta_{23} DRT_{i,t} + \beta_{24} DLLT_{i,t} + \beta_{25} DRT_{i,t} \times DLLT_{i,t} + \gamma' \boldsymbol{X}_{i,t} + \varepsilon_{i,t}, \quad (A.1)$$

In Panel B, we further decompose *PSENT* and *NSENT* into four non-overlapping intraday periods: before market open (BMO, 0:00:00–9:29:59), morning session (AM, 9:30:00–13:00:00), afternoon session (PM, 13:00:01–16:00:00), and after market close (AMC, 16:00:01–23:59:59). Following Boehmer et al. (2021) for retail trading activity, the sample period is from January, 2010 to December, 2015. For the other aspects of the tests, see the legends in Tables 2 to 4. In all estimations, day fixed effects are included in the estimation and standard errors of coefficients are clustered by day and firm. *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Daily *PSENT* and *NSENT*

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	-0.305**	-0.192	-0.517*
	(-2.03)	(-0.43)	(-1.94)
MISP x *DRT*	1.478***	2.207***	1.021*
	(5.47)	(2.88)	(1.96)
MISP x *DLLT*	-0.357	2.305	-0.350
	(-0.69)	(1.02)	(-0.41)
MISP x *DRT* x *DLLT*	0.655	-0.192	-1.280
	(0.62)	(-0.06)	(-0.68)
MISP x *SUE*	66.273***	65.708*	310.370*
	(3.23)	(1.92)	(1.92)
MISP x *PSENT*	0.807	9.522	-7.252***
	(0.56)	(1.60)	(-2.61)
MISP x *PSENT* x *DRT*	-11.750***	33.930***	-12.349*
	(-3.12)	(2.99)	(-1.68)
MISP x *PSENT* x *DLLT*	1.889	20.135	-2.698
	(0.69)	(1.33)	(-0.68)
MISP x *PSENT* x *DRT* x *DLLT*	-29.945***	-163.950***	43.884***
	(-4.74)	(-11.78)	(6.54)
MISP x *NSENT*	-1.354	-4.107	8.012**
	(-0.80)	(-0.59)	(2.42)

MISP x *NSENT* x *DRT*	13.505***	-44.085***	17.359*
	(2.94)	(-2.88)	(1.86)
MISP x *NSENT* x *DLLT*	-3.092	-23.119	2.933
	(-0.89)	(-1.51)	(0.52)
MISP x *NSENT* x *DRT* x *DLLT*	25.982***	137.320***	-44.629***
	(3.55)	(8.81)	(-5.42)
SUE	7.625***	8.117***	2.042
	(8.80)	(5.33)	(0.48)
PSENT	0.570***	0.792***	0.663***
	(14.16)	(6.12)	(6.85)
PSENT x *DRT*	2.731***	2.256***	1.901***
	(29.26)	(8.83)	(8.01)
PSENT x *DLLT*	-0.827***	-0.937**	-0.420**
	(-7.41)	(-2.20)	(-2.33)
PSENT x *DRT* x *DLLT*	0.960***	0.408	-0.687**
	(3.57)	(0.68)	(-2.44)
NSENT	-0.412***	-0.421***	-0.541***
	(-7.76)	(-2.69)	(-4.44)
NSENT x *DRT*	-3.020***	-3.159***	-2.280***
	(-24.57)	(-9.08)	(-7.33)
NSENT x *DLLT*	0.865***	0.687	0.367
	(5.67)	(1.50)	(1.53)
NSENT x *DRT* x *DLLT*	-0.928***	-0.169	0.717*
	(-2.72)	(-0.21)	(1.90)
DRT	-0.050***	-0.030	-0.048**
	(-4.23)	(-1.56)	(-2.41)
DLLT	0.038**	0.123**	0.034
	(2.35)	(2.32)	(1.08)
DRT x *DLLT*	-0.112***	-0.153*	-0.026
	(-3.02)	(-1.66)	(-0.38)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	3,804,475	1,084,497	1,433,084
Adj. R^2	0.014	0.018	0.013

Panel B: Intraday *PSENT* and *NSENT*

	(1) All Stocks	(2) Overpriced Stocks	(3) Underpriced Stocks
MISP	-0.086	0.242	-0.434
	(-0.53)	(0.50)	(-1.43)
*MISP*DRT*	1.127***	1.484*	0.825
	(3.95)	(1.80)	(1.46)
*MISP*DLLT*	-0.581	-1.109	0.230
	(-1.04)	(-0.43)	(0.25)

*MISP*DRT*DLLT*	1.610	4.418	-0.461
	(1.34)	(1.07)	(-0.22)
*MISP*SUE*	65.764***	65.171**	302.840*
	(3.37)	(1.99)	(1.95)
SUE	7.127***	7.593***	1.732
	(8.63)	(5.19)	(0.43)
DRT	-0.017	-0.005	-0.015
	(-1.45)	(-0.26)	(-0.73)
DLLT	0.038**	0.031	0.019
	(2.10)	(0.54)	(0.51)
*DRT*DLLT*	-0.080**	-0.011	-0.013
	(-2.07)	(-0.10)	(-0.17)
*MISP*PSENT_BMO*	-20.038***	-17.973	-28.561**
	(-3.02)	(-0.70)	(-2.16)
*MISP*PSENT_BMO*DRT*	-17.893	-22.610	43.704
	(-1.37)	(-0.56)	(1.51)
*MISP*PSENT_BMO*DLLT*	25.533*	87.822	18.645
	(1.87)	(1.09)	(0.83)
*MISP*PSENT_BMO*DRT*DLLT*	-48.745*	-128.920	-29.061
	(-1.81)	(-0.96)	(-0.61)
*MISP*NSENT_BMO*	-1.594	56.422	5.318
	(-0.15)	(1.10)	(0.27)
*MISP*NSENT_BMO*DRT*	84.425***	-11.946	31.978
	(3.38)	(-0.13)	(0.60)
*MISP*NSENT_BMO*DLLT*	-47.686*	-171.330	11.953
	(-1.75)	(-1.23)	(0.27)
*MISP*NSENT_BMO*DRT*DLLT*	91.794*	318.750	-39.740
	(1.72)	(1.38)	(-0.40)
PSENT_BMO	2.280***	2.519***	2.523***
	(12.61)	(4.28)	(5.34)
*PSENT_BMO*DRT*	4.730***	4.706***	2.776***
	(13.60)	(4.59)	(2.68)
*PSENT_BMO*DLLT*	-2.186***	-0.811	-2.120**
	(-5.33)	(-0.39)	(-2.52)
*PSENT_BMO*DRT*DLLT*	0.579	-0.792	0.102
	(0.71)	(-0.22)	(0.06)
NSENT_BMO	-1.019***	0.785	-1.052
	(-3.49)	(0.72)	(-1.44)
*NSENT_BMO*DRT*	-12.725***	-16.070***	-11.461***
	(-17.71)	(-6.64)	(-5.77)
*NSENT_BMO*DLLT*	2.487***	-0.784	0.500
	(3.07)	(-0.23)	(0.30)
*NSENT_BMO*DRT*DLLT*	0.770	6.629	5.299

	(0.47)	(1.05)	(1.43)
*MISP*PSENT_AM*	-11.755**	22.192	-12.990
	(-2.20)	(0.92)	(-1.25)
*MISP*PSENT_AM*DRT*	-47.681***	-148.570***	24.420
	(-3.82)	(-3.41)	(0.89)
*MISP*PSENT_AM*DLLT*	58.537***	243.890**	-47.016*
	(3.02)	(2.45)	(-1.79)
*MISP*PSENT_AM*DRT*DLLT*	-131.520***	-500.620***	89.000
	(-3.13)	(-2.75)	(1.41)
*MISP*NSENT_AM*	-15.510*	-28.571	21.638
	(-1.90)	(-0.75)	(1.30)
*MISP*NSENT_AM*DRT*	78.820***	125.000*	-81.042**
	(4.41)	(1.83)	(-2.03)
*MISP*NSENT_AM*DLLT*	-7.705	-26.199	13.231
	(-0.32)	(-0.21)	(0.35)
*MISP*NSENT_AM*DRT*DLLT*	52.261	90.263	-14.592
	(1.07)	(0.48)	(-0.17)
PSENT_AM	1.415***	2.387***	1.271***
	(9.05)	(4.50)	(3.55)
*PSENT_AM*DRT*	7.253***	5.039***	5.466***
	(21.12)	(4.88)	(5.76)
*PSENT_AM*DLLT*	-2.211***	1.894	1.113
	(-4.00)	(0.87)	(1.21)
*PSENT_AM*DRT*DLLT*	1.958	-6.375	-5.027**
	(1.60)	(-1.48)	(-2.34)
NSENT_AM	0.293	0.132	-0.840
	(1.19)	(0.16)	(-1.31)
*NSENT_AM*DRT*	-12.307***	-12.038***	-7.479***
	(-23.87)	(-7.27)	(-5.09)
*NSENT_AM*DLLT*	0.565	-0.392	0.070
	(0.80)	(-0.14)	(0.05)
*NSENT_AM*DRT*DLLT*	0.966	2.384	3.121
	(0.66)	(0.47)	(1.08)
*MISP*PSENT_PM*	-9.974	-19.221	-13.700
	(-1.43)	(-0.68)	(-1.02)
*MISP*PSENT_PM*DRT*	-71.849***	-128.320***	25.901
	(-4.82)	(-2.59)	(0.77)
*MISP*PSENT_PM*DLLT*	42.437***	134.090**	15.052
	(2.81)	(1.99)	(0.58)
*MISP*PSENT_PM*DRT*DLLT*	-52.588*	-223.390*	-40.465
	(-1.67)	(-1.92)	(-0.72)
*MISP*NSENT_PM*	3.189	-53.987*	21.521
	(0.41)	(-1.68)	(1.26)

*MISP*NSENT_PM*DRT*	78.328***	231.660***	-4.388
	(4.35)	(3.60)	(-0.12)
*MISP*NSENT_PM*DLLT*	-59.162***	-86.144	-10.368
	(-3.34)	(-0.92)	(-0.39)
*MISP*NSENT_PM*DRT*DLLT*	45.639	89.011	-75.637
	(1.27)	(0.59)	(-1.28)
PSENT_PM	2.290***	2.135***	2.290***
	(11.07)	(3.42)	(4.82)
*PSENT_PM*DRT*	5.110***	4.096***	2.513**
	(11.55)	(3.37)	(2.14)
*PSENT_PM*DLLT*	-2.525***	-0.470	-2.080**
	(-5.70)	(-0.32)	(-2.02)
*PSENT_PM*DRT*DLLT*	-0.978	-5.173*	-0.533
	(-1.04)	(-1.90)	(-0.24)
NSENT_PM	-1.189***	-2.436***	-1.395**
	(-5.11)	(-3.60)	(-2.36)
*NSENT_PM*DRT*	-7.469***	-4.049***	-5.712***
	(-14.45)	(-2.84)	(-4.31)
*NSENT_PM*DLLT*	2.291***	1.821	0.670
	(4.24)	(0.86)	(0.66)
*NSENT_PM*DRT*DLLT*	2.105*	2.736	6.367***
	(1.93)	(0.74)	(2.85)
*MISP*PSENT_AMC*	0.233	-28.195	-9.946
	(0.03)	(-0.93)	(-0.60)
*MISP*PSENT_AMC*DRT*	-12.020	-13.803	34.162
	(-0.87)	(-0.25)	(1.10)
*MISP*PSENT_AMC*DLLT*	-42.541**	0.757	-9.891
	(-2.11)	(0.01)	(-0.33)
*MISP*PSENT_AMC*DRT*DLLT*	55.743	114.270	17.031
	(1.47)	(0.53)	(0.29)
*MISP*NSENT_AMC*	-15.265*	22.753	-4.018
	(-1.70)	(0.55)	(-0.26)
*MISP*NSENT_AMC*DRT*	47.241***	4.335	24.936
	(2.84)	(0.07)	(0.75)
*MISP*NSENT_AMC*DLLT*	44.408**	136.320	-26.711
	(2.05)	(1.13)	(-0.53)
*MISP*NSENT_AMC*DRT*DLLT*	-84.179**	-229.210	18.721
	(-2.00)	(-1.18)	(0.20)
PSENT_AMC	1.005***	0.472	1.365**
	(5.84)	(0.73)	(2.49)
*PSENT_AMC*DRT*	1.136***	1.020	-0.412
	(3.32)	(0.84)	(-0.38)
*PSENT_AMC*DLLT*	0.432	1.515	-0.741

	(0.69)	(0.42)	(-0.71)
*PSENT_AMC*DRT*DLLT*	-1.029	1.277	0.249
	(-0.91)	(0.24)	(0.12)
NSENT_AMC	0.821***	1.983*	0.462
	(3.28)	(1.92)	(0.86)
*NSENT_AMC*DRT*	-3.311***	-4.875***	-2.824**
	(-6.94)	(-2.84)	(-2.23)
*NSENT_AMC*DLLT*	-0.940	0.642	1.860
	(-1.42)	(0.25)	(0.82)
*NSENT_AMC*DRT*DLLT*	0.593	-2.403	-3.362
	(0.49)	(-0.52)	(-0.82)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	3,804,475	1,084,497	1,433,084
Adj. R^2	0.028	0.034	0.027

Internet Appendix B

Tarun Chordia, Bin Miao, and Joonki Noh

June 2024

In this appendix, we reconcile our test results with those in Engelberg, McLean, and Pontiff (2018), where we report all test results based on equations (1) to (3) and the standard errors of coefficients are clustered by day for apple-to-apple comparisons.

Our analysis in the main text shows that the daily returns of overpriced stocks in the subsequent month are significantly higher on news release days, implying that non-earnings news releases slow down the price discovery process. This result, however, appears to be contrary to the findings in Engelberg et al. (2018) (henceforth EMP) that news facilitates the correction of existing mispricing as the returns for overpriced stocks are lower on news release days in their sample with their research design. In this section, we investigate the underlying reasons for this discrepancy between our results and those in EMP.

1. Choice of mispricing signals

EMP measure stock-level mispricing using a much broader set of 97 return predictors based on McLean and Pontiff (2016), compared to our mispricing measure based on the 11 SYY signals. To examine whether the choice of the empirical proxy for stock-level mispricing is responsible for the different results, we rerun the regressions (1) to (3) in Panels A through C, using an alternative measure of *MISP* based on the anomalies included in EMP.[27] The results in Column (1) of Table B.1 show that the choice of anomalies used to calculate *MISP* does not drive the difference in the results. Specifically, in Panel A, the coefficient of *MISP*NEWS* remains significantly negative, although the magnitude of the

[27] We obtain these return predictors from Chen and Zimmerman's Open Source Asset Pricing website at https://www.openassetpricing.com/. We drop Merger and SEO from the 97 anomalies used in EMP and employ the remaining 95 anomaly signals to calculate our *MISP* because the Merger and SEO data are not available in Chen and Zimmerman's dataset.

coefficient and its t-statistic are both smaller than those in Table 2. In Panel B, the coefficient of *MISP*SENT* is also negative and highly significant. Finally, Panel C shows that the increase in returns of overpriced stocks on news release days is driven by news articles with positive sentiment. These patterns are all consistent with those reported in Table 2 and suggest that our finding that positive sentiment news impedes the price efficiency of overpriced stocks is not specific to the 11 SYY mispricing signals, but generalizable to those extracted from a broader set of firm characteristics.

2. Choice of news sources

Another research design difference between our analysis and that in EMP is that we use RavenPack database to identify news release days and to measure the sentiment of news stories, while EMP use Dow Jones News Archive database.[28] If there are systematic differences in the coverage of news articles or how the sentiment of news stories is measured between these two news databases, the results obtained with one database may fail to replicate those from the other. To examine this possibility, we re-estimate equations (1) to (3) using the Dow Jones News Archive database for the sample period of January 2000 to December 2019. Specifically, we define $NEWS_{i,t}$ = 1 if there is at least one news article in the Dow Jones News Archive database about firm *i* published on day *t*, and 0 otherwise. We measure the sentiment of a news article by its net tone, which is calculated as the difference between the number of positive words and the number of negative words, divided by the total number of words in the news article.[29] Our measure of the information content of a news day, $SENT_{i,t}$, is calculated as the average net tone of all news articles about firm *i* released on day *t*. $SENT_{i,t}$ ranges between -1 and 1. The results are reported in Column (2) of Table B.1. In Column (2) of Panel A, the coefficient on *MISP*NEWS* is positive but statistically

[28] The news variables constructed using Dow Jones News Archive database can be different from those constructed using Dow Jones Newswire classified by RavenPack since these two databases are different in news coverage and in how news sentiment is measured.

[29] The lists of positive and negative words are obtained from Loughran and McDonald's master dictionary, which is available at https://sraf.nd.edu/loughranmcdonald-master-dictionary/

insignificant, which is different from the significantly negative coefficient in Column (1) and suggests that the differences between the two news databases might affect the estimated impact of news days on returns. However, consistent with the results in Column (1), Column (2) of Panel B shows that the coefficient on *MISP*SENT* is significantly negative, and Panel C also shows that the coefficient on *MISP*PSENT* is significantly negative and that on *MISP*NSENT* is significantly positive. Therefore, the opposing impacts of positive and negative sentiment news articles on returns continue to hold for the Dow Jones News Archive Database sample and are not sensitive to how the sentiments of the news stories are measured.

3. Sample period

Since we rely on RavenPack to identify news release days and to measure the sentiment of news articles, our sample is restricted by the coverage of the RavenPack database, which starts in 2000. In contrast, EMP use the Dow Jones News Archive database to identify news days and therefore their sample period began in June 1979. One possibility for the discrepancy between their results and ours could be that the average effect of news on returns evolves over time. This could be due to multiple factors including, e.g., the changes in the average quality of news stories, the efficiency of information dissemination channels driven by the advances in technology, and growing popularity of social media platforms. But more importantly, it could also be driven by changes in news sentiment over time. Since positive and negative news articles have opposite impacts on returns of mispriced stocks, the relative proportion of positive vs. negative news days and/or the relative strength of their average sentiment scores in a sample will determine the average effect of news release days and hence the sign of the coefficient on *MISP*NEWS*. In an effort to replicate the result in EMP, we use the Dow Jones News Archive database over the same sample period from June 1979 to December 2013 as in EMP.[30] Column (3) of Panel A

[30] For an additional robustness check, we also use the news dataset shared by Jeon et al. (2022), which covers Dow Jones'

shows that when *MISP* is calculated using the 95 return predictors used in EMP over their sample period, the coefficient of *MISP*NEWS* turns significantly positive at 0.308 (t-statistic = 4.55), suggesting that overpriced stocks earn lower returns on average news release days. This is consistent with the finding in EMP. We next examine the differential impact of positive and negative sentiment news for this earlier sample period. Column (3) of Panel B shows that the coefficient on *MISP*SENT* is significantly negative. Further, Column (3) of Panel C shows that the coefficient on *MISP*PSENT* is significantly negative, and that on *MISP*NSENT* is significantly positive. These results are consistent with our main finding reported in Table 2, that negative sentiment news articles tend to reduce overpricing, while positive sentiment news articles tend to exacerbate overpricing. As in our 2000-2019 sample, not all news days reduce mispricing, and the net effect of news days on returns largely depends on the composition of news articles with positive and negative sentiment and the relative strength of their sentiment scores.

In sum, in our effort to reconcile our main finding with that in EMP, we find robust evidence that the release of positive sentiment news exacerbates the existing overpricing. Our results also suggest that the news-day dummy variable, *NEWS,* is unable to capture the heterogeneity of the content of news articles and thus can mask important differences between positive and negative sentiment news articles in affecting the returns of mispriced stocks.

Factiva database for the sample period from 1980 to 1999. We find that our findings are robust. The dataset can be downloaded from https://www.dropbox.com/s/62lt6uq1t4n6gcr/MainData_Factiva_Public.zip?dl=0

Table B.1: Returns on Earnings Days and News Days – Using 95 Anomaly Signals from McLean and Pontiff (2016)

This table repeats the tests in Table 2 by constructing the stock-level mispricing variable *MISP* using the 95 (out of 97) return predictors studied in McLean and Pontiff (2016). We obtain these 95 predictors from Chen and Zimmerman's Open Source Asset Pricing at https://www.openassetpricing.com/. They drop Merger and SEO variables due to the lack of data from the original 97 predictors in McLean and Pontiff (2016). In Column (1), the news variables are constructed using RavenPack database, while in Columns (2) and (3), the same news variables are constructed using Dow Jones News Archive database. The sample periods are, respectively, January 2000 through December 2019 for Columns (1) and (2), and June 1979 through December 2013 for Column (3). In Columns (2) and (3), where Dow Jones News Archive database is used as the news source, $SENT_{i,t}$ is defined as the average sentiment of all news articles released on day *t* for firm *i,* where the sentiment is calculated as (number of positive words – number of negative words)/total number of words based on Loughran and McDonald's master dictionary. We assign *SENT* = 0 for non-news days that have *NEWS* = 0. All other variables are defined as in Table 2 and see the legend in Table 2 for the other aspects of the tests. Day fixed effects are included in estimating equations (1) to (3). The standard errors of coefficients are clustered by day, and *t*-statistics are reported in parenthesis. Significance level: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel A: Model (1)

	(1) RavenPack Database (Jan. 2000 ~ Dec. 2019)	(2) Dow Jones News Archive Database (Jan. 2000 ~ Dec. 2019)	(3) Dow Jones News Archive Database (Jun. 1979 ~ Dec. 2013)
MISP	0.414***	0.373***	0.467***
	(5.29)	(5.43)	(11.07)
MISP x *EARN*	2.410***	2.344***	2.056***
	(8.95)	(8.78)	(12.40)
MISP x *NEWS*	-0.136**	0.099	0.308***
	(-2.16)	(1.57)	(4.55)
EARN	0.075***	0.059***	0.160***
	(4.51)	(3.55)	(14.34)
NEWS	0.049***	0.091***	0.144***
	(9.44)	(19.32)	(28.22)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,643,730	28,643,644	43,664,459
Adj. R^2	0.004	0.004	0.007

Panel B: Model (2)

	(1) RavenPack Database (Jan. 2000 ~ Dec. 2019)	(2) Dow Jones News Archive Database (Jan. 2000 ~ Dec. 2019)	(3) Dow Jones News Archive Database (Jun. 1979 ~ Dec. 2013)
MISP	0.500***	0.370***	0.459***
	(7.21)	(5.36)	(10.46)
MISP x *SUE*	4.493**	10.667***	10.301***
	(2.17)	(3.63)	(6.95)
MISP x *SENT*	-4.466***	-28.831***	-32.42***
	(-35.37)	(-14.88)	(-19.50)
SUE	2.402***	4.222***	3.047***
	(18.48)	(14.57)	(26.35)
SENT	1.470***	5.482***	4.953***
	(111.63)	(40.26)	(42.71)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,643,730	28,643,644	43,664,459
Adj. R^2	0.008	0.004	0.008

Panel C: Model (3)

	RavenPack Database (Jan. 2000 ~ Dec. 2019)	Dow Jones News Archive Database (Jan. 2000 ~ Dec. 2019)	Dow Jones News Archive Database (Jun. 1979 ~ Dec. 2013)
MISP	0.453***	0.346***	0.455***
	(6.24)	(5.06)	(10.53)
MISP x *SUE*	4.546**	10.674***	10.262***
	(2.19)	(3.63)	(6.93)
MISP x *PSENT*	-3.839***	-15.380***	-22.579***
	(-20.63)	(-6.79)	(-8.46)
MISP x *NSENT*	5.787***	36.356***	37.498***
	(21.09)	(13.11)	(17.48)
SUE	2.399***	4.221***	3.046***
	(18.46)	(14.57)	(26.34)
PSENT	1.397***	6.277***	8.090***
	(82.59)	(39.06)	(41.25)
NSENT	-1.633***	-5.004***	-3.808***
	(-64.40)	(-24.39)	(-24.94)
Control variables	Yes	Yes	Yes
Fixed effects	Day	Day	Day
Num. of Obs.	28,643,730	28,643,644	43,664,459
Adj. R^2	0.008	0.004	0.008