Social Media-Driven Noise Trading: Liquidity Provision and Price Revelation Ahead of Earnings Announcements

Edna Lopez Avila, Charles Martineau, and Jordi Mondria*

August 29, 2024

ABSTRACT

Social media attention before earnings announcements is overly optimistic, fails to predict fundamentals, and generates buying pressure, leading to a 58 bps stock return as intermediaries seek higher returns for providing liquidity. Such price pressure distorts the price informativeness of fundamentals. A return reversal occurs immediately following announcements as markets correct mispricing. How stock prices respond to earning news is endogenous to the effect of social media in the pre-announcement price formation. A pre-announcement trading strategy based on expected social media attention yields 40 bps monthly alphas. When noise trading is systematically driven, it can deter liquidity provision and price revelation.

JEL Classification: G12, G14, G40.
Keywords: earnings announcements, inventory risks, investor attention, liquidity provision, price efficiency, return reversal, social media

*A previous version of this paper was circulated under the title "When Crowds Aren't Wise: Biased Social Networks and its Price Impact." We thank Pat Akey, Ing-Haw Cheng, Alex Chinco, Roberto Cram, Alex Corhay, Valeria Fedyk, Goutham Galapakrishna, Vincent Gregoire, Ryan Israelsen, Runjing Lu, Marina Niessner, Yoshio Nozawa, Terry Odean, Chay Ornthanalai, Mike Simutin, Christopher Shiller, and Xiaofei Zhao for valuable feedback. We thank the seminar participants at the 2023 NFA, 2023 Future of Financial Information, 2023 Trans-Atlantic Doctoral, 2023 Waseda Summer Workshop in Finance, 2024 MFA, 2024 SOFiE conferences. We acknowledge the financial support from the Canadian Securities Institute, the Social Sciences and Humanities Research Council (SSHRC), and TD MDAL. Jay Cao at TD MDAL and Kevin Li provided excellent research assistance. Lopez Avila: Rotman School of Management, University of Toronto, 105 St-George, Toronto ON, Canada, M5S 3E6 (edna.lopez.avila@rotman.utoronto.ca), Martineau: Rotman School of Management and UTSC Management, University of Toronto, 105 St-George, Toronto ON, Canada, M5S 3E6 (charles.martineau@rotman.utoronto.ca), Mondria: University of Toronto, 150 St. George Street, Toronto ON, Canada, M5S 3G7 (jordi.mondria@utoronto.ca).

1. Introduction

Classical frameworks of market efficiency (Friedman et al., 1953, Fama, 1965) consider only the presence of investors who possess valuable skills in evaluating a firm's prospects for future fundamentals as the main source of information affecting price efficiency. Investors that trade on irrelevant information ("noise traders") are assumed to be normally distributed, have no permanent price impact, and are met in the market by arbitrageurs who trade against them and quickly eliminate mispricing (Grossman and Stiglitz, 1980, Kyle, 1985). Noise trading has no role in price formation.

Social investment networks have become the primary source of investment advice for many retail investors (Cookson, Mullins, and Niessner, 2024), a driver of persistent flow of non-fundamental trading motives (Pedersen, 2022, Waters, Zhang, Zhou, and Santosh, 2024), and are an attraction device generating systematic noise trading (Barber, Huang, Odean, and Schwarz, 2022).[1] While Barber, Odean, and Zhu (2008, 2009) show that systematic noise trading can move prices, how it affects price efficiency and the informational content of stock prices remains unclear.[2]

This paper focuses on the friction that arises from social media-driven systematic noise trading to price revelation.[3] Models of price formation suggest that trades originating from social media should not affect price revelation if they reflect "noise." But when such noise is systematically driven, it can worsen liquidity provision as intermediaries face an imbalance between the quantity sought by buyers and sellers at a given price. Intermediaries may

[1]Heimer (2016) find that peer interactions on social networks influence investors' trading decisions. De Long, Shleifer, Summers, and Waldmann (1990) propose a model of correlated noise trading impacting informational efficiency.

[2]Han and Yang (2013) analyze a rational expectations equilibrium model with endogenous information acquisition and argue that social communication networks can worsen price revelation.

[3]We follow the terminology of Brunnermeier (2005) to distinguish between two components of price revelation: "price efficiency" and "price informativeness". Price efficiency relates to the price revelation of public information and how fast information is impounded into prices and price informativeness reflects the absolute level of information to future fundamentals (see also Biais, Hillion, and Spatt, 1999, Weller, 2018, Boguth, Fisher, Gregoire, and Martineau, 2024).

absorb the order imbalance into their own account but through price concession to minimize inventory risks by setting prices above (below) fundamental value in response to buy (sell) order imbalances (So and Wang, 2014). Such price concession will worsen the informational content of stock prices.

We exploit the cross-sectional variation in the coverage of stocks on social media in the days leading to earnings announcements. Such days are periods of high information uncertainty (Akey, Grégoire, and Martineau, 2022) and low liquidity where traders abstain from trading (Lee, Mucklow, and Ready, 1993). Liquidity providers' demand curves are steeper during that time due to increases in inventory risks associated with unexpected price changes following announcements (So and Wang, 2014). Consequently, stock prices are expected to be more sensitive to systematic noise trading driven by social media chatter ahead of earnings announcements.

The main punchline of our paper is that stocks widely discussed on social media are associated with retail buying pressure, generating a significant positive price drift ahead of earnings announcements as intermediaries require a higher rate of return to mitigate inventory risks. Such systematic noise trading worsens price revelation ahead of earnings announcements as the price concession required by intermediaries adds noise to prices, and a price reversal follows after the earnings news release. Conditioning on the fundamental news, i.e., earnings surprises, stock prices can diverge away from fundamentals ahead of announcements, but markets quickly correct the mispricing following the announcement. A key implication to future research is that theories should consider the role of systematic noise trading in price formation. Moreover, understanding how social media influences pre-announcement price dynamics is critical. Peng, Wang, and Zhou (2022), Campbell, Drake, Thornock, and Twedt (2023), Hirshleifer, Peng, and Wang (2024) have examined the role of social media in price formation following earnings announcements. We show that examining only the effect of social media on price formation following announcements can lead to biased inferences, as post-announcement price responses are endogenous to the effect of social media to pre-announcement price formation.

We focus on the social network *StockTwits*, the largest investor-focused social media platform, which averages more than one million monthly posts covering most stocks. Our sample covers the period of 2013 to 2022 and provides the widest coverage of stocks, more than any other social media platform. Cookson, Lu, Mullins, and Niessner (2022) find that StockTwits' coverage correlates more strongly with returns than other popular social media platforms (i.e., Twitter or SeekingAlpha) and that StockTwits fairly represents the aggregate information shared on social media. We find that StockTwits activity increases five days before earnings announcements. This increase in coverage is not observed for other news sources such as newswires. StockTwits activity is not only concentrated in large stocks (top NYSE breakpoint quintile). Small market capitalization (bottom quintile) stocks receive as much coverage as large firms.

Aggregating StockTwits users' self-labeled "bullish" and "bearish" sentiment tags reveal a positive bias among users. Over 60% of the stock-earnings announcement observations exhibit a bullish outcome, defined as having more than 80% of the sentiment-tagged posts about that specific stock-earnings announcement observation tagged bullish. We show that users' sentiments do not predict earnings announcement fundamentals, i.e., earnings surprises.[4] We further show that such excess positivism displayed on StockTwits transmits into correlated retail buying pressure ahead of announcements.[5]

[4]Previous studies find that the content of social media posts can predict earnings surprises, (e.g., Chen, De, Hu, and Hwang, 2014, Dim, 2020, Bartov, Faurel, and Mohanram, 2018). The first two papers use Seeking Alpha, a platform where non-anonymous users can write lengthy articles. Seeking Alpha provides limited coverage before announcements, and only a handful of stocks receive coverage. Bartov, Faurel, and Mohanram (2018) employ Twitter data for an earlier period, specifically 2009-2012.

[5]Bradley, Hanousek Jr, Jame, and Xiao (2023) and Hu, Jones, Zhang, and Zhang (2021) examines the social media platform Reddit/Wall Street Bets and find that greater Reddit attention relates to more retail trading. Kakhbod, Kazempour, Livdan, and Schuerhoff (2023) classify StockTwits users into "skilled", "unskilled", and "antiskilled" and find that 56% of users are "antiskilled" and create optimistic beliefs and are the most influential in changing followers' beliefs.

When intermediaries (market makers) face an uninformed and balanced buying and selling order flow, prices remain efficient (Grossman and Stiglitz, 1980). But when faced with systematic noise trading, the findings of So and Wang (2014) suggest that intermediaries will require a higher expected return to meet the demand as they seek liquidity in the opposite direction to hedge their inventory risk. We find that stocks with high StockTwits attention (i.e., top attention quintile) generate a 58 bps price pressure ahead of announcements and a 45 bps price reversal on announcement dates. By comparison, the average return to a comparable matched sample of stocks with low attention shows no evidence of price reversal. Additional tests show that the effect of StockTwits' attention to price pressure and reversals is robust to earnings surprises, firm size, the number of analyst following, and news coverage before announcements. We further show that buying pressure induced by social media activity ahead of announcements predicts the reversal.[6]

We then examine the cross-sectional variation in the reversal magnitudes. If the buying pressure is a result of inventory risk, the reversal magnitude will be greater for stocks with ex-ante higher anticipated return volatility on announcement dates. As argued by So and Wang (2014), intermediaries can take several days to unwind their net positions. Consequently, providing liquidity ahead of scheduled news events, characterized by high volatility, increases their exposure to inventory risks (Madhavan and Smidt, 1993) and the risk of margin calls (Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes, 2010). We use firm size[7], option implied volatility prior to announcements, and the average of previous earnings announcement dates' absolute returns as proxies of inventory risks (anticipated return volatility). For stocks with high StockTwits attention, the reversal magnitudes are higher for small firms and stocks with high option implied volatility and prior earnings announcement date absolute returns. Low-StockTwits attention stocks show little cross-sectional varia-

[6]Price pressure ahead of earnings announcements as a result of social media lends support to the findings of Laarits and Sammon (2023) that retail-heavy stocks are more expensive to trade ahead of announcements.

[7]As shown in Martineau (2022), smaller firms have larger absolute returns on announcement dates.

tion in reversals. The increased inventory risks arising from social media-driven systematic noise trading have a striking influence on the autocorrelation of returns and the information content of stock prices.

We next show that social media stock activity ahead of earnings announcements can distort the price informativeness of stocks' fundamentals. Conditioning on announcements' earnings surprises, the average price run-up before announcements pushes prices closer to fundamentals for positive earnings surprise announcements, albeit not a reflection of informative trading. However, for negative earnings surprises, the buying price pressure ahead of announcements deviates prices *away* from fundamentals. On the announcement date, markets quickly adjust prices to public information and correct any mispricing generated by social media information production. In other words, the release of public information quickly eliminates mispricing caused by the social-media-driven price pressure. We find no evidence that social media attention results in slower price formation following announcements.[8] This last result contrasts with the findings of Campbell, Drake, Thornock, and Twedt (2023). They find that social media activity on Twitter slows price formation following earnings announcements.[9] We show that post-announcement price formation are endogenous to pre-announcement price responses caused by social media attention.[10] Our paper illustrate why one would find social media to be associated with slower price discovery post-announcement if one disregards pre-announcement price dynamics.

Our results exhibit a *contemporaneous* relationship between attention and price pressure ahead of announcements. This raises the question of whether intermediaries can foresee which stocks will gather the most social media attention ahead of earnings announcements and if there is a trading strategy to exploit the price pressure.[11] We developed a trading

[8]See Martineau (2022) for a review on the evolution of stock return post-earnings announcement drifts.

[9]Other papers that examine the relationship between social media and post-earnings announcement price dynamics include Ding, Shi, and Zhou (2023) and Hirshleifer, Peng, and Wang (2024).

[10]See Boguth, Grégoire, and Martineau (2019) and Fisher, Martineau, and Sheng (2022) for evidence of investor attention increasing before scheduled macroeconomic announcements.

[11]We do not argue that this strategy is expected to be profitable when taking into account transaction

strategy that buys stocks expected to receive high attention, sells those expected to receive low attention, and closes the position the day before the announcement. We used posts from Earnings Whispers, a firm known for its earnings forecasts and strong social media following, to indicate which stocks might receive more social media attention. At the beginning of each week, Earnings Whispers posts on StockTwits and other social media platforms the week's most anticipated earnings releases. We find that stocks mentioned in these posts attract more attention from retail investors, generate more retail buying pressure, and earn higher returns in the days leading up to their earnings announcements compared to stocks not mentioned. A pre-announcement long-short strategy earns monthly abnormal returns of 45 basis points.

We present a simple theoretical framework explaining how social media may influence optimistic trading behaviors that can result in systematic noise trading. The model is based on the concept of "wishful thinking" from Caplin and Leahy (2019), which posits that individuals derive utility from their beliefs and thus tend to interpret information optimistically. The model predicts that investors will display positive (negative) optimism when seeking to buy (sell) stocks. It is well-known that retail investors are more inclined to buy than sell (Barber and Odean, 2008) and, consistent with our findings, we expect investors to display more positive optimism. Furthermore, our model indicates that this optimistic bias is stronger when information is easily manipulated, as evidenced by the significant engagement with Earnings Whispers' posts on StockTwits. This simple model provides interesting avenues for future research to explore the role of social media in shaping investors' beliefs and trading behaviors ahead of scheduled announcements.

costs.

1.1. Related literature and contributions

This paper contributes to the growing literature examining the role of social media in financial markets. Antweiler and Frank (2004) find a link between stock market volatility and messages on online investment message boards. Cookson and Niessner (2020) is the first paper to study the content of StockTwits messages. The authors find a relationship between users' disagreement and trade volume. Cookson, Fos, and Niessner (2021) find that greater investor disagreement measured from StockTwits facilitates informed trading and short sellers. Cookson, Engelberg, and Mullins (2020) study StocksTwits users' political partisanship and how it influences investors' expectations in the wake of COVID-19. More closely related to our paper, Cookson, Engelberg, and Mullins (2023) find that StockTwits users choose selective exposure to confirmatory information, i.e., *echo chambers*.[12] With retail traders more likely to buy than to sell a stock (Barber and Odean, 2008), such users are more likely to be influenced by positive information, which can explain the positive-biased sentiment of StockTwits posts ahead of earnings announcements. Our paper further shows that prices reflect the positive-biased sentiment, but markets correct mispricing following public information. Cookson, Niessner, and Schiller (2022) find that corporate managers are also influenced by the content of StockTwits when evaluating the prospects of a merger and acquisition. The recent episodes of GameStop and other "meme-stocks" provided compelling evidence of correlated noise trading resulting from social media (Bradley, Hanousek Jr, Jame, and Xiao, 2023). Our paper shows that episodes of correlated noise trading resulting from social media are much more pervasive than documented in the media and such noise trading impacts market markers inventory risks and price revelation. Jia, Redigolo, Shu, and Zhao (2020) find that social media posts from Twitter does not predict rumor realization of mergers. Similarly, we find that social media posts do not predict earnings fundamentals.

[12] Jiao, Veiga, and Walther (2020) provide evidence consistent with echo chambers from an aggregate source of social media platforms collected by MarketPsych Data.

Other research examines the impact of alternative social media outlets on asset prices and retail trading, such as Reddit/Wall Street Bets (Kang, Lou, Ozik, Sadka, and Shen, 2024), Seeking Alpha (Chen, De, Hu, and Hwang, 2014, Dim, 2020, Farrell, Green, Jame, and Markov, 2022), and Twitter (Jiao, Veiga, and Walther, 2020, Campbell, Drake, Thornock, and Twedt, 2023, Bianchi, Gómez-Cram, and Kung, 2024). In contrast to these existing papers, we document how social media influences retail trading ahead of earnings announcements, how it can impact inventory risks and price revelation, and how quickly markets correct social media-driven mispricing following the release of public information.

Our paper further relates to the growing literature on retail trading. Barber and Odean (2000), Barber and Odean (2008), Barber, Lee, Liu, and Odean (2009), and Barber, Huang, Odean, and Schwarz (2022) find that retail investors are generally uninformed and make systematic mistakes when selecting stocks. Another strand of the literature finds otherwise. Hvidkjaer (2008), Kaniel, Saar, and Titman (2008), Kaniel, Liu, Saar, and Titman (2012), Kelley and Tetlock (2013), Barrot, Kaniel, and Sraer (2016), Boehmer, Jones, Zhang, and Zhang (2021) show that retail order imbalance positively predicts future returns at short horizons. We also find that retail order imbalances are positively associated with returns but that such trades do not reflect fundamentals and lead to a price reversal on announcement dates. Peress and Schmidt (2021) find that noise trading follows the common assumption of i.i.d.-normal only at monthly and lower frequencies but at daily frequencies to be more correlated. We show that social media can be responsible for such correlated noise trading. Finally, Peress and Schmidt (2020) show that market makers are more concerned with trading against insiders than noise traders. When faced with systematic noise trading ahead of scheduled announcements, we show that market makers are concerned and require a higher expected return to meet the demand from noise traders.

2. Data and Methodology

2.1. StockTwits

We retrieve data on StockTwits posts from January 2013 to December 2022 through RapidAPI, starting our analysis from 2013 to ensure quality stock coverage as per Cookson and Niessner (2020). By identifying posts with \$-tagged tickers (e.g., \$AAPL), we matched these to tickers in the CRSP database, totaling 150,262,272 stock-specific posts. Figure 1 presents the total monthly number of stock-specific posts on StockTwits and reveals a significant increase in posting activity during the COVID-19 pandemic, with monthly posts exceeding four million, compared to approximately one million posts per month before the pandemic. This activity returned to pre-pandemic levels in 2022.

On StockTwits, users can label their posts as "bullish" or "bearish" to express their sentiment toward a stock.[13] They can also indicate their trading experience level as novice, intermediate, or professional.[14] We follow Cookson and Niessner (2020) in assigning posts made after 4 p.m. to the following trading day. This approach allows us to align our analysis with daily stock returns, calculated from 4 p.m. to 4 p.m. the next trading day. Weekend posts are similarly assigned to the next trading day.

A stock's "attention" (or coverage) and user sentiments on StockTwits are the main two variables that we construct. When measuring how much attention a stock receives on social media, the standard approach is to divide the total number of posts for a stock on a given day by the total number of posts on the platform on that day. Comparing a particular stock post activity to other stocks' posts controls for the non-stationary time-series in posts activity on StockTwits. We compute StockTwits attention as follows:

$$Att_{i,t} = \frac{post_{i,t}}{\sum_i^N post_{i,t}}, \tag{1}$$

[13] 56% of our stock-day observations have sentiment-tagged StockTwits posts.

[14] According to Cookson and Niessner (2020), 20% of users identify as professionals, 52% as intermediates, and 28% as novices.

where $post_{i,t}$ is the number of posts for stock i on date t. The denominator is the sum of posts for all stocks on StockTwits on date t.

To measure investors' sentiment associated with StockTwits messages on a given day, we calculate the daily proportion of bullish posts to the total number of bullish and bearish posts for stock i on the day t, as follows:

$$Sent_{i,t} = \frac{bull_{i,t}}{bull_{i,t} + bear_{i,t}}, \tag{2}$$

where "bull" and "bear" correspond to the number of posts with bullish and bearish tags, respectively. When calculating attention (sentiment) over a longer window, e.g., 5 days, we sum the number of posts (bullish posts) over the 5-day period and divide the total number of StockTwits posts over the 5-day period (sum of bearish and bullish posts).

2.2. Earnings, news, and stock-level data

We supplement our StockTwits data with analyst forecasts and earnings announcement dates from Thomson Reuters I/B/E/S. We include earnings announcements in IBES that meet the following criteria: the earnings date is reported in Compustat, the stock price five days before the announcement is available in CRSP, and the stock price is available in Compustat as of the end of the quarter. We calculate the *Surprise* in earnings announcements as the difference between the firm's earnings per share for the quarterly earnings announcement and the consensus analysts' forecast, divided by the stock price five days prior to the earnings announcement day. We compute analysts' forecasts by taking the median of all analysts' estimates issued within the 90 days preceding the earnings announcement date. Lastly, we winsorize the earnings surprise at the 1st and 99th percentiles.

Following Gregoire and Martineau (2022), we also gather analyst recommendation news events and other firm-level news for our sample of stocks from Ravenpack. All news events occurring after 4 p.m. or on weekends are attributed to the next trading day.

Additionally, we retrieve daily stock returns from CRSP, the five factors from Kenneth French's website, and intraday trading data from TAQ. With the trading data, we follow the methodology outlined in Boehmer, Jones, Zhang, and Zhang (2021) and the suggested adjustments in Barber, Huang, Jorion, Odean, and Schwarz (2024) to identify trades by retail investors and construct various retail trading measures such as order imbalances.

Table 1 reports summary statistics for high and low-attention stocks. We define high-attention stocks if *Att* belongs to the top quintile five days before earnings announcements for stocks with the same earnings announcement dates. High-attention stocks have a higher average market capitalization, stock price volatility, absolute abnormal returns on announcements, and analyst following. A key takeaway from this table is that when comparing high-attention stocks to low-attention stocks, it is important to use a set of *matched* low-attention stocks as pre-earnings liquidity and asset price dynamics vary across stocks (Liu, Wang, Yu, and Zhao, 2020). Also, stocks that receive the highest StockTwits activity ahead of announcements are more volatile, with higher absolute returns on the announcement date. These stock characteristics attract the most investor attention (Barber and Odean, 2008).

2.3. StockTwits activity around earnings announcements

Table 2 reports a breakdown of the coverage across NYSE market capitalization breakpoint quintile. It shows the count of stock-earnings observations with at least one StockTwits message, one analyst recommendation, or one newswire mention from five to one day before the announcements. Additionally, the table reports the number of observations without StockTwits posts, analyst recommendations, or newswire coverage. A key insight from this table is the broader scope of StockTwits in covering stocks before earnings announcements compared to analysts and newswires. Approximately 24% of pre-announcement StockTwits posts pertain to the smallest firms, and 37% to the largest. In contrast, analyst recommendations and newswire reports before earnings announcements predominantly focus on the largest stocks, accounting for 66% of recommendations and 73% of newswire, respectively. Only 7% of the earnings announcements in the sample lacked StockTwits posts in the five days leading up to the announcements. For the smallest stocks, only 13% lacked StockTwits messages. However, the absence of analyst recommendations and newswire coverage for the smallest stocks significantly jumps to 98% and 81%, respectively, highlighting a disparity in coverage based on firm size.

We plot in Figure 2 the abnormal activity in StockTwits posts and newswires articles using boxplots five days before to five days after earnings announcements. Abnormal post (newswire coverage) activity is computed as the daily log number of posts (newswire) minus the log of the average daily number of posts (newswire) from 20 to 6 days before the earnings announcements. The figure shows an increase in abnormal StockTwits post activity in the days approaching earnings announcements, with a notable 50% increase on the day before the announcement. In contrast, newswire activity shows a modest rise on the day before the earnings announcements, yet below the benchmark period ($t = [-20, -6]$) and increases following announcements.[15] This figure highlights the significant role of social media networks in disseminating information about stocks before earnings announcements, bridging a gap not covered by traditional news sources. Unlike newswires, which often report earnings results post-release, social media platforms enable investors to monitor real-time discussions and sentiments regarding a stock leading up to its earnings announcement.

We then examine users' sentiment on StockTwits in the 60 days leading up to earnings announcements and compare it to analysts' recommendations.[16] Figure 3 shows the fraction of stock-earnings observations according to StockTwits tagged-sentiment and analyst recommendation sentiment. We define sentiment as in equation (2) and split sentiment ratio into

[15] Gregoire and Martineau (2022) and Li, Ramesh, Shen, and Wu (2015) show that analyst recommendations typically follow earnings announcements. We find no increase in abnormal analyst recommendations ahead of earnings announcements.

[16] It has been shown that analysts' forecasts exhibit predictable biases (Kothari, So, and Verdi, 2016, Van Binsbergen, Han, and Lopez-Lira, 2023) and over-optimism (Cowen, Groysberg, and Healy, 2006).

five buckets: [0-20%], (20-40%], (40-60%], (60-80%], and (80-100%], i.e., from very bearish to very bullish. Sentiment on StockTwits regarding upcoming earnings is predominantly positive. Over 60% of the stock-earnings announcement observations exhibit a bullish outcome, defined as having more than 80% of the sentiment-tagged posts about that specific stock-earnings announcement observation tagged bullish, while fewer than 5% of observations show a similar dominance of bearish posts (bucket [0-20%]). In contrast, analyst recommendations display a less pronounced positive bias and exhibit a more balanced distribution. Approximately 55% of stock-earnings observations has over 80% bullish recommendations, and 30% of stock-earnings observations have a majority of bearish recommendations (bucket [0-20%]). This figure reveals a significant inclination among StockTwits users towards sharing and engaging with positively biased posts on StockTwits. Selecting posts five days before earnings announcements shows similar positive-biased sentiment on StockTwits.

3. How Informative Is Social Media Ahead of Earnings Announcements?

We first examine the informativeness of StockTwits' sentiment about earnings fundamentals ahead of earnings announcements. We then investigate the relationship between StockTwits' attention and retail trading.

3.1. StockTwits sentiment does not predict earnings fundamentals

Having determined that StockTwits users display a predominantly positive sentiment, questions arise regarding the informativeness of their posts about earnings fundamentals. We investigate this question by estimating the following regression:

$$Surp_{i,t} = \beta_1 Sent_{i,t} + \beta_2 \mathbb{1}_{i,t}^{Att} \times Sent_{i,t} + \beta_3 \mathbb{1}_{i,t}^{Att} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t}, \quad (3)$$

where $Surp_{i,t}$ is the earnings surprise for stock-earnings i announced on date t, $Sent_{i,t}$ represents the sentiment as defined in equation (2), and $\mathbb{1}_{i,t}^{Att}$ is a dummy variable set to one if the stock's StockTwits attention, as defined in equation (1), falls within the top quintile, otherwise it is set to zero. Both sentiment and attention are computed from posts made from five days to one day prior to earnings announcements. The regression also includes an interaction term between sentiment and attention to examine whether stocks with a higher volume of posts yield a more accurate sentiment prediction of earnings surprises. The control variables are the buy-and-hold abnormal returns, sentiment from analyst recommendations computed as in equation (2) based on recommendation outlook being bullish or bearish, and RavenPack newswire sentiment, all measured in the five days leading up to the earnings announcements. α_i and α_t correspond to firm- and year-fixed effects through the parameters.

Table 3 presents the results for the full sample, large caps (the top three NYSE market capitalization quintiles), and small caps (bottom two quintiles). The model specifications defined in columns (1)–(3) exclude stock-earnings observations with no tagged sentiment. In columns (4)–(6), we treat stock-earnings observations without sentiment data as having a neutral sentiment (i.e., $Sent = 0.5$). Across all model specifications, we find no statistically significant evidence that the sentiment ($Sent$) expressed in StockTwits posts predicts earnings surprises and similarly when interacting sentiment with attention ($\mathbb{1}^{Att} \times Sent$).

In a robustness check, we estimate equation (3) using the change rather than the level of sentiment. Table IA1 of the Internet Appendix reports the results and finds no statistically significant evidence that the change in sentiment predicts earnings surprises. Cookson and Niessner (2020) find that StockTwits' users that are self-labelled as *professionals* are indeed more sophisticated than *novice* users. They find that professional posts' sentiments are positively related to future returns. We examine whether sentiment posts for novice, intermediate, and professional self-labeled users predict earnings surprises and report the findings in Table IA2 of the Internet Appendix. Consistent with our previous findings, we find no statistically significant evidence that sentiment predicts earnings surprises across all user types.

3.2. StockTwits activity induces buying pressure

Barber and Odean (2008) find that retail investors are net buyers of attention-grabbing stocks, e.g., stocks in the news. We confirm this finding by examining the relationship between StockTwits attention and retail trading. We calculate retail trading orders following the methodology of Boehmer, Jones, Zhang, and Zhang (2021), incorporating the adjustments suggested by Barber, Huang, Jorion, Odean, and Schwarz (2024). We retrieve the number of retail trades, trading volume, and dollar volume and compute retail order imbalance measures. Barber, Lin, and Odean (2023) show that focusing on the number of trades rather than the volume provides a more accurate reflection of attention-induced retail trading.[17] We proceed to estimate the following regression model:

$$Retail\ OI_{i,t} = \beta \mathbb{1}^{Att}_{i,t} + \alpha_i + \alpha_t + \epsilon_{i,t},$$

where $Retail\ OI_{i,t}$ in the regression specification represents retail order imbalance computed five to one day before the stock's earnings announcement i released on date t using the number of trades, volume, dollar volume, and their corresponding detrending transformation. We choose 50 to six before announcements as the detrending window.

We present the findings in Table 4. In all model specifications, higher attention is associated with positive retail order imbalances and is statistically significant at the 1% level. The β estimate varies from 0.007 to 0.027, and computing retail order imbalance using trades indicates stronger buying pressure, consistent with the arguments of Barber, Lin, and Odean (2023) that trades best capture retail investor attention. These increases in retail order imbalances are economically significant. For example, the point estimates in Panels A and B of 0.027 and 0.013 correspond to a four and 1.3 times increase in retail order imbalance relative to the unconditional mean of 0.0068 and 0.017, respectively.

We also examine whether the component of the order imbalance driven by social media

[17] Specifically, Barber, Lin, and Odean (2023) find that smaller retail trades tend to focus on stocks that capture significant attention and are inversely related to future returns.

attention relates to earnings surprises. Table IA3 of the Internet Appendix reports that the fitted component of retail order imbalance, from regressing retail order imbalance on $\mathbb{1}^{Att}$, does not predict earnings surprises, but the orthogonal component does.

4. Social Media-Driven Systematic Noise Trading, Price Pressure, and Liquidity Provision

The preceding sections show that social media activity ahead of earnings announcements is associated with more retail buying pressure. We next examine how the price pressure relates to inventory risks for intermediaries and how it impacts the price revelation of earnings news.

4.1. Price pressure and return reversals

The evidence reported in Section 3 suggests that the content on StockTwits does not relate to earnings fundamentals, and retail investors are net buyers of stocks with high StockTwits attention. When facing buying pressure (a positive net order imbalance) ahead of announcements, intermediaries should be compensated via a price concession by setting prices above fundamental value, resulting in a positive expected return. As market markers unwind their net position following announcements and adjust the excess of price concession will result in a negative expected return. We follow So and Wang (2014) and use market-adjusted returns around announcements to proxy for intermediaries' inventory balances. We use the extent of negative autocorrelation (i.e., return reversal) from the pre-to-post announcement as a proxy for the expected returns that market makers demand to provide liquidity to net buyers of high-attention social media stocks.

Figure 4 plots the difference in the buy-and-hold abnormal returns for high-attention and matched-low-attention stocks. Matched stocks are assigned based on the firm size, industry (GIC), buy-and-hold abnormal returns from 30 to 6 days prior to the announcement, and earnings surprises for the same year-quarter. The figure shows a significant return divergence between high- and low-attention stocks of more than 1%, with the most significant increase occurring five days before the announcement. Following the announcement, we observed a reversal of more than 50 bps, and the difference between high- and low-attention stocks is not statistically different from zero one day after the announcement.

We next conduct a "diff-in-diff" analysis to validate the robustness of our findings in Table 5. The first difference compares the effect of price pressure before earnings announcements to using a randomly selected 'pseudo-earnings-announcement' date in place of the actual announcement date. Following So and Wang (2014), we select pseudo-announcement dates from randomly selecting a pseudo-date 50 to 20 days window prior to actual announcement dates. Columns (1) and (2) report the average BHAR, in percent, around earnings announcements (EA) and pseudo-earnings announcements for high-attention stocks, respectively, and the difference is reported in columns (3). Columns (4) and (5) report the difference between the low-attention-match stocks and for the full sample of low-attention stocks. Columns (6) and (7) report the "Diff-in-Diff".

For high-attention stocks, column (1) reports a 58 bps and -45 bps in BHAR[-5,-1] and BHAR[0,1], respectively. Column (3) reports a 49 bps increase (t-statistic of 4.68) in BHAR relative to the pseudo earnings dates in the five days leading earnings announcements. The "Diff-in-Diff" columns (6) and (7) report a 65 bps (t-statistic of 5.27) and 62 bps (t-statistic of 4.10) increase in BHAR relative to low attention stocks, respectively. On the announcement date ([0,1]), the pre-announcement increase in BHAR is reversed. Columns (3) report a decrease of 47 bps (t-statistic=5.42). Columns (4) and (5) report no significant decrease in abnormal returns for the low-attention stocks, and the diff-in-diff columns report a decrease of 58 bps (t-statistic=-5.04) and 45 bps (t-statistic=-3.72), respectively.

We examine the robustness of the return reversal to alternative explanations that can result in higher returns leading to earnings announcements. For example, it has been well-documented that around earnings announcements, stocks earn a risk premium (Barth and

So, 2014). Moreover, information leakage can result in pre-earnings announcement drifts (Akey, Grégoire, and Martineau, 2022). To further examine the robustness of price pressure to alternative explanations, we run the following regression

$$BHAR[-5,-1]_{i,t} = \beta \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_t + \epsilon_{i,t}, \tag{4}$$

where the control variables are earnings surprises, firm size (log market capitalization), analyst following, news sentiment, and abnormal newswire coverage. We report the findings in Table 6 for the sample comprised of high-attention StockTwits stocks and their corresponding matched low-attention stocks. We further report the results for the full sample in Table IA4 of the Internet Appendix. In the univariate analysis for the matched sample, column (1) reports an increase in BHAR of 79 bps for high-attention stocks. With the additional control variables and fixed effects, column (2) reports an increase in BHAR of 108 bps for high-attention stocks. We then repeat the same analysis with this time the announcement date return $BHAR[0,1]$ as the dependent variable and including the $BHAR[-5,-1]$ as an additional control variable. Column (3) reports a 69 bps decrease in BHAR for high-attention stocks and a 53 bps decrease in column (4) when including all the control variables and the fixed effects.

In column (5), we replace the attention dummy in equation (4) for the fitted and residual components from regressing retail order imbalance using trades five days before announcements onto attention as defined in equation (1). Column (5) reports that the fitted component (Retail OI^{fit}) predicts the reversal and not the orthogonal component. We conclude that social-media-induced buying pressure ahead of announcements results in a price reversal on announcement dates.

4.1.1. Inventory risks

If the reversal documented in Section 4.1 results from intermediaries requiring a higher compensation to manage inventories when facing social-media-driven buying retail pressure, the magnitude of the reversal is expected to be larger when there is greater uncertainty regarding the market's reaction to earnings news. We select a firm's market capitalization, historical absolute announcement returns, and the 3-day average implied volatility before announcements as proxies for inventory risks. With respect to firm size, Martineau (2022) shows that absolute returns on announcement date are higher for small firms.

Table 7 presents time-series average announcement-window returns after independently double-sorting observations into high and matched low-attention StockTwits stocks (rows) and a quintile sort for a corresponding inventory risk proxy (columns). Panels A to C present the results sorting on firm size, historical announcement returns, and implied volatility, respectively.

For high-attention stocks, the average return reversal for small stocks is -2.49 bps (0.26 bps for large stocks), -1.06 bps for stocks with high absolute announcement returns (0.08 bps for low absolute returns), and -0.72 bps for high implied volatility stocks (0.05 bps for low implied volatility). Column "High-Low" confirms statistical significance at the 1% between the high and low sorted groups. We do not find such relationship for low-attention stocks, except when sorting on size. However, column "Low" (bottom size quintile) of Panel A reports a reversal of -2.49 bps for high-attention stocks and -0.56 bps for low-attention stocks. The difference is statistically significant with a t-statistics of -6.91. Overall, the effect of social media-induced price pressure ahead of earnings announcements is more pronounced for stocks where intermediaries encounter limited risk-bearing capacity.

4.2. Social Media and Price Revelation

A return reversal on earnings announcement is evidence consistent with market efficiency. After announcements, markets correct for the "inefficiency" caused by temporary price deviation ahead of announcements as compensation for intermediaries to accommodate the buying pressure. This section examines how such buying pressure impacts price informativeness with respect to fundamentals ahead of announcements and price efficiency following announcements. We follow the terminology of Brunnermeier (2005) to distinguish between "price informativeness" and "price efficiency," the main two components of the price discovery process. Price informativeness reflects the absolute level of information to future fundamentals (see also Biais, Hillion, and Spatt, 1999, Weller, 2018, Boguth, Fisher, Gregoire, and Martineau, 2024) and price efficiency relates to the price revelation of public information and how fast information is impounded into prices.

To examine the impact of price pressure on price informativeness, we graphically depict in Figure 5 the buy-and-hold abnormal returns (BHAR) around earnings announcements for high and matched low-StockTwits attention stocks. Panels A and B show the BHAR for positive earnings surprise (top two quintiles) and negative earnings surprise (bottom two quintiles), respectively. We rescale the figure such that BHAR is equal to zero at $t = -6$. Both panels show positive upward price drifts leading to earnings announcements for stocks with high StockTwits attention. In contrast, stocks with low coverage show no price drifts before positive earnings surprises and downward price drifts for negative earnings surprises. Five days before the announcement, the difference between BHAR for high vs low attention is approximately 60 bps in both panels. This figure conveys that in days leading to earnings announcements, social media can diminish price informativeness, i.e., in the case of low earnings surprises, prices deviate from future fundamentals to be revealed on announcement dates. In the case of positive earnings surprises, prices converge to fundamentals, not because social-media-induced trading reflects fundamentals but because social-media induced trading results in buying pressure, which pushes prices toward fundamentals.

The second main insight from this figure is how markets correct mispricing upon the release of earnings announcements. For positive earnings surprises (Panel A), markets take into account the heightened pre-announcement level and adjust prices less than those with low attention such that there is no significant difference in total returns. In other words, markets are efficient at adjusting prices to fundamental news post-announcement, independently of the stock's social media popularity. In Panel B, the BHAR of stocks with high StockTwits attention deviate from fundamentals ahead of earnings announcements with negative earnings surprises, but at the time of the announcement, the BHAR quickly converge to those with low StockTwits attention. In the days that follow the earnings announcement, we do not observe significant price drifts, consistent with the findings of Martineau (2022) that markets quickly process earnings news.

These news findings are important in light of the results reported in Campbell, Drake, Thornock, and Twedt (2023) and Ding, Shi, and Zhou (2023). These authors conclude that stocks with more Twitter and Seeking Alpha coverage before earnings announcements lead to smaller price reactions to earnings surprises, i.e., a lower earnings response coefficient (ERC), and conclude that social media "slows down" the price discovery process. Their conclusion is much different from ours. We reexamine this premise that social media attention following earnings announcements dampens the price discovery process by running the following regression

$$AR_{i,t} = \beta_1 Surp_{i,t} + \beta_2 \mathbb{1}^{Att}_{i,t} + \beta_3 Surp_{i,t} \times \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

where AR corresponds to the abnormal return on the announcement date in columns (1)–(3), buy-and-hold abnormal return (BHAR) from the announcement date to the next trading day in column (4), and two to five days after the announcement in column (5). Table 8 reports the results. Columns (1) to (4) report a negative earnings response coefficient ($Surp_{i,t} \times \mathbb{1}^{Att}_{i,t}$) of -0.18 to -0.23 for stocks with high StockTwits attention, corresponding to approximately a decline of 25% to the total earning response. These results lend support to the findings of Campbell, Drake, Thornock, and Twedt (2023). However, simply examining the regression coefficient is misleading. We show in Figure 5 that stocks with more social media activity ahead of earnings announcement have the same efficient price level as stocks with low social media activity following announcements. The reason we obtain a negative earnings response coefficient is simple. More than 67% of earnings announcements are associated with positive earnings surprises. Therefore, the negative relationship between social media activity and earnings surprises results from the buying price pressure leading to earnings announcements, which *diminishes* the price response to positive earnings surprises as shown in Figure 5. Column (5) of Table 8 provides additional evidence that high StockTwits attention leading to earnings announcements does not result in a continuation of price drifts two to five days following announcements as the loading on $Surp_{i,t} \times \mathbb{1}^{Att}_{i,t}$ is not statistically different from zero.

5. Is Social Media-Driven Noise Trading Predictable?

Our results up to this point find a contemporaneous relationship between social media attention and price pressure leading to earnings announcements. We next show that social-media-induced systematic retail trading is *ex-ante* forecastable and highlight the economic significance of the price reversal. We propose a long-short strategy centered on investor attention using an ex-ante measure of expected StockTwits attention to upcoming earnings announcements.

The ex-ante measure we utilize is based on social media posts about the "most anticipated earnings announcements" by Earnings Whispers (EW). EW aggregates analysts' earnings forecasts and provides insights into which stocks are likely to experience significant movements after earnings announcements and are more likely to have post-earnings announcement drifts. Starting in January 2016, during the weekend, Earnings Whispers releases a post highlighting the most "anticipated" earnings announcements. Figure 6 presents two examples of such posts. Typically, these posts are shared across all social media platforms over the weekend. As of February 2024, EW has 150,000 followers on StockTwits, 450,000 on X (formerly known as Twitter), and 116,000 on Instagram. We leverage these posts as an ex-ante indicator of the stocks most likely to attract retail investors.

Before proceeding to our long-short strategy, we first document how stock returns, attention, sentiment, and retail order imbalance change conditioning on appearing on a post made by EW. Table 9 presents the results from the following regression:

$$y_{i,t}^{post} = \beta_1 \mathbb{1}_{i,t}^{Ewhispers} + \Gamma' Controls_{i,t}^{prior} + \alpha_i + \alpha_t + \varepsilon_{i,t}, \quad (5)$$

where $y_{i,t}^{post}$ corresponds to the stock i buy-and-hold abnormal return (in percent), StockTwits attention (in percent), StockTwits sentiment, and retail trade order imbalance from the beginning of the earnings week (Monday) to $t-1$, i.e., the day before the earnings are announced on date t. $\mathbb{1}_{i,t}^{Ewhispers}$ is a dummy variable equal to one if stock i appears on the EW "most anticipated earnings" and zero otherwise.[18] The control variables ($Controls^{prior}$) are the buy-and-hold abnormal return, StockTwits attention and sentiment, and retail order imbalance from the prior week (Monday to Friday). We also control for the upcoming earnings surprise ($Surp$) and the absolute surprise ($|Surp|$), as well as the abnormal news coverage and average newswire sentiment spanning the last ten days prior to the announcement. We assign a neutral score of 0.5 for stocks with no sentiment-tagged posts.

Table 9 reports a statistically significant (at the 1% and 5% level) positive impact of a stock appearing in an EW post on its return, attention, sentiment, and retail order imbalance. We find an increase in BHAR by 51 basis points, a 2.7 basis points increase in attention (a 75% increase relative to the unconditional mean of 3.6 basis points), 2.3% increase in sentiment (a 4% increase relative to the unconditional mean), a 6.8% increase in retail trading, and a 1.1% increase in net retail order imbalance (a 55% increase relative to the unconditional mean) for stocks appearing in an EW post.

Having demonstrated the impact of the EW posts on investor attention, we next form long portfolios of stocks that appear on the EW posts and short the other stocks with earnings for that particular week but that do not appear on the EW posts. We value-weight

[18] We make sure to select only Earnings Whispers posts occurring on the weekend and on Monday before 9:30 am. The number of stocks-earnings observations appearing on a EW post that have an earnings announcement on Monday is 6.8% (431 observations) of the total sample (6,301 observations).

the stocks to create the portfolios.[19] The portfolios are rebalanced weekly. Once we obtain the daily portfolio returns for the long and short sides, we accumulate the daily returns to the monthly frequency. The "Long-Short" portfolio buys the long portfolio and sells the short portfolio. We rebalance the portfolios weekly. We execute the buy (sell) order starting Monday at 4 p.m. and hold the stock until t-1, 4 p.m., i.e., the last session of regular trading hours before the earnings announcement.

Table 10 presents the average monthly value-weighted long-short portfolio returns in percentage points for the pre- and post-announcement period, where the post-announcement period consists of the announcement date and the following trading day. The long-short column reveals that based on this strategy, the long-short portfolio generates abnormal returns. The long-short value-weighted portfolio consistently earns significant abnormal returns, whether using CAPM alphas, three-factors alphas, five-factors alphas, or five-factors with momentum alphas, which abnormal returns ranging from 0.43% per month (t-statistics = 2.81) to 0.45% (t-statistics = 2.37). It is worth noting that the majority of this spread is attributable to the long side, where the abnormal returns for the long portfolio range from 0.51% to 0.55% per month (t-statistics=2.29 and t-statistics=3.21, respectively) whereas the abnormal returns for the short side range from 0.06% to 0.11%. If one forms a long-short portfolio using the same strategy and hold the portfolio over $t = [0, 1]$, columns "Post-announcement" report that it results in a negative alpha ranging from -0.22% to -0.34%. The negative alphas are due to the return reversal.

In the Internet Appendix, Figure IA1 shows the time series of the long, short, and long-short cumulative returns since the initial portfolio inception. The largest increase in the performance occurs at the outset of the COVID-19 pandemic, a period of growing retail trading (Ozik, Sadka, and Shen, 2021, Martineau and Zoican, 2023) and investor attention to social media (see Figure 1). This figure provides evidence that the dynamics in

[19]From 2016 to 2022, we were not able to retrieve the EW "most anticipated earnings" posts for a total of 15 weeks in our sample. When a post is missing, we replace the missing week with the risk-free rate.

social-media-induced retail trading impact the magnitude of price pressure ahead of earnings announcements.

6. Implications to Future Research

We next discuss alternative social media platforms, precisely, the Reddit forum WallStreetBets and Seeking Alpha, and demonstrate why StockTwits is the most appropriate platform to examine the impact of social media on stock prices ahead of earnings announcements. We then present a simple theoretical framework to rationalize why investors consume optimistically biased information and trade on such information. This simple model provides fruitful avenues for future theoretical work to explore the role of social media in shaping investors' beliefs and in generating systematic noise trading.

6.1. Alternative social media platforms

StockTwits is not the only social media platform examined in the literature. Seeking Alpha and the Reddit forum WallStreetBets are two other platforms that have been examined. Cookson, Lu, Mullins, and Niessner (2022) highlights the importance of distinguishing sentiment and attention across different investor social media platforms.[20] Chen, De, Hu, and Hwang (2014) and Dim (2020) find that post sentiment on the social media platform Seeking Alpha predicts earnings surprises. An important distinction between StockTwits and Seeking Alpha is that Seeking Alpha contributors are not anonymous and get compensated for their posts.[21] A more closely related paper is Kang, Lou, Ozik, Sadka, and Shen (2024), which examines social media content from WallStreetBets around earnings announcements from 2020 to 2021 and finds that increased social discussion reduced pre-earnings turnover,

[20] Pyun (2024) further demonstrate the importance of examining real-time (synchronous) group chats such as Discord and compare them with forum-style (asynchronous) postings on Reddit's WallStreetBets.

[21] Farrell, Green, Jame, and Markov (2022) find that the ability of retail order imbalances to predict stock returns increases in the intraday following a Seeking Alpha publication. Ding, Zhou, and Li (2020) find that Seeking Alpha coverage reduces individual stock return comovement with the market.

return drift, and higher earnings response coefficients.[22] These findings depart from ours. StockTwits activity ahead of earnings announcements is excessively optimistic, fails to predict earnings surprises, and induces attention-based buying pressure that increases stock returns followed by a reversal on announcement dates.

The difference between our findings and other social media papers can result from the differences in coverage across platforms. Table IA5 highlights the significant discrepancy in coverage ahead of earnings announcements for StockTwits, WallStreetBets, and Seeking Alpha from 2018 to 2021. The table reports the number of stock-earnings observations with at least one post and the number of posts five days ahead of earnings announcements by NYSE quintile breakpoints. StockTwits has the broadest coverage, with more than 38,000 stock-earnings announcement observations compared to 5,543 and 2,515 observations for WallStreetBets and Seeking Alpha, respectively. The number of posts exceeds 5.5 million for StockTwits, close to 33,000 for WallStreetBets, and 3,500 for Seeking Alpha. Small and large firms are widely discussed on StockTwits, whereas more than 50% of the posts are about the largest firms (top quintile) on the other platforms. Stock prices of small firms are more sensitive to price impact. Consequently, a social media platform that widely covers small stocks will play a more determinant role in understanding the impact of social media on aggregate price dynamics of small firms.

6.2. Theoretical implications

Our results raise the question: Why would investors trade on optimistically-biased information? We present a rational-based model based on a special case of Caplin and Leahy (2019) to demonstrate why investors might consume optimistically biased information and trade on such information in a systematic way.[23] The model is motivated by our empirical findings

[22] Bradley, Hanousek Jr, Jame, and Xiao (2023) find that recommendations shared on WallStreetBets are significant predictors of returns and cash-flow news.

[23] The importance of wishful thinking in financial markets is further highlighted in Cassella, Dim, and Karimli (2023). The authors find that investors who are optimistic about a stock's prospect react to negative

and has implications to future theoretical work to better understand social media-driven systematic noise trading.

Consider a wishful-thinking investor who is considering buying $q > 0$ shares of an asset with price p before the release of the company's earnings announcement. For simplicity, we will abstract how q and p are determined and take them as given. After the release of the earnings announcement, the asset payoff $\tilde{v}$ can take two values: a high value $v_H = p + v$ after a positive surprise or a low value $v_L = p - v$ after a negative surprise, where $v > 0$ and $v_H > v_L$. There is an objective probability for each value. With probability $\bar{\pi}_H$ there is a positive surprise and a high realization of the asset v_H, and with probability $\bar{\pi}_L$ there is a negative surprise and a low realization of the asset v_L. An alternative interpretation of the objective probabilities is that these probabilities represent the consensus or mainstream opinion in case there are agents with heterogeneous information.

The model assumes that wishful-thinking investors have subjective beliefs about the probability realization of $\tilde{v}$. We denote π_H as the subjective probability of a positive surprise v_H and π_L as the subjective probability of a negative surprise v_L. These subjective beliefs may differ from objective beliefs, but deviating from objective beliefs is costly. We represent the cost of deviating from objective beliefs by the Kullback-Leibler distance:

$$\frac{1}{\theta}\pi_H \ln \frac{\pi_H}{\bar{\pi}_H} + \frac{1}{\theta}\pi_L \ln \frac{\pi_L}{\bar{\pi}_L}.$$

The parameter θ represents the ease with which the agent can manipulate their beliefs. The larger is θ, the greater the amount of evidence the agent would need before they reject their chosen beliefs in favor of the objective ones. In other words, the larger θ, the more likely the investor is to opt for subjective beliefs. The lower the θ, the more costly it is to deviate from the objective beliefs.

The investor's expected utility of holding the asset and manipulating beliefs is then given

news by shifting their optimistic expectations to a longer forecast horizon.

by:

$$EU(\pi_H, \pi_L) = q(\pi_H v_H + \pi_L v_L - p) - \frac{1}{\theta}\pi_H \ln \frac{\pi_H}{\bar{\pi}_H} - \frac{1}{\theta}\pi_L \ln \frac{\pi_L}{\bar{\pi}_L}. \tag{6}$$

The investor understands the preferences and that the beliefs differ from the objective beliefs. The wishful thinking investor will choose subjective beliefs π_H and π_L by maximizing expected utility in (6), taking into account that $\pi_H + \pi_L = 1$. The optimization problem leads the investor to choose the following subjective beliefs:[24]

$$\pi_H = \frac{\bar{\pi}_H \exp(\theta q v_H)}{\bar{\pi}_H \exp(\theta q v_H) + \bar{\pi}_L \exp(\theta q v_L)}. \tag{7}$$

The investor chooses to distort beliefs towards states with positive surprises v_H so that $\pi_H > \bar{\pi}_H$ for $\bar{\pi}_H \in (0,1)$. The investor exhibits wishful thinking behavior by being over-optimistic about the high utility states. In other words, the wishful-thinking investor obtains utility from anticipating future events. At the extremes, when the objective probability is either zero or one, subjective probabilities are equal to objective probabilities, and the investor is rational. A wishful-thinking investor will not get any utility for dreaming about impossible events. As the cost of manipulating beliefs decreases (θ increases), the beliefs become even more distorted towards positive surprises. The same effect appears the more shares q the investor is considering to buy; as q increases the subjective probability π_H deviates more from the objective probability $\bar{\pi}_H$ and thus more positive optimistic biased are investors.

We can observe how a wishful-thinking investor distorts beliefs in a numerical example in Figure 7. In this figure, we set the following parameters: $v_H = 3$, $v_L = 1$, $\theta = .5$ and $q = 1, 3, 5$. The solid line represents the beliefs of a wishful-thinking investor given by (7). The dashed line represents the beliefs of a rational investor that uses the objective beliefs $\pi_H^{Rational} = \bar{\pi}_H$. The figure shows that the wishful-thinking investor distorts beliefs towards positive surprises. Even when the probability of a positive surprise is less likely

[24]See Section IA for derivations.

than a negative surprise $\bar{\pi}_H < 0.5$, the wishful thinking investor may distort beliefs so that $\pi_H > 0.5$. In words, even when the consensus is that there will be a negative surprise, the wishful-thinking investor may think that a positive surprise is more likely (for example, when $\bar{\pi}_H = 0.4$, then $\pi_H > 0.5$). As the consensus probabilities get closer to the extremes, when events are almost certain, then wishful-thinking investors resemble rational investors. Figure 7 shows how beliefs get distorted as the number of shares q increases. As the stakes increase, there is an increase in the distortion of beliefs.

The wishful thinking investor will choose to purchase q units of the asset at price p when the expected utility in equation (6) with subjective beliefs given by (7) is positive $EU(\pi_H, \pi_L) \geq 0$, which happens when:

$$\bar{\pi}_H \geq \frac{\exp(\theta q p) - \exp(\theta q v_L)}{\exp(\theta q v_H) - \exp(\theta q v_L)} = \frac{1}{1 + \exp(\theta q v)} = \bar{\pi}_H^{cutoff}.$$

Thus, a wishful thinking investor will choose to purchase the q shares of an asset at price p when $\bar{\pi}_H \geq \bar{\pi}_H^{cutoff}$. Instead a rational investor with $\pi_H^{Rational} = \bar{\pi}_H$ would choose to purchase the q shares of an asset at price p when $\bar{\pi}_H \geq 0.5$. We can see that a wishful-thinking investor would make the same choices as a rational investor only when it is infinitely costly to distort beliefs ($\theta = 0$). For any $\theta > 0$, the wishful thinking investor will have a lower cutoff to purchase the asset than a rational investor such that $\bar{\pi}_H^{cutoff} < 0.5$.

We believe that more theoretical work on wishful thinking could shed some light on the role of social media in financial markets. In Banerjee, Davis, and Gondhi (2024), wishful thinking leads to endogenous disagreement. Their findings show a connection between wishful thinking and market outcomes, including return volatility, price informativeness, trading volume, and return predictability, which match empirical evidence presented in this paper.

7. Conclusion

Social media activity during the week leading up to announcements is overly optimistic and does not forecast the earnings fundamentals. This optimism leads to systematic noise trading, resulting in a buying pressure that increases stock returns by more than 50 bps ahead of earnings announcements as market makers seek higher returns to provide liquidity. Noise trading, when correlated, can increase inventory risk for market makers when providing liquidity before anticipated information events.

Our findings show that social media activity ahead of earnings announcements can distort price revelation because of the higher required rate of return by market makers adds noise to prices. Conditioning on announcements' earnings surprises, the average price run-up pushes prices closer to fundamentals for positive earnings surprise announcements, albeit not a reflection of informative trading, and away from fundamentals ahead of announcements with negative earnings surprises. On the announcement date, markets efficiently adjust prices quickly to new information and correct any mispricing generated by social media information production. We further show that social media-driven noise trading is predictable.

A key implication of our paper to future research is that theories should consider the role of systematic noise trading in price formation. Moreover, analyzing the effect of social media on post-announcement price formation without considering pre-announcement price dynamics can lead to biased inferences on the role of social media in price discovery. Finally, our paper does not claim that there is no useful information about upcoming earnings on StockTwits, but that, on average, the information content is uninformative about earnings fundamentals. Differentiating between "skilled" and "unskilled" social media content creators, or "finfluencer," is a fruitful avenue for future research. Kakhbod, Kazempour, Livdan, and Schuerhoff (2023) have already made admirable progress in that regard.

References

Akey, Pat, Vincent Grégoire, and Charles Martineau, 2022, Price revelation from insider trading: Evidence from hacked earnings news, *Journal of Financial Economics* 143, 1162–1184.

Antweiler, Werner, and Murray Z. Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *Journal of Finance* 59, 1259–1294.

Banerjee, Snehal, Jesse Davis, and Naveen Gondhi, 2024, Choosing to disagree: Endogenous dismissiveness and overconfidence in financial markets, *Journal of Finance* 79, 1635–1695.

Barber, Brad M., Xing Huang, Philippe Jorion, Terrance Odean, and Christopher Schwarz, 2024, A (sub) penny for your thoughts: Tracking retail investor activity in TAQ, *Journal of Finance* 79, 2403–2427.

Barber, Brad M., Xing Huang, Terrance Odean, and Christopher Schwarz, 2022, Attention-induced trading and returns: Evidence from robinhood users, *Journal of Finance* 77, 3141–3190.

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean, 2009, Just how much do individual investors lose by trading?, *Review of Financial Studies* 22, 609–632.

Barber, Brad M., Shengle Lin, and Terrance Odean, 2023, Resolving a paradox: Retail trades positively predict returns but are not profitable, *Journal of Financial and Quantitative Analysis* pp. 1–35.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

———, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

———, and Ning Zhu, 2008, Do retail trades move markets?, *Review of Financial Studies* 22, 151–186.

———, 2009, Systematic noise, *Journal of Financial Markets* 12, 547–569.

Barrot, Jean-Noel, Ron Kaniel, and David Sraer, 2016, Are retail traders compensated for providing liquidity?, *Journal of Financial Economics* 120, 146–168.

Barth, Mary E., and Eric C. So, 2014, Non-diversifiable volatility risk and risk premiums at earnings announcements, *The Accounting Review* 89, 1579–1607.

Bartov, Eli, Lucile Faurel, and Partha S. Mohanram, 2018, Can twitter help predict firm-level earnings and stock returns?, *The Accounting Review* 93, 25–57.

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1999, Price discovery and learning during the preopening period in the paris bourse, *Journal of Political Economy* 107, 1218–1248.

Bianchi, Francesco, Roberto Gómez-Cram, and Howard Kung, 2024, Using social media to identify the effects of congressional viewpoints on asset prices, *Review of Financial Studies* 37, 2244–2272.

Boehmer, Ekkehart, Charles M Jones, Xiaoyan Zhang, and Xinran Zhang, 2021, Tracking retail investor activity, *Journal of Finance* 76, 2249–2305.

Boguth, Oliver, Adlai J. Fisher, Vincent Gregoire, and Charles Martineau, 2024, Noisy FOMC returns? Information, price pressure, and post-announcement reversals, *Working paper*.

Boguth, Oliver, Vincent Grégoire, and Charles Martineau, 2019, Shaping expectations and coordinating attention: The unintended consequences of fomc press conferences, *Journal of Financial and Quantitative Analysis* 54, 2327–2353.

Bradley, Daniel, Jan Hanousek Jr, Russell Jame, and Zicheng Xiao, 2023, Place your bets? The market consequences of investment research on reddit's wallstreetbets, *Working paper*.

Brunnermeier, Markus K., 2005, Information leakage and market efficiency, *Review of Financial Studies* 18, 417–457.

Campbell, Brett, Michael Drake, Jacob Thornock, and Brady Twedt, 2023, Earnings virality, *Journal of Accounting and Economics* 75, 101517.

Caplin, Andrew, and John V. Leahy, 2019, Wishful thinking, *NBER Working Paper*.

Cassella, Stefano, Chukwuma Dim, and Tural Karimli, 2023, Optimism shifting, *Working paper*.

Chen, Hailiang, Prabuddha De, Yu Jeffrey Hu, and Byoung-Hyoun Hwang, 2014, Wisdom of crowds: The value of stock opinions transmitted through social media, *Review of Financial Studies* 27, 1367–1403.

Comerton-Forde, Carole, Terrence Hendershott, Charles M. Jones, Pamela C. Moulton, and Mark S. Seasholes, 2010, Time variation in liquidity: The role of market-maker inventories and revenues, *Journal of Finance* 65, 295–331.

Cookson, J. Anthony, Joseph E. Engelberg, and William Mullins, 2020, Does partisanship shape investor beliefs? Evidence from the COVID-19 pandemic, *Review of Asset Pricing Studies* 10, 863–893.

———, 2023, Echo chambers, *Review of Financial Studies* 36, 450–500.

Cookson, J. Anthony, Vyacheslav Fos, and Marina Niessner, 2021, Does disagreement facilitate informed trading? evidence from activist investors, *Working paper* p. 45.

Cookson, J. Anthony, Runjing Lu, William Mullins, and Marina Niessner, 2022, The social signal, *Journal of Financial Economics, forthcoming.*

Cookson, J. Anthony, William Mullins, and Marina Niessner, 2024, Social media and finance, *Working paper*.

Cookson, J. Anthony, and Marina Niessner, 2020, Why don't we agree? Evidence from a social network of investors, *Journal of Finance* 75, 173–228.

———, and Christoph Schiller, 2022, Can social media inform corporate decisions? Evidence from merger withdrawals, *Working paper*.

Cowen, Amanda, Boris Groysberg, and Paul Healy, 2006, Which types of analyst firms are more optimistic?, *Journal of Accounting and Economics* 41, 119–146.

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

Dim, Chukwuma, 2020, Should retail investors listen to social media analysts? Evidence from text-implied beliefs, *Working paper.*

Ding, Rong, Yukun Shi, and Hang Zhou, 2023, Social media coverage and post-earnings announcement drift: Evidence from Seeking Alpha, *The European Journal of Finance* 29, 207–227.

Ding, Rong, Hang Zhou, and Yifan Li, 2020, Social media, financial reporting opacity, and return comovement: Evidence from Seeking Alpha, *Journal of Financial Markets* 50, 100511.

Fama, Eugene F., 1965, The behavior of stock-market prices, *Journal of Business* 38, 34–105.

Farrell, Michael, T. Clifton Green, Russell Jame, and Stanimir Markov, 2022, The democratization of investment research and the informativeness of retail investor trading, *Journal of Financial Economics* 145, 616–641.

Fisher, Adlai, Charles Martineau, and Jinfei Sheng, 2022, Macroeconomic attention and announcement risk premia, *The Review of Financial Studies* 35, 5057–5093.

Friedman, Milton, et al., 1953, The case for flexible exchange rates, *Essays in positive economics* 157, 33.

Gregoire, Vincent, and Charles Martineau, 2022, How is earnings news transmitted to stock prices?, *Journal of Accounting Research* 60, 261–297.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Han, Bing, and Liyan Yang, 2013, Social networks, information acquisition, and asset prices, *Management Science* 59, 1444–1457.

Heimer, Rawley Z, 2016, Peer pressure: Social interaction and the disposition effect, *Review of Financial Studies* 29, 3177–3209.

Hirshleifer, David, Lin Peng, and Qiguang Wang, 2024, News diffusion in social networks and stock market reactions, *Review of Financial Studies, forthcoming.*

Hu, Danqi, Charles M. Jones, Valerie Zhang, and Xiaoyan Zhang, 2021, The rise of Reddit: How social media affects retail investors and short-sellers' roles in price discovery, *Working paper.*

Hvidkjaer, Soeren, 2008, Small trades and the cross-section of stock returns, *Review of Financial Studies* 21, 1123–1151.

Jia, Weishi, Giulia Redigolo, Susan Shu, and Jingran Zhao, 2020, Can social media distort price discovery? Evidence from merger rumors, *Journal of Accounting and Economics* 70, 101334.

Jiao, Peiran, Andre Veiga, and Ansgar Walther, 2020, Social media, news media and the stock market, *Journal of Economic Behavior & Organization* 176, 63–90.

Kakhbod, Ali, Seyed Mohammad Kazempour, Dmitry Livdan, and Norman Schuerhoff, 2023, Finfluencers, *Working paper.*

Kang, Namho, Xiaoxia Lou, Gideon Ozik, Ronnie Sadka, and Siyi Shen, 2024, Innocuous noise? Social media and asset prices, *Working paper.*

Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman, 2012, Individual investor trading and return patterns around earnings announcements, *Journal of Finance* 67, 639–680.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *Journal of Finance* 63, 273–310.

Kelley, Eric K., and Paul C. Tetlock, 2013, How wise are crowds? Insights from retail orders and stock returns, *Journal of Finance* 68, 1229–1265.

Kothari, Sagar P., Eric So, and Rodrigo Verdi, 2016, Analysts' forecasts and asset pricing: A survey, *Annual Review of Financial Economics* 8, 197–219.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* pp. 1315–1335.

Laarits, Toomas, and Marco Sammon, 2023, The retail habitat, *Working paper.*

Lee, Charles MC, Belinda Mucklow, and Mark J. Ready, 1993, Spreads, depths, and the impact of earnings information: An intraday analysis, *Review of Financial Studies* 6, 345–374.

Li, Edward Xuejun, K. Ramesh, Min Shen, and Joanna Shuang Wu, 2015, Do analyst stock recommendations piggyback on recent corporate news? An analysis of regular-hour and after-hours revisions, *Journal of Accounting Research* 53, 821–861.

Liu, Bibo, Huijun Wang, Jianfeng Yu, and Shen Zhao, 2020, Time-varying demand for lottery: Speculation ahead of earnings announcements, *Journal of Financial Economics* 138, 789–817.

Madhavan, Ananth, and Seymour Smidt, 1993, An analysis of changes in specialist inventories and quotations, *Journal of Finance* 48, 1595–1628.

Martineau, Charles, 2022, Rest in peace post-earnings announcement drift, *Critical Finance Review* 11, 613–646.

———, and Marius Zoican, 2023, Retail trading and analyst coverage, *Journal of Financial Markets* 66, 100849.

Ozik, Gideon, Ronnie Sadka, and Siyi Shen, 2021, Flattening the illiquidity curve: Retail trading during the covid-19 lockdown, *Journal of Financial and Quantitative Analysis* 56, 2356–2388.

Pedersen, Lasse Heje, 2022, Game on: Social networks and markets, *Journal of Financial Economics* 146, 1097–1119.

Peng, Lin, Qiguang Wang, and Dexin Zhou, 2022, Social networks, trading, and liquidity, *Journal of Portfolio Management* 48, 196–215.

Peress, Joel, and Daniel Schmidt, 2020, Glued to the tv: Distracted noise traders and stock market liquidity, *Journal of Finance* 75, 1083–1133.

——— , 2021, Noise traders incarnate: Describing a realistic noise trading process, *Journal of Financial Markets* 54, 100618.

Pyun, Chaehyun, 2024, Synchronous social media and the stock market, *Journal of Financial Markets* p. 100915.

So, Eric C., and Sean Wang, 2014, News-driven return reversals: Liquidity provision ahead of earnings announcements, *Journal of Financial Economics* 114, 20–35.

Van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira, 2023, Man versus machine learning: The term structure of earnings expectations and conditional biases, *Review of Financial Studies* 36, 2361–2396.

Waters, Brian, Xingtan Zhang, Wei Zhou, and Shrihari Santosh, 2024, Exploiting non-fundamental flows, *Working paper*.

Weller, Brian M., 2018, Does algorithmic trading reduce information acquisition?, *Review of Financial Studies* 31, 2184–2226.

Figure 1. StockTwits Activity Over Time

This figure shows the monthly number of stock-specific posts on StockTwits. The sample period is from January 1, 2013, to December 31, 2022.

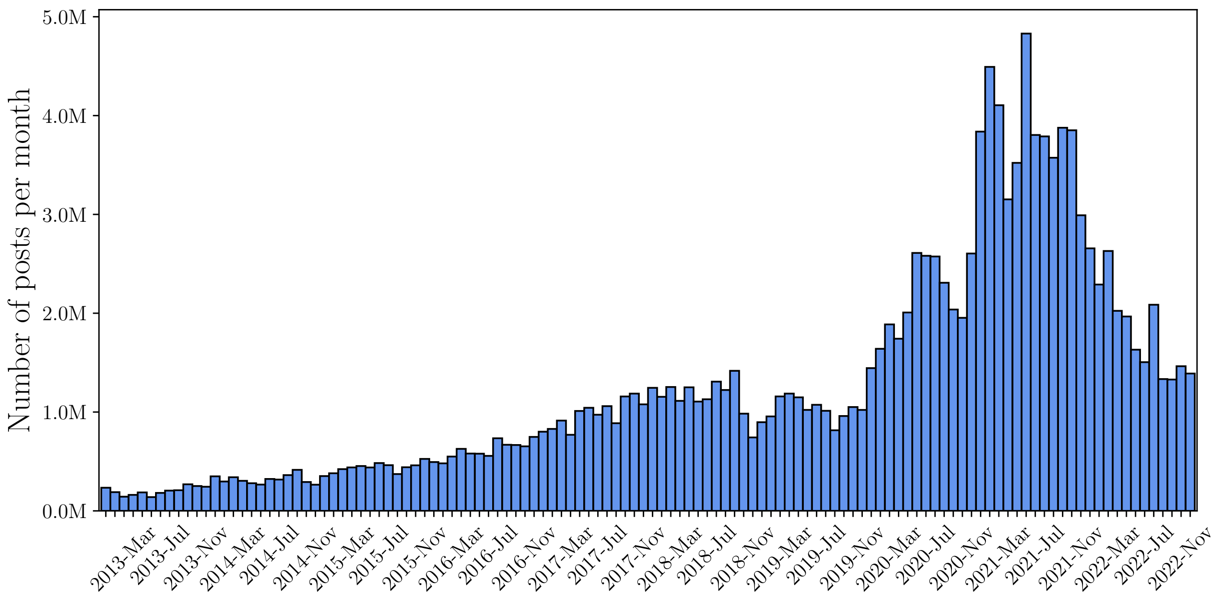


Figure 2. Abnormal Attention Around Earnings Announcements

This figure shows a boxplot representing the distribution of the number of abnormal StockTwits posts and newswires coverage five days around earnings announcements. The whiskers correspond to the 5th and 95th percentiles. Abnormal attention (newswire coverage) is computed as the daily log number of StockTwits posts (newswire articles) minus the log of the average daily number of posts (newswire articles) from 20 to 6 days before the earnings announcements. The sample period is from January 1, 2013, to December 31, 2022.

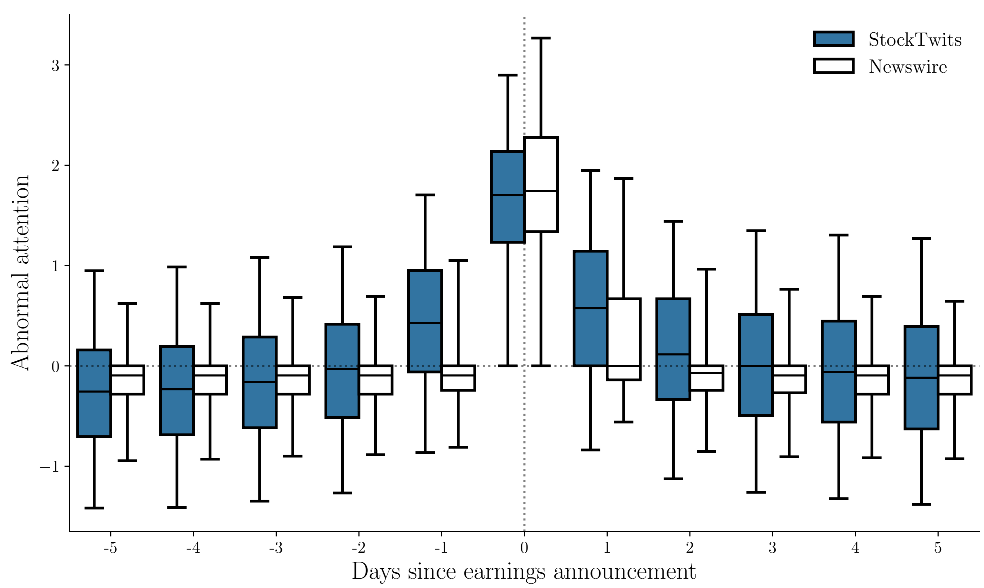


Figure 3. StockTwits Sentiment is Optimistic

This figure shows the fraction of stock-earnings observations by (1) the fraction of bullish StockTwits posts and (2) the fraction of bullish (positive) analyst recommendations sixty days before earnings announcements. The sample period is from January 1, 2013, to December 31, 2022.

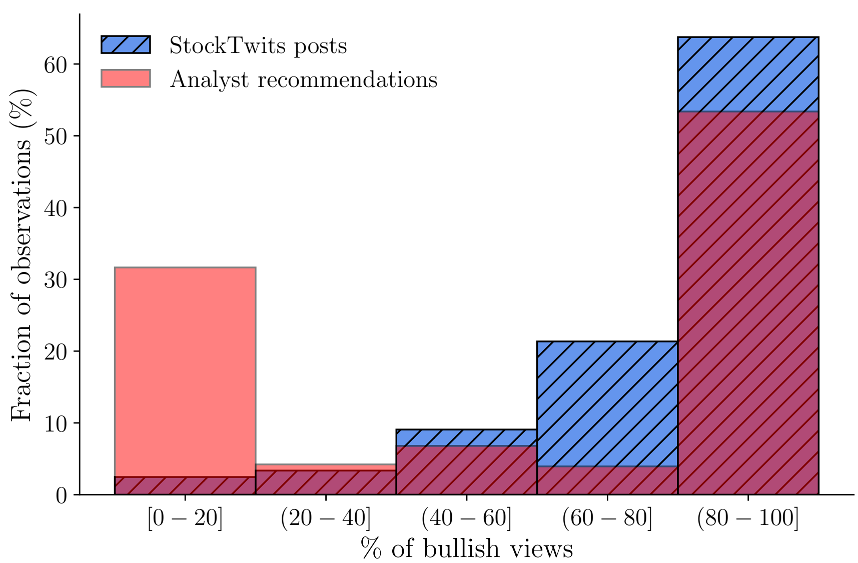


Figure 4. Return Reversals After Earnings Announcements

This figure plots the average difference in buy-and-hold abnormal return (BHAR, in %) between high-StockTwits attention stocks and matched low-StockTwits attention stocks around earnings announcements ($t = 0$). High-attention stocks correspond to the top quintile stock-earnings announcements with the highest coverage on StockTwits five days before announcements. Matched stocks are assigned based on the same industry, the BHAR 30 to six days before announcements, firm size, and earnings surprises. The 95% confidence intervals are represented by the error bars. The sample period is from January 1, 2013, to December 31, 2022.

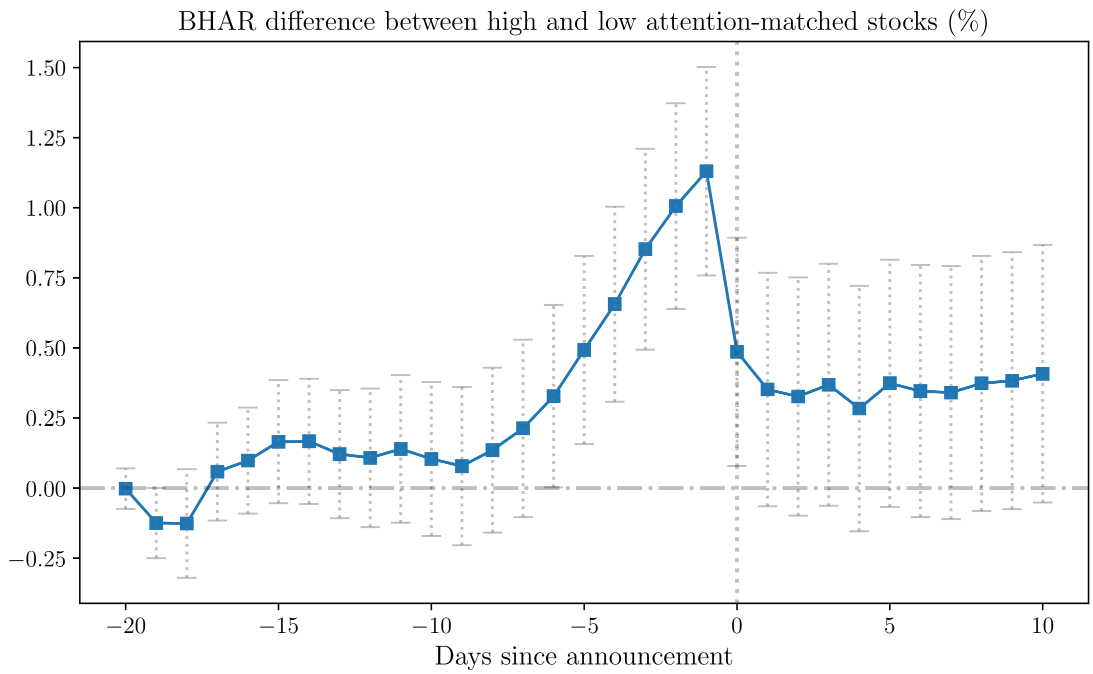
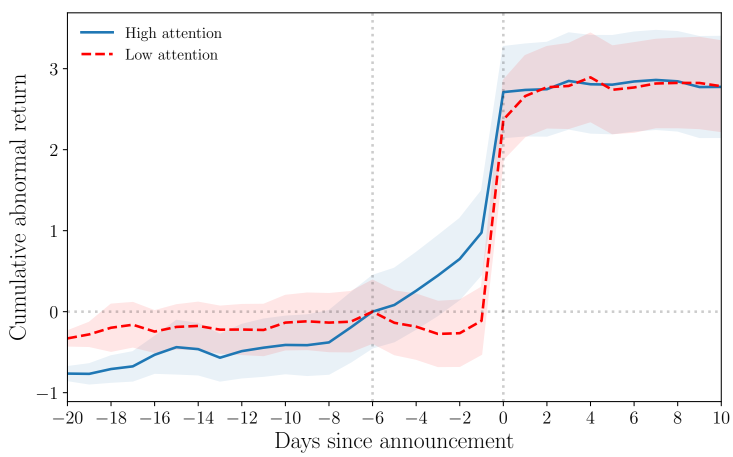


Figure 5. Cumulative Returns Around Earnings Announcements Conditioning on StockTwits Attention and Earnings Surprises

This figure shows the buy-and-hold abnormal returns (BHAR, in %) around earnings announcements for stocks with high-StockTwits attention and matched-low StockTwits attention 20 days before to 10 days after earnings announcements. High attention stocks correspond to the top quintile stock-earnings announcements with the highest coverage on StockTwits five days before announcements. Matched stocks are assigned based on the same industry, the BHAR 20 to six days before announcement, firm size, and on earnings surprises. Panels A and B show the cumulative returns around earnings announcements with positive (top two quintiles) and negative earnings surprises (bottom two quintiles), respectively. The shaded area corresponds to the 95% confidence intervals. The plots are rescaled to zero at $t = -6$. The sample period is from January 1, 2013, to December 31, 2022.

Panel A. Positive earnings surprises



Panel B. Negative earnings surprises

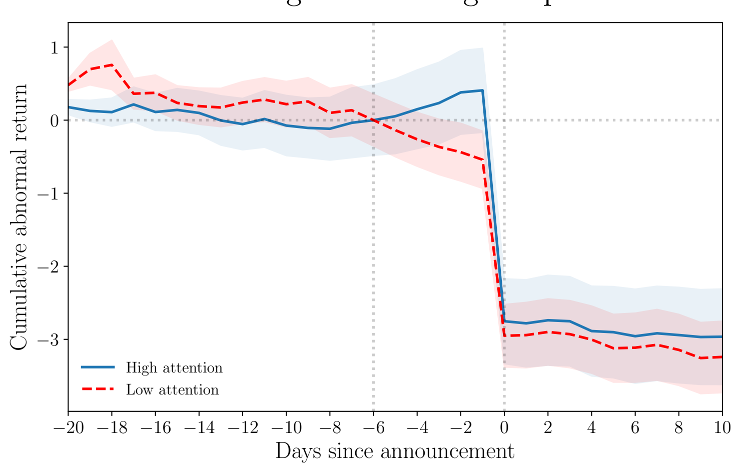


Figure 6. Earnings Whispers Social Media Posts

This figure shows two examples of Earnings Whispers StockTwits posts about upcoming earnings announcements for the week of August 29, 2022 and September 26, 2022.

Panel A. August 29, 2022

Panel B. September 26, 2022

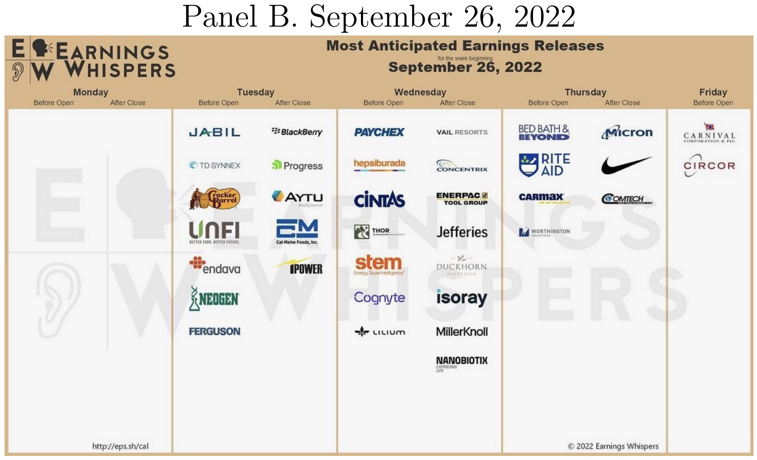


Figure 7. Plots of π_H for Wishful Thinking Versus Rational Investors

This figure presents the relationship between π_H and $\bar{\pi}_H$ for a rational and wishful thinking agent. Dashed black line represents a rational agent. Solid lines represent wishful-thinking investors for different quantities q. We set $v_H = 3, v_L = 1, \theta = 0.5$.

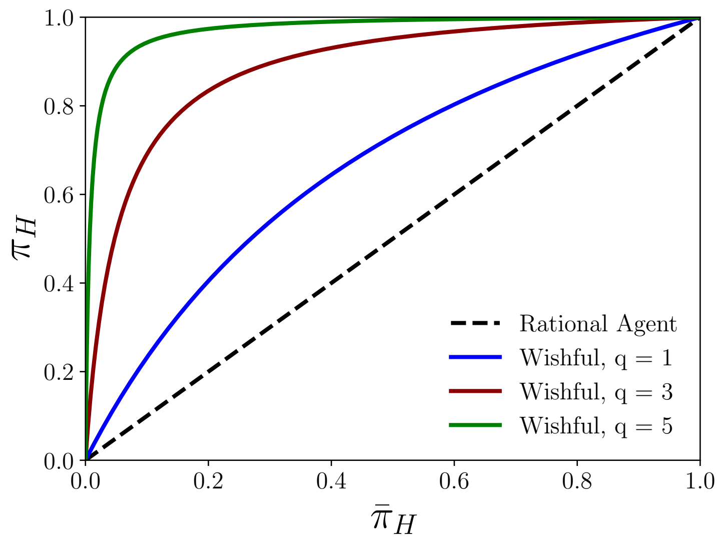


Table 1
Summary Statistics on High and Low StockTwits Coverage

This table reports summary statistics of the stock-earnings announcement sample for high and low StockTwits attention stocks. High-attention stocks correspond to the top quintile stock-earnings announcements with the highest coverage on StockTwits five days before the announcement. Volatility is the standard deviation of 30 daily returns ending ten days before announcements. *Abn ret.* and *Abn* $|ret.|$ corresponds to the abnormal and absolute abnormal returns on the earnings announcement date, respectively. The sample period is from January 2013 to December 2022.

	High attention				Low attention			
	Mean	25th	Median	75th	Mean	25th	Median	75th
Market cap (mill.$)	35,179	849	5,482	28,335	4,720	351	1,204	3,758
Volatility (%)	3.23	1.50	2.36	3.87	2.57	1.41	1.99	3.05
Abn ret. (%)	-0.43	-4.66	-0.35	3.77	-0.03	-3.59	0.04	3.61
Abn |ret.| (%)	6.41	1.84	4.23	8.47	5.57	1.49	3.60	7.35
Surprise (%)	-1.90	-0.05	0.06	0.26	-0.23	-0.09	0.06	0.29
N. analysts	10.16	4.00	9.00	15.00	5.33	2.00	4.00	7.00

Table 2
Summary Statistics on Coverage

This table reports summary statistics on StockTwits coverage, analyst recommendations, and newswire coverage in Panels A to C, respectively. N. posts, N. rec., and N. news correspond to the total number of StockTwits posts, analyst recommendations, and Dow Jones newswires, respectively, five to one day before earnings announcements. The sample period is from January 2013 to December 2022.

A. StockTwits coverage

	StockTwits coverage			No StockTwits coverage	
NYSE qnt	Stock-EA obs.	N. posts	% of posts	Stock-EA obs.	% with no coverage
1 (small)	31,145	2,213,564	24	4,820	13
2	19,662	1,204,046	13	1,218	6
3	15,406	1,132,371	12	678	4
4	13,953	1,200,506	13	479	3
5 (large)	13,780	3,320,317	37	315	2
Total	93,946	9,070,804	100	7,510	

B. Analyst recommendation

	Analyst recommendation			No analyst recommendation	
NYSE qnt	Stock-EA obs.	N. rec.	% of rec.	Stock-EA obs.	% with no rec.
1 (small)	779	1,350	5	35,186	98
2	999	2,326	8	19,881	95
3	1,231	2,031	7	14,853	92
4	1,592	3,497	13	12,840	89
5 (large)	3,129	18,185	66	10,966	78
Total	7,730	27,389	100	93,726	

C. Newswire

	News			No news	
NYSE qnt	Stock-EA obs.	N. news	% of news	Stock-EA obs.	% with no news
1 (small)	6,837	24,520	4	29,128	81
2	5,372	28,763	5	15,508	74
3	5,527	37,890	6	10,557	66
4	6,890	75,524	12	7,542	52
5 (large)	11,120	461,581	73	2,975	21
Total	35,746	628,278	100	65,710	

Table 3
StockTwits' Sentiment Does Not Predict Fundamentals

This table reports estimates for the full sample, large caps (the top three NYSE market capitalization quintiles), and small caps (bottom two quintiles) of the following regression:

$$Surp_{i,t} = \beta_1 Sent_{i,t} + \beta_2 \mathbb{1}^{Att}_{i,t} \times Sent_{i,t} + \beta_3 \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

where $Surp_{i,t}$ is the earnings surprise (%) for stock i on earnings announcement t. $Sent$ is the sentiment on StockTwits as defined in equation (2) over five days before earnings announcements, and $\mathbb{1}^{Att}_{i,t}$ is a dummy variable equal to one if the stock attention on StockTwits belongs to the top quintile. $Controls_{i,t}$ is a vector of control variables corresponding to the buy-and-hold abnormal returns (BHAR[-5,-1]), sentiment from analyst recommendations (Analysts sent), and RavenPack newswire sentiment (News sent), all measured in the five days leading up to the earnings announcements. α_i and α_t are stock and time-fixed effects, respectively. Columns(1)–(3) exclude observations for which posts have no sentiment tags. Columns (4)–(6) include non-tagged sentiment posts with an assigned neutral sentiment score of 0.5. The sample period is from January 2013 to December 2022. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Sample:	Excl. non-tagged posts			Incl. non-tagged posts as neutral		
	Full sample (1)	Large caps (2)	Small caps (3)	Full sample (4)	Large caps (5)	Small caps (6)
Sent	0.015	-0.005	0.046	-0.010	-0.011	-0.015
	(0.032)	(0.014)	(0.069)	(0.026)	(0.012)	(0.048)
$\mathbb{1}^{Att} \times$ Sent	0.039	0.032	0.144	0.022	0.053	0.088
	(0.074)	(0.037)	(0.273)	(0.070)	(0.035)	(0.237)
$\mathbb{1}^{Att}$	-0.045	-0.048	-0.099	-0.059	-0.061**	-0.127
	(0.063)	(0.030)	(0.236)	(0.056)	(0.027)	(0.199)
$\text{BHAR}_{[-5,-1]}$	0.577**	0.415**	0.549*	0.497**	0.363**	0.541**
	(0.249)	(0.175)	(0.326)	(0.200)	(0.147)	(0.253)
Analysts sent	0.019	0.029*	0.024	0.017	0.024*	0.015
	(0.028)	(0.015)	(0.132)	(0.024)	(0.013)	(0.087)
News sent	0.025	0.016	0.200	0.094	-0.006	0.360
	(0.099)	(0.069)	(0.436)	(0.081)	(0.055)	(0.280)
N	60,105	30,736	29,369	101,393	44,598	56,795
R^2	0.000	0.001	0.000	0.000	0.001	0.000
Date FE	Y	Y	Y	Y	Y	Y

Table 4
StockTwits Attention and Retail Buying Pressure

This table reports the β estimate, t-statistic, and R^2 of the following regression:

$$Retail\ OI_{i,t} = \beta \mathbb{1}_{i,t}^{Att} + \alpha_i + \alpha_t + \epsilon_{i,t},$$

where *Retail OI* is the retail order imbalance over five days before earnings announcements. $\mathbb{1}_{i,t}^{Att}$ is a dummy variable equal to one if the stock belongs to the top attention quintile. α_i and α_t are stock and time-fixed effects, respectively. We employ the method of Barber, Huang, Jorion, Odean, and Schwarz (2024) to label retail trades from TAQ, and we compute retail order imbalance using trades, volume, dollar volume, and their corresponding detrending counterpart. We detrend the retail order imbalances by the retail order imbalance 50 to 6 days before earnings announcements. We compute retail order imbalance using trades, volume, and dollar volume in Panels A–B, C–D, and E–F, respectively. The table reports the coefficient loading β and its corresponding t-statistic and the regression's R^2. The sample period is from January 2013 to December 2022.

Coeff	t-stat	R^2	Fixed effects
A. Trades			
0.027	14.27	0.24	firm-date
B. Trades detrended			
0.013	9.83	0.10	date
C. Volume			
0.011	5.76	0.02	firm-date
D. Volume detrended			
0.007	5.31	0.02	date
E. Dollar volume			
0.011	5.72	0.02	firm-date
F. Dollar volume detrended			
0.007	5.23	0.02	date

Table 5
Social Media Attention and Earnings Announcement Reversals

This table reports the average buy-and-hold abnormal return (BHAR, in %) over day intervals, $[t_1, t_2]$, around earnings announcement defined in the first column for high and low StockTwits attention stocks. Columns *EA* and *Pseudo-EA* correspond to the BHAR for the actual earnings announcement and a pseudo-earnings announcement, respectively. We select pseudo-announcement dates by randomly selecting a pseudo-date 50 to 20 days window prior to actual announcement dates. *Diff* in column (3) corresponds to the difference between (1) and (2), and $Diff^{Matched}$ and $Diff^{Full}$ correspond to the same difference but for matched low attention stocks and for the full sample of low attention stocks. Columns (6) and (7) report the diff-in-diff estimate where the diff-in-diff estimate in column (6) is the difference between columns (3) and (4), and the diff-in-diff estimate in column (7) is the difference between columns (3) and (5). The t-statistic of the difference is reported in square brackets. Bold t-statistic indicates statistical significance at the 10% level. The sample period is from January 2013 to December 2022.

	High attention			Low attention		Diff-in-Diff	
	EA (1)	Pseudo EA (2)	Diff (3)	Diff^{Matched} (4)	Diff^{Full} (5)	Diff-in-Diff^{Matched} (6)	Diff-in-Diff^{Full} (7)
$[-5, -1]$	0.575	0.087	0.488	-0.161	-0.129	0.649	0.616
			[4.68]	**[-2.44]**	**[-4.31]**	**[5.27]**	**[4.10]**
$[0, 1]$	-0.448	0.025	-0.474	0.111	-0.026	-0.584	-0.448
			[-5.42]	[1.46]	[-0.68]	**[-5.04]**	**[-3.72]**
$[2, 5]$	0.051	0.088	-0.036	0.059	0.039	-0.095	-0.075
			[-0.40]	[0.87]	[1.28]	[-0.84]	[0.39]
$[6, 10]$	0.085	0.060	0.027	0.089	0.293	-0.061	-0.266
			[0.28]	[1.27]	**[8.96]**	[-0.47]	**[-3.07]**

Table 6
Pre-earnings Announcement Returns and StockTwits Activity

This table reports estimates of the following regression:

$$BHAR[t_1, t_2]_{i,t} = \beta_1 \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_t + \epsilon_{i,t}$$

where $BHAR[t_1, t_2]_{i,t}$ is the buy-and-hold abnormal return for stock i from t_1 to t_2 around earnings announcement t. The dependent variables in columns (1)–(2) and (3)–(5) are BHAR[-5,-1] and BHAR[0,1], respectively. $\mathbb{1}^{Att}_{i,t}$ is a dummy variable equal to one if the stock belongs to the top StockTwits attention quintile and zero otherwise. $Controls_{i,t}$ is a vector of control variables and corresponds to analyst sentiment from analyst recommendations, the average newswire sentiment over five days before earnings announcements, the log change in the mean newswire coverage from days [-40,-6] to [-5,-1], and the earnings surprise on date t. α_i and α_t are stock and time-fixed effects, respectively. In column (5), we replace $\mathbb{1}^{Att}_{i,t}$ for $Retail\ OI^{fit}$ and $Retail\ OI^{resid}$, which corresponds to the fitted and residual components from regressing retail order imbalance using trades five days before announcements onto $\mathbb{1}^{Att}_{i,t}$. The sample of stock-earnings observations is comprised of high-attention StockTwits stocks and their corresponding matched low-attention stocks. The sample period is from January 2013 to December 2022. *$p < .1$; **$p < .05$; ***$p < .01$.

	Dependent variable:				
	BHAR[-5,-1]		BHAR[0,1]		
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}^{Att}$	0.788***	1.080***	-0.694***	-0.532***	
	(0.127)	(0.183)	(0.130)	(0.202)	
Surp		0.061		0.725***	0.740***
		(0.041)		(0.058)	(0.063)
Ln(mcap)		0.333**		1.120***	1.126***
		(0.139)		(0.162)	(0.173)
Num analysts		-0.043***		-0.026	-0.030
		(0.011)		(0.019)	(0.019)
News sent		9.914***		-1.030	-1.598**
		(0.847)		(0.730)	(0.747)
News cov		6.174		-58.357*	10.000
		(24.548)		(30.808)	(31.761)
$BHAR_{[-5,-1]}$				-0.031**	-0.030**
				(0.012)	(0.013)
Retail OI^{fit}					-0.967**
					(0.412)
Retail OI^{resid}					0.016**
					(0.007)
Intercept	-0.214**		0.246**		
	(0.087)		(0.097)		
N	38,744	38,744	38,744	38,744	34,232
R^2	0.002	0.015	0.001	0.032	0.036
Firm and date FE	N	Y	N	Y	Y

Table 7
Testing the impact of inventory risks on reversals

This table reports the abnormal announcement date returns from independently sorting high and matched low StockTwits attention stocks (rows) and in quintiles for a corresponding proxy for inventory risk (columns). The proxies are market capitalization, the average absolute announcement returns from the previous three quarters, and the 3-day average implied volatility before announcements in Panels A to C, respectively. The t-statistics are in brackets, and bold t-statistics indicate statistical significance at the 10% level. The sample period is from January 2013 to December 2022.

	Sort on the inventory risk proxy						
	Low	Q2	Q3	Q4	High	High-Low	t-stat
A. Market capitalization							
High att.	-2.487	-0.267	0.133	0.288	0.260	2.747	**[11.92]**
Low att.	-0.556	0.860	0.533	0.374	0.301	0.858	**[4.66]**
High-Low	-1.930	-1.127	-0.400	-0.085	-0.041	1.890	**[6.41]**
t-stat	**[-6.91]**	**[-3.35]**	[-1.35]	[-0.39]	[-0.43]		
B. Absolute announcement returns							
High att.	0.076	-0.361	-0.174	-0.460	-1.062	-1.137	**[-4.74]**
Low att.	0.201	0.281	-0.019	0.224	0.059	-0.142	[-0.56]
High-Low	-0.125	-0.643	-0.155	-0.684	-1.121	-0.995	**[-2.84]**
t-stat	[-0.91]	**[-3.87]**	[-0.75]	**[-2.71]**	**[-3.49]**		
C. Implied volatility							
High att.	0.047	-0.051	0.291	-0.379	-0.718	-0.765	**[-3.48]**
Low att.	0.281	0.154	0.053	0.500	0.046	-0.235	[-0.96]
High-Low	-0.234	-0.205	0.239	-0.878	-0.764	-0.530	[-1.61]
t-stat	**[-2.57]**	[-1.24]	[1.00]	**[-2.84]**	**[-2.42]**		

Table 8
StockTwits Attention and Stock Return Response to Earnings Surprises

This table reports estimates of the following regression:

$$AR_{i,t} = \beta_1 Surp_{i,t} + \beta_2 \mathbb{1}^{Att}_{i,t} + \beta_3 Surp_{i,t} \times \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

where AR corresponds to the abnormal return on the announcement date in columns (1)–(3), buy-and-hold abnormal return (BHAR) from the announcement date to the next trading day in column (4), and two to five days after the announcement in column (5). The control variables are the average analyst recommendation sentiment, news sentiment, and abnormal news coverage computed five days before the earnings announcement. α_i and α_t correspond to the firm and earnings-date fixed effects.

	Dependent variable:				
		AR[0]		BHAR[0,1]	BHAR[2,5]
	(1)	(2)	(3)	(4)	(5)
Surp	0.827***	0.827***	0.849***	0.929***	0.068***
	(0.035)	(0.035)	(0.038)	(0.046)	(0.025)
$\mathbb{1}^{Att}$	-0.004***	-0.004***	-0.003***	-0.004***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Surp $\times \mathbb{1}^{Att}$	-0.175***	-0.175***	-0.205***	-0.232***	-0.035
	(0.051)	(0.051)	(0.055)	(0.064)	(0.055)
Analysts sent		0.001	0.002	0.002	0.001
		(0.001)	(0.002)	(0.002)	(0.001)
News sent		-0.004	-0.009*	-0.012**	-0.000
		(0.004)	(0.005)	(0.005)	(0.004)
News cov		0.070	-0.660***	-0.623***	-0.232
		(0.110)	(0.202)	(0.210)	(0.192)
N	101,393	101,393	101,393	101,393	101,367
R^2	0.031	0.031	0.029	0.027	0.000
Firm and date FE	N	N	Y	Y	Y

Table 9
The Influence of Earnings Whispers

This table reports estimates of the following regression:

$$y_{i,t}^{post} = \beta_1 \mathbb{1}_{i,t}^{Ewhispers} + \Gamma' Controls_{i,t}^{prior} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

$y_{i,t}$ corresponds to the buy-and-hold abnormal return ($BHAR^{post}$, in percent), StockTwits attention (Att^{post}, in percent) defined in equation (1), sentiment defined in equation (2), the log number of retail trades (Log retail trd^{post}), and retail trade order imbalance ($Retail\ OI^{post}$) in columns (1)–(5), respectively, for stock i. The dependent variables are computed from Monday following the Earnings Whispers post to one day before earnings announcements on date t. The control variables are the buy-and-hold abnormal returns ($BHAR^{prior}$), StockTwits attention (Att^{prior}), sentiment (Sent^{prior}), log retail trades (Log retail trd^{prior}) and retail order imbalance ($Retail\ OI^{prior}$) over the five trading days prior to the earnings whispers posts. The additional control variables not reported in the table are the upcoming earnings surprises, the absolute earnings surprise, abnormal newswire coverage and average newswire sentiment. α_i and α_t are stock and time-fixed effects, respectively. The sample period is from January 2016 to December 2022. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	Dependent variable:				
	BHAR^{post}	Att^{post}	Sent^{post}	Log Retail trd^{post}	Retail OI^{post}
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}^{Ewhispers}$	0.511***	0.027***	0.023***	0.068***	1.082**
	(0.176)	(0.005)	(0.007)	(0.010)	(0.435)
BHAR^{prior}	0.242***	-0.000	-0.000	-0.000	-0.082***
	(0.044)	(0.000)	(0.000)	(0.000)	(0.014)
Att^{prior}	-0.446	0.976***	0.004	0.056**	1.411**
	(0.818)	(0.085)	(0.006)	(0.026)	(0.690)
Sent^{prior}	0.216	0.004***	0.182***	0.021***	1.261***
	(0.139)	(0.002)	(0.006)	(0.007)	(0.305)
Log retail trd^{prior}	-0.179*	0.000	0.023***	0.769***	0.165
	(0.099)	(0.003)	(0.002)	(0.006)	(0.175)
Retail OI^{prior}	0.011***	0.000**	0.000**	0.001***	0.236***
	(0.002)	(0.000)	(0.000)	(0.000)	(0.009)
N	60,910	60,910	60,910	60,910	60,882
R^2	0.094	0.622	0.042	0.575	0.038
Firm and date FE	Y	Y	Y	Y	Y

Table 10
Long-Short Portfolio Alpha with Earnings Whispers

This table presents monthly mean returns and monthly factor alphas for the long (short) portfolio that buys (sells) stock i if the corresponding stock's earnings announcement appears (does not appear) in the Earnings Whispers post. The portfolios are value-weighted and rebalanced weekly. Once we obtain the daily portfolio returns for the long and short sides, we accumulate the daily returns at the monthly frequency. The "Long-Short" portfolio buys the long portfolio and sells the short portfolio. The table reports the strategy *pre*-announcement and *post*-announcement. The pre-announcement consists of the period from the publication of the Earnings Whispers post to one day before the announcement. The post-announcement period consists of the announcement date and the following trading day. The alphas represent the intercepts from time series regressions of the portfolio excess returns on factor alphas. The five factors include the aggregate market excess return, the size factor, the value factor, the investor factor, and the profitability factor. Mom corresponds to the momentum factor. Standard errors adjust for heteroskedasticity and autocorrelation. Returns are in percent. We report the t-statistics in brackets. The sample period is from January 2016 to December 2022.

	Pre-announcement			Post-announcement		
	Short	Long	Long-Short	Short	Long	Long-Short
Average return	0.2566	0.6024	0.3458	0.0089	-0.3904	-0.3993
	[1.37]	**[2.82]**	**[1.80]**	[0.07]	**[-2.22]**	**[-1.65]**
Standard deviation	1.634	1.682	1.574	1.289	1.648	2.118
CAPM alpha	0.0589	0.5115	0.4526	-0.0959	-0.4314	-0.3355
	[0.42]	**[2.29]**	**[2.37]**	[-0.69]	**[-2.20]**	[-1.28]
FF3 alpha	0.0950	0.5390	0.4440	-0.0918	-0.4118	-0.3200
	[0.76]	**[2.92]**	**[2.61]**	[-0.70]	**[-2.17]**	[-1.35]
FF5 alpha	0.0994	0.5429	0.4436	-0.1302	-0.3799	-0.2496
	[0.83]	**[3.22]**	**[2.91]**	[-1.08]	**[-1.86]**	[-1.02]
FF5+mom alpha	0.1119	0.5462	0.4342	-0.1297	-0.3537	-0.2239
	[0.87]	**[3.21]**	**[2.81]**	[-1.09]	**[-1.70]**	[-0.92]

Internet Appendix to

Social Media-Driven Noise Trading: Liquidity Provision and Price Revelation Ahead of Earnings Announcements

Intended for online publication.

IA. Model derivation

The wishful thinking investor will choose subjective beliefs π_H and π_L by maximizing expected utility in (6) taking into account that $\pi_H + \pi_L = 1$. The Lagrangian of the investor is given by

$$\mathcal{L} = q(\pi_H v_H + \pi_L v_L - p) - \frac{1}{\theta}\pi_H \ln \frac{\pi_H}{\bar{\pi}_H} - \frac{1}{\theta}\pi_L \ln \frac{\pi_L}{\bar{\pi}_L} - \mu(\pi_H + \pi_L - 1)$$

where μ is a Lagrange multiplier. The first order condition with respect to π_H is given by

$$q v_H - \frac{1}{\theta} \ln \frac{\pi_H}{\bar{\pi}_H} - \frac{1}{\theta} - \mu = 0.$$

A similar first order condition can be found for π_L. The first order conditions can be rearranged to yield

$$\pi_H = \bar{\pi}_H \exp\left(\theta q v_H - \theta\mu - 1\right) \qquad \text{and} \qquad \pi_L = \bar{\pi}_L \exp\left(\theta q v_L - \theta\mu - 1\right). \tag{8}$$

Plugging (8) into $\pi_H + \pi_L = 1$, we obtain

$$\exp\left(\theta\mu + 1\right) = \bar{\pi}_H \exp\left(\theta q v_H\right) + \bar{\pi}_L \exp\left(\theta q v_L\right)$$

If we plug this expression back into (8), we get (7).

IA. Tables and Figures

Table IA1
StockTwits Change in Sentiment Does Not Predict Fundamentals

This table reports estimates for the full sample, large caps (the top three NYSE market capitalization quintiles), and small caps (bottom two quintiles) of the following regression:

$$Surp_{i,t} = \beta_1 \Delta Sent_{i,t} + \beta_2 Ln(N\ post)_{i,t} \times \Delta Sent_{i,t} + \beta_3 Ln(N\ post)_{i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

where $Surp_{i,t}$ is the earnings surprise (%) for stock i on earnings announcement date t. $\Delta Sent$ is the change in sentiment on StockTwits five days before earnings announcements minus sentiment from 20 to six days before announcements. $Controls_{i,t}$ is a vector of control variables and corresponds to the buy-and-hold abnormal returns, analyst sentiment from analyst recommendations and average newswire sentiment over five days before earnings announcements. α_i and α_t are stock and time-fixed effects, respectively. The regression includes non-tagged sentiment posts with an assigned neutral sentiment score of 0.5. The sample period is from January 2013 to December 2022. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	Dependent variable: Surprise (%)					
	Excl. non-tagged posts			Incl. non-tagged posts as neutral		
	Full sample (1)	Large caps (2)	Small caps (3)	Full sample (4)	Large caps (5)	Small caps (6)
ΔSent	0.022	-0.014	0.061	0.027	0.003	0.059
	(0.030)	(0.013)	(0.070)	(0.023)	(0.010)	(0.045)
$\mathbb{1}^{Att} \times$ ΔSent	-0.016	-0.017	0.098	-0.057	-0.028	-0.062
	(0.061)	(0.028)	(0.263)	(0.047)	(0.026)	(0.176)
$\mathbb{1}^{Att}$	-0.010	-0.024	0.045	-0.037	-0.019	-0.045
	(0.037)	(0.015)	(0.085)	(0.032)	(0.013)	(0.069)
$\text{BHAR}_{[-5,-1]}$	0.620**	0.404**	0.605*	0.498**	0.376**	0.542**
	(0.250)	(0.181)	(0.333)	(0.199)	(0.147)	(0.252)
Analysts sent	0.020	0.030**	0.036	0.017	0.024*	0.014
	(0.030)	(0.015)	(0.145)	(0.024)	(0.013)	(0.087)
News sent	0.052	0.019	0.210	0.096	-0.002	0.365
	(0.104)	(0.072)	(0.478)	(0.080)	(0.055)	(0.280)
N	55,297	28,771	26,526	101,393	44,598	56,795
R^2	0.001	0.001	0.001	0.000	0.001	0.000
Date FE	Y	Y	Y	Y	Y	Y

Table IA2
StockTwits User Types' Sentiment Does Not Predict Fundamentals

This table reports estimates of the following regression:

$$Surp_{i,t} = \beta_1 Sent_{i,t} + \beta_2 \mathbb{1}^{Att}_{i,t} \times Sent_{i,t} + \beta_3 \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

where $Surp_{i,t}$ is the earnings surprise (%) for stock i on earnings announcement date t. $Sent$ is the sentiment on StockTwits as defined in equation (2) over five days before earnings announcements and $Ln(N\ posts)_{i,t}$ is the natural logarithm of the number of StockTwits posts for self-labeled *Novice*, *Intermediate*, and *Professional* users in columns (1)–(2), (3)–(4), and (5)–(6), respectively. $Controls_{i,t}$ is a vector of control variables and corresponds to the buy-and-hold abnormal returns, analyst sentiment from analyst recommendations and average newswire sentiment over five days before earnings announcements. α_i and α_t are stock and time-fixed effects, respectively. The regression includes non-tagged sentiment posts with an assigned neutral sentiment score of 0.5. The sample period is from January 2013 to December 2022. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	User type:					
	Novice		Intermediate		Professional	
	(1)	(2)	(3)	(4)	(5)	(6)
Sent	0.052	0.059	-0.020	-0.016	-0.016	-0.008
	(0.112)	(0.114)	(0.049)	(0.049)	(0.038)	(0.039)
Ln(N posts)		0.001		0.000		0.000
		(0.002)		(0.000)		(0.000)
Sent × Ln(N posts)		-0.001		-0.000		-0.001
		(0.002)		(0.000)		(0.001)
$BHAR_{[-5,-1]}$	0.542	0.553	0.656**	0.684**	0.703**	0.730**
	(0.469)	(0.472)	(0.312)	(0.315)	(0.304)	(0.308)
Analysts sent	-0.042	-0.044	0.029	0.025	-0.000	-0.003
	(0.075)	(0.075)	(0.041)	(0.042)	(0.037)	(0.037)
News sent	0.007	0.021	-0.035	-0.015	0.046	0.064
	(0.297)	(0.297)	(0.152)	(0.153)	(0.136)	(0.138)
N	14,620	14,620	31,894	31,894	35,560	35,560
R^2	0.000	0.001	0.001	0.001	0.001	0.001
Date FE	Y	Y	Y	Y	Y	Y

Table IA3
Social-Media Induced Retail Order Imbalance Does Not Predict Fundamentals

This table reports estimates of the following regression:

$$Surp_{i,t} = \beta_1 Retail\ OI_{i,t}^{fit} + \beta_2 Retail\ OI_{i,t}^{resid} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

where $Surp_{i,t}$ is the earnings surprise (%) for stock i on earnings announcement date t. $Retail\ OI_{i,t}^{fit}$ and $Retail\ OI_{i,t}^{resid}$ correspond to the fitted and residual components from regressing retail order imbalance on $\mathbb{1}^{Att}$. $Controls_{i,t}$ is a vector of control variables and corresponds to the buy -and-hold abnormal returns, analyst sentiment from analyst recommendations and average newswire sentiment over five days before earnings announcements. α_i and α_t are stock and time-fixed effects, respectively. The results are reported for order imbalances computed using trades, volume, and dollar volume in columns (1)–(2), (3)–(4), and (5)–(6), respectively. The sample period is from January 2013 to December 2022. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Order imbalance:	Trade		Volume		Dollar volume	
	(1)	(2)	(3)	(4)	(5)	(6)
Retail OI^{fit}	-0.110	-0.535	-0.038	-0.703	-0.034	-0.693
	(0.953)	(0.968)	(1.404)	(1.415)	(1.393)	(1.403)
Retail OI^{resid}	0.136***	0.123**	0.049	0.041	0.049	0.041
	(0.049)	(0.049)	(0.032)	(0.032)	(0.032)	(0.032)
$\text{BHAR}_{[-5,-1]}$		0.534**		0.543**		0.543**
		(0.216)		(0.215)		(0.215)
Analysts sent		0.013		0.014		0.014
		(0.025)		(0.025)		(0.025)
News sent		0.163*		0.159*		0.159*
		(0.088)		(0.088)		(0.088)
N	91,199	91,199	91,199	91,199	91,199	91,199
R^2	0.000	0.001	0.000	0.000	0.000	0.000
Date FE	Y	Y	Y	Y	Y	Y

Table IA4
Pre-earnings Announcement Returns and StockTwits Activity for the Full Sample

This table reports estimates of the following regression:

$$BHAR[t_1,t_2]_{i,t} = \beta_1 \mathbb{1}^{Att}_{i,t} + \Gamma' Controls_{i,t} + \alpha_t + \epsilon_{i,t},$$

where $BHAR[t_1,t_2]_{i,t}$ is the buy-and-hold abnormal return for stock i from t_1 to t_2 around earnings announcement t. The dependent variables in columns (1)–(2) and (3)–(5) are BHAR[-5,-1] and BHAR[0,1], respectively. $\mathbb{1}^{Att}_{i,t}$ is a dummy variable equal to one if the stock belongs to the top attention quintile and zero otherwise. $Controls_{i,t}$ is a vector of control variables and corresponds to analyst sentiment from analyst recommendations, the average newswire sentiment over five days before earnings announcements, the log change in the mean newswire coverage from days [-40,-6] to [-5,-1], and the earnings surprise on date t. In column (5), we replace $\mathbb{1}^{Att}_{i,t}$ for $Retail\ OI^{fit}$ and $Retail\ OI^{resid}$, which corresponds to the fitted and residual components from regressing retail order imbalance using trades five days before announcements onto $\mathbb{1}^{Att}_{i,t}$. α_i and α_t are stock and time-fixed effects, respectively. The sample period is from January 2013 to December 2022. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

	Dependent variable:				
	BHAR[-5,-1]		BHAR[0,1]		
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}^{Att}$	0.712***	1.201***	-0.443***	-0.330**	
	(0.113)	(0.144)	(0.103)	(0.129)	
Surp		0.056**		0.861***	0.876***
		(0.024)		(0.039)	(0.042)
Ln(mcap)		0.362***		1.204***	1.236***
		(0.090)		(0.103)	(0.106)
Num analysts		-0.043***		-0.067***	-0.066***
		(0.008)		(0.012)	(0.012)
News sent		9.253***		-1.012**	-1.325**
		(0.510)		(0.488)	(0.521)
News cov		43.483*		-68.067***	-16.384
		(23.555)		(21.599)	(26.900)
$\text{BHAR}_{[-5,-1]}$				-0.053***	-0.052***
				(0.009)	(0.009)
Retail OI^{fit}					-0.969**
					(0.426)
Retail OI^{resid}					0.003
					(0.002)
Intercept	-0.138**		-0.005		
	(0.066)		(0.054)		
N	101,393	101,393	101,393	101,393	91,198
R^2	0.002	0.013	0.000	0.031	0.033
Firm and date FE	N	Y	N	Y	Y

Table IA5
Summary Statistics: Alternative Platforms

This table reports the number of observations with at least one post, the number of posts, and their corresponding sample proportion for StockTwits, WallStreet Bets, and Seeking Alpha, five to one day before earnings announcements. The sample period is from January 2018 to December 2021.

	StockTwits				WallStreetBets				Seeking Alpha			
	Stock-EA obs		Posts		Stock-EA obs		Posts		Stock-EA obs		Posts	
	N	%	N	%	N	%	N	%	N	%	N	%
1 (small)	12,649	32.8	1,342,286	24.1	809	14.6	1,873	5.7	331	13.2	359	10.3
2	7,939	20.6	794,874	14.3	723	13.0	3,034	9.2	286	11.4	316	9.1
3	6,630	17.2	771,888	13.9	818	14.8	4,303	13.0	259	10.3	290	8.4
4	5,651	14.6	700,676	12.6	970	17.5	5,837	17.7	405	16.1	467	13.4
5 (large)	5,746	14.9	1,960,798	35.2	2,223	40.1	17,949	54.4	1,234	49.1	2,041	58.8
Total	38,615	100.0	5,570,522	100.0	5,543	100.0	32,996	100.0	2,515	100.0	3,473	100.0

Figure IA1. Cumulative Returns of Long and Short Portfolios

This figure presents the cumulative excess returns for the long and short portfolios. The long (short) portfolio buys (sells) stock i if the corresponding stock's earnings announcement appears (does not appears) in the Earnings Whispers post. We value-weight weights the stocks to create the portfolios. The portfolios are rebalanced weekly. Once we obtain the daily portfolio returns for the long and short sides, we accumulate the daily returns at the monthly frequency. The "Long-Short" portfolio buys the long portfolio and sells the short portfolio. The sample period is from January 2016 to December 2022.