The Role of Social Media-Driven Daily Disagreement in Market Dynamics and Trading Behaviour

Farzaneh Ghandehari
Associate Professor Lina El-Jahel
Dr Dulani Jayasuriya
Associate Professor Helen Lu

Abstract

Disagreement among investors is a fundamental aspect of financial markets, influencing market dynamics and trading behaviour. Traditionally, measuring disagreement has been challenging, often relying on proxies like analyst forecast dispersion, which suffer from biases and infrequent updates. Recent movements in social media indicate that investors actively seek financial advice online and can shape stock market movements. The evolution of the investing landscape, particularly the rise of social media as a leading source of financial advice, offers an alternative for measuring investor sentiment and disagreement. Platforms like Reddit provide rich, anonymous, community-driven discussions that reflect genuine investor opinions. This research examines how social media empowers investors and explores the potential of leveraging textual analysis of social media content to capture daily fluctuations in investor disagreement. Specifically, this study investigates the link between daily disagreement, abnormal trading volume, and daily short-selling activities among institutional and retail investors. Using data from WallStreetBets (WSB) on Reddit from 2020 to 2023, the study analyses 2,784 firms with sufficient social media activity to construct stock-day-level disagreement proxies. The findings show that social media-driven disagreement measures significantly correlate with increased trading volume and institutional and retail short-sale turnover, supporting traditional theories that disagreement induces trading activity. While both retail and institutional short sellers react to disagreement proxies, institutional investors respond more strongly. Their heightened short-sale turnover suggests that institutional investors may interpret social media-driven disagreement as a signal for contrarian trading strategies and an arbitrage opportunity.

Keywords: Social Media, Sentiment Analysis, Investor Disagreement, Short Sale, Behavioural Finance, Reddit, FINRA

I. Introduction

> *"Investing...is a social activity. Investors spend a substantial part of their leisure time discussing investments, reading about investments, or gossiping about others' successes or failures in investing".* - Shiller, R.J. 1989. (Eliner & Kobilov, 2022).

Disagreement is inherent in financial markets and significantly impacts trading behaviour. Prior literature identifies investor disagreement as a key driver of market outcomes such as trading volume, asset mispricing, and short-sale constraints (Miller, 1977; Hong & Stein, 2007). Hong and Stein (2007) argue that properly measuring disagreement can explain much of the puzzle about trading volume. While extensive theoretical literature explores its implications (Miller, 1977; Varian, 1985; Karpoff, 1987; Scheinkman & Xiong, 2003; Banerjee & Kremer, 2010; Banerjee et al., 2018, 2024), empirical measurement remains challenging. Traditional proxies, such as analyst forecast dispersion, suffer from biases and slow responsiveness and are subject to strategic distortion, thus failing to reflect genuine differences in beliefs (Kandel & Pearson, 1995; Diether et al., 2002; Garfinkel & Sokobin, 2006). Additionally, while disagreement is believed to contribute to the high levels of daily trading volume (Hong & Stein, 2007), these measures, typically available monthly or quarterly, fail to capture daily fluctuations, hindering a deep understanding of the link between beliefs and trading (Giglio et al., 2021; Charles et al., 2023).

Regarding technological advancements, the investing landscape has evolved significantly, and social media has become a key source for investors to seek and exchange financial advice. A Forbes Advisor survey[1] in March 2023 found that 79% of millennials and Gen Z seek financial advice on platforms like Reddit and YouTube. This shift underscores the growing influence of social media on investment (Barber et al., 2009; Kaniel et al., 2012; Kelley & Tetlock, 2013, 2017; Boehmer et al., 2021; Baig et al., 2023). Considering recent technological developments and leveraging textual analysis techniques on the wealth of data generated by users in social media platforms for sentiment analysis enables to gauge sentiment and disagreement in real-time (Antweiler & Frank, 2004; Sprenger et al., 2014; Li et al., 2018; Giannini et al., 2018, 2019; Fan et al., 2020; Cookson & Niessner, 2020, 2023).

Prior studies commonly use labelled data from Yahoo! Finance, Raging Bull, and StockTwits (Antweiler & Frank, 2004; Cookson & Niessner, 2020, 2023), yet these sources may skew sentiment positively due to users' incentives to attract followers and gain Internet fame (Cookson & Niessner, 2020). Unlike Twitter and Facebook, Reddit's structure emphasises anonymous, community-driven discussions and fosters discussions focusing on content rather than individual identities. This anonymity encourages authentic expression without fear of repercussions, making Reddit a more reliable source for gauging investor sentiment (Srinivasan, 2023), and this social media can be counted as the most realistic sound of investors. The WallStreetBets (WSB) subreddit, known for its impact on financial markets, gained mainstream attention in 2020 after being featured on the cover page of Bloomberg's Businessweek and reaching a pinnacle during the GameStop saga of 2021, highlighting

1. https://www.forbes.com/advisor/investing/financial-advisor/adults-financial-advice-social-media/

investors' collective power (Hu et al., 2021; Bradley et al., 2021, 2023). However, research on daily investor disagreement within such platforms remains limited.

Prior studies link investor disagreement to trading volume. Theoretically, Harris and Raviv (1993) and Karpoff (1986) argue that differing investor views drive trading activity (Fan et al., 2020). Empirically, Ajinkya et al. (1991) provide empirical support, showing a positive relationship between trading volume and diversity of investor beliefs. Antweiler and Frank (2004) confirm this relationship using online stock forum data. More recent studies find similar patterns in social media: Giannini et al. (2019) show that disagreement on Twitter around earnings announcements increases trading volume, while Cookson and Niessner (2020, 2023) document a strong link on StockTwits. However, the 'no-trade theorem' (Milgrom & Stokey, 1982) posits that rational agents with common knowledge and Bayesian updating should not trade solely based on differences in beliefs, challenging the notion that disagreement alone induces volume. These conflicting views underscore the theoretical ambiguity surrounding the relationship between disagreement and trading activity, highlighting the importance of empirical testing to identify which mechanisms dominate in practice.

Investor disagreement plays a key role in short-selling activity (Miller, 1977). Short sellers—who account for over 20% of trading volume—are typically informed traders who exploit mispricing and contribute to price discovery (Boehmer et al., 2008; Engelberg et al., 2012; Christophe et al., 2010; Karpoff & Lou, 2010). Given their role in incorporating information into prices, understanding how disagreement influences short-selling behaviour is crucial for assessing market efficiency.

This research explores the connection between daily disagreement on WSB, abnormal trading volume, and daily short-selling activities of institutional and retail investors. Specifically, it is considered that higher levels of daily disagreement will correlate with increased trading volume and short sellers' activities. This study examines 2,784 firms with enough social media activity in WSB to define stock-day level disagreement proxies from 2020 to 2023 through panel regression. Off-exchange short sales data from the FINRA[2] and BJZZ algorithm allows differentiation between retail and institutional short selling (Boehmer et al., 2021). Adjusted VADER sentiment analysis (Hutto & Gilbert, 2014) by Long et al. (2023) for WSB is considered to classify message sentiment, which underpins the disagreement measures of this study. To the best of our knowledge, this study is the first to analyse daily sentiment-driven disagreement on social media alongside retail and institutional short-selling behaviour using high-frequency FINRA TRF data. This study makes multiple contributions to the literature on investor disagreement and market efficiency.

Findings indicate that all disagreement measures are significantly associated with increased abnormal trading volume and short-sale turnover, supporting theoretical models linking belief dispersion to trading behaviour (Miller, 1977; Hong & Stein, 2007). Institutional investors react more strongly, and this suggests that social media-driven disagreement influences institutional trading strategies, potentially serving as an arbitrage signal. Notably, disagreement encompassing both submissions and comments has the strongest effect,

2. FINRA (Financial Industry Regulatory Authority) publicly provides short-sale trade data for off-exchange (OTC) transactions involving exchange-listed securities that are reported to a FINRA Trade Reporting Facility (TRF) or the Alternative Display Facility (ADF). Additionally, it shares data on trades in OTC securities reported to FINRA's Over-the-Counter Reporting Facility (ORF). https://www.finra.org/finra-data/browse-catalog/ short-sale-volume-data/daily-short-sale-volume-files.

particularly when employing the direct sentiment scoring method. The results underscore the informational value of social media sentiment rather than noise, contributing to a deeper understanding of how high-frequency investor disagreement, mainly as expressed through social media platforms, plays an increasingly prominent role in price discovery and market efficiency (Antweiler & Frank, 2004; Sprenger et al., 2014; Cookson & Niessner, 2020).

A growing body of empirical work (Antweiler & Frank, 2004; Sprenger et al., 2014; Cookson & Niessner, 2020) demonstrate that social media platforms provide rich, real-time data for understanding investor sentiment and disagreement. However, most of these studies predominantly focus on aggregate sentiment polarity, categorising content as broadly bullish or bearish, rather than capturing the degree of disagreement, the accurate dispersion of beliefs among investors. While the utility of social media for daily sentiment tracking is well established, its potential to capture belief heterogeneity at high granularity remains underutilised, despite its central role in theoretical models of disagreement-driven trading (Hong & Stein, 2007). Addressing this gap, the present study applies the commonly used empirical approach of measuring disagreement via the standard deviation of analyst forecasts (Diether et al., 2002) to a novel setting: direct sentiment scores derived from post-comment interactions on social media.

Behavioural finance research (Shiller, 1984; Barberis et al., 2018) emphasises that social interactions and sentiment dynamics among investors significantly influence trading behaviour. This insight is particularly relevant for measuring disagreement in online settings such as Reddit's WallStreetBets (WSB), where belief heterogeneity emerges from original posts and dynamic, interactive comment threads. By incorporating the posts and their associated comment discussions, this study offers a more granular and nuanced measure of belief heterogeneity that aligns with behavioural and information aggregation theories.

Additionally, while prior research underscores the crucial role of short sellers in price discovery and enhancing market efficiency (Boehmer et al., 2008; Engelberg et al., 2012), limited attention has been paid to how investor disagreement specifically influences short-selling activity, especially across different trader types (i.e., retail vs. institutional). Much of the extant literature treats short sellers as a homogenous group, overlooking how they may respond differently to belief dispersion in social media-driven information environments. In line with market efficiency theory, this study posits that social media is a meaningful source of information rather than noise and sophisticated traders, such as short sellers, who are skilled at processing public information (Boehmer et al., 2008, 2020; Engelberg et al., 2012), are likely to interpret social media sentiment as a signal for contrarian strategies or arbitrage and thus integrate it into their trading decisions.

This study advances the disagreement literature by providing high-frequency firm-daily sentiment proxies and direct measures of disagreement using social media interactions, capturing the full spectrum of investor beliefs, contrasting with prior research that relies on low-frequency, biased and indirect proxies (Diether et al., 2002; Nagel, 2005; Fischer et al., 2022). The research enhances understanding of social media's influence on financial markets by analysing daily disagreement on WSB and its effect on trading activities.

Finally, the study has practical implications for investors, regulators, and algorithmic traders by highlighting the role of online discussions in shaping market behaviour.

Understanding this relationship is crucial for market efficiency, regulatory oversight, and risk management, given the growing impact of social media on financial markets.

The remainder of the paper is organised as follows: Section II is about developing the study's hypotheses; Section III discusses the data; Section IV reviews the definition of variables and methodology; Section V reports on the empirical results; and Section VI concludes.

II. Hypothesis Development

Abnormal Trading Volume

Prior financial literature suggests that investors' heterogeneous beliefs are a key driver of trading activity. Theoretically, Hirshleifer (1977), Karpoff (1986), and Harris and Raviv (1993) argue that trading volume increases when investors hold differing opinions about an asset's value. Empirical evidence supports this view: Ajinkya et al. (1991) find a positive relationship between investor disagreement and trading volume, while Antweiler and Frank (2004) confirm this using online stock forums. Chang et al. (2013) further show that diverse information sources contribute to opinion divergence, leading to higher trading volume.

More recent studies reinforce this connection. Giannini et al. (2019) find disagreement in Twitter discussions around earnings announcements correlates with increased trading, while Cookson and Niessner (2020, 2023) document a similar pattern on StockTwits. However, the no-trade theorem (Milgrom & Stokey, 1982) challenges this view, arguing that rational expectations should prevent trading based solely on belief differences. Despite theoretical debates, empirical evidence largely supports a positive association between disagreement and trading activity. Thus, the following hypothesis is proposed:

Hypothesis 1: The higher level of daily disagreement in r/WallStreetBets is related to higher trading volume.

Short Sale Turnover

Investor disagreement influences market dynamics by driving trading activity. According to the differences-of-opinion theory (Hong & Stein, 2003), investors interpret the same information differently due to variations in beliefs, risk tolerance, and biases, leading to increased trading volume. Miller's (1977) overpricing hypothesis further argues that disagreement amplifies trading, particularly under short-sale constraints, potentially causing overpricing. Chen et al. (2002) suggest that these constraints delay price corrections by limiting the incorporation of negative information.

However, the role of short-sale constraints remains debated. While Diamond and Verrecchia (1987) argue that markets adjust for these restrictions, mitigating their impact on mispricing, conflicting views highlight the complexity of disagreement and short-selling mechanisms. Investor behaviour also varies institutional investors often engage in contrarian short selling to correct mispricing (Nagel, 2005; Boehmer et al., 2008), while retail investors tend to be more sentiment-driven and reactive (De Long et al., 1989; Barber & Odean, 2008).

Social media now plays a key role in shaping investor sentiment and trading decisions. Online platforms provide real-time access to information, influencing market perceptions. Since short sellers rely on sophisticated information processing, they enhance price efficiency by incorporating new data (Tetlock, 2007; Antweiler & Frank, 2004). Building on these insights, this study expects higher investor disagreement on WSB to be positively associated with institutional and retail short-sale turnover. This leads to the following hypotheses:

Hypothesis 2a: The higher level of daily disagreement on r/WallStreetBets is positively associated with all types of short-sale turnover.

Hypothesis 2b: The higher level of daily disagreement on r/WallStreetBets is positively associated with retail short-sale turnover.

Hypothesis 2c: The higher level of daily disagreement on r/WallStreetBets is positively associated with institutional short-sale turnover.

III. Data

WallStreetBets

WSB has gained immense popularity within its ecosystem, amassing over 19 million subscribers since its inception on April 11, 2012. In June 2025, it was ranked as the 41st largest subreddit. From its inception until December 31, 2023, a total of 81,144,555 comments and submissions were made. Since 2020, WSB has come into the spotlight, so the study sample is restricted to messages posted between 2020 and 2023. As a result, the sample retains 89.5% of the messages in WSB, totalling 72,600,100 submissions and comments.

This study focuses only on submissions that mention a single ticker to ensure that the sentiment is directly linked to a specific stock and includes only comments posted on the same date as their corresponding original posts. There were 637,805 submissions, including those referencing one or multiple tickers, with 528,143 (83%) mentioning only one ticker. Additionally, since capturing daily investor opinions on individual firms is essential for constructing a daily disagreement measure, the analysis includes only firms with at least two submissions per day. The final sample consists of 2,784 tickers. The summary statistics for this sample are presented in Table 1.

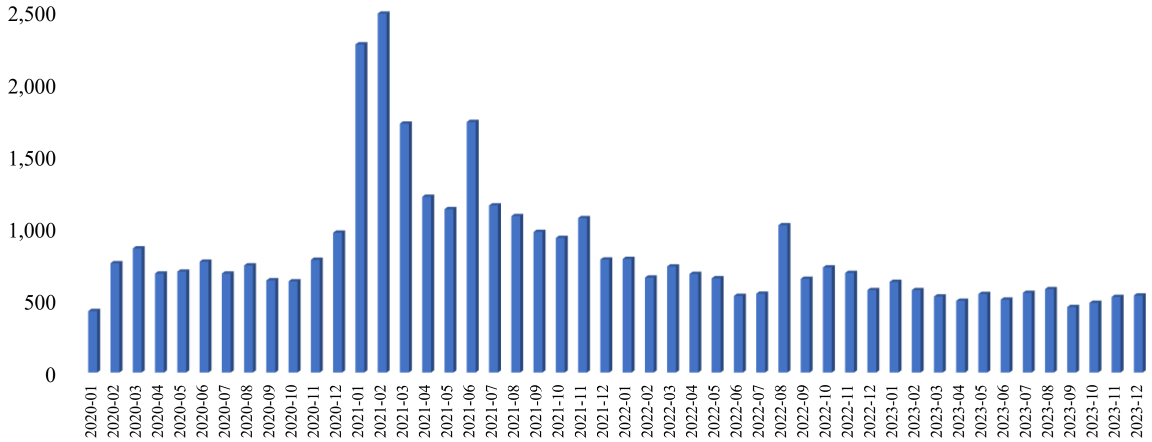
Table 1: The summary statistics of the sample

	2020	**2021**	**2022**	**2023**
Number of Comments	18,194,372	30,115,922	14,235,271	10,357,532
Number of submissions	304,337	1,407,544	244,349	170,146
Number of submissions with only one mentioned ticker	69,547	382,591	49,613	34,309
Number of comments under the submissions with only one mentioned ticker	1,401,054	6,235,848	7,871,571	2,555,948
Number of unique tickers	2,708	4,302	2,892	2,444
Average messages for each ticker	543	1,538	2,739	1,059
Firm day observation	903,485	1,044,838	1,118,547	1,170,248

The data collection workflow involved using the Python Reddit API Wrapper (PRAW) to collect historical posts and comments from the subreddit WSB between January 2020 and December 2023. The default time zone of the Reddit database is UTC, which is converted to Eastern Time to align with the NYSE, NASDAQ, and AMEX time zones.

Figure 1 shows the number of firms covered in a given month, with a slight increase observed around the Gamestop event.

Figure (1): Number of unique ticker mention per month



Each message is timestamped at the moment it is posted (Eastern Time), allowing for an analysis of whether investors share messages while actively updating their beliefs in response to news or during their free time in the evenings after work. The figure below illustrates the distribution of messages by day of the week and time of day.

Investors tend to post messages during market hours and in the evenings after work and dinner (Monday–Friday, between 8 a.m. and 8 p.m.). Database statistics reveal that 98% of investor discussions occur on non-holiday days. This pattern suggests that investors frequently update their messages in real time as financial events unfold.

4o

Figure (2): Disrtibution of Messages Time

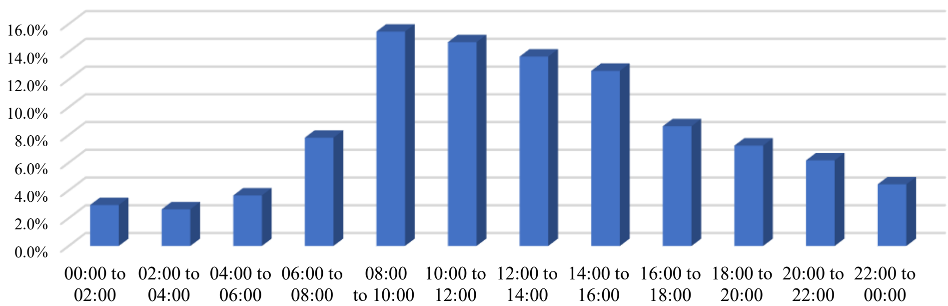
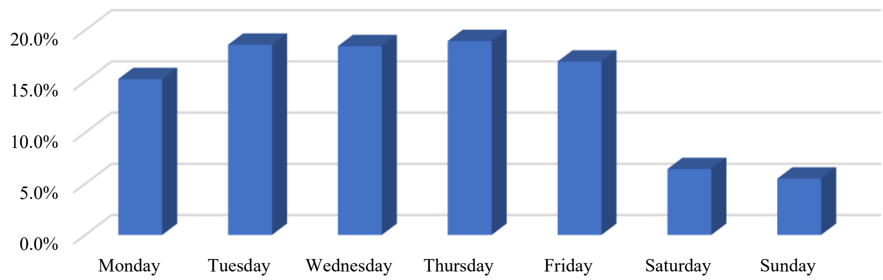


Figure (3): Disrtibution of Messages Week day



WSB does not cover a uniformly distributed range of sectors; rather, it focuses on those that consumers can relate to because they are regularly exposed to the products or services.

Special attention is given to stocks of relatively young companies, often categorised as growth stocks. These stocks are typically riskier due to greater volatility but offer a higher potential for value appreciation.

FINRA

The intraday short-selling data used in this study is obtained from the FINRA TRF transaction level. In response to the financial crisis, the Securities and Exchange Commission (SEC) mandated that self-regulatory organisations (SROs)[3] publish short-sale transactions on their websites to enhance transparency. The SEC specifically requested that FINRA report stock-level aggregate short-sale volumes daily[4]. This includes data from the FINRA Trade Reporting Facility (TRF)[5], that facilitates the reporting of off-exchange (OTC) short-sale transactions. Since September 2009[6], these transactions have been mandated for public reporting. This database is publicly accessible and free of charge. As noted by Hu et al. (2021), short sales reported by FINRA make up approximately 13.6% of the total trading volume.

The dataset contains transaction time (in seconds), share price, and the number of shares for each off-exchange short-sale transaction in exchange-listed stocks. It is a subset of the transactions reported in the daily TAQ (with the exchange code "D"). Additionally, an analysis of order routing disclosures mandated by Rule 606 of Regulation NMS reveals that almost all retail orders are directed to off-exchange market makers[7] and reported to FINRA TRFs (Boehmer & Song, 2020). This study uses the algorithm developed by Boehmer et al. (2021) to identify traders based on sub-penny price improvements, helping infer off-exchange trades made by retail and institutional short sellers. The metric is based on U.S. regulatory constraints and the resulting institutional structures. Unlike institutional order flow, retail order flow can benefit from sub-penny price improvements (Boehmer et al., 2021). For the analysis, the total number of short-sell shares or short-selling trades for a stock on a given trading day is aggregated for both institutional and retail investors. This data is combined with CRSP data (Boehmer & Song, 2020). Since short-sale transactions can occur across multiple markets, the short volume from different markets is combined daily for each stock. If the total short volume is missing, it is replaced with zero (Wang et al., 2020).

IV. Methodology of Measuring Daily Disagreement

Disagreement Proxies

1. Measuring Disagreement based on the Literature (Bullishness/Bearishness Method)

A. Sentiment Classification

To construct measures of disagreement, the methodology proposed by Cookson and Niessner (2020, 2023) is adopted, which involves discretising continuous sentiment scores obtained from the adjusted VADER tool developed by Long et al. (2023) specifically for

3. SROs include registered exchanges (i.e., NYSE, NASDAQ, or Bats), and FIRNA.
4. FINRA usually releases the data on the same day as the trading (after regular trading hours) and, in rare cases, on the following business day.
5. TRFs are managed by a registered national exchange, such as NASDAQ or NYSE, and are all under the oversight of FINRA.
6. FINRA's website provides two types of data: the "daily file," which contains the total daily short volume for each firm, and the "monthly file," which offers detailed, transaction-by-transaction information on short sale trades reported to the consolidated tape. To examine the transaction price of each off-exchange short sale, we use the monthly file.
7. See "SEC 2010 Concept Release on Equity Market Structure".

WallStreetBets (WSB) data analysis (Appendix A). The average sentiment of comments is calculated using their score (upvotes minus downvotes) as a weight. Similarly, the average sentiment of each original post and its related comments (on the same date) is also weighted by their respective scores.

Sentiment scores are categorised as follows: a score greater than 0.5 is labelled as +1 (Bullish), indicating a positive sentiment, while a score less than -0.5 is labelled as -1 (Bearish), indicating a negative sentiment. Observations with sentiment scores between -0.5 and 0.5 (neutral) are excluded from the analysis. This systematic approach enables the quantification of disagreement within the context of sentiment analysis. Neutral messages are not used in the analysis conducted by Antweiler and Frank (2004) in their study on Internet stock message boards. They mentioned that the "neutral" messages contain both "noise" and neutral opinions, with the noise dominating. Therefore, including "neutral" messages would lead to potentially noisier and distorted bullishness signals.

Table (3) presents the distribution of sentiment score and the number of bullishness\bearishness messages in the sample. According to these summary statistics,75.7% of classified messages are bullish, and 24.3% are bearish.

Table 2: Distribution of sentiment score and the number of bullishness\bearishness messages

	Submissions	**Submissions and Comments**
Number of Bullish	46,293	11,961
Number of Bearish	14,455	3,929
Average of Sentiment Score	0.135	0.0992
Standard Deviation of Sentiment Score	0.3732	0.2234

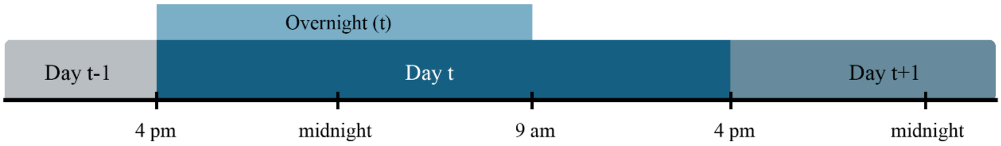
B. Average Sentiment Measure

In line with the methodology proposed by Antweiler and Frank (2004) for constructing a sentiment measure from bearish and bullish message data, the VADER compound score is first utilised to categorise each message. Messages are segmented into three categories: Bearish (compound score between -1 to -0.5), Bullish (compound score between 1 to 0.5), and Neutral (compound score between 0.5 to -0.5). Subsequently, each bearish message is assigned a value of -1, while each bullish message is assigned a value of 1. The average of these classifications is then calculated for each firm × day:

$$AvgSentiment_{it} = \frac{N_{it}^{bullish} - N_{it}^{bearish}}{N_{it}^{bullish} + N_{it}^{bearish}}$$

The $AvgSentiment_{it}$ measure ranges from -1 (all bearish) to +1 (all bullish). This measure is weighted by the number of comments associated with each submission. To establish the base measure, the average sentiment measure is computed for day t by aggregating messages posted between the market close of the preceding day *(t - 1)* at 4:00 PM and the market close of day t at 4:00 PM. Alternatively, overnight messages are from 4:00 PM of the day *(t-1)* to 9:00 AM day t. The figure (4) presents a timeline that illustrates this measurement. This average is measured in two distinct ways, once solely among submissions and another time considering submissions along with their related comments. The average sentiment measure is determined in the primary analysis by assigning equal weight to each message.

Figure 4: Timeline for calculating disagreement



This figure presents a timeline illustrating how disagreement is computed. Messages posted on day *t - 1* after 4 PM up to trading day *t* are assigned, as trading stops at 4 PM on day *t - 1*. Similarly, messages posted after 4 PM on day *t* up to day *t + 1* are assigned. To calculate "overnight" changes in disagreement before the market opens on day *t*, messages posted after 4 PM on day *t - 1* and before 9 AM on day *t* are included.

C. Measuring Disagreement

The overall disagreement measure is constructed by computing the standard deviation of expressed sentiment across daily messages related to a ticker, following Antweiler and Frank (2004) and Cookson and Niesner (2020, 2023). Because the underlying variable is binary (-1/1), the variance of the sentiment measure for period t equals $(1 - AvgSentiment_{it}^2)$. Disagreement measure for a given firm × day is computed as:

$$Disagreement_{it} = \sqrt{1 - AvgSentiment_{it}^2}$$

The $AvgSentiment_{it}$ measure ranges from −1 (all bearish) to +1 (all bullish), while the disagreement measure ranges from 0 to 1, with 1 signifying maximal disagreement. The formula is applied to firm-day observations that have non-zero messages. When there are no messages for a firm day, it is impossible to compute the standard deviation of sentiment across messages. For this corner case, it is assumed that non-posting means traders do not wish to buy or sell in the near term. Accordingly, disagreement is normalised in the no-message case to 0, consistent with latent agreement, following the definition in Cookson and Niessner (2020). This choice regarding normalising the no-message case (latent agreement) is consistent with the idea that minimal disagreement should correspond to minimal trading.

2. Alternative Measure of Disagreement (Sentiment Score Method)

Researchers often use labelled datasets from platforms like Yahoo! Finance, Raging Bull, and StockTwits to measure investor disagreement, where users classify their sentiment as bullish, bearish, or unclassified. However, as Cookson and Niessner (2020) note, investors on these platforms may be motivated by personal gain, influencing the authenticity of their sentiment. Their findings show a strong bullish bias, with 81.7% of messages classified as bullish, suggesting a tendency to broadcast optimism rather than genuine beliefs; this portion for WSB is 75.7%.

In contrast, Reddit emphasises content over user identity, allowing anonymous discussions. This structure fosters authentic sentiment expression, reducing biases linked to self-promotion or reputational concerns (Srinivasan, 2023), and all this content is controlled by moderators[8]. Given this, WSB provides a more accurate measure of investor sentiment, making it a valuable dataset for studying disagreement.

This study proposes measuring disagreement using sentiment scores from both submissions and comments rather than simply counting bullish and bearish messages. In empirical works, it is common to measure disagreement as the standard deviation of analyst forecasts (Diether et al., 2002); this study follows this approach and computes disagreement as the standard deviation of the sentiment of submissions and comments for each firm × day.

$$Disagreement_{it} = Standard\ Deviation\ of\ sentiment\ score_{it}$$

The average sentiment of comments is calculated using their score (upvotes minus downvotes) as a weight. Similarly, the average sentiment of each original post and its related comments (on the same date) is also weighted by their respective scores. This disagreement measure is weighted by the number of comments associated with each submission.

Short Sale Turnover

The short-sale turnover is calculated as the volume of shares sold divided by shares outstanding for day *t*, for each investor type (institutional and retail) and the total short-sale turnover (i.e., not differentiated between institutional and retail investors) based on FINRA data (Wang et al., 2020; Kot et al., 2020). To facilitate comparison across investor types (retail and institutional), short sale turnover is standardised for each group to have a mean of zero and a standard deviation of one. This enables meaningful comparison across variables by placing them on a common scale and allows the regression coefficients to be interpreted in terms of standard deviation changes.

Short turnover is a strong dependent variable that reflects investor sentiment and market inefficiencies. It differentiates between retail and institutional behaviours, with institutional investors often engaging in contrarian short-selling and retail investors reacting to market movements. Short turnover offers insights into market liquidity and how investors adjust to price discrepancies, making it a valuable measure of market efficiency.

V. Empirical Results

Descriptive Statistics of Variables

Table (4) presents descriptive statistics on the abnormal volume trading, short sale turnover, control variables and eight measurements of daily disagreement in WSB, respectively, from 2020 to 2023.

8. Moderators on Reddit manage and enforce the rules of a specific subreddit, ensuring that content follows the community guidelines and maintaining a positive environment for discussion.

Table 3: Descriptive Statistics of Variables

Variables	N	Mean	SD	Min	Max
Panel A: Abnormal Trading Volume					
Abnormal Log Volume (t)	2,889,260	0.009	0.709	-1.295	1.559
Abnormal Log Volume (t-1)	2,889,260	0.009	0.709	-1.295	1.558
Panel B: Short Sale Turnover (All Type of Investors)					
Shorsale Turnover_All (t)	3,033,278	0.004	0.012	0.000	0.093
Shor sale Turnover_All (t-1)	2,502,693	0.005	0.014	0.000	0.107
Past short sale turnover_All	557,514	0.007	0.017	0.000	0.125
Panel C: Short Sale Turnover (Retail Investors)					
Shorsale Turnover_Ret (t)	3,033,278	0.001	0.005	0.000	0.038
Shor sale Turnover_Ret (t-1)	3,033,278	0.001	0.005	0.000	0.038
Past short sale turnover_Ret	3,033,278	0.001	0.005	0.000	0.048
Panel D: Short Sale Turnover (Institutional Investors)					
Shorsale Turnover_Inst (t)	3,033,623	0.003	0.007	0.000	0.054
Shor sale Turnover_Inst (t-1)	3,033,623	0.003	0.007	0.000	0.054
Past short sale turnover_Inst	3,033,623	0.003	0.007	0.000	0.065
Panel E: Other Control Variables					
Volatility (t-5 to t-1)	3,027,229	0.032	0.029	0.003	0.175
CAbRet (t-5 to t-1)	3,027,415	-0.002	0.079	-0.249	0.299
CAbRet (t-30 to t-6)	3,002,285	-0.009	0.205	-0.622	0.787
Size	3,024,436	20.320	2.493	16.057	24.744
Illiquidity	3,031,246	0.072	0.171	0.000	0.677
Return on S&P	3,033,278	0.0005	0.011	-0.035	0.031
Media Coverage	3,033,278	0.043	0.202	0.000	1.000
Panel F: Disagreement Measures Based on Bullish/Bearishness					
Disagreement in Submission+Comments (All day) Number of Bullish\Bearish messages (t)	3,033,278	0.073	0.250	0.000	1.000
Disagreement in Submission (All day) Number of Bullish\Bearish messages (t)	3,033,278	0.117	0.309	0.000	1.000
Disagreement in Submission+Comments (Overnight) Number of Bullish\Bearish messages (t)	3,033,278	0.215	0.053	0.000	1.000
Disagreement in Submission (Overnight) Number of Bullish\Bearish messages (t)	3,033,278	0.269	0.085	0.000	1.000
Panel G: Disagreement Measures Based on Sentiment Score					
Disagreement in Submission+Comments (All day) Sentiment Score (t)	3,033,278	0.167	0.144	0.000	1.676
Disagreement in Submission (All day) Sentiment Score (t)	3,033,278	0.191	0.174	0.000	0.963
Disagreement in Submission+Comments (Overnight) Sentiment Score (t)	3,033,278	0.144	0.124	0.000	1.592
Disagreement in Submission (Overnight) Sentiment Score (t)	3,033,278	0.171	0.142	0.000	0.963

This table presents the mean, standard deviation, minimum, and maximum values for the dependent and explanatory variables used in this study. Panel A displays the statistics for the variables that analyse abnormal trading volume. Panels B, C, and D show the statistics for short-selling turnover among all sample retail and institutional investors. Panel E presents the statistics for another control variable in the regression models. Panels F and G report the statistics for disagreement measures based on Bullish/Bearishness and Sentiment Score. The sample includes 2,784 unique firms of common stocks listed on NASDAQ, NYSE, and AMEX from January 2020 to December 2023. For variable definitions, refer to Appendix B.

Main Results

This section addresses the research hypothesis: "The higher level of daily disagreement in WSB is related to higher trading volume", and "The higher level of daily disagreement on WSB is positively associated with short-sale turnover."

Abnormal Trading Volume and Disagreement

This section assesses the relationship between trading volume and various disagreement measures. The factors driving trading volume and its fluctuations over time continue to be topics of debate in finance literature (e.g., Hong & Stein, 2007). To explore how the disagreement measures relate to abnormal trading volume and in line with the existing literature (Cookson & Neissner, 2020, 2023), the empirical link between disagreement and abnormal trading volume is estimated using the following regression model:

$$AbLogVol_{it} = \alpha_t + \gamma_i + \beta_1 DisagreementMeasure_{it} + \beta_2 AbLogVol_{it-1} + \gamma Controls_{it} + \epsilon_{it}$$

The dependent variable $AbLogVol_{it}$ is the abnormal log trading volume on date *t* for firm *i*. It is calculated as the difference between the log volume on date *t* and the average log volume from trading days *t - 140* to *t - 20* (six-month period, skipping a month). The logarithm is used to normalise the data and reduce the impact of extreme values. The reason for skipping the most recent month is to avoid the immediate effects of short-term events or fluctuations. The variable $DisagreementMeasure_{it}$ is one of the eight disagreement measures described in Section IV (overall, overnight, only submissions, submissions and related comments, based on sentiment score and bullishness/bearishness of messages) for a given firm *i* on day *t*.

To account for alternative interpretations, all specifications include the date and firm fixed effects (α_t and γ_i). Control variables include abnormal trading volume on day *t-1* to account for persistence in abnormal trading volume, volatility (measured from *t-5* to *t-1*), the standard deviation of abnormal returns over days *t-5* to *t-1*, and cumulative abnormal returns (CAR) over days *t-5* to *t-1* and *t-30* to *t-6*. The size variable, used to control for the size effect (Fama & French, 1993), is the logarithmic value of the stock's market capitalisation, calculated as shares outstanding multiplied by the closing price at the end of day *t*. Illiquidity is measured based on Amihud (2002) as the absolute daily return divided by dollar trading volume (ILLIQ × 10^6). This measure is included to control for the impact of liquidity, as higher illiquidity can affect market behaviour and trading dynamics (Fernando et al., 2024). The return on the Standard & Poor's Index is included to control for overall market-wide effects, but is omitted because it is probably collinear with the fixed effects. Media coverage is accounted for to capture firm-date-specific surges in attention. This dummy variable equals 1 if firm *i* has events recorded in the RavenPack database on day *t*. Across specifications, standard errors are double-clustered by date and firm to account for within-firm autocorrelation and common daily shocks. To control for outliers, the top 0.1% of observations are winsorised.

Table (5) and Table (6) present the panel regression fixed effect results on the link between disagreement measures (the Sentiment Score method), disagreement measures (the Bullishness/Bearishness method) and trading volume. In columns (1) and (2) of the Table (5), $DisagreementMeasure_{it}$ is the disagreement all-day (between 4 PM on day *t-1* and 4 PM on day *t*) among submissions and related comments and only submissions based on sentiment score. Columns (3) and (4) show overnight disagreement (between 4 PM on day *t-1* and 9 AM on day *t*) across submissions and related comments, and only submissions based on sentiment score. Columns (5) and (6) in Table (6) indicate all-day disagreement among submissions and related comments, and only submissions based on the number of bullishness/bearishness messages. Columns (7) and (8) show overnight disagreement across submissions and related comments, and only submissions based on the number of bullishness/bearishness messages. Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Table 5: Disagreement Measures (the sentiment score method) and Trading Volume

		Abnormal Log Volume (t)			
Disagreement Measures		(1)	(2)	(3)	(4)
Disagreement in Submission+Comments (All day) Sentiment Score (t)	(1)	**0.571***** **(0.052)**			
Disagreement in Submission (All day) Sentiment Score (t)	(2)		**0.494***** **(0.042)**		
Disagreement in Submission+Comments (Overnight) Sentiment Score (t)	(3)			**0.444***** **(0.045)**	
Disagreement in Submission (Overnight) Sentiment Score (t)	(4)				**0.400***** **(0.038)**
Abnormal Log Volume (t-1)		0.641*** (0.004)	0.641*** (0.004)	0.641*** (0.004)	0.641*** (0.004)
Volatility (t-5 to t-1)		0.384*** (0.093)	0.384*** (0.092)	0.386*** (0.093)	0.385*** (0.093)
CAbRet (t-5 to t-1)		0.062*** (0.013)	0.062*** (0.013)	0.062*** (0.013)	0.062*** (0.013)
CAbRet (t-30 to t-6)		0.068*** (0.005)	0.068*** (0.005)	0.068*** (0.005)	0.068*** (0.005)
Size		-0.037*** (0.002)	-0.037*** (0.002)	-0.037*** (0.002)	-0.037*** (0.002)
Illiquidity		-0.760*** (0.017)	-0.760*** (0.017)	-0.760*** (0.017)	-0.760*** (0.017)
Media Coverage		0.022*** (0.005)	0.022*** (0.005)	0.022*** (0.005)	0.022*** (0.005)
_cons		0.796*** (0.052)	0.796*** (0.052)	0.793*** (0.052)	0.794*** (0.052)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,879,986	2,879,986	2,879,986	2,879,986
Number of Firms		2,784	2,784	2,784	2,784
R-Squred		0.524	0.524	0.524	0.524

Table 6: Disagreement Measures (the bullishness/bearishness method) and Trading Volume

		Abnormal Log Volume (t)			
Disagreement Measures		(5)	(6)	(7)	(8)
Disagreement in Submission+Comments (All day) Number of Bullish\Bearish messages (t)	(5)	**0.221*** (0.027)**			
Disagreement in Submission (All day) Number of Bullish\Bearish messages (t)	(6)		**0.210*** (0.022)**		
Disagreement in Submission+Comments (Overnight) Number of Bullish\Bearish messages (t)	(7)			**0.186*** (0.024)**	
Disagreement in Submission (Overnight) Number of Bullish\Bearish messages (t)	(8)				**0.178*** (0.021)**
Abnormal Log Volume (t-1)		0.641*** (0.004)	0.641*** (0.004)	0.641*** (0.004)	0.641*** (0.004)
Volatility (t-5 to t-1)		0.387*** (0.093)	0.387*** (0.093)	0.387*** (0.093)	0.387*** (0.093)
CAbRet (t-5 to t-1)		0.062*** (0.013)	0.062*** (0.013)	0.062*** (0.013)	0.062*** (0.013)
CAbRet (t-30 to t-6)		0.068*** (0.005)	0.068*** (0.005)	0.068*** (0.005)	0.068*** (0.005)
Size		-0.037*** (0.002)	-0.037*** (0.002)	-0.037*** (0.002)	-0.037*** (0.002)
Illiquidity		-0.760*** (0.017)	-0.760*** (0.017)	-0.760*** (0.017)	-0.760*** (0.017)
Media Coverage		0.022*** (0.005)	0.022*** (0.005)	0.022*** (0.005)	0.022*** (0.005)
_cons		0.792*** (0.052)	0.792*** (0.052)	0.791*** (0.052)	0.791*** (0.052)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,879,986	2,879,986	2,879,986	2,879,986
Number of Firms		2,784	2,784	2,784	2,784
R-Squred		0.524	0.524	0.524	0.524

The tables present the relationship between disagreement proxies and trading volume. Column (1) shows that a one-unit increase in all-day disagreement, measured using sentiment scores in submissions and related comments, is associated with a 0.571 increase in abnormal trading volume, statistically significant at the 1% level. This effect remains robust after controlling for firm and date fixed effects, lagged abnormal log volume, volatility, and abnormal returns, suggesting that this measure effectively captures broader market disagreement.

Column (2) isolates disagreement from submissions alone, which precedes trading volume by construction. A one-unit increase in all-day disagreement leads to a 0.494 increase in abnormal trading volume. Similarly, overnight disagreement is significantly linked to trading volume (columns (3) and (4)), with a 0.444 increase for sentiment-based measures in submissions and comments and 0.400 when only submissions are considered.

Columns (5) and (6) explore disagreement using bullish/bearish classifications. A one-unit increase in all-day disagreement (submissions and comments) is associated with a 0.221 rise in abnormal trading volume, while disagreement from submissions alone results in a 0.210 increase. Overnight disagreement (columns (7) and (8)) maintains a positive relation, with estimates of 0.186 (submissions and comments) and 0.178 (submissions alone), all statistically significant at the 1% level.

Overall, the results confirm that disagreement on WSB, both all day and overnight, significantly influences abnormal trading volume. Sentiment-based measures exhibit stronger effects than simple bullish/bearish classifications, emphasising the role of sentiment dynamics in investor decision-making. These findings highlight the rapid transmission of social media disagreement into trading activity, reinforcing its role in price discovery and market efficiency.

Short Selling Activities and Disagreement

A) All Types of Investors

This section evaluates the relationship between short-selling activities among all types of investors and various disagreement measures. To examine how the disagreement measures relate to short sale turnover and following the existing literature (e.g., Boehmer et al., 2008; Boehmer & Wu, 2013; Dechow et al., 2001; Engelberg et al., 2012), the empirical link between disagreement and short sale turnover is estimated using the following regression specification:

$$ShrtTurnover_{it} = \alpha_t + \gamma_i + \beta_1 DisagreementMeasure_{it} + \beta_2 ShrtTurnover_{it-1} + \gamma Controls_{it} + \epsilon_{it}$$

The dependent variable $ShrtTurnover_{it}$ is calculated as the volume of shares sold divided by shares outstanding for day *t* for the total short-sale turnover (i.e., not differentiated between institutional and retail investors) based on FINRA data (Wang et al., 2020; Kot et al., 2020).

The variable $DisagreementMeasure_{it}$ is one of the eight disagreement measures described in Section IV (overall, overnight, only submissions, submissions and related comments, based on sentiment score and bullishness/bearishness of messages) for a given firm *i* on day *t*.

To account for alternative interpretations, all specifications include the date and firm fixed effects (α_t and γ_i). Control variables include short-sale turnover on day *t-1* is also included to account for the persistence of short-selling behaviour. Another control variable in these models is past turnover, which may influence or correlate with the turnover observed on the current day. To address this, past short-sale turnover is included as an independent variable to account for the autocorrelation of short turnover (Diether et al., 2009). This control variable is the average short turnover from *t-5* to *t-1*. Volatility (measured from *t-5* to *t-1*), the standard deviation of abnormal returns over days *t-5* to *t-1*, and cumulative abnormal returns (CAR) over days *t-5* to *t-1* and *t-30* to *t-6*. The size variable, used to control for the size effect (Fama & French, 1993), is the logarithmic value of the stock's market capitalisation, calculated as shares outstanding multiplied by the closing price at the end of day *t*. Illiquidity is measured based on Amihud (2002) as the absolute daily return divided by dollar trading volume (ILLIQ × 10^6). This measure is included to control for the impact of liquidity, as higher illiquidity can affect market behaviour and trading dynamics (Fernando et al., 2024). The return on the Standard & Poor's Index is included to control for overall market-wide effects, but is omitted because it is probably collinear with the fixed effects. Media coverage is accounted for to capture firm-date-specific surges in attention. This dummy variable equals 1 if firm *i* has events recorded in the RavenPack database on day *t*. Across specifications, standard errors are double-clustered by date and firm to account for within-firm autocorrelation and common daily shocks. To control for outliers, the top 0.1% of observations are winsorised.

Table (7) and Table (8) present the panel regression fixed effect results on the link between disagreement measures (the Sentiment Score method), disagreement measures (the Bullishness/Bearishness method) and short-sale turnover among all types of investors. In columns (1) and (2) of the Table (7), $DisagreementMeasure_{it}$ is the disagreement all-day (between 4 PM on day *t-1* and 4 PM on day *t*) among submissions and related comments and only submissions based on sentiment score. Columns (3) and (4) show overnight disagreement (between 4 PM on day *t-1* and 9 AM on day *t*) across submissions and related comments and only submissions based on sentiment score. Columns (5) and (6) in Table (8) indicate all-day disagreement among submissions and related comments and only submissions based on the number of bullishness/bearishness messages. Columns (7) and (8) show overnight disagreement across submissions and related comments and only submissions based on the number of bullishness/bearishness messages. Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Table 7: Disagreement Measures (the sentiment score method) and Short Sale Turnover (All Investors)

		Short Sale Turnover_All (t)			
Disagreement Measures		(1)	(2)	(3)	(4)
Disagreement in Submission+Comments (All day) Sentiment Score (t)	(1)	**0.041*** (0.004)**			
Disagreement in Submission (All day) Sentiment Score (t)	(2)		**0.035*** (0.004)**		
Disagreement in Submission+Comments (Overnight) Sentiment Score (t)	(3)			**0.039*** (0.004)**	
Disagreement in Submission (Overnight) Sentiment Score (t)	(4)				**0.035*** (0.004)**
Short Sale Turnover_All (t-1)		-0.135*** (0.006)	-0.135*** (0.006)	-0.135*** (0.006)	-0.135*** (0.006)
Past Short Sale Turnover_All		0.316*** (0.008)	0.316*** (0.008)	0.317*** (0.008)	0.316*** (0.008)
Volatility (t-5 to t-1)		0.105*** (0.003)	0.105*** (0.003)	0.105*** (0.003)	0.105*** (0.003)
CAbRet (t-5 to t-1)		0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)
CAbRet (t-30 to t-6)		0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Size		-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Illiquidity		-0.015*** (0.000)	-0.015*** (0.000)	-0.015*** (0.000)	-0.015*** (0.000)
Media Coverage		-0.0001 (0.000)	-0.0001 (0.000)	-0.0001 (0.000)	-0.0001 (0.000)
_cons		0.033*** (0.003)	0.033*** (0.003)	0.033*** (0.003)	0.033*** (0.003)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,349,896	2,349,896	2,349,896	2,349,896
Number of Firms		2,669	2,669	2,669	2,669
R-Squred		0.521	0.521	0.520	0.520

Table 8: Disagreement Measures (the bullishness/bearishness method) and Short Sale Turnover (All Investors)

		Short Sale Turnover_All (t)			
Disagreement Measures		(5)	(6)	(7)	(8)
Disagreement in Submission+Comments (All day) Number of Bullish\Bearish messages (t)	(5)	**0.017***** **(0.003)**			
Disagreement in Submission (All day) Number of Bullish\Bearish messages (t)	(6)		**0.016***** **(0.003)**		
Disagreement in Submission+Comments (Overnight) Number of Bullish\Bearish messages (t)	(7)			**0.018***** **(0.004)**	
Disagreement in Submission (Overnight) Number of Bullish\Bearish messages (t)	(8)				**0.015***** **(0.003)**
Short Sale Turnover_All (t-1)		-0.135*** (0.006)	-0.135*** (0.006)	-0.135*** (0.006)	-0.135*** (0.006)
Past Short Sale Turnover_All		0.318*** (0.008)	0.317*** (0.008)	0.318*** (0.008)	0.318*** (0.008)
Volatility (t-5 to t-1)		0.106*** (0.003)	0.106*** (0.003)	0.105*** (0.003)	0.105*** (0.003)
CAbRet (t-5 to t-1)		0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)
CAbRet (t-30 to t-6)		0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Size		-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Illiquidity		-0.015*** (0.000)	-0.015*** (0.000)	-0.015*** (0.000)	-0.015*** (0.000)
Media Coverage		-0.0001 (0.000)	-0.0001 (0.000)	-0.0001 (0.000)	-0.0001 (0.000)
_cons		0.033*** (0.003)	0.033*** (0.003)	0.033*** (0.003)	0.033*** (0.003)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,349,896	2,349,896	2,349,896	2,349,896
Number of Firms		2,669	2,669	2,669	2,669
R-Squred		0.519	0.520	0.519	0.519

The tables examine the relationship between disagreement proxies and short sale turnover among all types of investors. Column (1) shows that a one-unit increase in all-day disagreement, measured by sentiment scores from submissions and comments, is associated with a 0.041 rise in short sale turnover, statistically significant at the 1% level. This effect remains robust after controlling for firm and date fixed effects, lagged short sale turnover, volatility, and abnormal returns. In column (2), where disagreement is derived solely from submissions, the impact is smaller but still significant, with a 0.035 increase in short sale turnover.

Overnight disagreement measures also exhibit a significant positive relationship with short-selling activities. Column (3) shows that a one-unit increase in overnight disagreement based on sentiment scores in submissions and comments is associated with a 0.039 increase in short sale turnover, while column (4) reports a smaller but statistically significant estimate of 0.035 when only submissions are considered. Similarly, columns (5) and (6) examine disagreement using bullish/bearish classifications, finding that a one-unit increase in all-day disagreement from submissions and comments corresponds to a 0.017 increase in short sale turnover, whereas the effect from submissions alone is 0.016. Overnight disagreement in this classification, as shown in columns (7) and (8), maintains a positive and significant relationship with short sale turnover, with effects of 0.018 and 0.015, respectively.

These findings provide strong evidence that social media-driven disagreement significantly influences short-sale turnover among all types of investors, both throughout the trading day and overnight. The effect is particularly pronounced when disagreement is measured using sentiment scores rather than bullish/bearish classifications, suggesting that short sellers are more responsive to nuanced sentiment dynamics than broad market outlook categorisations. The results further indicate that overnight disagreement remains a strong predictor of short-selling activity, reinforcing differences of opinion formed before market opening play a role in shaping trading behaviour. Overall, these findings highlight the growing importance of investor sentiment in short-selling decisions, supporting its role in price discovery and market efficiency. This aligns with behavioural finance theories, which suggest that sentiment-induced mispricing and overreaction to public signals can drive market participation, particularly in short-selling activity.

B) Retail Investors

Table (9) and Table (10) present the panel regression fixed effect results on the link between disagreement measures (the Sentiment Score method), disagreement measures (the Bullishness/Bearishness method) and retail short-sale turnover. In columns (1) and (2) of the Table (9), $DisagreementMeasure_{it}$ is the disagreement all-day (between 4 PM on day *t-1* and 4 PM on day *t*) among submissions and related comments and only submissions based on sentiment score. Columns (3) and (4) show overnight disagreement (between 4 PM on day *t-1* and 9 AM on day *t*) across submissions and related comments and only submissions based on sentiment score. Columns (5) and (6) in Table (10) indicate all-day disagreement among submissions and related comments and only submissions based on the number of bullishness/bearishness messages. Columns (7) and (8) show overnight disagreement across submissions and related comments and only submissions based on the number of bullishness/bearishness messages. Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Table 9: Disagreement Measures (the sentiment score method) and Retail Short Sale Turnover

		Short Sale Turnover_Ret (t)			
Disagreement Measures		(1)	(2)	(3)	(4)
Disagreement in Submission+Comments (All day) Sentiment Score (t)	(1)	**0.057*** (0.006)**			
Disagreement in Submission (All day) Sentiment Score (t)	(2)		**0.052*** (0.005)**		
Disagreement in Submission+Comments (Overnight) Sentiment Score (t)	(3)			**0.047*** (0.005)**	
Disagreement in Submission (Overnight) Sentiment Score (t)	(4)				**0.044*** (0.004)**
Short Sale Turnover_Ret (t-1)		0.582*** (0.008)	0.581*** (0.008)	0.582*** (0.008)	0.582*** (0.008)
Past Short Sale Turnover_Ret		0.020*** (0.002)	0.020*** (0.002)	0.020*** (0.002)	0.020*** (0.002)
Volatility (t-5 to t-1)		0.041*** (0.003)	0.041*** (0.003)	0.041*** (0.003)	0.041*** (0.003)
CAbRet (t-5 to t-1)		-0.008*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)
CAbRet (t-30 to t-6)		0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)
Size		-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)
Illiquidity		-0.031*** (0.001)	-0.031*** (0.001)	-0.031*** (0.001)	-0.031*** (0.001)
Media Coverage		-0.0002* (0.000)	-0.0002* (0.000)	-0.0002* (0.000)	-0.0002* (0.000)
_cons		0.055*** (0.004)	0.055*** (0.004)	0.054*** (0.004)	0.054*** (0.004)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,294,582	2,294,582	2,294,582	2,294,582
Number of Firms		2,776	2,776	2,776	2,776
R-Squred		0.675	0.675	0.675	0.675

Table 10: Disagreement Measures (the bullishness/bearishness method) and Retail Short Sale Turnover

		Short Sale Turnover_Ret (t)			
Disagreement Measures		(5)	(6)	(7)	(8)
Disagreement in Submission+Comments (All day) Number of Bullish\Bearish messages (t)	(5)	**0.026*** (0.004)**			
Disagreement in Submission (All day) Number of Bullish\Bearish messages (t)	(6)		**0.026*** (0.003)**		
Disagreement in Submission+Comments (Overnight) Number of Bullish\Bearish messages (t)	(7)			**0.025*** (0.011)**	
Disagreement in Submission (Overnight) Number of Bullish\Bearish messages (t)	(8)				**0.023*** (0.003)**
Short Sale Turnover_Ret (t-1)		0.582*** (0.008)	0.582*** (0.008)	0.582*** (0.008)	0.582*** (0.008)
Past Short Sale Turnover_Ret		0.020*** (0.002)	0.020*** (0.002)	0.020*** (0.002)	0.020*** (0.002)
Volatility (t-5 to t-1)		0.042*** (0.003)	0.041*** (0.003)	0.041*** (0.003)	0.041*** (0.003)
CAbRet (t-5 to t-1)		-0.008*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)
CAbRet (t-30 to t-6)		0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)
Size		-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)
Illiquidity		-0.031*** (0.001)	-0.031*** (0.001)	-0.031*** (0.001)	-0.031*** (0.001)
Media Coverage		-0.0002* (0.000)	-0.0002* (0.000)	-0.0002* (0.000)	-0.0002* (0.000)
_cons		0.054*** (0.004)	0.054*** (0.004)	0.055*** (0.004)	0.055*** (0.004)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,294,582	2,294,582	2,294,582	2,294,582
Number of Firms		2,776	2,776	2,776	2,776
R-Squred		0.675	0.675	0.675	0.675

The tables analyse the relationship between disagreement proxies and short sale turnover among retail investors. Column (1) shows that a one-unit increase in all-day disagreement, based on sentiment scores from submissions and comments, leads to a 0.057 increase in short sale turnover, statistically significant at the 1% level. This effect remains robust after controlling for firm and date fixed effects, lagged short sale turnover, volatility, and abnormal returns. Column (2) reports a smaller but significant effect of 0.052 when disagreement is measured solely from submissions.

Overnight disagreement measures also show a significant positive relationship with short sale turnover. Column (3) finds that a one-unit increase in overnight disagreement from sentiment scores in submissions and comments corresponds to a 0.047 increase in short sale turnover, while column (4) reports a smaller estimate of 0.044 when considering only submissions. Similarly, disagreement based on bullish/bearish classifications, analysed in columns (5) and (6), shows a positive effect, with all-day disagreement increasing short sale turnover by 0.026 when derived from both submissions and comments and by 0.026 when derived solely from submissions. Overnight disagreement in this classification (columns (7) and (8)) also significantly predicts short sale turnover, with estimates of 0.025 and 0.023, respectively.

These findings prove that disagreement from WSB submissions and comments, both during the trading day and overnight, is significantly associated with increased short-selling activity among retail investors. The effect is stronger when disagreement is measured using sentiment scores rather than bullish/bearish classifications, suggesting that retail short sellers react more to nuanced sentiment shifts than broad optimism or pessimism. Moreover, disagreement derived solely from submissions remains a significant predictor of short-selling activity, reinforcing that retail investors actively process sentiment signals before executing trades rather than reacting in real time. These results highlight the growing role of social media-driven sentiment in shaping retail trading behaviour, supporting its influence on price discovery and market efficiency. This aligns with behavioural finance theories, which suggest that investor disagreement, sentiment-driven overreaction, and noise amplification through social media contribute to short-selling activity and market dynamics.

C) Institutional Investors

Table (11) and Table (12) present the panel regression fixed effect results on the link between disagreement measures (the Sentiment Score method), disagreement measures (the Bullishness/Bearishness method) and institutional short-sale turnover. In columns (1) and (2) of the Table (11), $DisagreementMeasure_{it}$ is the disagreement all-day (between 4 PM on day *t-1* and 4 PM on day *t*) among submissions and related comments and only submissions based on sentiment score. Columns (3) and (4) show overnight disagreement (between 4 PM on day *t-1* and 9 AM on day *t*) across submissions and related comments and only submissions based on sentiment score. Columns (5) and (6) in Table (12) indicate all-day disagreement among submissions and related comments and only submissions based on the number of bullishness/bearishness messages. Columns (7) and (8) show overnight disagreement across submissions and related comments and only submissions based on the number of bullishness/bearishness messages. Standard errors are clustered by firm and date. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Standard errors are in parentheses.

Table 11: Disagreement Measures (the sentiment score method) and Institutional Short Sale Turnover

		Short Sale Turnover_Inst (t)			
Disagreement Measures		(1)	(2)	(3)	(4)
Disagreement in Submission+Comments (All day) Sentiment Score (t)	(1)	**0.080*** (0.008)**			
Disagreement in Submission (All day) Sentiment Score (t)	(2)		**0.072*** (0.007)**		
Disagreement in Submission+Comments (Overnight) Sentiment Score (t)	(3)			**0.064*** (0.008)**	
Disagreement in Submission (Overnight) Sentiment Score (t)	(4)				**0.059*** (0.007)**
Short Sale Turnover_Inst (t-1)		0.542*** (0.009)	0.542*** (0.009)	0.542*** (0.009)	0.542*** (0.009)
Past Short Sale Turnover_Inst		0.022*** (0.004)	0.022*** (0.004)	0.022*** (0.004)	0.022*** (0.004)
Volatility (t-5 to t-1)		0.073*** (0.005)	0.073*** (0.005)	0.073*** (0.005)	0.073*** (0.005)
CAbRet (t-5 to t-1)		-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)
CAbRet (t-30 to t-6)		0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)
Size		-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Illiquidity		-0.038*** (0.001)	-0.038*** (0.001)	-0.038*** (0.001)	-0.038*** (0.001)
Media Coverage		-0.0004* (0.000)	-0.0004* (0.000)	-0.0004* (0.000)	-0.0004* (0.000)
_cons		0.035*** (0.005)	0.035*** (0.005)	0.035*** (0.005)	0.035*** (0.005)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,349,895	2,349,895	2,349,895	2,349,895
Number of Firms		2,784	2,784	2,784	2,784
R-Squred		0.664	0.665	0.665	0.665

Table 12: Disagreement Measures (the bullishness/bearishness method) and Institutional Short Sale Turnover

		Short Sale Turnover_Inst (t)			
Disagreement Measures		(5)	(6)	(7)	(8)
Disagreement in Submission+Comments (All day) Number of Bullish\Bearish messages (t)	(5)	**0.037***** **(0.006)**			
Disagreement in Submission (All day) Number of Bullish\Bearish messages (t)	(6)		**0.035***** **(0.005)**		
Disagreement in Submission+Comments (Overnight) Number of Bullish\Bearish messages (t)	(7)			**0.036***** **(0.007)**	
Disagreement in Submission (Overnight) Number of Bullish\Bearish messages (t)	(8)				**0.031***** **(0.005)**
Short Sale Turnover_Inst (t-1)		0.543*** (0.009)	0.543*** (0.009)	0.543*** (0.009)	0.542*** (0.009)
Past Short Sale Turnover_Inst		0.022*** (0.004)	0.022*** (0.004)	0.022*** (0.004)	0.022*** (0.004)
Volatility (t-5 to t-1)		0.073*** (0.005)	0.073*** (0.005)	0.073*** (0.005)	0.073*** (0.005)
CAbRet (t-5 to t-1)		-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)
CAbRet (t-30 to t-6)		0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)
Size		-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Illiquidity		-0.038*** (0.001)	-0.038*** (0.001)	-0.038*** (0.001)	-0.038*** (0.001)
Media Coverage		-0.0004* (0.000)	-0.0004* (0.000)	-0.0004* (0.000)	-0.0003* (0.000)
_cons		0.034*** (0.005)	0.035*** (0.005)	0.034*** (0.005)	0.034*** (0.005)
Date Fixed Effect		Yes	Yes	Yes	Yes
Firm Fixed Effect		Yes	Yes	Yes	Yes
Observations		2,349,895	2,349,895	2,349,895	2,349,895
Number of Firms		2,784	2,784	2,784	2,784
R-Squred		0.665	0.665	0.664	0.664

The tables examine the relationship between disagreement proxies and short sale turnover among institutional investors. Column (1) shows that a one-unit increase in all-day disagreement, based on sentiment scores from submissions and comments, leads to a 0.080 increase in short sale turnover, statistically significant at the 1% level. This effect remains robust after controlling for firm and date fixed effects, lagged short sale turnover, volatility, and abnormal returns. Column (2) reports a smaller but significant effect of 0.072 when disagreement is measured solely from submissions.

Overnight disagreement also has a significant positive effect on institutional short-selling activity. Column (3) finds that a one-unit increase in overnight disagreement from sentiment scores in submissions and comments corresponds to a 0.054 increase in short sale turnover. In contrast, column (4) reports a smaller estimate of 0.059 when considering only submissions. Disagreement based on bullish/bearish classifications, analysed in columns (5) and (6), also positively affects short sale turnover, with all-day disagreement increasing it by 0.037 when derived from both submissions and comments and by 0.035 when derived solely from submissions. Overnight disagreement in this classification (columns (7) and (8)) remains significant, with estimates of 0.036 and 0.031, respectively.

These findings demonstrate that disagreement from WSB submissions and comments, both throughout the trading day and overnight, significantly influences institutional short-selling activity. The effect is stronger when disagreement is measured using sentiment scores rather than broad bullish/bearish classifications, suggesting that institutional investors respond more to nuanced sentiment shifts. The results highlight the increasing role of social media-driven sentiment in institutional trading, reinforcing its impact on price discovery and market efficiency. These findings suggest that institutional investors integrate online sentiment into their strategies. This aligns with behavioural finance theories that investor disagreement, sentiment-driven mispricing, and noise amplification through social media contribute to short-selling activity and broader market dynamics. These insights further challenge the assumption that institutional investors are immune to sentiment-driven trading, indicating they may strategically exploit disagreement signals when making short-selling decisions.

VI. Conclusion

Disagreement is a well-established driver of trading activity. This study examines the relationship between daily investor disagreement, trading volume, and short-sale activity, focusing on social media's growing role in market dynamics. By analysing disagreement measures from r/WallStreetBets (WSB), this research provides new insights into how sentiment divergence influences retail and institutional trading behaviour.

The findings confirm that disagreement stimulates trading. Both sentiment scores and message-labelling methods predict trading volume and short-sale turnover, with sentiment scores proving more effective. Disagreement from submissions and comments throughout the trading day significantly impacts trading, while overnight disagreement predicts next-day activity, particularly from comment-driven discussions, underscoring the persistence of sentiment effects.

A key contribution is the distinction between retail and institutional responses. Both groups increase short-selling amid heightened disagreement, but institutional investors react more strongly. Using FINRA off-exchange short-sales data, the results suggest that institutional investors, often acting as informed contrarians, exploit disagreement as an arbitrage opportunity, reinforcing their role in price discovery and market efficiency. These findings align with behavioural finance theories, which posit that disagreement-driven mispricing offers opportunities for informed traders to capitalise on market overreactions.

The study further highlights that disagreement effects persist throughout the trading day and overnight. The correlation between WSB disagreement, abnormal trading volume, and short-sale turnover is strongest during market hours. This suggests that social media sentiment, especially when captured through sentiment scoring, translates rapidly into market decisions, reinforcing its role in price formation.

This study contributes to the literature by demonstrating that social media-driven investor disagreement is significantly associated with increased abnormal trading volume and short-sale turnover, particularly among institutional investors. By applying a novel, high-frequency disagreement measure based on direct sentiment scores from both posts and comments on Reddit's WallStreetBets, the paper captures belief heterogeneity with greater granularity than previous studies. The findings underscore the informational value of social media sentiment and its role in price discovery, suggesting that sophisticated traders use such disagreement as a signal for arbitrage. This research also highlights the heterogeneity in short-seller behaviour and offers practical implications for investors, regulators, and algorithmic traders in understanding the impact of online discourse on market dynamics.

However, some limitations must be acknowledged. The accuracy of disagreement metrics, particularly sentiment aggregation from posts and comments, requires further validation. While WSB provides a unique setting for retail investors, the findings may not fully generalise to other investor populations or platforms. Additionally, using observational data raises concerns about endogeneity and potential biases. Future research could address these by incorporating data from multiple social media sources, examining disagreement effects during major market events, and exploring the regulatory implications of sentiment-driven trading.

Overall, this study demonstrates that higher investor disagreement on WSB is positively associated with increased trading volume and short-sale turnover, with institutional investors responding more strongly. These findings underscore the growing influence of social media sentiment on trading behaviour and contribute to the broader literature on behavioural finance, market efficiency, and sentiment's role in asset pricing. By linking empirical market data with behavioural theories, this research deepens the understanding of how real-time investor sentiment drives market movements in the digital age.

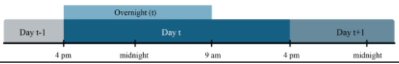
VII. Appendix

Appendix A: Adjusted VADER lexicon by Long et al. (2023)

This table presents a list of words from the new lexicon and their corresponding valence scores, as proposed by Long et al. (2023):

Word	Score	Word	Score	Word	Score	Word	Score
available	0.8	diamond_hand	3	cash	0.6	advice	1.3
awesome	3.7	dip	-0.4	concern	-1.3	alternative	0.9
baby	1.2	dumb	-1.9	crash	-3.2	amazing	3.2
bad	-2.7	earning	1.8	crazy	0.7	ass	-1.9
ball	0.4	easy	1.6	crypto	0.5	attack	-1.9
bull	2.8	end	-0.8	damn	-1.7	capital	1
bullshit	-2.4	enough	0.1	diamond	2.9	fact	0.3
buy	1.9	hype	1.2	hard	-1.1	fake	-2.3
call	0.9	idiot	-2.6	hedge	0.5	fight	-1.2
future	1.1	illegal	-3.2	hell	-2.5	fine	1.3
gain	2.2	interest	1.1	high	2.4	flair	1.4
gamma	0	issue	-1.1	hodl	2.8	fuck	-2.8
gang	-0.3	joke	-0.5	hold	1.5	fucking	-2.7
gold	2	jump	1.4	holding	1.6	fun	1.9
good	2.5	least	-0.4	hope	1.5	funny	1.9
great	3.1	legal	1.9	limit	-0.4	problem	-2.3
green	2	manipulation	-2.3	lmao	2.6	profit	2.5
hand	0.1	margin	-0.1	lol	1	proud	2.1
party	0.8	moment	0.7	long	1.8	pump	-0.5
penny	-0.2	moon	2.1	loss	-2.5	purchase	1.3
poor	-1.9	movement	0.9	love	2.3	push	0.5
possible	0.8	naked	-1.1	low	-1.7	quick	0.8
potential	1.4	nice	2	luck	2.1	retard	-2.2
power	2.2	order	0.4	revolution	2	share	0.8
pretty	2.3	panic	-3	rich	2.5	shit	-2.6
probably	0.4	straight	1	ride	1	short	-1.8
top	2.4	strong	2.1	rocket	2.8	silver	-0.2
trade	0.6	stupid	-2.1	sale	-0.7	small	-0.3
value	1.3	support	2.2	scare	-2.3	squeeze	-1.6
win	2.7	target	1.3	scared	-2.6	star	2.4
worth	1.9	tendie	1.7	sell	-1.8	stonk	1.5
wrong	-1.8	to_the_moon	3.5	seller	-1.3	stop	-0.8
yolo	2.4			selling	-1.9		

Appendix B: Definitions of Variables

Variable Name	Description	Source
Dependent Variables:		
Abnormal Log Trading Volume	The abnormal log trading volume on date *t* for firm *i* is calculated as the **difference** between the **log volume on date t** and **the average log volume from trading days t - 140 to t – 20**.	CRSP
Short Sale Turnover	The short turnover is calculated as the **volume of shares sold by each investor type divided by shares outstanding for day t**. The short turnover is defined separately for institutions and retailers based on the BJZZ (2021) algorithm.	FINRA short sale
Short Sale Turnover_Ret	Boehmer et al.'s (2021) (BJZZ) algorithm estimates institutional and retail-initiated short-sale trades at the stock-day level. This method identifies retail short sales through **sub-penny price improvements**.	FINRA short sale
Short Sale Turnover_Inst	**Retail investors** account for approximately **25%** of **short-sale trades**, and around **96%** of the tickers mentioned on WSB are involved in short-sale trades on FINRA.	FINRA short sale
Daily Disagreement Variables:		
Bullishness/Bearishness Method of measuring disagreement (based on the literature)	Antweiler and Frank (2004) and Cookson and Niessner (2020, 2023) proposed a methodology to construct disagreement measures. This approach involves discretising continuous sentiment scores, which, in this study, are obtained using the **adjusted VADER** tool developed by Long et al. (2023) specifically for WallStreetBets (WSB) data analysis. Sentiment scores are categorised as follows: a score greater than 0.5 is labelled as +1 (**Bullish**), indicating positive sentiment, while a score less than -0.5 is labelled as -1 (**Bearish**), indicating negative sentiment. Observations with sentiment scores between -0.5 and 0.5 (**Neutral**) are excluded from the analysis. When there are no messages, disagreement is normalised to zero, assuming latent agreement, following Antweiler and Frank's (2004) and Cookson and Niessner's (2020) definitions.	Reddit/WSB
Sentiment Score Method	To indicate differences in investors' opinions, and following Diether et al. (2002), the sentiment score directly is used and the standard deviation of sentiment scores obtained from the adjusted VADER tool is applied to WSB posts and comments for each firm × day as an **alternative measure** of disagreement.	Reddit/WSB
Disagreement (Allday)	The disagreement (based on the two previously mentioned methods) is computed for day *t* by aggregating messages posted between the market close of the preceding **day *(t - 1)* at 4:00 PM and the market close of day t at 4:00 PM**.	Reddit/WSB
Disagreement (Overnight)	Overnight messages are from 4:00 PM of the day *(t-1)* to 9:00 AM day *t*. 	Reddit/WSB
Disagreement (Among Submissions)	As one way to measure disagreement, this study focuses only on disagreement among submissions.	Reddit/WSB
Disagreement (Among Submissions and related Comments)	Another approach to measuring disagreement involves considering both submissions and their related comments.	Reddit/WSB
8 Disagreement Variables	Overall disagreement in this study is one of eight measures for a given firm i on day t, including: allday, overnight, only submissions, submissions and related comments, as well as measures based on sentiment scores and the bullishness/bearishness of messages.	Reddit/WSB
Control Variables:		
Abnormal Trading volume on day t – 1	Abnormal trading volume on day t - 1 to account for persistence in abnormal trading volume.	CRSP
Short Sale Turnover on day t-1	To account for the persistence of short-selling behaviour, past turnover may influence or correlate with the turnover observed on the current day (t).	FINRA short sale
Past Short Sale Turnover	It includes past short turnover as an independent variable to account for the autocorrelation of short turnover (Diether et al., 2009). The past short turnover is the average short turnover from t-5 to t-1, calculated separately for institutions and retailers.	FINRA short sale
Volatility	The standard deviation of abnormal returns over days *t - 5* to *t - 1*.	CRSP
Cumulative Abnormal Returns (CAR)	Cumulative abnormal returns over days t - 5 to t – 1 and over days t - 30 to t – 6.	CRSP
Size	The size is to control for the size effect (Fama and French, 1993) is the log value of market capitalisation (shares outstanding into the closing price) of stock i at the end of day t.	CRSP
Illiquidity	Illiquidity, constructed based on Amihud (2002) as the **absolute daily return divided by dollar trading volume** (ILLIQ × 10^6), is included to control for the impact of liquidity on short sale turnover, as higher illiquidity can affect market behaviour and trading dynamics.	CRSP
RetS&P	Return on the Standard & Poor's Composite Index to control for overall market-wide effects, but is omitted because it is probably collinear with the fixed effects.	CRSP
Media Coverage	To account for firm-date-specific spikes in attention, an indicator variable is used. This dummy variable equals 1 if firm i was mentioned in the RavenPack database on day t.	RavenPack

VIII. References

Ajinkya, B. B., Atiase, R. K., & Gift, M. J. (1991). Volume of trading and the dispersion in financial analysts' earnings forecasts. *Accounting Review*, 389–401.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31-56.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*, *59*(3), 1259–1294.

Baig, A. S., Blau, B. M., Butt, H. A., & Yasin, A. (2023). Reprint of: Do retail traders destabilize financial markets? An investigation surrounding the COVID-19 pandemic. *Journal of Banking & Finance*, *147*, 106744. https://doi.org/10.1016/j.jbankfin.2022.106744

Banerjee, S., Davis, J., & Gondhi, N. (2018). When transparency improves, must prices reflect fundamentals better? *The Review of Financial Studies*, *31*(6), 2377–2414.

Banerjee, S., Davis, J., & Gondhi, N. (2024). Choosing to disagree: Endogenous dismissiveness and overconfidence in financial markets. *The Journal of Finance*, *79*(2), 1635–1695.

Banerjee, S., & Kremer, I. (2010). Disagreement and Learning: Dynamic Patterns of Trade. *The Journal of Finance*, *65*(4), 1269–1302. https://doi.org/10.1111/j.1540-6261.2010.01570.x

Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, *21*(2), 785–818. https://doi.org/10.1093/rfs/hhm079

Barber, B. M., Odean, T., & Zhu, N. (2009). Do Retail Trades Move Markets? *Review of Financial Studies*, *22*(1), 151–186. https://doi.org/10.1093/rfs/hhn035

Boehmer, E., Jones, C. M., & Zhang, X. (2008). Which Shorts Are Informed? *The Journal of Finance*, *63*(2), 491–527. https://doi.org/10.1111/j.1540-6261.2008.01324.x

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2021). Tracking Retail Investor Activity. *The Journal of Finance*, *76*(5), 2249–2305. https://doi.org/10.1111/jofi.13033

Boehmer, E., & Song, W. (2020). Smart Retail Traders, Short Sellers, and Stock Returns. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3723096

Boehmer, E., & Wu, J. (Julie). (2013). Short Selling and the Price Discovery Process. *Review of Financial Studies*, *26*(2), 287–322. https://doi.org/10.1093/rfs/hhs097

Bradley, D., Hanousek, J., Jame, R., & Xiao, Z. (2023). Place Your Bets? The Value of Investment Research on Reddit's Wallstreetbets. *The Review of Financial Studies*, hhad098. https://doi.org/10.1093/rfs/hhad098

Bradley, D., Hanousek Jr., J., Jame, R., & Xiao, Z. (2021). Place Your Bets? The Market Consequences of Investment Advice on Reddit's Wallstreetbets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3806065

Buz, T., & de Melo, G. (2021). Should you take investment advice from WallStreetBets? A data-driven approach. *arXiv Preprint arXiv:2105.02728*.

Chang, Y.-C., Hong, H. G., Tiedens, L., Wang, N., & Zhao, B. (2013). *Does Diversity Lead to Diverse Opinions?: Evidence from Languages and Stock Markets*. Univ., Rock Center for Corporate Governance.

Charles, C., Frydman, C., & Kilic, M. (2023). Insensitive investors. *Journal of Finance Forthcoming*.

Chen, J., Hong, H., & Stein, J. C. (2002). Breadth of ownership and stock returns. *Journal of Financial Economics*, *66*(2–3), 171–205. https://doi.org/10.1016/S0304-405X(02)00223-4

Christophe, S. E., Ferri, M. G., & Hsieh, J. (2010). Informed trading before analyst downgrades: Evidence from short sellers. *Journal of Financial Economics*, *95*(1), 85–106. https://doi.org/10.1016/j.jfineco.2009.02.004

Cookson, J. A., & Niessner, M. (2020). Why don't we agree? Evidence from a social network of investors. *The Journal of Finance*, *75*(1), 173–228.

Cookson, J. A., & Niessner, M. (2023). Investor Disagreement: Daily Measures from Social Media. *Available at SSRN 4529594*.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1989). The Size and Incidence of the Losses from Noise Trading. *The Journal of Finance*, *44*(3), 681–696. https://doi.org/10.1111/j.1540-6261.1989.tb04385.x

Dechow, P. (2001). Short-sellers, fundamental analysis, and stock returns. *Journal of Financial Economics*, *61*(1), 77–106. https://doi.org/10.1016/S0304-405X(01)00056-3

Diamond, D. W., & Verrecchia, R. E. (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics*, *18*(2), 277–311. https://doi.org/10.1016/0304-405X(87)90042-0

Diether, K. B., Lee, K.-H., & Werner, I. M. (2009). Short-Sale Strategies and Return Predictability. *Review of Financial Studies*, *22*(2), 575–607. https://doi.org/10.1093/rfs/hhn047

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, *57*(5), 2113–2141.

Eliner, L., & Kobilov, B. (2022). *To the moon or bust: Do retail investors profit from social media induced trading*. Working Paper.

Engelberg, J. E., Reed, A. V., & Ringgenberg, M. C. (2012). How are shorts informed? *Journal of Financial Economics*, *105*(2), 260–278. https://doi.org/10.1016/j.jfineco.2012.03.001

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fan, R., Talavera, O., & Tran, V. (2020). Social media, political uncertainty, and stock markets. *Review of Quantitative Finance and Accounting*, *55*(3), 1137–1153. https://doi.org/10.1007/s11156-020-00870-4

Fernando, S., Onishchenko, O., & Kuruppuarachchi, D. (2024). Do short sellers amplify extreme market declines? *Pacific-Basin Finance Journal*, *87*, 102498. https://doi.org/10.1016/j.pacfin.2024.102498

Fischer, P., Kim, C., & Zhou, F. (2022). Disagreement about fundamentals: Measurement and consequences. *Review of Accounting Studies*, *27*(4), 1423–1456.

Garfinkel, J. A., & Sokobin, J. (2006). Volume, Opinion Divergence, and Returns: A Study of Post-Earnings Announcement Drift. *Journal of Accounting Research*, *44*(1), 85–112. https://doi.org/10.1111/j.1475-679X.2006.00193.x

Giannini, R., Irvine, P., & Shu, T. (2018). Nonlocal disadvantage: An examination of social media sentiment. *The Review of Asset Pricing Studies*, *8*(2), 293–336.

Giannini, R., Irvine, P., & Shu, T. (2019). The convergence and divergence of investors' opinions around earnings news: Evidence from a social network. *Journal of Financial Markets*, *42*, 94–120. https://doi.org/10.1016/j.finmar.2018.12.003

Giglio, S., Maggiori, M., Stroebel, J., & Utkus, S. (2021). Five facts about beliefs and portfolios. *American Economic Review*, *111*(5), 1481–1522.

Harris, M., & Raviv, A. (1993). Differences of Opinion Make a Horse Race. *Review of Financial Studies*, *6*(3), 473–506. https://doi.org/10.1093/rfs/5.3.473

Hasso, T., Müller, D., Pelster, M., & Warkulat, S. (2022). Who participated in the GameStop frenzy? Evidence from brokerage accounts. *Finance Research Letters*, *45*, 102140.

Hirshleifer, J. (1977). The theory of speculation under alternative regimes of markets. *The Journal of Finance*, *32*(4), 975–999.

Hong, H., & Stein, J. C. (2003). Differences of Opinion, Short-Sales Constraints, and Market Crashes. *Review of Financial Studies*, *16*(2), 487–525. https://doi.org/10.1093/rfs/hhg006

Hong, H., & Stein, J. C. (2007). Disagreement and the Stock Market. *Journal of Economic Perspectives*, *21*(2), 109–128. https://doi.org/10.1257/jep.21.2.109

Hu, D., Jones, C. M., Zhang, V., & Zhang, X. (2021). The Rise of Reddit: How Social Media Affects Retail Investors and Short-sellers' Roles in Price Discovery. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3807655

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Kandel, E., & Pearson, N. D. (1995). Differential Interpretation of Public Signals and Trade in Speculative Markets. *Journal of Political Economy*, *103*(4), 831–872. https://doi.org/10.1086/262005

Kaniel, R., Liu, S., Saar, G., & Titman, S. (2012). Individual Investor Trading and Return Patterns around Earnings Announcements. *The Journal of Finance*, *67*(2), 639–680. https://doi.org/10.1111/j.1540-6261.2012.01727.x

Karpoff, J. M. (1986). A theory of trading volume. *The Journal of Finance*, *41*(5), 1069–1087.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, *22*(1), 109–126.

Karpoff, J. M., & Lou, X. (2010). Short Sellers and Financial Misconduct. *The Journal of Finance*, *65*(5), 1879–1913. https://doi.org/10.1111/j.1540-6261.2010.01597.x

Kelley, E. K., & Tetlock, P. C. (2013). How Wise Are Crowds? Insights from Retail Orders and Stock Returns. *The Journal of Finance*, *68*(3), 1229–1265. https://doi.org/10.1111/jofi.12028

Kot, H. W., Li, F. W., Liu, M., & Wei, K.-C. (John). (2020). Do Short Sellers Use Textual Information? Evidence from Annual Reports. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3683086

Li, T., Van Dalen, J., & Van Rees, P. J. (2018). More than just Noise? Examining the Information Content of Stock Microblogs on Financial Markets. *Journal of Information Technology*, *33*(1), 50–69. https://doi.org/10.1057/s41265-016-0034-2

Long, S. (Cheng), Lucey, B., Xie, Y., & Yarovaya, L. (2023). "I just like the stock": The role of Reddit sentiment in the GameStop share rally. *Financial Review*, *58*(1), 19–37. https://doi.org/10.1111/fire.12328

Lyócsa, Š., Baumöhl, E., & Výrost, T. (2022). YOLO trading: Riding with the herd during the GameStop episode. *Finance Research Letters*, *46*, 102359.

Milgrom, P., & Stokey, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory*, *26*(1), 17–27.

Miller, E. M. (1977). RISK, UNCERTAINTY, AND DIVERGENCE OF OPINION. *The Journal of Finance*, *32*(4), 1151–1168. https://doi.org/10.1111/j.1540-6261.1977.tb03317.x

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*, *78*(2), 277–309.

Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and Speculative Bubbles. *Journal of Political Economy*, *111*(6), 1183–1220. https://doi.org/10.1086/378531

Semenova, V., & Winkler, J. (2021). Social contagion and asset prices: Reddit's self-organised bull runs. *arXiv Preprint arXiv:2104.01847*.

Shiller, R. J., & Pound, J. (1989). Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior & Organization*, *12*(1), 47–66.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and Trades: The Information Content of Stock Microblogs. *European Financial Management*, *20*(5), 926–957. https://doi.org/10.1111/j.1468-036X.2013.12007.x

Srinivasan, K. (2023). *Paying Attention*. Technical Report 2023. mimeo, University of Chicago.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Varian, H. R. (1985). Divergence of Opinion in Complete Markets: A Note. *The Journal of Finance*, *40*(1), 309–317. https://doi.org/10.1111/j.1540-6261.1985.tb04951.x

Wang, T., Lu, K., Chow, K. P., & Zhu, Q. (2020). COVID-19 sensing: Negative sentiment analysis on social media in China via BERT model. *Ieee Access*, *8*, 138162–138169.

Welch, I. (2022). The wisdom of the Robinhood crowd. *The Journal of Finance*, *77*(3), 1489–1527.