Social Informedness and Investor Sentiment in the GameStop Short Squeeze

Kwansoo, Kim; Lee, Sang-Yong Tom; Kauffman, Robert J.

Document Version
Accepted author manuscript

Published in:
Electronic Markets

DOI:
10.1007/s12525-023-00632-9

Publication date:
2023

License
Unspecified

Citation for published version (APA):
Kwansoo, K., Lee, S.-Y. T., & Kauffman, R. J. (2023). Social Informedness and Investor Sentiment in the GameStop Short Squeeze. *Electronic Markets*, *33*(1), Article 23. https://doi.org/10.1007/s12525-023-00632-9

Electronic Markets – The International Journal on Networked Business

Full title of article:	**Social Informedness and Investor Sentiment in the GameStop Short Squeeze**
Subtitle (optional):	None
Preferred abbreviated title for running head (max. 65 characters including spaces)	Social Informedness and Investor Sentiment
Keywords (for indexing and abstract services – up to 6 words):	Collective behavior, informedness theory, investor sentiment, irrational trading, short squeeze, social informedness
JEL classification	JEL: D8, D82, G14, G41, L86
Word count	15776 words in this file
Word processing program name and version number:	Word Microsoft 365 MSO 2022

Abstract

We examine investor behavior on social media platforms related to the GameStop (GME) short squeeze in early 2021. Individual investors stimulated the stock market via Reddit social posts in the presence of institutional investors who bet against GME's success as short sellers. We analyzed r/WallStreetBets subreddit posts related to GME's trading patterns. We performed text-based sentiment analysis and compared the social informedness of posting users for GME trading on two social media platforms. The short squeeze occurred due to coordinated trading by individual investors, who discussed trading strategies on the platforms and drove collective social informedness-based trading behavior. Our findings suggest that the valence and number of submissions influenced GME's intraday transaction volumes and precursors for irrational trading behavior patterns to have emerged. We provide a theoretical interpretation of what occurred and call for tighter monitoring of social news platforms. We also encourage effort to create an in-depth understanding of the observed patterns and the linkages between them and the larger equity markets. (163 words)

Introduction

At the beginning of 2021, an event occurred in the United States stock market that has drawn attention from observers around the world. Individual investors used to be viewed as a weak force who competed for influence with big players in the market, such as institutional investors and hedge funds (Krantz, 2021). The former group demonstrated collective behavior by exchanging stock trading-related information through social media (via Reddit's r/WallStreetBets, or r/WSB for short), and invested heavily in specific stocks. By coordinating their stock purchases through social media, they created a sharp rise on the New York Stock Exchange Euronext (NYSE Euronext) in GameStop (GME) stock's price with no company fundamentals or newly released information to support higher valuation.

This led to a situation known as a *short squeeze* that affected professional investors.[1] The short-sellers' problem was that GME shares exhibited a 1,700% price increase, reaching US$347.51 per share on January 21, 2021 – up from US$17.25 several weeks earlier on January 4 (*CNBC*, 2021). The price rise was driven by numerous individual investors who had purchased the stock, resulting in an all-time high intraday price for GME at US$483 on January 28, 2021 (Allen et al., 2022). This collapsed to less than US$69 by early February 2021, though the damage to short sellers was done (Angel, 2021).[2]

A decline in stock markets worldwide occurred due to the COVID-19 pandemic and related economic uncertainty. The U.S. market subsequently experienced a shock related to herd behavior induced by a social media platform. It seems to have had its largest effects in early 2021 but persisted even after mid-year (Lyócsa et al., 2021). Some stock prices appear to have gone up and down as many individual day-traders shared information about stocks through social media and bought or sold them in coordinated acts of collective behavior.[3] This caused the stock price to spike. Individual investors concentrated on the subreddit, r/WSB, to initiate a short squeeze of institutional investors who were betting on the decline of several underrated stocks. Beyond GME, they focused on AMC Entertainment (AMC), BlackBerry (BB), Nokia (NOK), and a few others. Meanwhile, hedge fund investors paid attention to the sharp rise in stock prices and individual investors began to focus less on corporate fundamentals and to actively deny evidence of the lower *intrinsic value of traded firms*. They also were willing to be counterparties in short-sales contracts. This is because institutional investors assumed that the

[1] A *short squeeze* occurs with "*rapidly rising prices in a stock or other tradable [security, when] many short sellers [hold] positions in it. The short squeeze begins when the price jumps higher unexpectedly. The condition plays out as a significant measure of the short sellers coincidentally decide to cut losses and exit their positions*" (Mitchell, 2022).
[2] The price of GME stock followed a similar trajectory to Phunware on the NASDAQ Exchange, which went from US$10 in December 2018 to more than US$300 in February 2019 (McEnery, 2021).
[3] A *day-trader* is a type of trader who executes a relatively large volume of short and long trades to capitalize on intraday market price movements to profit from very short-term price volatility. Day traders also use leverage to amplify their returns, but this can result in heightened losses (Chen, 2022).

prices of specific stocks would fall in the future. Meanwhile though, individual investors coordinated their collective effort to move prices up. Some hedge funds, as a result, suffered large losses in the process, and individual investors made large profits – though this did not persist for long.

Until recently, individual investors have been recognized as only having dispersed power and relatively little capital compared to institutional investors. Thus, it has been reasoned, they cannot have much influence on the market value of stocks. The short squeeze incident overturned this notion, however. Individual investors, the market learned, were able to collectively gather their opinions and act as a group by exchanging information on social media platforms and exert a powerful influence on the market. Their collective effort was accompanied by an increase in the stock price of GME from US$5 in July 2020 to US$10 in October 2020, before larger value jumps were observed in January 2021. Short sellers booked significant losses. A "David versus Goliath" narrative on individuals versus institutional investors became a key story line for the market dynamics of the short squeeze.

Moreover, the popularity and use of social media platforms has increased tremendously (See-To & Yang, 2017). They generally are used to share opinions and information, but their presence has been exploited to spread false news and bogus stories that influence stock market valuation. The platforms have become sources for news-seekers to obtain updates, track comments, and sample expert opinions about events to understand socioeconomic issues. Differences across the platforms may permit them to be exploited differently though.

In Reddit, users create and follow several subreddits. A subreddit is an interest-based group or forum where members post and discuss specific topics of interest. Twitter, in contrast, supports its users to follow others based on common topics between users, including personal or professional relationships. Another difference is that Twitter limits tweets to a maximum of 280 characters. Reddit comments have no length limit though. Reddit users also can share additional analysis of the posted content. Thus, users exhibit different behavioral aspects in their posting patterns that reflect the characteristics of each platform. Reddit and Twitter have differences in their characteristics, so the social connections among users across these platforms is affected (Priya et al., 2019).

The global spread of social media connects people to others, and many enjoy having conversations on common topics without being limited by space or time. This suggests that social networks can influence their thinking and collective behavior. People may see only the information they want on social media and exclude unwanted messages, however. Because social media emphasizes communication with peers, active users can fall into the *trap of infallibility*, thinking that the information around them is unconditionally correct. When a political or socio-economic purpose is suggested in the content of a message, there is a high risk of groupthink

due to information distortion and group member conformity with others. Some studies have identified related *situational echo chamber effects* and *predatory trading* such that social media users tend to respond the most to information and views that are alike and reinforce their own (Brunnermeier & Pedersen, 2005; Pedersen, 2021a). They also may have concerns about causing social fragmentation among their group members, while avoiding views that contradict their own by segmenting members of different groups with filter bubble search algorithms that deliver more palatable information and news based on their preferences (Gray et al., 2019).

An interesting trend in the stock market has been observed in recent years. Online stock trading platforms such as Robinhood (robinhood.com) have created new forces to prompt greater democratization, by lowering the fees and minimum balances the firm requires for account holders. This has made it possible to trade fractional shares economically and not only in large lots, while gamifying the investment process. This has been attractive to young investors, who want to trade stocks but lack the money needed to do this via the more expensive, full-service brokers online. The rapid influx into the stock market of those born from the 1980s to the 2010s with the emergence of free trading apps has brought the collective power of individual investors to the fore. With the recent development of mobile platforms and trading apps, stock investing is no longer the domain of the older and wealthier generation. Instead, we observe new technological support for participation of young people in the market (Boorstin, 2021).

The collective decisions of younger individual investors – in parallel with mobile phone and social media platform usage – appear to have only begun to impact the market. Eaton et al. (2021) exploited data on platform outages for brokers used by many GME traders on Robinhood to identify their causal effects on financial markets. Their research showed that exogenous negative shocks from mobile trading platform participation led to lower return volatility among stocks favored by the firm's investors.

Our research questions (RQs) in this work are: (1) What patterns of trading did individual investors demonstrate in the presence of social media platforms such as Reddit and Twitter during the GME short squeeze? (2) Did social sentiment trigger the collective trading behavior of individual investors? And (3) what can we learn from their observed behavior that suggests the basis for a new theoretical interpretation based on a deeper understanding of social media's information effects that occur in social trading?

Prior studies have classified individual investors' opinions and preferences and verified that their sentiment can influence their market behavior (Allen et al., 2022; Caron et al., 2021; Hu et al., 2021; Pedersen 2021b; Vasileiou, 2021). Such findings regarding social media, individual investors, and short sellers sought to explain whether the sentiment aroused by social platforms can immediately affect their decisions. None demonstrated

how such sentiment formed on social platforms (e.g., Reddit vs. Twitter) though. Instead, they focused on discussing topics related to investors' aggregate trading activities and transaction patterns. Nor did they report what kinds of decisions were observed when investors had a good understanding of the fundamentals that reflect how market mechanisms work. Our study is novel in this respect since it provides insight into a new dimension of social media activities: *collective social informedness.* We report on how it can become a dominant market force.

Our study further shows that the collective behavior of individual investors as *socially-informed traders* was triggered by social media. We also learned that social media platform communication effects reinforce individual investors' propensity to use new sources of social media information that may result in information-driven irrationality in stock market trading. This is likely to diminish investor trust as a primary tenet of high-quality financial market mechanisms: they must be reliable in terms of how they can effectively process newly-discovered information, support trustworthy valuation, and make market prices reliable.

The next section presents literature and background theory related to economic exchange, the concept of social informedness, the emergence and relevance of narrative economics, and prior studies on collective investor behavior. We also offer our perspective on investor sentiment and trading momentum in the GME short squeeze. Thereafter, we establish proposed hypotheses to represent the theoretical perspective of this study and present the research methodology used in our data collection and quantitative empirical approach. We then report on the estimation results for our models and test results for the hypotheses and discuss the relationships we discovered in our hypothesis test findings and some broader issues that warrant comment. We conclude with the new knowledge contributions of this research, its scientific limitations, and future research.

Literature and Theory to Understand the Short Squeeze

We next discuss the theoretical background for understanding the GME short squeeze, how collective behavior can be triggered by social sentiment, how sentiment effects may lead to momentum in trade prices and market followers, and how momentum and investor narratives create the basis for new behavior patterns in trading.

Our literature review identifies the key terms and definitions in its text, and the theoretical challenges of investigating collective trading behavior of individual investors based on social sentiment throughout the section. Investors seem to have behaved randomly and irrationally sometimes since they could not conduct systematic analyses. They were influenced by external factors and public discussion. Their decisions were often made on their own or with media information. And, when they traded collectively, their behavior influenced the stock

market. Thus, we sought to understand their individual informedness. We also tried to verify their collective behavior. So, we assessed the informedness of related economic agents involved in social media. We also levered investor sentiment and momentum that was developed and formed through their collective communication. This is rooted in narrative economics for the collective behavior of individual investors (Shiller 2013, 2020).

Social informedness and the GME short squeeze

The informedness of economic agents has been examined in the *information systems* (IS) literature. Smith et al. (2011), for example, defined *being informed* about a product as a consumer's perceived awareness of the consequences of using, consuming, or owning it, based on interacting with product-related data. We further define *social informedness* to indicate that investors perceive they are informed about a firm's market value by collecting or receiving information through social media. This pertains to members of a social subgroup in financial market activities such that they are in the state of being collectively informed about asset prices and what factors may explain them. Li et al. (2014) and Hoang and Kauffman (2018) have referred more broadly to this related body of knowledge as the emerging *theory of informedness*.

The theory suggests that individual investors can collectively tap into social sentiment from social media platforms to gather information on financial investments to make. Informedness through social media creates the perception of being informed because investors may think they are more aware of stock market developments than others are – regardless of whether the acquired information is new or true. Reddit has proven to be most popular in this respect in recent years, while Twitter is more suitable for the spread of updated information on emergency and disaster situations (Ruan et al., 2019). If social media interactions on the platforms are different, trader social informedness from them may be different too. Hence, there may be differential impacts of the two platforms on investors' behavior.

Recent studies have analyzed the GME short squeeze. Pedersen (2021b) and Allen et al. (2022) explored the mechanisms at work and the impediments to stock market price discovery from a financial theory perspective. Vasileiou (2021) showed that a short squeeze leads to a market anomaly and an *anti-leverage effect* that violates the efficient market hypothesis.[4] Angel (2021) also proposed that short squeezes may reveal weaknesses in the U.S. equity market. He presented a model that unified the views of naïve, fanatical, and rational short-term investors, as well as long-term investors in a social network. They support study of observed phenomena related to the GME short-squeeze events. Long et al. (2021) found that comments expressing observers' and investors' sentiment from r/WSB also influenced GME's intraday returns. And Lyócsa et al. (2021) utilized the activity

[4] The *leverage effect* is frequently used to describe the increasing volatility when stock prices fall. In contrast, the *anti-leverage effect* is used to describe increasing volatility that occurs as stock prices rise (Vasileiou, 2021).

level on r/WSB and related searches on Google to explain variations in GME's price.

A unique feature of the GME short squeeze was the role of investor sentiment and social media in attracting attention. Caron et al. (2021) showed that user activity on r/WSB was associated with the trading volume of GME stock. The effect of emotions on investment decisions also is well known. The social media revolution since the 1990s has enabled investors to express sentiment on different social media platforms at an unprecedented level.

Narrative economics and collective behavior

Tversky and Kahneman (1981) introduced human-thinking systems based on *dual process theory* – how the same thought can arise due to different processes or stimuli. One is an immediate and faster approach; the other is a logical and effort-intensive slower approach. The GME event is an example of what triggered faster thinking for a lot of people. There are many instances in which investors make quick decisions about large sums of money – even when fortunes at stake. Such decisions exemplify judgment that is influenced by the sentiment of others with little deliberation or logic. Individual investors sometimes also act irrationally, stemming from their anxiety, level of confidence, and illusion – rather than rationality, as well as based on intuition and emotion. A key aspect of their behavioral impulses in GME trading was the "story behind the story" – the GME narrative.

Shiller (2013, 2020), a proponent of *narrative economics*, has argued that the act of investing in speculative assets is not simply a decision based on individual psychology but instead is a *social activity*.[5] Individual traders tend to invest based on information they receive from others, and this tendency is recognized by investment professionals, institutional investors, and fund managers. Shiller further suggested that conveying stories that have elements of fact or fiction often results in ambiguous meanings and equivocal interpretations – as may be intended by those wishing to manipulate social responses.

When such stories morph into financial market narratives through viral transmission in a population, individual stock prices tied to them may not exhibit the typical adjustments observers expect to see based on firm fundamentals. Instead, asset prices can be pushed to extremes by the power of the related story. Social media makes it easy for fan groups to not only shape the narratives, but also to actively disseminate them. This gives them enormous influence in the market. Such narrative-driven momentum combined with a stock trading platform (e.g., Robinhood's influence on GME) can shock share prices, pushing them to irrationally high levels, and social media-coordinated buying signals are believed to have affected GME prices and trading volumes.

[5] A *narrative* is a story that is told in a way to reflect and inspire a particular point of view or value. Unlike traditional economics that emphasizes data, *narrative economics* emerged to analyze the economic impact of the contagious spread of popular narratives and coming up with solutions that address their potential impact.

Collective behavior through investor sentiment and momentum

Related to the collective communication in r/WSB, Hirshleifer (2020) argued that the sharing of the human dramas was weak at first. People were likely to speak with bias about their tearful experiences of being evicted from their homes due to a financial crisis created by speculators. However, over time, *social transmission bias* grew became self-reinforcing, accumulating and amplifying the original impact.[6] Collective behavior-driven rational bubbles, as viewed by Surowiecki (2004), has produced bad judgment by investors. All along, the *wisdom of crowds* has been an unconscious aggregation of the opinions of many, such that the participants begin to emulate each other and conform rather than think independently as individuals. The author's experimental study indicated that crowds are collectively swayed by persuasive arguments, so groups of people may be observed to conform. This leads to a system for making decisions that exhibits systemic problems that are hard to remedy.

Hong and Stein (1999) proposed establishing a conceptual link between sentiment and momentum in stock trading and the related prices. In their view, news diffuses slowly through the actions of different observers who are *news-watchers.* They react to news they see or hear when they trade, which creates market momentum. Investor momentum trading may reflect their perceptions of value based on its effects on economic fundamentals, which loops back to influence investors' valuation perceptions.

Barber et al. (2022) examined herding behavior by Robinhood users in mid-2021. Their findings are consistent with Kelley and Tetlock (2013): individual investors used a profitable momentum-based buying strategy when they were informed, or a contrarian-based buying strategy that yielded liquidity when the market declined. Antoniou et al. (2013) also wrote that Hong and Stein's news-watchers underreacted more when they received information that contradicted their sentiment due to cognitive dissonance. This implies that bad news among less attractive stocks diffuses more slowly when sentiment is optimistic. This led to trading momentum, which is driven by a less well-performing stock portfolio in periods of optimistic sentiment and by a more well-performing portfolio in periods of pessimistic sentiment. Although this argument predicts *symmetric momentum* across periods, this may be more pronounced when sentiment is optimistic because arbitraging cognitive dissonance requires costly short-selling of shares for a lower-performing stock portfolio.

Asness et al. (2013) argued that social network effects can lead to price momentum and subsequent rever-

[6] Related to the GME short squeeze, the Reddit and Twitter discussion forums' participants appear to have viewed the events as comprising a "heroic attack" by small investors on Wall Street's leading financial hedge-fund capitalists – if only for a limited time. This aspect was more firmly planted in small investors' memories than their failed GME investment experiences, as its stock price crashed. Nevertheless, this was an unusual group-wide event: a manifestation of *collective behavior*.

sals toward fundamentals – essentially a *value effect*.[7] They are seen across asset classes and in global markets. The build-up of naïve individual investors' demand also can lead to momentum effects and price reversal, as fundamental value is revealed and leads to perceptions of a value premium. Pedersen (2021b) asserted that it would be worthwhile to link these patterns to behavioral patterns in social networks when stock trading is involved.

Hypotheses Development

We assessed the background theory for this research in two ways. One was our effort in the case study of GME to lever relevant perspectives to identify generalizable explanations for what we observed, based on past case study work (Eisenhardt, 1989; Eisenhardt and Graebner, 2007) and research knowledge related to financial markets, stock performance, and social sentiment. A second effort was related to the findings from our dataset on social postings and financial history for this study. Our data analytics let us understand investor sentiment for forming social informedness, shaping collective behavior, and triggering trading momentum. Our study uses Reddit and Twitter data to analyze investor sentiment (Bollen et al. (2011); and Caron et al. (2021)). We further discuss this study's theoretical foundations with reference to several themes prior to presenting our main hypotheses.

Investment sentiment and collective trading behavior. We investigated whether coordinated trading by individual investors, who discussed their trading strategies on social platforms, drove collective behavior that was based on their informedness. We leveraged theories of financial behavior in electronic markets (Alt & Klein, 2011; See-To & Yang, 2017) by measuring and comparing investor sentiment on two social platforms – Reddit vs. Twitter. This enabled us to evaluate the impact of social informedness linked to investor sentiment. We also assessed how the prices of stock investments reacted to information on specific topics discussed on social media platforms. We will posit hypotheses to test the relationship between investor sentiment and collective trading behavior on different platforms beyond what others have done, as we noted earlier.

The theory of informedness. *Informedness* in market settings arises due to the low costs of sharing information on platforms for use in all kinds of e-markets, and irrespective of whether the information is true or false news (Clarke et al. 2022; Pedersen 2021b). The nature of e-markets makes them more likely to operate on an open access basis with increasingly intelligent algorithmic identification of what information to share with whom in response to what events and when. This way, their role has been transformed and their relationship

[7] This is the excess return of a portfolio of *value stocks* (of low market value relative to their fundamentals), earned by a portfolio of *growth stocks* (with a high market value relative to their fundamentals).

with platform and e-market participants is no longer the same as two decades ago (Clemons et al. 2017; Gomber et al., 2018).

The role of e-market innovations. Alt and Klein (2011) indicated two origins of electronic markets: the technological side with its roots in *electronic data interchange* (EDI) and the economic side which analyzes the cost of economic exchange activity between agents. Otto et al. (2011) demonstrated that e-markets represent a form of business networking and information sharing through value networks. The relevance of information and the quality of data in networking have grown, as social networks and information goods have proliferated.

Additionally, Bitter and Grabner-Kräuter (2016) emphasized that information that builds strong ties is perceived as highly diagnostic and influential in the context of user interaction on social media platforms. Sharing product information has become integral in today's online social networking world as a result. Customer engagement behavior in online social networks supports the understanding of how online social connections impact decision-making as well. This is in line with social capital theory, which considers strong ties to represent trusted relationships that provide in-depth information that agents can rely upon (Gubbins & McCurtain, 2008).

From e-markets to e-platforms. Alt and Zimmermann (2018) wrote that electronic platforms for collaboration, publishing, and distribution often go hand in hand with the availability of modularized information content. *Modularization* is the process of separating content into discrete topics and storing them for reuse as modules in a content management system or database (Haramundanis 2009). The power of modularization and disaggregation are well known in e-commerce. The impacts have been observed in the transformation of value creation structures due to reduced transaction costs, resulting in disintermediation and reintermediation effects and the emergence of new actors, networks, and ecosystems (Chircu et al., 1999; Sarkar et al., 1995).

Hein et al (2020) argued that *digital platform ecosystems* have more recently been built on modular architectures that enable platform owners to implement governance mechanisms to facilitate value-creating mechanisms between a platform owner and an ecosystem of autonomous technology and service providers as *complementors* and consumers. They suggested that a platform owner can provide affordances that complementors can create value with based on their individual innovation capabilities. Complementors also can interact with each other to leverage the generativity of the digital platform ecosystem for the co-creation of additional innovations.

Institutional investors are known to monitor stocks through financial reports and portfolio risk management intermediaries who are connected to the digital ecosystem of the financial market. In contrast, individual investors may seek to sense the "mood of the market" by coordinating with one another via discussions on social media. Stocks with a cult following involve investors who are sensitive to other people's stories, especially those

they encounter over the Internet. Sometimes, an extreme event may not correspond to any shock to stock's fundamentals buy still may impact its price and trading volume. Costola et al. (2021) investigated the momentum when synchronized buying signals that originate on social media influence a stock's price and trading volume. They revealed that the coordinated efforts of individual investors on social media may allow them to act as a single large trader able to manipulate prices. To proxy for coordinated activity on social media for our focal firm (GME) and its matching firm's stock (BBY), we will create a daily-count time-series based on posts referring to the stocks.

Collective behavior on market transactions. Hu et al. (2021) found that trading through social news platforms creates social attention that contains information for predicting stock returns and increases trading volumes in the traded stocks. Pedersen (2021b) showed that fake news distributed through social media can lead to disagreement over time. Interestingly though, such disagreement can generate a turnover spike in trading, high volatility, and price momentum as a bubble grows. Allen et al. (2022) indicated that greater market participation increases during speculation, as investors with limited expertise are attracted by market events. They showed that short squeezes affect the market quality of the impacted stocks. To assess how market activities changed during the squeeze, we will examine the effects of social media activities on stock market transactions. So, we posit:

- **Hypotheses 1a and 1b (Individual Investors' Collective Behavior for Market Transactions).** There are two sub-hypotheses asserted:
  (1a) *The number of social media posts is correlated with a firm's trading volume during a short squeeze.*
  (1b) *The impact of the number of a firm's stock trades is stronger during than before or after a short squeeze.*

Social sentiment and momentum. De Long et al. (1990) explained why institutional investors may incur financial losses in the short run if investor sentiment turns extreme and prices move away from their fundamental levels. In other settings, individual investors may underestimate how mispricing becomes more pronounced over time, leading to a positive correlation between investor sentiment and short-term returns. Banerjee (1992) showed that individual investors do not weight their own information properly though and instead weight information from existing market participants too much. If herd behavior occurs, it implies that trend-chasing activities are present. This widens sentiment-induced mispricing in subsequent periods. Trend-chasing behavior suggests a positive effect of sentiment on future returns. Han and Li (2017) demonstrated that misperception about future market movements may become more pronounced over time. This suggests there will be momentum-driven predictability of investor sentiment for subsequent returns should be positive. So, we further assert:

- **Hypothesis 2 (Momentum-Based Prediction for Investor Sentiment).** *Investor sentiment metrics are*

correlated with the focal firm's stock returns during a short squeeze.

Social media platform effects. When a public narrative becomes viral and has market consequences, it can become an economic narrative. By modeling the underlying relationships, economic events can be predicted and prepared for in advance. Advanced technologies such as social media have the potential to change the way narratives are transmitted. As a result, on social media platforms, they can act as a disruptive factor in analyzing and understanding short squeeze-related economic phenomena. They also can act as a channel for information dissemination, which may benefit small firms with less analyst coverage (Al Guindy, 2021).

As noted earlier, Reddit and Twitter have differences in their characteristics and key functionality, which may affect the way and degree that users share information. Long et al. (2021) examined Reddit's discussion forum, with its intraday changes of sentiment valences and prices, and linkages between investor sentiment and GME stock that are often bidirectional. They showed that Reddit sentiment varies over time and that active discussions on r/WSB that occur during trading hours may lead to stronger social media effects on GME's shares. Umar et al. (2021) used Twitter post counts as a proxy for social sentiment in the short squeeze to understand the role of Reddit's amateur investors. The Twitter count was not an effective proxy for Reddit sentiment posts though. Its comments were organized in threads with titles, and some threads became more popular and received higher scores than others from forum participants. But the number of comments in a thread could proxy for the popularity of a discussion among participants, while the comment scores only indicate the popularity of specific comments.

Priya et al. (2019) revealed that, depending on the information requirements, bursty traffic in various stages of a thread's development made Reddit suitable for providing an overview of opinions in a short span of time. Reddit's structure makes it more suitable for finding experts' opinions. Further, its unconstrained post length is important since posters can share more background information. In contrast, the shorter post length reduces the occurrence of biased and extreme views on Twitter compared to Reddit. In addition, an event on Twitter is viewable for a longer time, which is useful for analysis of evolving sentiment streams. Shorter inter-arrival times also make Twitter suitable for frequent updates in an emergency or live event. Thus, we could compare sentiment and comments streamed on the r/WSB forum with those from Twitter. This leads us to offer a third two-part hypothesis:

- **Hypotheses 3a and 3b (Differential Social Informedness by Platform).** *We test two assertions*:

  (3a) *The impacts of the number of posts on a focal firm's trading volume are different based on whether an interest-based discussion platform versus a common-topics discussion platform during the short squeeze was used for posting social sentiment.*

  (3b) *The impacts of social sentiment valences on a focal firm's trading volume are different on whether*

an interest-based discussion platform versus a common-topics discussion platform during the short squeeze was used for posting social sentiment.

Collective behavior between the focal and matching firm. GME has become a prominent instance of retail investment activism and social discussion-driven collective behavior. Many investors dismissed the short squeeze as a one-time event in early January 2021. However, its recurrence in May and June that year forced observers to take a deeper look at the mechanism behind short squeezes to make sense out of Reddit-driven collective behavior. The uncertainty surrounding GME stock made it difficult for longer-term investors to take a position in the stock, for example. GME's valuation then was based on the implementation of a pricing strategy involving a strategic adjustment to an online business model. This case illustrates that emotion-based social media sentiment campaigns involving individual investors have the potential to influence stock prices. The increasing role of social sentiment platforms also has led to new literature that focuses on the impact of media-driven sentiment on investment decisions (Tetlock, 2007; Duz Tan & Tas, 2020). Individual investors' use of information obtained from social media for decision-making has become increasingly important so they can be more informed.

Compared to GME, stocks that were not the target of the short squeeze were less likely to be discussed on social media. These relatively low-profile stocks might have been viewed as a better buy than GME from a fundamental valuation perspective. But individual investors who followed collective trading behavior based on social media sentiment would tend to be less interested in such stocks than GME. Thus, comparing aggregate individual investor trading of GME and a non-short squeeze firm on similar social media platforms made it possible to assess how investors differed in their views of the stocks. We chose Best Buy (BBY), a U.S. e-gaming product seller, as a matching firm. It represented a stock that was not unduly affected by the GME short squeeze. Thus, we posit:

- **Hypothesis 4 (Sentiment Impact for Focal and Matching Firms in the Short Squeeze).** S*entiment derived from a social media platform has a stronger impact on a focal firm's than its matching firm's trading volume during a short squeeze of its stock.*

Table 1 presents theoretical background in six different topical buckets that guided our research and notes the four hypotheses we formulated and tested in the research.

Theory Aspects	**Literature**
- Social informedness	- Smith et al. (2011); Li et al. (2014); Hoang and Kauffman (2018)
- Theory of informedness	- Clemons et al. (2017); Gomber et al. (2018); Li et al. (2014)
- GME short squeeze	- Allen et al. (2022); Caron et al. (2021); Pedersen (2021b); Vasileiou (2021)
- Narrative economics and collective behavior	- Tversky and Kahneman (1981); Shiller (2013, 2020)
- Collective behavior via sentiment and momentum	- Hirshleifer (2020); Hong and Stein (1999); Barber et al. (2022); Kelley and Tetlock (2013); Antoniou et al. (2013)

- Role of e-market innovation	- Gubbins and MacCurtain (2008); Alt and Klein (2011); Otto et al. (2011); Bitter and Grabner-Kräuter (2016)
- From e-markets to platforms	- Sarkar et al. (1995); Haramundanis (2009); Alt and Zimmermann (2018); Hein et al. (2020)
Hypotheses	
- H1: Collaborative behavior	- Individual Investors' Collective Behavior for Market Transactions
- H2: Investor sentiment	- Momentum-Based Prediction for Investor Sentiment
- H3: Social informedness	- Differential Social Informedness by Platform: an interest-based discussion platform versus a common-topics discussion platform
- H4: Social sentiment	- Sentiment Impact for a focal firm and a matching firm in the short squeeze

Table 1. Overview: Theoretical aspects, literature, and hypotheses for theories leveraged in this research

Data and Empirical Methods

Through our earlier literature review, we identified the role of investor sentiment for creating social informedness, shaping collective trading behavior, and triggering trading momentum. By leveraging big data and advances in text mining techniques, the difficulty of collecting data and performing an accurate measurement has been reduced, although some fine-tuning was still necessary to achieve our research and data analytics purposes.

A prior study reported that Twitter contains information about the stock market, but our study mainly uses both Reddit and Twitter data to analyze investor sentiment (Bollen et al. 2011). Reddit makes it possible to obtain data for separate submissions and comments, which provide more elaborate information on individual investors to support the data analytics (Caron et al. 2021). We further compared the impact of investor sentiment in the two social media platforms related to the market behavior of GME and BBY in the online marketing domain. This enabled us to perform two regressions in a combined platform-level analysis with *seemingly unrelated regression* (SUR). This captures the information structure of the analysis setting and treats the problem of cross-correlation between the residuals for each platform's individual estimation model. We next present an explanation of our data collection and the related social sentiment analysis.

Primary data sources. For the GME dataset, we used Reddit and Twitter. For our empirical study, we brought together a dataset comprised of user posts, spanning the period from January 1, 2020 to June 31, 2021.

Collection of Reddit data. We performed the following steps to collect the data from Reddit.

(1) ***Step 1: GME submissions, Reddit posts, and GME comments.*** Since Reddit's PRAW API allows requests of up to only 100 items at once, we collected GME-related submission data from an online archive of historical posts (PushShift.io, 2022). We downloaded and decompressed the related archive in

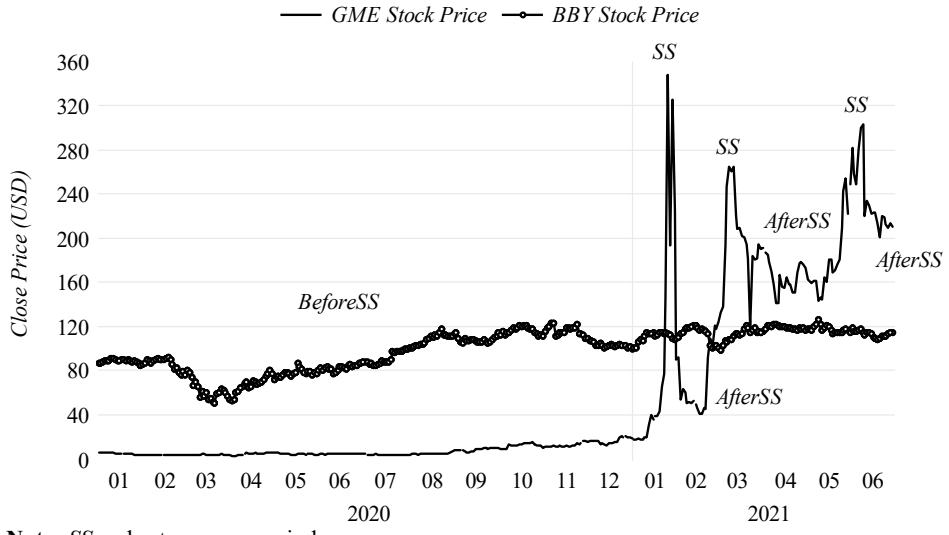
the .zst file format.[8] We obtained post observations from r/WSB for those which containing at least one of these keywords: "gme," "gamestop," or "gamestonk." For the collection of comments from data on each GME submission page, we used Python's BeautifulSoup library. We also created a web crawler to gather comments and collected submissions. (See Appendix Figure A1.)

(2) ***Step 2 – Data cleaning and preprocessing.*** For this, we first used Python's Selenium library and deleted the Reddit bot's comments. (See Appendix Figure A2 again.) Then, we went through data preprocessing to perform lowercase conversion, tokenization, punctuation removal, stop word elimination, lemmatization, and so on. (See Figure A3, and Tables A1a-A1b for our data preprocessing methods terms and code.)

(3) ***Step 3 – Collection of Twitter data.*** These were obtained using the *SNScrape* open library from Github (2022). We included BBY as a *baseline matching firm* to represent a similar stock that was not unduly affected by the GME short squeeze. To address the concern of repeated ads in Twitter, we removed all tweets for a user ID that included any word such as bitcoin, bot, or crypto. We also removed tweets in languages other than English and did the same preprocessing sequence for Reddit.

(4) ***Step 4 – Final Reddit and Twitter dataset.*** The final Reddit dataset that resulted contained 2,280 submissions and 294,377 comments about GME, while the Twitter dataset had 2,025,592 tweets about GME and 1,236,888 tweets on BBY. Sentiment analysis on Reddit and Twitter data was done through a pre-trained *TweetEval* model from GitHub (https://github.com/cardiffnlp/tweeteval), as in Caron et al. (2021). It reports the daily proportion of neutral, negative, and positive sentiments. We provide a graphical example of the daily sentiment analysis results. We added historical data on daily closing prices, trading volumes, and returns for the two stocks from Yahoo! Finance for 376 business days between January 2020 and June 2021. (See Figure 1 and Appendix Table A2.)

[8] We downloaded the archived file through the Pushshift Reddit API and used Linux Ubuntu to decompress it.



Note: *SS* = short-squeeze period

Figure 1. Daily closing prices for GME and BBY stocks, January 2, 2020 to June 29, 2021

Investor sentiment mined from social media platforms provides insights on stocks, cryptocurrencies, and commodities that investors may not be able to acquire through traditional sources. Sentiment analytics are used to make sense of data-at-scale from social media all over the world now. Moreover, the *sentiment score* metric is popularly measured for social network communications by comparing positive and negative sentiment over time about a particular subject to which it applies. Investor sentiment has been recognized to have predictive power for stock price movement forecasting and for retrospectively fitting price and other curves. For instance, Bollen et al. (2011) used Twitter to measure public mood to forecast the daily movement of the Dow Jones Industrial Average (DJIA) and reported noteworthy accuracy. Similarly, Bing et al. (2014) used Twitter sentiment to forecast stock prices for 30 NYSE Euronext and NASDAQ firms, and were able to achieve an approximate accuracy of 76%.

We distinguished three stages as separate periods: prior to it (*BeforeSS*), the short-squeeze itself (*SS*), and after it (*AfterSS*). Since there were three short-squeeze periods, we broke the actual dates into seven sub-periods. (See Table 2.) We next describe the variables: Vol_{it} represents one of the dependent variables, the logged daily trading volume, and $Returns_{it}$, the daily stock returns. Also, i is an index for focal or matching firm's stock, p is the social media platform (Reddit or Twitter) on which sentiment was expressed, and t denotes a business day by its number. *SS* is a dummy variable to represent if the day occurred during a short squeeze, while *BeforeSS* and *AfterSS* have values of 1 if the day was before or after a short squeeze. (See Table 3.)

Period	*BeforeSS*	*SS*	*AfterSS*
1	Jan 2, 2020 to Jan 21, 2021		
2		Jan 22, 2021 to Feb 4, 2021	
3			Feb 5, 2021 to Feb 23, 2021
4		Feb 24, 2021 to Apr 8, 2021	
5			Apr 9, 2021 to May 24, 2021

6 7		May 25, 2021 to Jun 9, 2021	Jun. 10, 2021 to Jun 29, 2021

Note: The duration of the period that our data tracked was 376 business days, excluding weekends and holidays. Individual days are denoted by *t* in our empirical model.

Table 2. Dates for the periods before (1), during (2, 4, 6), and after the short squeezes occurred (3, 5, 7)

Variables	Definitions
• **Dependent Variables**	
Vol_{it}	Day *t* log transaction volume for stock *i*
$Returns_{it}$	Day *t* returns for stock *i* ($ClosingPrice_{\ t} - OpeningPrice_{\ t}$) / $OpeningPrice_{\ t}$
• **Social Sentiment**	
$\#Posts_{ipt}$	Total # of posts about stock *i* in social platform *p* on day *t*
$\#Pos_{ipt}$	Positive sentiment % about stock *i* in platform *p* on day *t*
$\#Neg_{ipt}$	Negative sentiment % about stock *i* in platform *p* on day *t*
$\#Neu_{ipt}$	Neutral sentiment % about stock *i* in platform *p* on day *t*
• **Periods**	
$BeforeSS_i$	Day *t* is before one of stock *i*'s short-squeeze periods = 1; 0 otherwise
SS_i	Day *t* is during one of stock *i*'s short-squeeze periods = 1; 0 otherwise
$AfterSS_i$	Day *t* is after one of stock *i*'s short-squeeze periods = 1; 0 otherwise

Notes. Social platforms *p*: Reddit or Twitter; calendar time in business days *t* = (1, 376). Obs.: the Twitter
GME dataset had 2,280 submissions and 294,377 comments; the Reddit GME dataset contained 2,025,592 tweets with 1,236,888 tweets were about BBY, the matching firm. Sentiment analysis: performed using *polarity scores.* Vars.: *Neg* = negative sentiment index; *Neu* = neutral sentiment; *Pos* = positive sentiment. Each score codes for the proportion of text post in these categories (e.g., a post may be 53% positive, 33% neutral, 14% negative, with the total equal to ~100%.

Table 3. Variables and Definitions

To test how trading momentum predictions and social informedness via the social media platforms can change over time in the different squeeze-related periods. We constructed the five estimation models below. However, the first two models (M1 and M2) were used to estimate the remaining three (M3, M4, and M5), which all include period-related dummy variables. They could not have been built without our use of the first two models. To obtain our findings on the hypotheses, we estimated the final three models, M3, M4, and M5:

$$\text{(M1)}\quad Y_{it} = \alpha + \beta_1 Pos_{ipt} + \beta_2 Neg_{ipt} + \varepsilon_{it}$$

$$\text{(M2)}\quad Y_{it} = \alpha + \beta_1 SS + \beta_2 AfterSS + \varepsilon_{it}$$

$$\text{(M3)}\quad Y_{it} = \alpha + \beta_1 (\#Posts_{ipt} \times BeforeSS_i) + \beta_2 (\#Posts_{ipt} \times SS_i) + \beta_3 (\#Posts_{ipt} \times AfterSS_i) + \varepsilon_{it}$$

$$\text{(M4)}\quad Y_{it} = \alpha + \beta_1 (Pos_{ipt} \times BeforeSS_i) + \beta_2 (Pos_{ipt} \times SS_i) + \beta_3 (Pos_{ipt} \times AfterSS_i) + \varepsilon_{it}$$

$$\text{(M5)}\quad Y_{it} = \alpha + \beta_1 (Neg_{ipt} \times BeforeSS_i) + \beta_2 (Neg_{ipt} \times SS_i) + \beta_3 (Neg_{ipt} \times AfterSS_i) + \varepsilon_{it}$$

In the models, Y_{it} represents one of two dependent variables – *daily trading volume* (*Vol*) or *daily stock return* (*Return*). Also, *i* is the firm's index for its stock, *p* is the social media platform (Reddit or Twitter) on which the sentiment was expressed, *t* is a specific business day by its sequence number in the dataset, and ε is the estimation residual. *SS* is a dummy to represent if the day occurred during one of several short-squeeze periods, while *BeforeSS* and *AfterSS* are each 1 if the day was before or after a short squeeze, else 0.

Estimation Results: Tests of the Hypotheses

In recent research, Allen et al. (2022) provided evidence on coordinated trading by many individual investors. They found that in a short squeeze, market quality was lower for the stocks despite the continuous processing of information. They also documented negative spillover effects on market quality from the competitors of the focal firms whose stocks were shorted. Our study considers how sentiment extracted from GME-related posts and comments on Reddit and Twitter are associated with social informedness-based behavior, as market quality changed.

Short squeeze model estimation results

Table 4 presents the transactions for GME stock and its matching firm, BBY, for different periods.

Variables	***Vol***		***Returns***	
Firms	**GME**	**BBY**	**GME**	**BBY**
BeforeSS	6.708***	6.408***	0.011	0.002
	(0.022)	(0.010)	(0.008)	(0.002)
SS	7.496***	6.458***	0.103***	0.001
	(0.053)	(0.025)	(0.022)	(0.004)
AfterSS	6.876***	6.230***	-0.005	-0.000
	(0.043)	(0.021)	(0.018)	(0.003)
Adj R^2	0.331	0.073	0.039	0.005

Notes. Obs: 365 business days. Signif.: * = $p < 0.10$; *** = $p < 0.01$. GME – GameStop; BBY = Best Buy.

Table 4. Transactions of GME and its matching firm for the three short-squeeze periods

The regression for the period dummies without a constant implies the mean values of the logged trading volume and return effects for *BeforeSS*, *SS*, and *AfterSS*. Trading volume and returns during the short squeezes increased for GME, while no meaningful differences were found for BBY. These results are consistent with Allen et al.'s (2022) that provided evidence for GME's deteriorating market quality during the short squeeze.

Since we observed declines in shorted shares concurrent with the sudden spikes in GME's stock price, we assessed the changing impacts of investor sentiment on GME's market activities and its matching firm over time. For this, we focused on interactions between sentiment extracted from social news platforms across the different time periods. We obtained further estimation results for the social media-related posting and sentiment variables on the trading volumes and returns of GME and BBY. (See Table 5 for the results from models M3 to M5.)

Test details

Hypothesis 1 test results. The Individual Investors' Collective Behavior for Market Transactions Hypothesis (H1) asserts that individual investors' collective behavior driven by social informedness became stronger during the short-squeeze period. Our empirical findings show that the number of social media posts was posi-

tively correlated with GME's trading volume during the short squeeze: *#Posts* × *SS* = {0.772***, 0.338***, 0.635***}. Also, the coefficients of Reddit were larger during the short-squeeze period than before or after it for Reddit submissions: (*#Posts* × *SS* = 0.772***) > (*#Posts* × *BeforeSS* = 0.629***) > (*#Posts* × *AfterSS* = 0.523***). Qualitatively similar results were obtained for the Reddit comments. The before and after-squeeze results estimates were still less than the results during the squeeze (though the order was inverted as denoted by the underlined portions in the following relationships): (*#Posts* × *SS* = 0.338***) > (*#Posts* × *AfterSS*) = 0.152***) > (*#Posts* × *BeforeSS* = 0.148***). The results suggest postings on Reddit strengthened individual investors' collective behavior and H1 was supported.

Variables	***Vol***				***Returns***			
Firms	**GME**			**BBY**	**GME**			**BBY**
Platform	**Reddit Submits**	**Reddit Comments**	**Twitter Tweets**	**Twitter Posts**	**Reddit Submits**	**Reddit Comments**	**Twitter Tweets**	**Twitter Tweets**
• **Model 3: Impact of *#Posts***								
#Posts × *BeforeSS*	0.629*** (0.073)	0.148*** (0.023)	0.653*** (0.049)	0.163*** (0.037)	0.077** (0.034)	0.006 (0.010)	0.080*** (0.024)	-0.005 (0.007)
#Posts × *SS*	**0.772*** (0.046)**	**0.338*** (0.020)**	**0.635*** (0.036)**	0.178*** (0.039)	**0.084*** (0.021)**	**0.035*** (0.009)**	**0.077*** (0.018)**	-0.005 (0.007)
#Posts × *AfterSS*	0.523*** (0.073)	0.152*** (0.020)	0.543*** (0.039)	0.132*** (0.040)	0.001 (0.033)	-0.004 (0.009)	0.057*** (0.019)	-0.006 (0.007)
Constant	6.633*** (0.021)	6.514*** (0.034)	4.840*** (0.141)	5.868*** (0.126)	0.000 (0.009)	0.003 (0.015)	-0.22*** (0.070)	0.018 (0.022)
Adj R^2	0.467	0.441	0.547	0.110	0.041	0.050	0.063	0.000
• **Model 4: Impact of *Positive* sentiment valence**								
Pos × *BeforeSS*	0.591*** (0.195)	0.029 (0.264)	-1.348** (0.645)	-0.001 (0.302)	0.013 (0.082)	0.025 (0.106)	-0.392 (0.264)	-0.043 (0.052)
Pos x SS	**4.209*** (0.301)**	**3.374*** (0.338)**	**1.716*** (0.602)**	0.165 (0.282)	**0.437*** (0.127)**	**0.443*** (0.136)**	-0.058 (0.246)	-0.041 (0.048)
Pos × *AfterSS*	1.287*** (0.273)	0.647** (0.293)	-0.594 (0.631)	-0.506* (0.296)	-0.073 (0.115)	-0.040 (0.118)	-0.437* (0.258)	-0.048 (0.051)
Constant	6.625*** (0.035)	6.713*** (0.052)	6.998*** (0.139)	6.407*** (0.065)	0.009 (0.014)	0.006 (0.021)	0.095* (0.057)	0.011 (0.011)
Adj R^2	0.356	0.302	0.311	0.063	0.034	0.045	0.027	0.005
• **Model 5: Impact of *Negative* sentiment valence**								
Neg × *BeforeSS*	0.788*** (0.205)	0.483* (0.275)	-0.795 (0.816)	-0.018 (0.377)	0.047 (0.085)	0.029 (0.111)	0.312 (0.329)	0.007 (0.065)
Neg × *SS*	**4.178*** (0.300)**	**3.807*** (0.354)**	**2.189*** (0.754)**	0.127 (0.348)	**0.463*** (0.125)**	**0.406*** (0.143)**	**0.585** (0.304)**	0.002 (0.060)
Neg × *AfterSS*	1.339*** (0.266)	1.001*** (0.299)	-0.086 (0.745)	-0.525 (0.344)	-0.051 (0.111)	-0.051 (0.121)	0.208 (0.301)	-0.001 (0.059)
Constant	6.604 (0.035)	6.634*** (0.053)	6.880*** (0.174)	6.413*** (0.080)	0.004 (0.015)	0.007 (0.021)	-0.055 (0.070)	0.000 (0.014)
Adj R^2	0.347	0.299	0.293	0.068	0.036	0.036	0.028	0.007

Notes. Model: OLS. Obs.: 376 business days. Variables: *Vol* and *#Posts* are log values. Signif.: * = $p < 0.1$; ** = $p < 0.5$; *** = $p < 0.01$. *Submit* = submission volume and *Comm* = # of comments. GME = GameStop; BBY = Best Buy

Table 5. Number of posts and sentiment-driven market activities by short-squeeze periods

Hypothesis 2 test results. Table 4, presented earlier, showed the results for our momentum-based predictions for the Investor Sentiment Hypothesis (H2). They suggest that investor sentiment can be a momentum predictor. Note that for the Reddit platform that the estimated values for the social sentiment valences on daily returns were the highest and, more importantly, were only statistically significant during the short squeeze. This is

shown by (*Pos* = 0.437***, *Neg*: 0.463***) on Reddit comments, and (*Pos* = 0.443***, *Neg* = 0.406***) on Reddit comments. The empirical findings further show the dominance of investor sentiment's influence on daily stock returns, but only during the short squeeze – a surprising result from this analysis. Thus, we conclude that H2 is generally supported, though both the *Pos* and *Neg* sentiment variables involved had positive coefficients.

Hypotheses 3 and 4 test results. To obtain the results for our tests of the Differential Social Informedness by Platform Hypothesis (H3) and the Sentiment Impact for the Focal and Matching Firms in the Short Squeeze Hypothesis (H4), it was necessary to estimate the joint impacts of the platform and the focal firm. For this, we jointly estimated the effects of Reddit versus Twitter for the platforms and GME and BBY as the focal stocks with the SUR procedure. This addresses the cross-correlated information structure of the residual terms for each platform and the firms. Different processes were in operation during the short-squeeze period, we believe.

GME was targeted by Reddit's individual investors, who found a new platform for market manipulation in the process. In comparison to Twitter, Reddit is more chaotic, and extracting sentiment from its subreddits is a challenging semantic problem. Considering the design of its platform, Reddit may contain unique instances of social sentiment that are unnatural and were not captured by Twitter, due to their contrasting functionality.

Further, the vocabulary of Reddit users differs since it plays to meme culture, where messages are often false, uninformed, or offensive. Individual investors were immersed in the meme culture, which spread across society, resulting in herding behavior being coordinated via the r/WallStreetBets subreddit. Individual investors who used Robinhood could obtain easy access to the financial markets. At that time, they were awareness of the social injustice that, they alleged, had been caused by financial institutions. We examine the impact of meme-driven culture and investor sentiment from Reddit posts, related submissions, and comments on the market. Before that time though, Reddit was recognized to have influence on the markets and was ignored by finance scholars. But during the GME short squeeze, it became an influential social media platform (Long et al., 2021).

The text messages in Reddit contained a large volume of data that was challenging to analyze directly without manipulation and preprocessing, and the use of powerful data analytics tools. We collected all the messages that could have explained GME's upward price movement, and the common terms used in the related subreddits. Then, we extracted the sentiment posts they contained. This enabled us to analyze its impact from Reddit, based on what we could extract from GME's submitted posts and user comments. We also compared the GME content with social sentiment from Twitter. (Table 6 presents the results.)

Variables	***Vol***			***Vol***	
Firm	**GME**			**GME**	**BBY**
Platform	**Reddit Submissions**	**Reddit Comments**	**Twitter Tweets**	**Twitter Tweets**	**Twitter Tweets**

• **Model 3: Post Influence**					
#Posts × BeforeSS	0.077*	0.017	0.188***	0.636***	0.160***
	(0.043)	(0.012)	(0.031)	(0.049)	(0.038)
#Posts × SS	**0.460*** **	**0.177*** **	**0.263*** **	**0.630*** **	**0.179*** **
	(0.037)	**(0.015)**	**(0.023)**	**(0.036)**	**(0.039)**
#Posts × AfterSS	0.137***	0.035***	0.168***	0.535***	0.131***
	(0.051)	(0.013)	(0.024)	(0.039)	(0.039)
#Constant	6.760***	6.738***	6.212***	4.890***	5.878***
	(0.181)	(0.023)	(0.089)	(0.142)	(0.125)
Adj R^2	0.329	0.033	0.042	0.553	0.120
• **Model 4: Positive Sentiment Influence**					
Pos × BeforeSS	0.014	-0.027	-0.137	-1.447**	-0.021
	(0.045)	(0.043)	(0.102)	(0.631)	(0.300)
Pos × SS	**1.358*** **	**1.106*** **	**0.963*** **	**1.935*** **	**0.217**
	(0.161)	**(0.152)**	**(0.151)**	**(0.591)**	**(0.281)**
Pos × AfterSS	0.146	0.063	-0.026	-0.627	-0.493*
	(0.103)	(0.094)	(0.112)	(0.612)	(0.291)
Constant	6.791***	6.802***	6.824***	7.019***	6.411***
	(0.019)	(0.021)	(0.028)	(0.136)	(0.064)
Adj R^2	0.198	0.167	0.186	0.340	0.077
• **Model 5: Negative Sentiment Influence**					
Neg × BeforeSS	0.023	-0.039	-0.037	-0.832	-0.029
	(0.051)	(0.043)	(0.118)	(0.798)	(0.374)
Neg × SS	**1.773*** **	**1.580*** **	**1.445*** **	**2.483*** **	**0.184**
	(0.187)	**(0.186)**	**(0.192)**	**(0.741)**	**(0.347)**
Neg × AfterSS	0.165	0.074	0.078	-0.064	-0.503
	(0.117)	(0.113)	(0.138)	(0.726)	(0.341)
Constant	6.778***	6.792***	6.791***	6.887***	6.415***
	(0.020)	(0.021)	(0.031)	(0.169)	(0.079)
Adj R^2	0.234	0.207	0.224	0.323	0.080

Notes. Obs: 376 business days. Signif.: * = $p < 0.10$; ** = $p < 0.50$*** = $p < 0.01$.

Table 6. Results of models 3, 4 & 5: Number of posts and sentiment-driven trading volume in short squeeze

The number of posts and sentiment comments related to GME had a larger impact on its trading volume during the short squeeze. For negative sentiment influence, the variables' interaction results were (*Neg* × *SS* = 1.773*** for Reddit submissions) > (*Neg* × *SS* = 1.580*** for Reddit comments) > (*Neg* × *SS* = 1.445*** for Twitter).

Table 7 shows the *t*-test results for H3 and compares the post and sentiment effects between the social media platforms, Reddit and Twitter. The *#Posts* and *Sentiment* variables' coefficients extracted from GME submissions on Reddit were significantly higher than those on Twitter. The sentiment coefficients retrieved for the GME comments on Reddit did not show consistently larger impacts compared to Twitter's though. Overall, it seems that H3 was supported, except in the case with positive sentiment extracted from Reddit comments. Thus, social informedness from Reddit apparently had a larger impact on investors' GME transactions than it did from Twitter.

Hypotheses	Coef. Compared	*p*-values	Findings

3a (*Posts*)	*#Posts* effect on Reddit submissions	$0.460^{***} > 0.263^{***}$ (Reddit > Twitter)	0.00^{***}	*#Posts* effect on Reddit >*#Posts* effect on Twitter
3a (*Posts*)	*#Posts* effect for each social media	$0.177^{***} < 0.263^{***}$ (Reddit > Twitter)	0.00^{***}	(*#Posts* effect on Twitter) > *#Posts effect on* Reddit)
3b (*Pos*)	Sentiment effect on Twitter > on Reddit	$1.358^{***} > 0.963^{***}$ (Reddit > Twitter)	0.00^{***}	(*Pos* effect on Reddit) > (*Pos* effect on Twitter)
3b (*Pos*)	Sentiment effect on Twitter < on Reddit	$1.106^{***} > 0.963^{***}$ (Reddit > Twitter	0.195	(*Pos* effect on Reddit) = (*Pos* effect on Twitter)
3b (*Neg*)	*Neg* sentiment effect for Reddit submissions	$1.773^{***} > 1.445^{***}$ (N*eg* for Reddit > (*Neg* for Twitter)	0.000^{***}	*Neg* effect on Reddit) > (*Neg* effect on Twitter)
3b (*Neg*)	*Neg* sentiment effect for Reddit comments	$1.580^{***} > 1.445^{***}$ (*Neg* for Reddit) = (*Neg* for Twitter)	0.242	(*Neg* effect on Reddit) = (*Neg* effect on Twitter)

Notes. *t*-tests used for hypothesis results. Signif.: $^{*} = p < 0.10$; $^{**} = p < 0.05$; $^{***} = p < 0.01$. Social media: Reddit and Twitter. Variables assessed: (a) *#Posts* and (b) *Neg* and *Pos* sentiment valences for post content.

Table 7. Hypothesis 3 test results with comparisons between Reddit and Twitter

We further compared the collective market behavior that individual investors' stock trading exhibited for the focal firm (GME) and the matching firm (BBY) during the short squeeze. We considered the issue of data quality-led bias (as well as sample selection bias), as possibly distorting the results of the study that could result in erroneous conclusions. We dealt with this type of bias with our data diagnostics and analytics though. An obvious method to use was a random sample selection process. Our data set includes all observations after the data preprocessing was done. So, this resulted in nearly the total populations of social media posts for each posting platform since our data collection covered several periods that could have affected the short squeeze event.

The statistical methods we used involved estimation via linear regression, with *t*-tests to check if there were significant relationships between the predictor and response variables. From an econometric viewpoint, it was interesting to investigate whether the predictors, *#Posts* and *Sentiment*, had the same effect on GME trading volume during the study period for the two social media platforms, Reddit and Twitter. For this, we performed a test for the null (H_0) and alternative (H_A) hypotheses (Breusch & Pagan, 1979), as follows:

H_0: *The coefficient values of posts and sentiment on Reddit didn't differ from those on the Twitter platform.*

H_A: *The coefficient values of posts and sentiment on Reddit differed from those on the Twitter platform.*

The results reported earlier confirm that investor sentiment was correlated with individual investor trading volume for GME during the short-squeeze period, but not with that of BBY. This conclusion is based on (*Pos* × *SS* = 1.935^{***} for GME) > (*Pos* × *SS* = $0.217^{n.s.}$ for BBY on Twitter) and (*Neg* × *SS* = 2.483^{***} for GME) > (*Neg* × *SS* = $0.184^{n.s.}$ for BBY on Twitter). Next, Table 8 reports the results of pairwise *t*-tests for H4.

Our analysis shows that individual investors were more affected by sentiment when they bought and sold GME's stock than when they did the same for BBY's stock. This shows that H4 was supported. Thus, the results

indicate social informedness-related behavioral patterns for GME stock trades may apply to other electronics retailers' stocks as well.

Hypotheses		Coefficient Comparisons	*p*-value	Findings
4 (*#Posts*)	*#Posts* effect on GME	0.630^{***} (*#Posts* for GME) > 0.179^{***} (*#Posts* for BBY)	0.00^{***}	*#Posts* effect on GME larger than *#Posts* effect on BBY
4 (*Pos*)	*Pos* sentiment effect on GME	1.935^{***} (*Pos* for GME) > 0.217 (*Pos* for BBY)	0.00^{***}	*Pos* effect on GME larger than *Pos* effect on BBY
4 (*Neg*)	*Neg* sentiment effect on GME	2.483^{***} (*Neg* for GME) > 0.184 (*Neg* for BBY)	0.00^{***}	*Neg* effect on GME larger than *Neg* effect on BBY

Notes. *t*-tests were used to assess hypothesis results. Signif.: $^{*} = p < 0.10$; $^{**} = p < 0.05$; $^{***} = p < 0.01$.

Table 8. Hypothesis 4 test results (comparison between GME and BBY coefficients)

Discussion

In this study, we have shown that individual investors exhibited our hypothesized trading patterns through social media platforms during the GME short squeeze for the most part. Social sentiment appears to have enhanced the collective behavior of individual investors. Subsequent analysis showed that we can interpret the controversial informational effect of social media on collective stock trading in a meaningful way. Contrary to previous studies, our empirical results suggest that Reddit's social information offerings solidified their collective behavior and had a greater impact on GME trading than it did due to Twitter during the short squeeze. Our study further indicates that individual investors were more affected by sentiment when participating in GME stock trading than the BBY matching firm's stock trading.

Differentiated measures of investor sentiment on the different platforms helped to address current research gaps and understand prior theories of financial behavior in the GME short squeeze setting more fully. The results reveal behavioral patterns related to potential social informedness and offer new perspectives for assessing the impact of investor sentiment on the stock market. Our evidence suggests that the social network effect drove social informedness-based collective trading behavior for the firm's stock – a new contribution in IS research.

Trading behavior and contemporary issues

Individual investor trading, behavior patterns, and links to hypotheses. These issues were worthwhile to study since social media and investor sentiment are thought to be influential in the stock market. Barber et al. (2009) have stated that group behavior of individual investors may drive stock prices away from the fundamental value of the firms they represent. Hu et al. (2021) argued in addition that lengthy deviations of a stock's price from what its fundamentals indicate depends on the balance between the demand of rational long-term and irrational short-term investors, as well as the relative importance of their respective thought leadership in the market. Our study has confirmed that investor sentiment, in the presence of relatively more market expertise with

discussing stock trading strategies on a specialized platform, affects the collective trading behavior and social informedness of investors. As such, it will be worthwhile to integrate our empirical findings into the present understanding of price movement to develop new theory for investors' collective behavior.

Our findings also provide evidence regarding what patterns of individual investor trading are demonstrated when social media platforms supply them with abundant information, as observed during the GME short squeeze. We examined whether the collective trading activity of individual investors was related to GME's trading volume and its predictive power for price movements. It turned out that these variables were stronger during than before or after the short squeeze. We investigated whether investor sentiment affected GME stock returns during the short squeeze, too. Our empirical results suggest that social media sentiment strengthened individual investors' collective behavior and was a momentum predictor during the short squeeze. So, our study provides support for new theory based on understanding social media's controversial information effects on stock trading. We also examined how platform which host discussions of specialized topics in social media may create variation in the related social informedness of the stakeholders involved.

Other issues. Since the Covid-19 pandemic occurred, individual investors have purchased a considerable amount of stock, pushing the overall value of the market higher, based on investors' post-Covid economic growth expectations (despite the apparent move toward economic recession in the third quarter of 2022). Along the way, they experienced lower buyer's market prices immediately after it became clear that the pandemic would result in diminished economic activity in the U.S. (Domm, 2021). Also, online individual investors' market interest peaked as their outdoor activities were curtailed due to the Covid lockdown, and they had to adjust to spending more time indoors. Further, online stock trading apps (e.g., Robinhood and its competitors, Coinbase, Webull, M1 Finance, and Public.com, among others) were more actively used across most individual investors, especially in the U.S. Recently, individual investors have shown a clear a preference for *participatory impact investing*, in which they select stocks themselves and decide the timing of stock purchases and sales, rather than indirect investment using fund managers, to directly project the influence .the influence of the social communities (O'Flynn & Higdon, 2019). Individual investors have been observed to trade actively – though frequently at a loss – because they believe that they have good information and think they should be able to realize a profit through their investments – a debatable proposition.[9]

Today, many individual investors are caught between the two poles of investment thinking and social communication. *Rational investors* read the news, try to understand the facts to the best of their ability, and develop

[9] *Overconfidence* refers to a psychological bias in which you believe that the information you know is more certain than the information others know (Barber & Odean, 2000).

a view of financial markets and their traded instruments that is within the bounds of the rational expectations framework of how the market economy works (Muth, 1961). *Irrational investors*, in contrast, update their views over time in response to social media-supplied information. Reddit and its sub-communities, as we have shown in this research, offer opinion-shaping input to build social informedness which – if not "correct" in the usual sense that theory suggests – still provides participants with a socially-biased sense of the market's valuation, and confidence to engage in collective action for stock purchases or sales – even if their decisions are less rational.[10]

Interpretation

In the 2010s, investors frequently discussed the notion of *meme stocks*. This is a stock that has become a hot topic in Internet communities or on social media and has shown a price increase unrelated to such fundamentals as a higher sale margin or improved profitability. GME in this research is an example of a meme stock. David Hirshleifer (2020), a past president of the American Finance Association, argued that social interactions in which individuals communicate with and observe each other, influence the collective economic thinking and behavior that appears in society. Information distortion through social transmission occurs while people are interacting with one another though. *Social transmission bias* refers to systematic distortion of information that occurs in the process of information transfer between people. An amount of exaggeration may occur when sharing a pleasant story with others or adorning one's past to be more glamorous or compelling are common social transmission biases.

Hirshleifer applied these properties to economic systems and stock markets, and emphasized two key characteristics related to individual preferences, strategies, and investment performance. One is *signal distortion* that occurs through exaggerating or reducing content in the information transfer process. A second is *selection bias* that involves being proud of others when their performance is good but being silent when it is not. This is typically expressed based on similar aspects of social transmission bias. For example, stock investments by individual investors are not only caused by individual preferences but are a by-product of individual priorities and concerns that are independently determined, socially manifested, and widely expressed. Conventional economics has explained the related phenomenon as economic agents' pursuit of their individual preferences and the consequent increase in utility they experience. Even though individuals may not think they create problems, any

[10] Other social media (e.g., Twitter, Facebook, etc.) have focused on individual connections. What sets Reddit users apart is that they seem more open to engaging in meme-stock trading. As Reddit offers community-based social media, users can talk freely about topics they are interested in. They also understand the platform operator's emphasis on *freedom of expression* – and the unlimited communication opportunities it offers. There is a dark side though: harmful and inflammatory comments can be shared, and not screened out. So, it becomes hard to distinguish one person's truth from another's fallacy – like fake news in the 2000s. This platform-specific social network effect created *individual investor social informedness* wrought by social sentiment. It ultimately became a dangerous form of *groupthink* for many who eventually lost a lot of their invested funds.

specific preference they do have can be strong enough to create social transmission bias in public communication.

In recent years, short selling has spread into the public arena, with growing influence on the political economy of nations where it occurs, due to the collective behavior of individual investors (Lunde et al., 2021). This occurred, for example, when individual domestic investors in each of two countries – South Korea and the U.S. – coordinated their respective efforts on social media platforms to defend selected stocks' prices against investors who were selling the stocks (Han 2020). They rattled the market with their abundant liquidity. In fact, for stocks with a high proportion of short-selling by the influential hedge funds, the collective slogan in the Korean market became "*Let's buy stocks. Let's strangle the short-selling forces*."

In this instance, individual investors formed groups on social media to purchase specific stocks (or buy *call options*, which gave them the right but not the obligation to buy the targeted equities via their call options at prespecified prices). They also engaged in algorithmic trading. Some stock prices surged upward for a week or more. This suggests the extent of the momentum that social media was able to generate. Such momentum did not last for very long though. For GME, its stock price surge may not have been simply due to individual investors' collective buying based on social media-driven momentum. Rather, the hedge funds also made *short-cover* trade actions to close out their suddenly loss-creating open short-selling positions. They recognized it was necessary to buy shares again to reduce their losses as GME's price rose unexpectedly – awareness that was important.

For informedness-based behavior arising due to social media-driven momentum, over-reliance on social media encourages individual investors to believe they have market wisdom and should exercise their own judgment – whether such knowledge is real or only illusory. People who are constantly exposed to the beliefs of others about the market must bear the risk that social media interpretations of how things work may not be correct.[11]

McPherson et al. (2001) noted that the widely-accepted *homophily principle* suggests that we tend to associate with others who are like us in social systems. Thus, our *ego network* will involve others with whom we have social relationships. This, in turn, results in *homophilic similarities* in social network members' social, demographic, behavioral, and relational characteristics.[12]

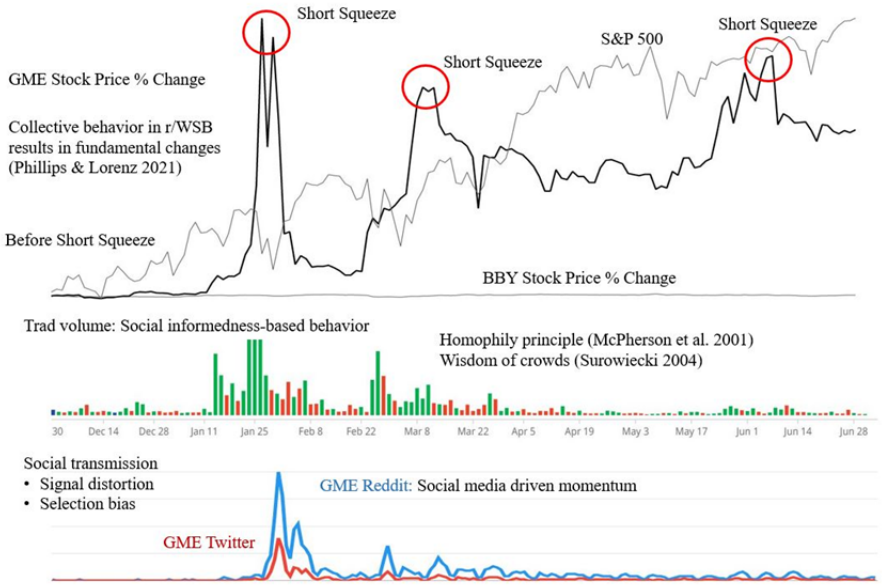
---

[11] For example, Wikipedia has successfully enhanced collective social intelligence by promoting user participation through content creation. But the apparent outcome that has emerged also has led to the problem of *collective bias*. This occurs in social media platforms such as Reddit, Twitter, and others. We are often influenced by others through their *social influence* on us rather than our own independent thinking. The result is that our understanding – right or wrong –converges socially.

[12] Echoing this, Carr (2008) noted that Internet search reduces academic research diversity. It narrows the scope of citations from which authors' knowledge is drawn by filtering varied information – and has similar implications in stock trading.

Surowiecki (2004) studied the *wisdom of crowds*. For wisdom to be formed socially, independent opinions and opposing views are subject to crowd processes such as sharing, cooperation, and coordination. Yet the independence and diversity of individual opinions on social media have been viewed as foundational for such wisdom to emerge – an apparent contradiction. Due to extensive networking in social media though, user groups with the same tastes and thoughts typically form, strengthening homophily. This collective bias becomes stronger due to the structure of forums such as r/WSB and other investment blogs. Other individual investors' thoughts are transparently visible, but it is difficult for meaningful discussion and conversation to take place. Thus, homophily and collective bias dampen the healthy exchange of diverse opinions, diminishing the realization of crowd wisdom.

Related to the GME short squeeze, the *New York Times* (Phillips & Lorenz, 2021) warned that the power shown by the collective behavior of individual investors participating in r/WSB and social media will result in fundamental changes in future financial markets, diminishing market quality. Other observers were concerned that the market's valuation bubble would collapse due to this incident – a warning of the possible repeat of the dotcom bubble burst from 2000 to 2002. In contrast, market regulators today are concerned that individual investors, regardless of firms' intrinsic value or fundamentals, will engage in extreme stock speculation. Further, the market may also be disrupted by the collusion of individual investors related to *initial public offerings* (IPOs) of stock. (See Figure 2, which illustrates the insights shared in this discussion mapped to the price curves of GME and BBY stock.)



Source. In the first short squeeze from January 22 to February 4, 2021, the S&P 500 showed that the U.S. stock markets were falling, whereas the stock price of GME soared during the same period. We adapted the above figure from the following sources and then expanded its contents based on the information that we need to present to the reader. See: BarChart.com (2021) and GoogleTrends (2021). https://www.barchart.com/stocks/quotes/GME/ and https://trends.google.com/trends/.

Figure 2. Daily indicator of changes for GME stock, December 1, 2020 – June 29, 2021

Conclusion

Research findings and digital platform implications

Our study investigated the trading patterns of individual investors and their collective behavior that was set off by social sentiment during the GameStop short squeeze. Based on our understanding of the more controversial informational effects of social media on social stock trading, the related behavior we observed the basis for applying a new perspective in IS research. To this end, we established hypotheses for Reddit, Twitter and GameStop-related stock trade data with daily closing prices, trading volumes, and returns for the two stocks. The theorized empirical outcomes in prior research were useful to understand the collective actions of individual investors on social platforms. The GME short squeeze event is relevant for studying the impact of investors in the current context because of commission-free trading and social media platforms. We sought to interpret these phenomena in the GME case context. We highlighted the role of investors' social sentiment. We did this by contrasting the focal firm with a matching firm for two platforms that represent an interest-based discussion platform and a common-topics discussion platform.

In response to RQ1 on investors' trading patterns, we found that social postings on Reddit solidified individual investors' collective behavior, which apparently was being steered by their extensive social media-based communication. In response to RQ2, we learned that investor sentiment triggered momentum trading based on daily stock returns and trade volumes during the short squeeze. With further refinement of our approach to understanding the applied empirical research setting, we were able to characterize the apparently irrational stock trade-related decision-making behavior of individual investors. They believed they were socially-informed due to their extensive participation in Reddit's highly-active social media that focused on GameStop stock trading.

In response to RQ3, we found that social informedness from communication via Reddit had a greater impact on individual investor GameStop trades than Twitter had, though one might have guessed there would be similar impacts. We further confirmed that individual investors were more affected by sentiment when they made trading decisions with respect to GameStop's stock as opposed to Best Buy's stock. Thus, we unexpectedly found that social informedness-related behavioral patterns for GameStop's stock trades apply not only to other electronics retailers, but very likely to smaller firm equities traded in the market-at-large.

Limitations

Behind Reddit's retail users' investments, there was a strong sense of antipathy toward Wall Street invest-

ment experts – especially from young investors (e.g., social media and Reddit users, including Gen-Y millennials, Gen-Z zoomers). Many of them believe that Wall Street capitalists caused the 2008 financial crisis and make money the "wrong way" in the market by controlling information dissemination, rigging market access, and creating rapid stock trade capabilities that are unavailable to individual investors. This suggests possible reasons for what we have interpreted as individual investor "irrationality" which may instead be tied to their social circumstances, the micro-level sequence of various aspects of the buy-sell stock trading process, and how social media information has evolved and been shared. Reddit's r/WSB bulletin board seems to have been central to having triggered the GameStop short squeeze and related events – despite its typically immediate (albeit slower and less automated) availability, access to information, and 24 x 7 market updates.

There is another possible narrative line as well though: that individual investors' actions were less a product of the individual investors' social informedness, but instead of their growing social rage with the "members only" club nature of Wall Street's investment practices, their lack of access, and the disadvantages and structural barriers they faced for participating in fair investing.[13] Though we focused mostly on one firm's stock, GameStop, it may be unsuitable for us to aver that we have produced generalizable explanations of other firms' and other markets' events and other traders' experiences with apparently irrational social trading. We encourage others to extend our study with different firms, stocks, market events, and periods to deepen the theoretical views we have advanced, or to challenge-test them by assessing the explanatory benefits of using other theories.

We performed econometric tests to quantify the direction and strength of the relationships between Reddit's social sentiment among individual investors and GameStop's market activities. We also compared and analyzed social informedness for this firm and its matching firm, Best Buy, on two social media platforms, Reddit and Twitter, during a period of short-squeeze events. Depending on the characteristics of the data sample, such as the period spanned, the number of observations, the variables' distributional normality, and other basic diagnostics, some tests were more effective than others in yielding use information. Our dataset is not very large, though it is comprehensive in its coverage of the focal events in social media posts for Reddit and Best Buy over the periods of time that we studied. This enabled us to distinguish among the pre-short, short, and post-short-squeeze periods, but we are still several steps short of being able to assert that a causal loop is truly at work in

[13] A parallel interpretation for retail investing has appeared in articles on the American race riots of the 1960s. For example, what some may have considered to be irrational violence and inappropriate damage to urban residential areas, others have viewed as collective actions that have been understood with numerous alternative explanations. They include time-wise geographic contagion, structural strains between social groups, relative and absolute deprivation of access, and unresponsive institutional structures (Myers, 1997). It also has been suggested that to understand the nature of social contagion and collective action requires micro-level assessment with event history methods to achieve a more causal inference-based understanding (DiPascuale & Glaeser, 1998). The GME short-squeeze setting also can be assessed with quasi-experimental methods that emphasize the sequence of events and investor differences, to get a clearer understanding of the key drivers of causality.

our study's setting.

Future research

Further exploration can be done into the abnormalities of the market short-squeeze events we have chronicled by examining the nature of the more normal pre-short squeeze market conditions and how they contrasted with those present in the short-squeeze and post-short-squeeze periods. It may be possible to identify the presence of an anti-leverage effect, involving the co-occurrence of a higher stock price and greater volatility made possible by social media-based coordination for collective individual investor stock purchase actions, for example. Vasileiou (2021) has argued that studying an extended period may provide the basis for assessing the validity of the assertion that the market quality was similar before, during, and after the short squeeze. We currently are considering how to conduct other tests to support causal inferences between the modeling variables in the presence of the relationship between stocks (with a focal firm vs. a matching firm) and platforms using empirical models with data that are becoming available.

References

Allen, F., Haas, M., Nowak, E., Pirovano, M. & Tengulov, A. (2022). Squeezing shorts through social news platforms. Research paper 21-31, Swiss Finance Institute, Zurich, Switzerland, February 3. Retrieved December 26, 2022, from https://bit.ly/3K8SfT2.

Al Guindy, M. (2021). Corporate Twitter use and cost of equity capital. *J. Corp. Fin., 68*, 101926. DOI: https://doi.org/10.1016/j.jcorpfin.2021.101926.

Alt, R. & Klein, S. (2011). Twenty years of electronic markets research: Looking backwards towards the future. *Electron. Mkts*., *21*(1), 41–51. DOI: https://doi.org/10.1007/s12525-011-0057-z.

Alt, R. & Zimmermann, H.D. (2018). Electronic Markets on networked media. *Electron. Mkts*., *28*(1), 1–6. DOI: https://doi.org/10.1007/s12525-018-0291-8.

Angel, J.J. (2021). Gamestonk: What happened and what to do about it. SSRN 3782195. Retrieved December 26, 2022, from https://bit.ly/3psOG24.

Antoniou, C., Doukas, J.A. & Subrahmanyam, A. (2013). Cognitive dissonance, sentiment, and momentum. *J. Fin. Quant. Anal., 48*(1), 245–275. DOI: https://doi.org/10.1017/S0022109012000592.

Asness, C., Moskowitz, T. & Pedersen L.H. (2013). Value and momentum everywhere. *J. Fin.*, *68*(3), 929–985. DOI: https://doi.org/10.1111/jofi.12021.

Banerjee, A.V. (1992). A simple model of herd behavior. *Qtrly. J. Econ*., *107*(3), 797–817. DOI: https://doi.org/10.2307/2118364.

Barber, B.M., Huang, X. Odean, T. & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *J. Fin.*, in press, September 30. Retrieved December 26, 2022, from https://bit.ly/3hqVVmK. DOI: https://doi.org/10.1111/jofi.13183.

Barber, B.M. & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *J. Fin. 55*(2),773-806. DOI: https://doi.org/10.1111/0022-1082.00226.

Barber, B.M., Odean, T., & Zhu, N. (2009). Do retail trades move markets? *Rev. Fin. Stud*., *22*(1), 151–186. DOI: https://doi.org/10.1093/rfs/hhn035.

BarChart.com (2021). GameStop Corp. (GME), from December 1, 2020, to June 30, 2021. Available: https://www.barchart.com/stocks/quotes/GME/ (last accessed December 22, 2022).

Bing, L., Chan, K.C. & Ou, C. (2014). Public sentiment analysis in Twitter data for prediction of a company's stock price movements. In *2014 Intl. Conf. in E-Bus & Eng.* (pp. 232-239), IEEE, Washington, DC. DOI: https://doi.org/10.1109/ICEBE.2014.47.

Bitter, S. & Grabner-Kräuter, S. (2016). Consequences of customer engagement behavior: When negative Facebook posts have positive effects. *Electron. Mkts*., *26*, 219–231. DOI: https://doi.org/10.1007/s12525-016-0220-7.

Bollen, J., Mao, H. & Zeng, X. (2011). Twitter mood predicts the stock market. *J. Comput. Sci*., *2*(1), 1–8. DOI:

https://doi.org/10.1016/j.jocs.2010.12.007.
Boorstin, J. (2021). Robinhood's disruptive force: The good, the bad, and the controversy. *CNBC*, May 25. Retrieved December 26, 2022, from https://cnb.cx/3sz6ryB.
Breusch, T.S. & Pagan, A.R. (1979). A simple test for heteroskedasticity and random coefficient variation. *Econometrica, 47*(5), 1287-1294. DOI: https://doi.org/10.2307/1911963.
Brunnermeier, M.K. & Pedersen, L.H. (2005). Predatory trading. *J. Fin*., *60*(4), 1825-1863. DOI: https://doi.org/10.1111/j.1540-6261.2005.00781.x.
Caron, M., Gulenko, M. & Muller, O. (2021). To the moon! Analyzing the community of degenerates engaged in the surge of the GME stock. In *Proc. 2021 Intl. Conf. Info. Sys*., Austin, TX.
Carr, N.G. (2008). Is Google making us stupid? What the Internet is doing to our brains! *The Atlantic*, July 1.
Chen, J. (2022). What is a day trader? *Investopedia*, February 2. Retrieved December 26, 2022, from https://bit.ly/35eP8u1.
Chircu, A.M. & Kauffman, R.J. (1999). Strategies for Internet middleman in the intermediation-disintermediation-reintermediation cycle. *Electron. Mkts*., *9*,(2), 109-117. DOI: https://doi.org.10.1080/101967899359337.
Clarke, J., Chen, H., Du, D. & Hu, Y.J. (2020). Fake news, investor attention, and market reaction. *Info. Sys. Res.,* 32(1),35-52. DOI: https://doi.org/10.1287/isre.2019.0910.
Clemons, E.K., Dewan, R.M., Kauffman, R.J. & Weber, T.A. (2017) Understanding the information-based transformation of strategy and society. *J. Manage. Info. Sys.*, *34*(2), 425-456. DOI: https://doi.org/10.1080/07421222.2017.1334474.
CNBC. (2022). GameStop skyrockets as retail investors force short squeeze. January 28. Retrieved December 26, 2022, from https://cnb.cx/3touDCU.
Costola, M., Iacopini, M. & Santagiustina, C. (2021). On the "mementum" of meme stocks. *Econ. Lett.*, *207*, 110021. DOI: https://doi.org/10.1016/j.econlet.2021.110021.
De Long, J.B., Shleifer, A., Summers, L.H. & Waldmann, R.J. (1990). Noise trader risk in financial markets. *J. Pol. Econ.*, *98*(4), 703–738. DOI: https://doi.org/10.1086/261703.
DiPascuale, D. & Glaeser, E. (1998). The Los Angeles riot and the economics of urban unrest. *J. Urb. Econ.*, *43*(1), 52-78. DOI: https://doi.org/10.1006/juec.1996.2035.
Domm, P. (2021). How the pandemic drove massive stock market gains, and what happens next. *CNBC*, December 30. Retrieved December 26, 2022, from https://cnb.cx/3wtUHjp.
Duz Tan, S. & Tas, O. (2020). Social media sentiment in international stock returns and trading activity. *J. Beh. Fin., 22*(2), 1–14. DOI: https://doi.org/10.1080/15427560.2020.1772261.
Eaton, G.W., Green, T.C., Roseman, B. & Wu, Y. (2021). Zero-commission individual investors, high frequency traders, and stock market quality. Working paper, Oklahoma State University.
Eisenhardt, K.M. (1989). Building theories from case study research. *Acad. Mgmt. Rev.*, *14*, 532-550. DOI: https://doi.org/10.5465/amr.1989.4308385.
Eisenhardt, K.M. & Graebner, M.E. (2007). Theory building from cases: Opportunities and challenges. *Acad. Mgmt. J.*, *50*(1), 25-32. DOI: https://doi.org/10.5465/amj.2007.24160888.
Github. (2022). JustAnotherArchivist/snscrape. Description of social network services packages for Reddit, Twitter, and other service providers. Retrieved December 26, 2022, from https://bit.ly/3iFkQTZ.
Gomber, P., Kauffman, R.J., Parker, C. & Weber, B.W. (2018). On the Fintech Revolution: Interpreting the forces of innovation, disruption, and transformation in financial services. *J. Mgt. Info. Sys.*, *35*(1), 220-265. DOI: https://doi.org/10.1080/07421222.2018.1440766.
GoogleTrends.com (2021). GameStop Corp (GME) Twitter and Reddit, from December 1, 2020, to June 30, 2021. Available at: https://trends.google.com/trends/ (last accessed December 22, 2022).
Gray, P., Johnson, S.L. & Kitchens, B. (2020). Understanding echo chambers and filter bubbles: The impact of social media on diversification and partisan shifts in news consumption. *MIS Qrtly.*, *44*(4), 1619-1649. DOI: https://doi.org/10.25300/MISQ/2020/16371.
Gubbins, C. & MacCurtain, S. (2008). Understanding the dynamics of collective learning: The role of trust and social capital. *Adv. Devel. Hum. Res*., *10*(4), 578–599. DOI: https://doi.org/10.1177/1523422308320372.
Han, A.R. (2020). Patriotic approach to market. *Korea JoongAng Daily,* December 10. Retrieved December 26, 2022, from https://bit.ly/3IZNSKp.
Han, X. & Li, Y. (2017). Can investor sentiment be a momentum time-series predictor? Evidence from China. *J. Empir. Fin*., *42*, 212-239. DOI: https://doi.org/10.1016/j.jempfin.2017.04.001.
Haramundanis, K. (2009). Experience report – Modularization: The new paradigm for the information engineer. In *Proc. 27th ACM Intl. Conf. Design of Communication* (pp. 151-154), ACM, New York, NY. DOI: https://doi.org/10.1145/1621995.1622023.
Hein, A., Schreieck, M., Riasanow, T., Setzke, D.S., Wiesche, M., Böhm, M. & Krcmar, H. (2020). Digital platform ecosystems. *Electron. Mkts*., *30*(1), 87–98. DOI: https://doi.org/10.1007/s12525-019-00377-4.
Hirshleifer, D. (2020). Presidential address: Social transmission bias in economics and finance. *J. Fin.*,

75(4), 1779-1831. DOI: https://doi.org/10.1111/jofi.12906.

Hoang, A.P. & Kauffman, R.J. (2018). Content sampling, household informedness, and the consumption of digital information goods. *J. Manage. Inform. Sys., 35*(2), 575-609. DOI: https://doi.org/10.1080/07421222.2018. 1451958.

Hong, H. & Stein, J. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *J. Fin.*, *54*(6), 2143-2184. DOI: https://doi.org/10.1111/0022-1082.00184.

Hsu, P.Y., Lei, H.T., Hoang, S.H., Liao, T.H., Lo, Y,.C. & Lo, C.C. (2019). Effects of sentiment on recommendations in social networks. *Electron. Mkts., 29*, 253-262. DOI: https://doi.org/10.1007/s12525-018-0314-5.

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. *Behav. and Exper. Fin. eJournal*, March 21.

Kelley, E.K. & Tetlock, P.C. (2013). How wise are crowds? Insights from retail orders and stock returns. *J. Fin.*, *68*(3), 1229–1265. DOI: https://doi.org/10.1111/jofi.12028.

Krantz, M. (2021). It's game over for the Reddit GameStop 'short-squeeze' rebellion. *Investor's Bus. Daily*. April 9. Retrieved December 26, from https://bit.ly/3y6U6Sq.

Li, T., Kauffman, R.J., van Heck, E., Vervest, P., & Dellaert, B.G. (2014). Consumer informedness and firm information strategy. *Inform. Sys. Res*., *25*(2), 345-363. DOI: https://doi.org/10.1287/isre.2014.0521.

Long C., Lucey, B. & Yarovaya, Y. (2021). 'I just like the stock' vs. 'fear and loathing on Main Street': The role of Reddit sentiment in GameStop short squeeze. Retrieved December 26, 2022, from https://bit.ly/3tngkhT.

Lunde, A., Jensen, J.B., Hauschultz, F.P. & Tizik, S. (2021). Market impact of short sales position disclosure. Copenhagen Economics, July 15. Retrieved December 26, 2022, from https://bit.ly/3v9fyH4.

Lyócsa, Š., Baumohl, E. & Výrost, T. (2021). YOLO trading: Riding with the herd during the GameStop episode. *Fin. Res. Lett.,* 102359. DOI: https://doi.org/10.1016/j.frl.2021.102359.

McEnery, T. (2021). Phunware's stock was up 1,000% on Friday. What in the world is Phunware? *MarketWatch*, October 21. Retrieved December 26, 2022, from https://on.mktw.net/3vwSFhQ**.**

McPherson, M., Lovin, L.S. & Cook. J.M. (2001). Birds of a feather: Homophily in social networks. *Ann. Rev. Soc.*, *27*(1) 415-444. DOI: https://doi.org/10.1146/annurev.soc.27.1.415.

Mitchell, C. (2022). Short squeeze definition. *Investopedia*, January 22. Retrieved December 26, 2022, from https://bit.ly/3IzA9t1.

Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica*, *29*(3), 315-335. DOI: https://doi.org/10.2307/1909635.

Myers, D.J. (1997). Racial rioting in the 1960s: An event history analysis of local conditions. *Amer. Soc. Rev.*, *62*(1), 94-112. DOI: https://doi.org/10.4135/9781446261613.

O'Flynn, P. & Higdon, G.L. (2019). Is participatory impact investing the antidote to impact washing? Institute of Development Studies, September. Retrieved December 26, 2022, from https://bit.ly/3PFwhK2.

Otto, B., Lee, Y.W. & Caballero, I. (2011). Information and data quality in business networking: A key concept for enterprises in its early stages of development. *Electron. Mkts*., *21*(2), 83-97. DOI: https://doi.org/10.1007/s12525-011-0063-1.

Pedersen, L.H. (2021a). GameStop, short squeezes, and predatory trading. Presentation, Princeton University, February 19. Retrieved December 26, 2022, from https://bit.ly/3K2kMKp.

Pedersen, L.H. (2021b). Game on: Social networks and markets. SSRN 3794616, February 28. Retrieved December 26, 2022, from https://bit.ly/35FBBv2.

Phillips, M. & Lorenz, T. (2021) 'Dumb money' is on GameStop, and it's beating Wall Street at its own game. *New York Times*, January 27 (updated October 18).

Priya, S., Sequeira, R., Chandra, J. & Dandapat, S.K. (2019). Where should one get news updates: Twitter or Reddit? *Online Social Network. Med.*, *9*, 17-29. DOI: https://doi.org/10.1016/j.osnem/2018.11.001.

PushShift.io. (2022). Directory archive of Reddit submissions (JSON objects) for selected 2021 posts. Retrieved December 26, 2022, from https://bit.ly/3JM0qVj.

Ruan, T., Kong, Q., McBride S.K., Sethjiwala, A., Lv, Q. (2022). Cross-platform analysis of public responses to the 2019 Ridgecrest earthquake sequence on Twitter and Reddit. *Sci Rep.,* 12, 1634. DOI: https://doi.org/10.1038/s41598-022-05359-9.

Sarkar, M.B., Butler, B. & Steinfield, C. (1995). Intermediaries and cybermediaries. *J. Comp.-Mediated Comm.*, *1*(3), 3. DOI: https://doi.org/10.1111/j.1083-6101.1995.tb00167.x.

See-To, E.W.K. & Yang, Y. (2017). Market sentiment dispersion and its effects on stock return and volatility. *Electron. Mkts., 27*, 293-296. DOI: https://doi.org/10.1007/s12525-017-0254-5.

Shiller, R.J. (2013). *Irrational exuberance*, 3rd ed. Princeton University Press. DOI: https://doi.org/10.1515/9781400865536.

Shiller, R.J. (2020). *Narrative economics: How stories go viral and drive major economic events.* Princeton University Press. DOI: https://doi.org/10.2307/j.ctvdf0jm5.

Smith S.P., Johnston R.B. & Howard, S. (2011) Putting yourself in the picture: An evaluation of virtual model technology as an online shopping tool. *Inform. Sys. Res*., *22*(3), 331-369. DOI: https://doi.org/10.

1287/isre.1090.0279.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations.* Doubleday, Anchor Press. DOI: https://doi.org/10. 4324/9781912453184.

Tetlock, P.C. (2007). Giving content to investor sentiment: The role of media in the stock market. *J. Fin*, *62*(3), 1139–1168. DOI: https://doi.org/10.1111/j.1540-6261.2007.01232.x.

Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Sci.*, *211*(4481), 453-458. DOI: https://doi.org/10.1126/science.7455683.

Umar, Z., Gubareva, M., Yousaf, I. & Ali, S. (2021). A tale of company fundamentals vs. sentiment-driven pricing: The case of GameStop. *J. Beh. Exper. Fin.*, 30, 100501. DOI: https:doi.org/10.1016/j.jbef.2021.100501.

Vasileiou, E. (2021). Does the short squeeze lead to market abnormality and anti-leverage effect? Evidence from the GameStop case. SSRN 3831169. Retrieved December 26, 2022, from https://bit.ly/3HD3T6O.

Appendix

```
from bs4 import BeautifulSoup as bs
from pprint import pprint
from datetime import datetime, timedelta
import requests, re, os, json, sys
from urllib.request import urlretrieve
import pandas as pd
from selenium import webdriver
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.keys import Keys
import time , os
import warnings
warnings.filterwarnings("ignore")
```

```
init_url = 'https://np.reddit.com/r/wallstreetbets/comments/nwjw4u/gme_q1_earni
driver = webdriver.Chrome()
driver.get(init_url)
time.sleep(3)
```

time

```
currentpage = driver.page_source
soup = bs(currentpage, "html.parser")
#comments_ =soup.find_all('div', attrs={'class':'md'})
date=soup.find_all('time')
df = pd.DataFrame({'time' : [str(i) for i in date]})
df=df['time'].apply(lambda x : re.search('datetime="(.*)00"', x).group(1))
#df['time'].apply(lambda x : re.search('title="(.*)>"', x).group(1))
df
```

```
df.to_excel('202106_agian_date.xlsx', encoding = 'utf-8-sig')
```

comment

```

#soup_body = soup.find("div", {"class":"usertext-body may-blank-within md-conta
comments_ =soup.find_all('div', attrs={'class':'md'})
df2 = pd.DataFrame({'comments':comments_})
#date = soup.findAll("div",attrs={"class":"_3XoW0oYd5806Xi0r24gGdb"})

df2

```

```
df2.to_excel('202106_e06_again.xlsx', encoding = 'utf-8-sig')
```

Figure A1. Comment-Scraping Code

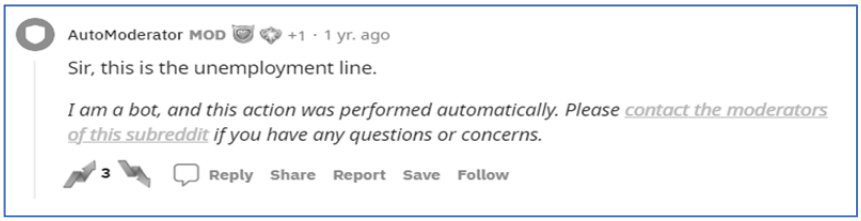


Figure A2. Reddit Bot Comments: Example

```
cmt1_text['cmt2'] = cmt1_text['comments'].str.encode('ascii', 'ignore').str.decode('ascii')
```

```
def text_str(text):
    return str(text)

cmt1_text['cmt2'] = cmt1_text['cmt2'].apply(lambda x:text_str(x))
```

```
def text_lowercase(text):
    return text.lower()

cmt1_text['lower_text'] = cmt1_text['cmt2'].apply(lambda x:text_lowercase(x))
```

```
def remove_numbers(text):
    result = re.sub(r'\d+','',text)
    return result

cmt1_text['rm_text'] = cmt1_text['lower_text'].apply(lambda x:remove_numbers(x))
```

```
def remove_punctuation(text):
    translator = str.maketrans('','',string.punctuation)
    return text.translate(translator)

cmt1_text['punc_text'] = cmt1_text['rm_text'].apply(lambda x:remove_punctuation(x))
```

```
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

def rm_stopwords(text):
    stop_words = set(stopwords.words("english"))
    word_tokens = word_tokenize(text)
    filtered_text = [word for word in word_tokens if word not in stop_words]
    return filtered_text

cmt1_text['tk_text'] = cmt1_text['rmws_text'].apply(lambda x:rm_stopwords(x))
```

```
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()

def lmt_word(text):
    lemmas = [lemmatizer.lemmatize(word, pos='v') for word in text]
    return lemmas

cmt1_text['lmt_text'] = cmt1_text['tk_text'].apply(lambda x:lmt_word(x))
```

Figure A3. Data preprocessing code

Data Preprocessing	Definition
Expand contractions	A contraction is a shortened form of a word; an expansion action is `don't` → `do not`
Lower case interpretation	It is easy for a machine or NLP software to parse words in different cases.
Remove words	*Word-number combinations* in text create problems for text mining to interpret. *Stop words* often occur but don't add information (e.g., "`the`," "`is`," "`or`"), so are filtered out too.
Remove punctuation	Text processing technique that removes punctuation as it parses the meaning of some text.
Tokenization	Splits long strings of text into smaller pieces called *tokens* to identify the meaning of its units.
Stemming & lemmatization	Process to reduce a word to its root (e.g., "`running`" or "`ran`" result in the root "`run`."

Table A1a. Text preprocessing actions and definitions

Action	Source and Text Samples	
Expand contractions	Source:\nhttps://www.reddit.com/r/gaming/comme … #Good morning traders and investors of the r/w…	**edit:tldr-buy amzn get tendies - See comm … Just wanted to hear why you are particularly b …
Lower case interpretation	source:\nhttps://www.reddit.com/r/gaming/comme ... #good morning traders and investors of the r/w ...	**edit:tldr-buy amzn get tendies -see comm ... just wanted to hear why you are particularly b ...
Remove words	source:\nhttps://www.reddit.com/r/gaming/comme … #good morning traders and investors of the r/w …	**edit:tldr-buy amzn get tendies -see comm … just wanted to hear why you are particularly b …
Remove punctuation	Source:\nhttpswww.reddit.comrgamingcommentsetnxtw... good morning traders and investors of the rwal ...	edit tldr-buy amzn get tendies see comments ... just wanted to hear why you are particularly b ...
Tokenization	[source, httpswww.reddit.comrgamingcommentsetnxtw... [good, morning, traders, investors rwallstree ...	[edit, tldr, buy, get, tendies, see, com ... [wanted, hear, particularly, bullish, amd, gai ...
Stemming and lemmatization	[source, httpswww.reddit.comrgamingcommentsetnxtw ... [good, morning, traders, investors rwallstree ...	[edit, tldr, buy, get, tendies, see, com ... [wanted, hear, particularly, bullish, amd, gain ...

Notes. In Reddit, "*edit tl-dr-*" codes for a posted messaged that was "too long" so the intended recipient(s) "didn't read" it.

Table A1b. Text preprocessing actions with data source URLs and text samples returned

Jan 1, 2020			Jan 5, 2020			Jan 9, 2020		
(1)	*Neu*	0.504	(1)	*Neu*	0.488	(1)	*Neu*	0.494
(2)	*Neg*	0.252	(2)	*Neg*	0.274	(2)	*Neg*	0.272
(3)	*Pos*	0.244	(3)	*Pos*	0.238	(3)	*Pos*	0.234
Jan 2, 2020			Jan 6, 2020			Jan 10, 2020		
(1)	*Neu*	0.497	(1)	*Neu*	0.501	(1)	*Neu*	0.58
(2)	*Neg*	0.279	(2)	*Neg*	0.276	(2)	*Neg*	0.232
(3)	*Pos*	0.225	(3)	*Pos*	0.222	(3)	*Pos*	0.190
Jan 3, 2020			Jan 7, 2020			Jan 11, 2020		
(1)	*Neu*	0.497	(1)	*Neu*	0.529	(1)	*Neu*	0.582
(2)	*Neg*	0.271	(2)	*Neg*	0.268	(2)	*Neg*	0.231
(3)	*Pos*	0.233	(3)	*Pos*	0.202	(3)	*Pos*	0.186
Jan 4, 2020			Jan 8, 2020					
(1)	*Neu*	0.500	(1)	*Neu*	0.518	–		
(2)	*Neg*	0.289	(2)	*Neg*	0.282			
(3)	*Pos*	0.211	(3)	*Pos*	0.201			

Table A2. Example of sentiment analysis results