IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 March 2025, accepted 28 March 2025, date of publication 3 April 2025, date of current version 16 April 2025.

*Digital Object Identifier 10.1109/ACCESS.2025.3557460*

RESEARCH ARTICLE

# A Statistical Analysis of the Relationship Between Meme Stocks and Social Media

**SEUNGJU LEE[1,*], YUNYOUNG LEE[2,*], JAEWOOK LEE[1,3], AND HOKI KIM[4]**

[1]Department of Industrial Engineering, Seoul National University, Seoul 08826, South Korea
[2]Department of Artificial Intelligence and Data Science, Sejong University, Seoul 05006, South Korea
[3]Institute of Engineering Research, Seoul National University, Seoul 08826, South Korea
[4]Department of Industrial Security, Chung-Ang University, Seoul 06974, South Korea

Corresponding author: Hoki Kim (hokikim@cau.ac.kr)

This work was supported in part by the Chung-Ang University Research Grants, in 2024; and in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) through the Artificial Intelligence Semiconductor support program to Nurture the Best Talents grant funded by Korean Government [Ministry of Science and Information and Communications Technology (MSIT)] under Grant IITP-2024 (2025)-RS-2023-00266605.

*Seungju Lee and Yunyoung Lee contributed equally to this work.

**ABSTRACT** Meme stocks, driven by viral social media trends, have added new complexities to financial markets. Prior studies have explored meme stock price dynamics and investor sentiment, but the interplay between social media activity and market movements, along with the structural and linguistic features of online discussions, remains understudied. To address this, we integrate econometric and NLP-based techniques, combining correlation analysis, Granger causality testing, BSADF-based bubble detection, and textual analysis. Our results reveal a strong correlation between trading volume and social media engagement, with Granger causality confirming a feedback loop between market fluctuations and online discussions. BSADF analysis demonstrates that social media-based detection complements price-based methods by identifying explosive periods they may miss. Additionally, network analysis indicates that meme stock discussions exhibit distinct structural patterns, while linguistic analysis highlights unique word choices and emoji usage. Sentiment analysis shows that bullish sentiment dominates during speculative surges, reinforcing the emotionally driven nature of meme stock trading. These findings provide investors with a complementary tool for risk assessment by integrating sentiment with traditional market indicators, while helping regulators monitor online sentiment to identify early signs of speculative excess and market instability.

**INDEX TERMS** Behavioral finance, investor sentiment, meme stocks, natural language processing, social media.

## I. INTRODUCTION

The financial market has traditionally relied on fundamental analysis, economic indicators, and institutional investors to guide investment decisions. Established valuation models and company performance metrics have played a crucial role in shaping investors' choices. However, the rise of big data and digital connectivity has reshaped market dynamics, empowering retail investors with unprecedented access to financial information. This shift has broadened the scope of investment decision-making, allowing individuals to consider alternative factors beyond traditional fundamental analysis [1], [2].

At the forefront of this transformation is social media's growing role in financial markets, where platforms such as Reddit, Twitter, and StockTwits enable rapid information dissemination and real-time investor engagement [3], [4], [5]. These platforms amplify the collective power of retail investors, facilitating coordinated actions that challenge traditional market structures. As a result, social media-driven trading has introduced new sources of volatility, disrupting conventional market dynamics and reshaping investor behavior [6].

This paradigm shift has given rise to the phenomenon of *meme stocks*, where online sentiment, rather than traditional

The associate editor coordinating the review of this manuscript and approving it for publication was Maria Chiara Caschera.

fundamentals, drives extreme price fluctuations. A prominent example is GameStop, which became the focal point of a stock market frenzy in early 2021, demonstrating the power of retail investors mobilized through social media. These price movements challenge conventional financial theories such as the Efficient Market Hypothesis (EMH) [7], which assumes asset prices fully reflect all available information. Instead, meme stock dynamics exhibit characteristics explained by Prospect Theory [8], as retail investors display herd-like trading, overconfidence, and a tendency to take speculative risks in pursuit of outsized gains. These deviations from rational investing indicate that traditional models may be insufficient to capture the speculative nature of meme stocks. As a result, alternative analytical approaches that incorporate social media signals are essential for understanding the dynamics and risks associated with meme stocks.

In this paper, we analyze the relationship between meme stocks and social media, focusing on five widely studied examples: AMC Entertainment Holdings, Inc. (AMC), BlackBerry Limited (BB), Bed Bath & Beyond Inc. (BBBY), GameStop Corp. (GME), and Nokia Oyj (NOK). While prior studies have examined meme stock volatility and the role of social media, they predominantly analyze price movements or sentiment-price relationships without fully capturing the mechanisms underlying meme stock dynamics. In contrast, our study employs a multifaceted approach by applying econometric techniques to social media data, extending bubble detection beyond price, and analyzing the linguistic and structural characteristics of meme stock discussions.

Our findings reveal several key insights. First, we establish a strong correlation between trading volume and social media engagement, further validated through Granger causality tests, which indicate a bidirectional influence between online discussions and market activity. Second, we introduce a novel application of the Backward Sup-ADF (BSADF) test to social media data, revealing explosive periods that price-based approaches fail to capture, highlighting the predictive power of social media trends and potential as early indicators of speculative interest. Third, we apply network visualization and NLP techniques to analyze meme stock discussions at a structural level, identifying distinct linguistic characteristics, thematic clusters, and the unique distribution of emojis in bullish and bearish discourse. Finally, we provide quantitative evidence of sentiment dominance during explosive periods, revealing how collective bullishness fuels speculation. By integrating these approaches, our study provides a quantitative framework for understanding how meme stocks are shaped by social media, offering insights that extend beyond the existing literature.

## II. RELATED WORKS

Research on meme stocks has primarily examined their volatility, speculative nature, and detachment from traditional market structures. Specifically, Costola et al. [9] examined the momentum effects of meme stocks, while Yousaf et al. [10] and Aloosh et al. [11] detected price bubbles and interconnectedness across speculative assets. Other studies highlight their weak ties to broader financial markets. Elsayed et al. [12] found meme stocks to be largely disconnected from sectoral markets, while Vasileiou [13] provided evidence that GameStop's short squeeze violated the Efficient Market Hypothesis (EMH), exhibiting abnormal returns and an anti-leverage effect where volatility increased with rising prices. Furthermore, Umar et al. [14] demonstrated that retail investor herding causes bubble behavior to spread across heavily shorted stocks, amplifying systemic risk. While these studies underscore the destabilizing effects of meme stocks, they focus primarily on price contagion and volatility patterns, leaving the role of social sentiment as a driver of speculation largely unexplored.

Growing evidence suggests that social media plays a pivotal role in meme stock rallies, particularly through platforms like Reddit's r/WallStreetBets (WSB). Studies have shown that heightened discussions on WSB contribute to increased trading volume and price volatility in GameStop [15], with retail investors displaying greater risk-taking behavior [16], [17]. Additionally, sentiment and engagement surrounding GME intensified prior to major price surges, suggesting that social media activity may act as an early indicator of speculative trading [18]. Zheng et al. [19] further demonstrated that WSB interactions became more centralized and sentiment grew increasingly positive during the short squeeze, reinforcing the connection between collective behavior and price dynamics.

Beyond trading activity, sentiment-driven factors, including social media activity, put–call ratios, and short-sale volume, have been shown to significantly influence GameStop's returns, contributing to price surges and market inefficiencies [20]. However, the influence of social media varies among investors. Pandey and Guillemette [21] found that its impact is strongest among inexperienced traders but diminishes for investors with greater financial knowledge, suggesting that seasoned investors rely less on sentiment-driven signals.

The broader effects of social-media-fueled speculation have also been explored. Klein et al. [22] observed that retail traders became more familiar with financial instruments during the GameStop short squeeze, suggesting that social media facilitated rapid self-education in trading mechanics. In contrast, Bradley et al. [23] found that WSB's culture shifted following the event, leading to a decline in the quality of investment discussions and disproportionately disadvantaging smaller investors. As retail trading communities evolved post-squeeze, concerns emerged about the sustainability of sentiment-driven investing. This raises questions about whether social media-based trading strategies can consistently generate excess returns.

Meanwhile, the effectiveness of social media-based trading strategies remains contested. Chacon et al. [24] found no consistent trading advantage in strategies based on WSB discussions, questioning their long-term viability. Similarly, Wang et al. [25] demonstrated that while sentiment data

**TABLE 1.** Summary statistics of price, volume, and sentiment data for each meme stock.

| | AMC | BB | BBBY | GME | NOK |
|---|---|---|---|---|---|
| **Highest/Lowest Price** | $62.55/$1.98 | $25.10/$3.18 | $52.89/$1.31 | $86.88/$1.04 | $6.55/$3.27 |
| **Average Daily Volume** | 65,809,336 | 17,004,875 | 15,914,153 | 34,006,311 | 33,919,338 |
| **Number of Twits (#Twits)** | 7,531,954 | 471,327 | 367,345 | 1,385,418 | 199,400 |

improves price prediction for traditional stocks, its predictive power weakens for meme stocks due to extreme volatility and the dominance of a few influential voices within the WSB community, underscoring the complexity of sentiment-driven speculation. These findings highlight the limitations of sentiment as a predictive tool and underscore the need for a deeper analysis of the structural and linguistic patterns of meme stock discussions, beyond simple sentiment correlations.

While these studies underscore the significant role of social media in meme stock activity, they primarily examine price correlations and sentiment-based trading behaviors. Limited attention has been given to the underlying structural and linguistic patterns of meme stock discourse or to the application of econometric techniques beyond price analysis to detect speculative bubbles. Our study fills this gap by integrating multiple analytical approaches, examining the distinct textual and sentiment characteristics of meme stock discussions, and extending financial bubble detection techniques beyond traditional price-based methods.

## III. METHODOLOGY

### A. DATA

We collected price and trading volume data from *Yahoo Finance* and sentiment data from *StockTwits* for each of the selected meme stocks. The dataset spans August 3, 2020, to February 28, 2023, offering a comprehensive perspective on stock performance and sentiment dynamics.

StockTwits was selected as our sentiment source due to its structured format, where users explicitly classify their posts as bullish or bearish. Unlike broader social media platforms such as Twitter or Reddit, which require NLP-based sentiment inference, StockTwits provides self-reported investor sentiment, reducing classification ambiguity. StockTwits posts, known as "Twits", are focused on stock discussions, with users tagging stocks using a "$" symbol (e.g., "$GME"). While sentiment labels may not always perfectly align with a user's overall outlook, we treat them as a direct measure of self-reported investor sentiment, ensuring a transparent and structured sentiment analysis. Furthermore, StockTwits has been widely used in financial research, demonstrating its validity as a sentiment-based market analysis tool. Prior studies have leveraged StockTwits data to analyze sentiment-driven price movements, investor behavior, and market reactions, reinforcing its credibility [26], [27], [28].

By analyzing a total of 9.95 million Twits, we conducted an in-depth examination of sentiment dynamics across meme stocks. The detailed statistics, including price, trading volume, and sentiment distribution, are presented in Table 1.

### B. MARKET-SOCIAL MEDIA INTERACTION ANALYSIS

To investigate the relationship between meme stock activity and social media engagement, we employ correlation analysis and Granger causality testing. These methods allow us to quantify associations and assess the direction of influence, providing insights into how online discussions interact with market movements.

We first compute Pearson correlation coefficients to measure the relationship between closing price, trading volume, and the number of Twits (#Twits). To evaluate whether meme stocks exhibit a stronger link between trading activity and social media engagement compared to traditional markets, we compare their correlation coefficients with those of the Standard & Poor's 500 index (SPY). This comparison is conducted using the Fisher test [29], which assesses the statistical significance of differences between correlation coefficients, and the Zou test [30], which evaluates the confidence intervals of these differences at a significance level of $\alpha = 0.05$.

To establish causal relationships, we apply the Granger causality test [31], which determines whether past values of one time series improve the prediction of another. Given the results of the Augmented Dickey-Fuller (ADF) test [32], we difference the price series to ensure stationarity, analyzing price changes rather than raw prices. We estimate autoregressive models for price changes, trading volume, and the number of Twits, selecting the optimal lag length using the Bayesian Information Criterion (BIC). This approach allows us to test whether market movements drive social media discussions or vice versa. A statistically significant F-statistic indicates the presence of Granger causality at the 0.05 level.

By comparing results between meme stocks and SPY, we assess whether social media plays a disproportionately larger role in meme stock trading dynamics. This analysis provides a systematic approach to understanding the interplay between online engagement and speculative market behavior.

### C. BACKWARD SUP-ADF

The Backward Sup-Augmented Dickey-Fuller (BSADF) statistic, introduced by Phillips et al. [33], extends the conventional ADF test to detect multiple episodes of explosive behavior in financial time series. Unlike the standard ADF test and the Phillips-Perron test [34], which evaluate stationarity over the full sample, the BSADF test dynamically identifies time-localized explosive behavior, making it more effective at detecting speculative bubbles. Compared to Markov-switching models [35], which identify broad regime shifts based on discrete state probabilities, the BSADF test pinpoints short-lived and recurrent explosive dynamics. Its non-parametric, data-driven approach allows for real-time detection of speculative surges, making it better suited for capturing meme stock bubbles. Additionally, compared to alternative approaches such as log-periodic power law (LPPL) models [36], the BSADF test does not impose a specific parametric structure on bubble formation, offering a more flexible and data-driven methodology. This makes it a valuable tool for analyzing financial markets, where speculative bubbles can emerge and collapse over short time spans. Several studies have successfully applied this method to detect bubbles in various financial markets [37], [38], [39], [40].

Following the methodology of Phillips et al. [33], we implement the BSADF test by repeatedly estimating the ADF regression within rolling subsamples to assess whether explosive dynamics emerge at different points in time. This recursive approach ensures that explosive episodes are identified locally rather than assuming a single structural break across the entire sample. We begin by defining the log price process at time $t$ as $P_t = \log S_t$, where $S_t$ represents the observed stock price. The ADF regression model is then estimated as follows:

$$\Delta P_t = \alpha + \beta P_{t-1} + \sum_{i=1}^{k} \psi^i \Delta P_{t-i} + \epsilon_t, \tag{1}$$

where $\Delta P_t = P_t - P_{t-1}$ represents the first difference of the log price, capturing the stock's log return. The term $\alpha$ is a constant, while $\beta$ is the coefficient that determines whether the process exhibits explosive behavior. The number of lag terms, $k$, is selected using the Bayesian Information Criterion (BIC) to control for serial correlation. The coefficients $\psi^i$ capture the effects of lagged first differences, and the error term $\epsilon_t$ is assumed to be normally distributed. The parameters $\alpha$, $\beta$, and $\psi^i$ are estimated using ordinary least squares.

To detect explosive behavior, we compute the BSADF statistic by applying the ADF test over backward-expanding rolling windows. Let $r_1$ and $r_2$ denote the start and end points of a rolling window, respectively. The BSADF statistic at time $r_2$ is given by:

$$\text{BSADF}_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \text{ADF}_{r_1}^{r_2}, \tag{2}$$

where $r_0$ represents the minimum fraction of the total sample size used in the recursive estimation process, ensuring sufficient window length for stable estimation. The term $\text{ADF}_{r_1}^{r_2}$ represents the ADF statistic estimated over the subsample with the start point $r_1$ and the end point $r_2$. By iteratively estimating the ADF statistic over backward-expanding rolling windows and selecting the maximum value, the BSADF test identifies the strongest evidence of explosiveness at each time point, capturing localized speculative bubbles. In this study, we determine $r_0$ as follows, following the style of Phillips et al. [33]:

$$r_0 = 0.01 + \frac{1.8}{\sqrt{T}}, \tag{3}$$

where $T$ represents the sample size. This specification ensures that the estimation window is sufficiently large to maintain statistical power while allowing the test to detect localized explosive behavior.

To determine whether a period is explosive, a critical value is calculated based on the distribution of the maximum ADF statistic obtained from a Monte Carlo simulation using 2,000 replications. A period $t$ is classified as explosive if the computed BSADF statistic exceeds the simulated critical value at the chosen significance level.

### D. NATURAL LANGUAGE PROCESSING

Natural language processing (NLP) encompasses systems that interpret human language and extract valuable information from text or speech. Deep learning-based models have recently shown promising performance in NLP tasks [41]. One of the key components of NLP is text embedding, which refers to the process of transforming sentences into numerical vectors in a high-dimensional space. Recent transformer-based models, which achieve near-human-level performance in various tasks, include text embedding techniques that capture semantic and syntactic relationships between words and enable machines to understand and reason about textual data [42], [43], [44]. For our study, we utilize a transformer-based model called Twitter-RoBERTa [45], which has demonstrated high accuracy in sentiment classification tasks. The model is publicly accessible on *Hugging-Face*.

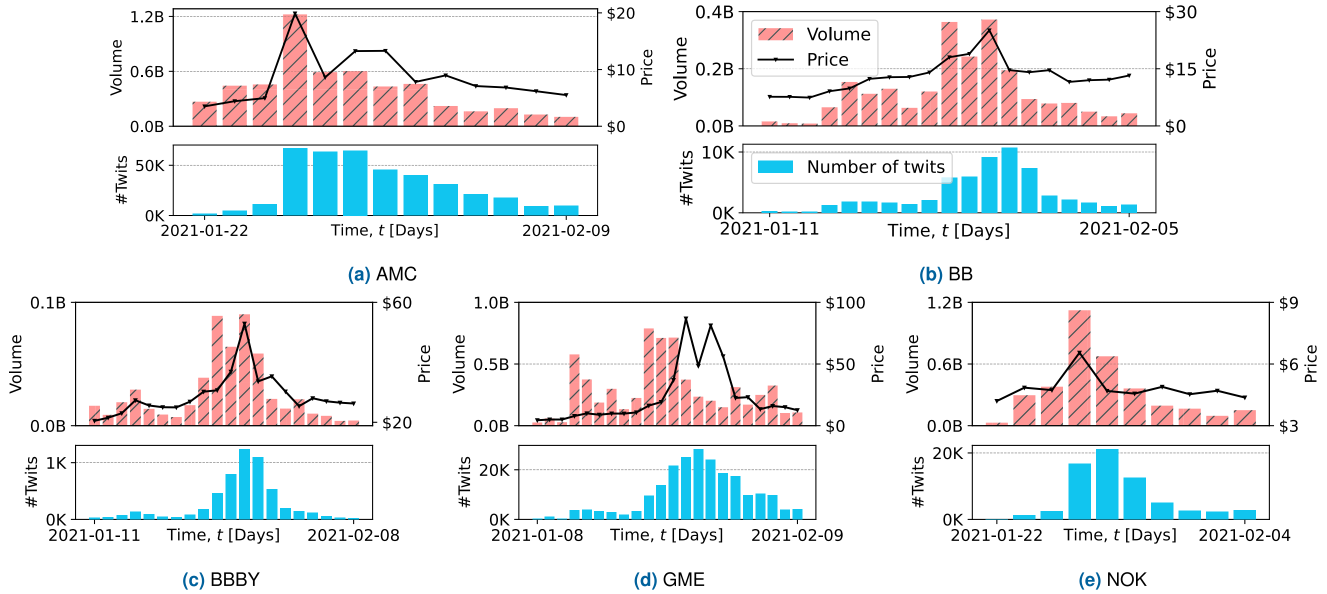
### E. NETWORK VISUALIZATION

Network visualization refers to the process of graphically representing the structure and connections of a network. It is widely utilized in fields like social network analysis and transportation to understand the patterns, relationships, and properties of complex networks [46], [47], [48]. To facilitate network visualization, we utilize the Netwulf Python library [49], which is specifically designed for the interactive visualization of networks. Within the context of natural language processing, we represent each feature of the Twits as a node in the network. The edges between nodes are constructed based on the calculated cosine similarity between each pair of Twits. This approach allows us to visually explore the relationships and similarities among the Twits' features.

## IV. EXPERIMENTAL RESULTS

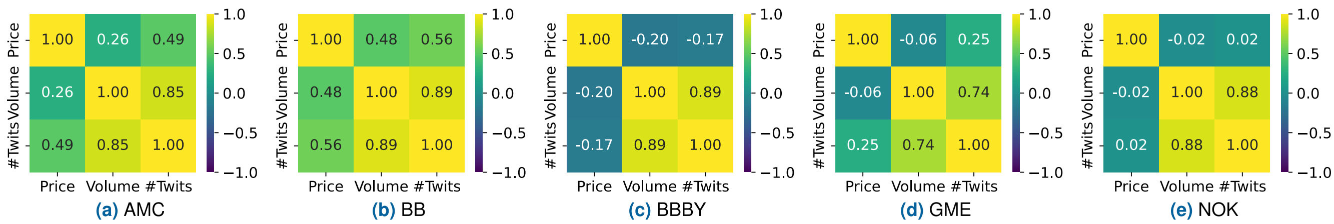
### A. MARKET-SOCIAL MEDIA INTERACTION ANALYSIS

Figure 1 illustrates the daily trends of closing price, volume, and the number of Twits during the period from January 2021 to February 2021. The specific timeframe for each stock is derived from our analysis of explosiveness, which will




(a) AMC (b) BB (c) BBBY (d) GME (e) NOK

**FIGURE 1. Daily trends of closing price, trading volume, and the number of Twits for each meme stock during their explosive period from January 2021 to February 2021. The specific timeframe for each stock is determined based on our analysis, calculating the BSADF statistic on price movements. We discover a pronounced correlation between trading volume and the number of Twits.**




(a) AMC (b) BB (c) BBBY (d) GME (e) NOK

**FIGURE 2. Pearson correlation coefficients between daily closing price, trading volume, and the number of Twits for meme stocks. While price shows inconsistent and low correlations with other variables, a strong positive correlation is consistently observed between trading volume and the number of Twits (#Twits) for all stocks.**

be described in more detail later. While price and volume may not exhibit a strong correlation, we observe similar fluctuations between trading volume and the number of Twits (#Twits). This suggests a potential relationship between market activity and the level of online engagement within the meme stock community.

In Figure 2, we further explore the correlations between different variables. The Pearson correlation coefficients reveal that the correlations between price and other factors remain relatively weak. However, the correlation between trading volume and the number of Twits displays a stronger positive relationship, with correlation coefficients exceeding 0.7. This indicates a notable association between the level of online discussions and the trading volume of meme stocks.

To evaluate whether meme stocks exhibit a stronger correlation between trading volume and the number of Twits than SPY, we compare their correlation coefficients using statistical tests. The results of these tests are summarized in Table 2. The Fisher test reveals that, with the exception of GME, meme stocks demonstrate a statistically significant higher correlation coefficient. Additionally, the Zou test indicates that the calculated differences for all meme stocks fall within the confidence intervals, further supporting the conclusion of a statistically higher correlation compared to SPY. These findings highlight the pronounced influence of social media on meme stock trading behavior.

To further investigate the relationship between meme stock trading activity and social media engagement, we apply the Granger causality test. Table 3 presents the F-statistics for each causal direction, with statistically significant results (p-value $< 0.05$) marked by an asterisk.

The results reveal a strong bidirectional relationship between meme stock trading activity and social media discussions. In most cases, trading volume, price changes, and the number of Twits Granger-cause one another, suggesting a reinforcing feedback loop where market fluctuations drive online engagement and heightened social media activity influences trading behavior. This aligns with the correlation findings, further emphasizing that meme stock dynamics are closely tied to investor sentiment expressed through social media.

In contrast, SPY follows a more conventional market pattern. While price changes and trading volume Granger-cause social media discussions, the reverse relationship is absent,

**TABLE 2. Statistical comparison of correlation coefficients between trading volume and the number of twits for meme stocks versus SPY (S&P 500 index). The Fisher test [29] evaluates whether each meme stock has a higher correlation coefficient than SPY, while the Zou test [30] calculates the difference between the correlation coefficient (Diff.) and the corresponding confidence intervals (CI) at a significance level of $\alpha = 0.05$. Results indicate that all meme stocks exhibit significantly higher correlations between trading volume and the number of twits compared to SPY according to both the Fisher test (except GME) and the Zou test.**

| | AMC | | BB | | BBBY | | GME | | NOK | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Statistic | Sig. | Statistic | Sig. | Statistic | Sig. | Statistic | Sig. | Statistic | Sig. |
| **Fisher** [29] | 9.358 | 0.000* | 9.686 | 0.000* | 10.445 | 0.000* | 1.611 | 0.107 | 7.389 | 0.000* |
| | Diff. | CI | Diff. | CI | Diff. | CI | Diff. | CI | Diff. | CI |
| **Zou** [30] | 0.194 | (0.15, 0.24) | 0.198 | (0.16, 0.24) | 0.207 | (0.17, 0.25) | 0.053 | (0.00, 0.11) | 0.167 | (0.12, 0.21) |

* This is a lower bound of the true significance with a significance level of $\alpha = 0.05$.

**TABLE 3. Granger causality test results for meme stocks and SPY. Each cell reports the F-statistic from the Granger causality test. An asterisk (*) indicates statistical significance at the 0.05 level. The results show that meme stocks exhibit a stronger bidirectional relationship with social media engagement compared to SPY, suggesting a greater influence of online activity on their market dynamics.**

| Direction of Causality ($\longrightarrow$) | | AMC | BB | BBBY | GME | NOK | SPY |
|---|---|---|---|---|---|---|---|
| Price Change | Volume | 3.249* | 10.872* | 0.560 | 5.590* | 0.187 | 4.969* |
| Price Change | # Twits | 4.503* | 33.056* | 3.405* | 20.931* | 3.565* | 7.212* |
| Volume | Price Change | 19.326* | 14.496* | 6.492* | 24.968* | 14.208* | 0.391 |
| Volume | # Twits | 23.774* | 15.761* | 5.975* | 5.619* | 106.458* | 5.982* |
| # Twits | Price Change | 7.638* | 11.839* | 6.205* | 16.297* | 35.168* | 1.344 |
| # Twits | Volume | 2.137 | 5.123* | 0.930 | 10.082* | 58.901* | 1.784 |

indicating that social media reacts to market movements rather than shaping them. This distinction reinforces the idea that meme stocks are uniquely susceptible to speculative trading fueled by online engagement, whereas traditional assets like SPY remain more insulated from such influences.

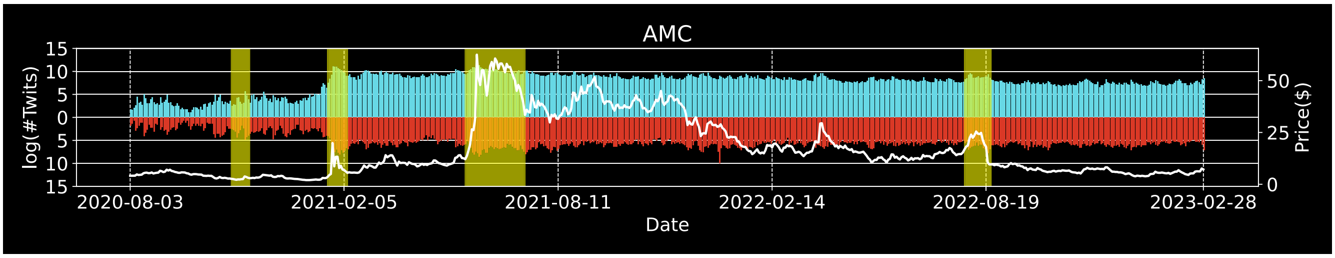
These findings underscore the reliability of social network data in capturing investor behavior and activities in the market for meme stocks. The strong correlation between the number of Twits and trading volume suggests that investors actively engage with social media as a key source of information when navigating meme stock volatility. Herbert Simon's theory of decision-making provides a useful framework for interpreting this behavior—when faced with uncertainty, investors seek additional information to refine their strategies. The bidirectional Granger causality results further reinforce this notion, revealing a feedback loop in which trading activity fuels online discussions, which in turn influence market movements. This interplay between market behavior and social sentiment distinguishes meme stocks from traditional assets like SPY, where social media reacts to market changes rather than driving them. Collectively, these results highlight the role of social media in shaping meme stock dynamics, reinforcing its function as both a reflection of and a catalyst for speculative trading activity.

### B. EXPLORING SOCIAL MEDIA'S ROLE IN IDENTIFYING PERIODS OF EXPLOSIVENESS IN MEME STOCKS

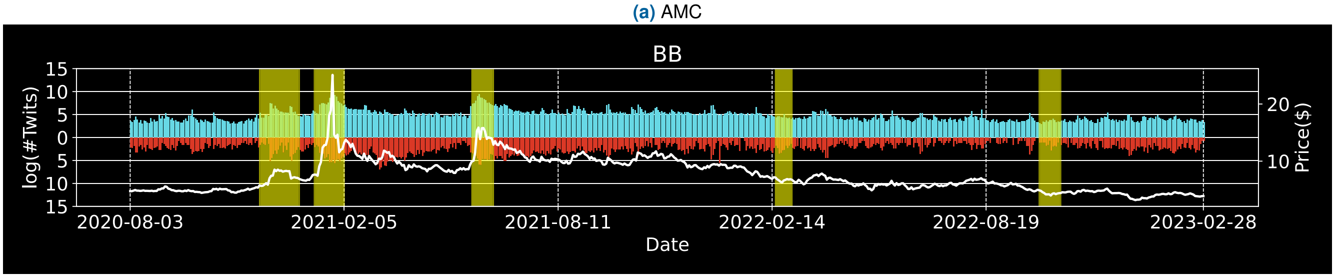
Given the strong correlation observed between the volume of social media content and the trading volume of meme stocks, along with the bidirectional causality identified through Granger tests, we delve deeper into the significance of social media in analyzing these stocks. In particular, we investigate the potential use of social media in identifying periods of explosiveness, which are a defining characteristic of meme stocks. Precisely identifying explosive periods is crucial for understanding meme stock behavior, as it enables investors to retrospectively understand the magnitude of price surges, assess the impact of market sentiment, and gain insights into the unique dynamics and characteristics of these stocks. As a first step in our analysis, we assess the effectiveness of the BSADF statistic-based method, applied to price movements, in detecting these periods across meme stocks.

Figure 3 visually presents the periods of explosiveness identified based on the BSADF statistic for price movements. The white line indicates the series of closing prices, while the blue and red bars represent the logarithm of the number of Bullish and Bearish Twits, respectively. The highlighted region in yellow indicates the explosive periods characterized by heightened price variance. The visualization of explosive periods includes consecutive dates connected by a maximum segment length of five, providing a clearer representation. Our experimental results are consistent with previous studies [10], [11], highlighting multiple instances of explosiveness in meme stocks and revealing periods of exceptionally high price volatility. Specifically, we observed widely known explosive periods across these stocks, such as January-February 2021 and May-June 2021 [50].

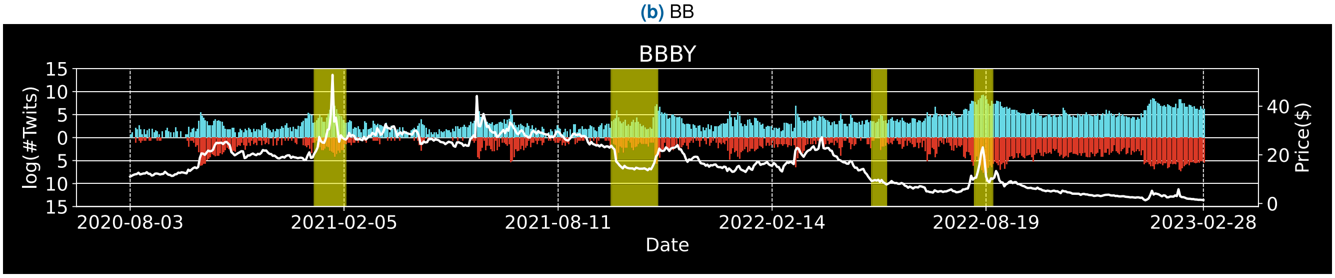
However, it is important to acknowledge that the BSADF statistic based on price movements did not capture certain periods of interest or notable events. For instance, the BSADF analysis on price movements failed to capture the bankruptcy warning for BBBY in January 2023, which resulted in a significant decline of over 22%, followed by a rapid rebound of up to 35% in share prices. Similarly, the explosive period



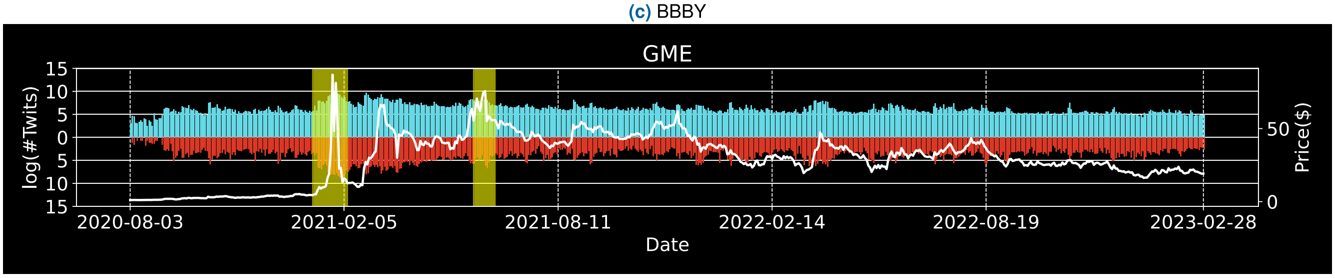

(a) AMC



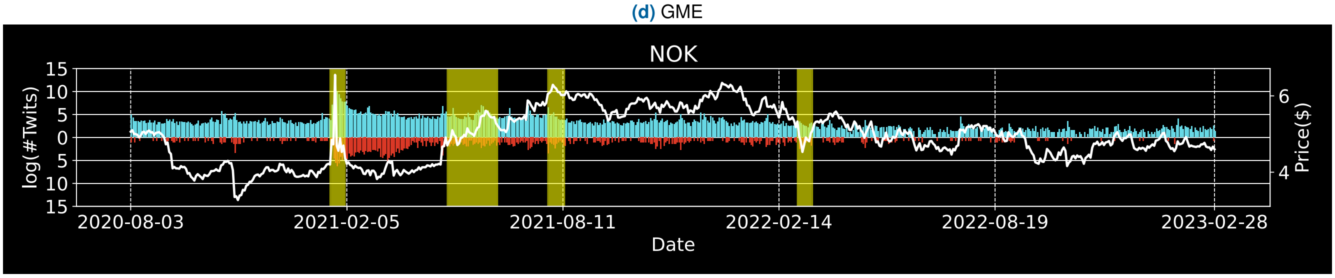

(b) BB




(c) BBBY




(d) GME

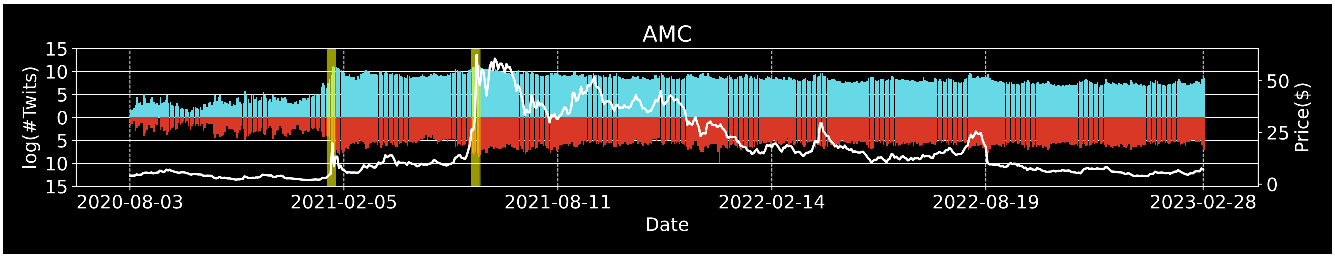



(e) NOK

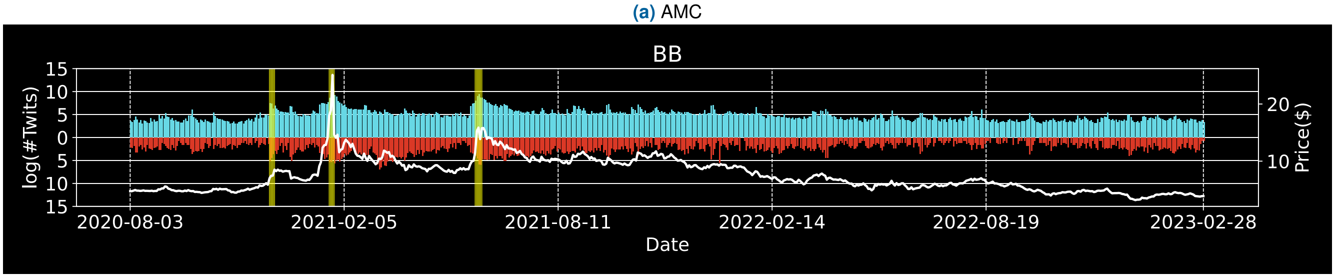
**FIGURE 3. Time series of the daily closing price (white line) and the number of Bullish (blue) and Bearish (red) Twits on StockTwits, with yellow regions highlighting explosive periods identified based on price movements. While this method captures some important periods, it fails to capture certain notable events, such as the bankruptcy warning for BBBY in January 2023.**

observed in May-June 2021 was also not captured for BBBY either. These instances highlight the limitations of relying solely on historical price data to capture and predict certain unique events and market dynamics.

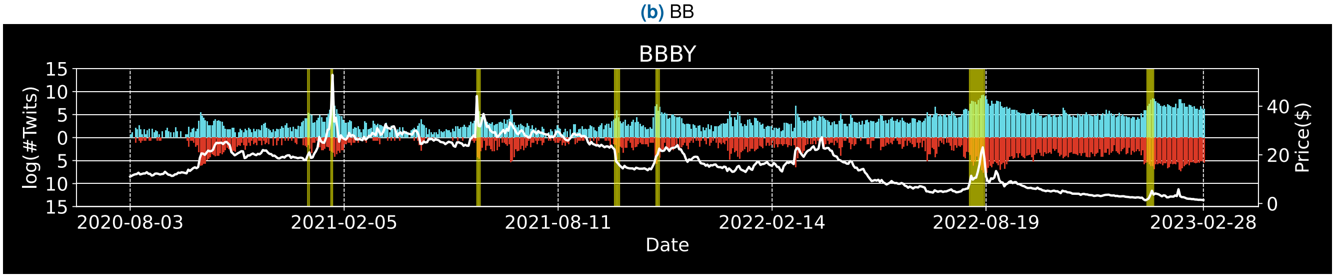
The limitations of the BSADF statistic based on price movements stem from its reliance on aggregated daily prices, such as the closing price, which may not fully capture the intraday dynamics and rapid movements in meme stocks. These stocks are characterized by sudden surges and declines within a single day, necessitating the inclusion of more granular and real-time data to accurately capture these explosive movements. Additionally, it is important to



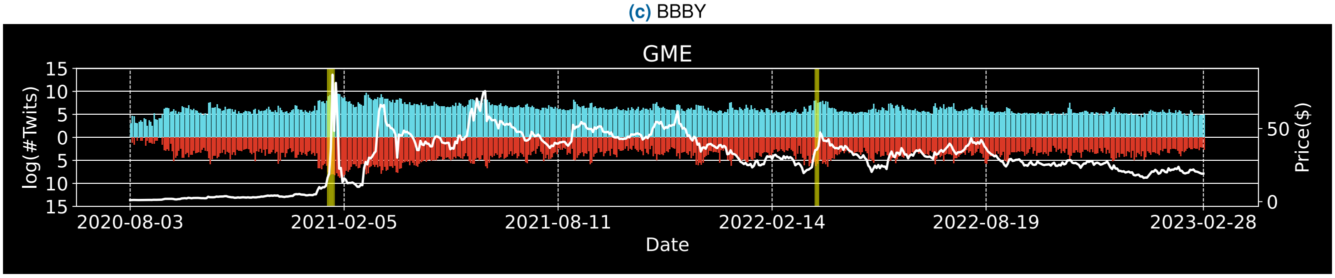

(a) AMC



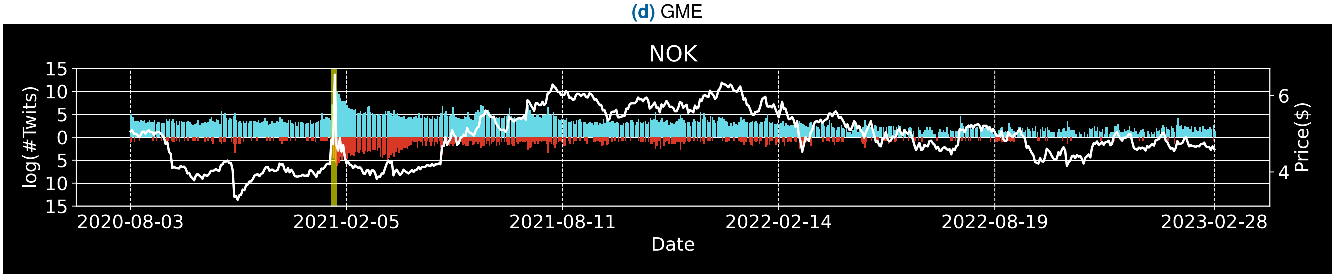

(b) BB




(c) BBBY




(d) GME




(e) NOK

**FIGURE 4. Time series of the daily closing price (white line) and the number of Bullish (blue) and Bearish (red) Twits on StockTwits, with yellow regions highlighting explosive periods identified based on the number of Twits. Analyzing the change in the number of Twits allows us to uncover previously unidentified periods associated with notable events, which cannot be observed solely through price movements.**

recognize that rapid stock price movements can occur without any specific news or events [51], [52], adding complexity to the understanding of market dynamics. Moreover, the presence of diverse market participants with varying positions can result in less dramatic price movements, even in the presence of significant news or events.

To address the limitations and further explore the influence of social media on meme stocks, we expand the application of the BSADF statistic to analyze the volume of Twits on social media. By considering the cumulative number of Twits throughout the trading day, our objective is to extract valuable insights into the popularity and dynamics of meme stocks, providing a complementary perspective to the analysis based on historical price data. This supplementary utilization of the BSADF statistic with Twit volume allows us to delve deeper into the explosive behavior exhibited by meme stocks in response to online discussions and social media activity.

Figure 4 presents the results of applying the BSADF statistic to Twit volumes, offering a visual representation of the observed explosive patterns in meme stocks based on social media activity. The results demonstrate the effectiveness of the model with Twit volumes in capturing previously unrecognized periods characterized by significant price movements, including the early 2023 period and the common explosive period in May 2021 for BBBY. This illustrates that the inclusion of Twit volumes in the analysis provides valuable insights into the popularity and dynamics of meme stocks, highlighting the substantial impact of social media activity on meme stock price movements. However, it is important to acknowledge that this extended model does not consistently identify all the explosive periods captured using price information alone. This limitation suggests that the dynamics of explosive behavior in meme stocks are influenced by various factors.

Therefore, our findings demonstrate that by analyzing social media as a supplementary measure, we can achieve a more comprehensive understanding of the underlying dynamics that contribute to explosive periods in meme stocks. This integrated approach allows us to capture a broader range of explosive periods and provides a clearer interpretation of the complex interplay between market sentiment, social media trends, and price movements in meme stocks. Consequently, this comprehensive analysis enhances our ability to detect and analyze explosive periods, offering a more robust framework for understanding the unique dynamics of meme stocks, with social media playing a pivotal role.

### C. CONTEXTUAL PATTERNS IN SOCIAL MEDIA DISCUSSIONS ABOUT MEME STOCKS

We now shift our focus to exploring how social media itself reflects the distinct characteristics of meme stocks. To gain deeper insights into the dynamics and interactions within the online community discussing meme stocks, we employ network analysis on the Twit messages. To construct the network, we utilize the Twits from the common explosive period (January 22, 2021, to February 4, 2021). In this analysis, we transform all Twits into embedding vectors using the transformer-based language model of Twitter-RoBERTa [45] (refer to the "Methodology" section for details) and extract the cosine similarity between the embedding vectors. To facilitate visualization, we include SPY as a non-meme stock and randomly sample 2,000 Twits for each stock, plotting only the top two edges for each node.

The results, as shown in Figure 5, illustrate that the Twits of meme stocks (represented by darker blue) exhibit a distinct distribution compared to the Twits of non-meme stocks (represented by bright blue). Generally, the Twits for meme stocks (darker blue) and non-meme stocks (bright blue) are concentrated in their respective areas. This concentrated structure suggests that the Twits of meme stocks may contain distinctive words or phrases that are more similar among themselves than with other stocks. This observation highlights the unique language and discourse surrounding meme stocks in the online community, indicating the presence of shared themes, sentiments, and trends within these discussions.

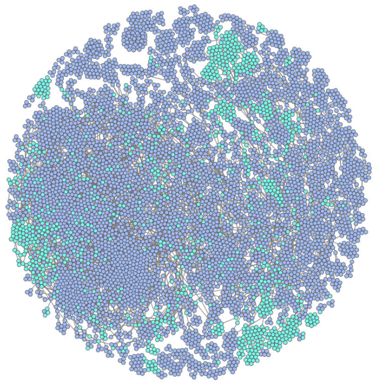


**FIGURE 5.** Network visualization of Twits for meme stocks (darker blue) and non-meme stocks (bright blue) during the explosive period. The Twits were transformed into embedding vectors using a transformer-based model (detailed in the "Methodology" section) and plotted based on cosine similarity between the embedding vectors. Distinct characteristics between meme stocks and non-meme stocks can be easily observed in terms of their expressions.

The subsequent analysis focuses on identifying the most frequently mentioned words in Twits related to various meme stocks. By examining the language used by market participants, we aim to uncover common themes, sentiments, and trends within these discussions. Table 4 presents the top 20 most frequently mentioned words in the Twits for each meme stock, after preprocessing (tokenization, stopword removal, punctuation removal, lemmatization, stemming). An intriguing observation is the frequent mention of other meme stocks within these Twits. Across various meme stocks, there is a substantial frequency of references to and mentions of other meme stocks within the same discussions. The mention of other meme stocks reflects the collective mindset and shared attention of investors in the online investing community, creating a network effect that amplifies the influence and volatility of these stocks. This interconnectedness underscores the unique dynamics and speculative nature of meme stock investing, driven by social media trends, community-driven narratives, and the desire to participate in the excitement surrounding these stocks.

Another noteworthy phenomenon that emerged from the analysis of Twits related to meme stocks is the repeated use of the phrase "I just like the stock." This phrase has become emblematic of meme stock investing, representing a distinctive characteristic of this investment trend [53]. Unlike traditional approaches that rely heavily on financial analysis, investors in meme stocks express their affinity for

**TABLE 4. Top 20 most mentioned words in meme stock twits. This table presents the top 20 most frequently mentioned words in the twits for each meme stock after preprocessing. Stock tickers are bolded for easy identification.**

| | AMC | | BB | | BBBY | | GME | | NOK | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Bearish | Bullish | Bearish | Bullish | Bearish | Bullish | Bearish | Bullish | Bearish | Bullish |
| **Words** | *ape* | *buy* | *go* | ***amc*** | *go* | *short* | *go* | ***amc*** | *go* | ***amc*** |
| | *go* | *hold* | *stock* | *buy* | *get* | *buy* | ***amc*** | *short* | ***amc*** | *buy* |
| | *get* | *ape* | *sell* | *go* | *buy* | *go* | *buy* | *buy* | *stock* | ***gme*** |
| | *quot* | *go* | *short* | ***gme*** | *short* | ***gme*** | *get* | *go* | ***gme*** | *go* |
| | *sell* | *get* | *get* | *get* | *sell* | *get* | *short* | *get* | *sell* | *hold* |
| | *buy* | *share* | *buy* | *hold* | *bull* | ***amc*** | *sell* | *hold* | *get* | ***bb*** |
| | *short* | *sell* | ***amc*** | *let* | *put* | *share* | *stock* | *share* | *short* | *let* |
| | *stock* | *short* | ***gme*** | *short* | *stock* | *squeez* | *money* | *stock* | *buy* | *get* |
| | *like* | *let* | *like* | *stock* | *like* | *sell* | *put* | *sell* | ***bb*** | *stock* |
| | *share* | ***gme*** | *quot* | *share* | *money* | *let* | *like* | *like* | *like* | *nokia* |
| | *money* | *like* | *chen* | *sell* | *quot* | *today* | *bull* | *squeez* | *money* | *sell* |
| | ***gme*** | *u* | *compani* | *blackberri* | ***amc*** | *like* | *hold* | *let* | *back* | *share* |
| | *make* | *day* | *blackberri* | *quot* | *compani* | *hold* | *make* | *bear* | *share* | *short* |
| | *hold* | *today* | *market* | *like* | *share* | *bear* | *quot* | *come* | *hold* | *like* |
| | *squeez* | *stock* | *today* | *today* | *today* | *stock* | *back* | *today* | *market* | *u* |
| | *back* | *time* | *share* | *day* | *day* | *day* | *peopl* | *ape* | *make* | *g* |
| | *day* | *see* | *day* | *time* | *ape* | *see* | *share* | *time* | *nokia* | *today* |
| | *lol* | *f**** | *bull* | *come* | *see* | *come* | *lol* | *day* | *think* | *make* |
| | *see* | *bear* | *back* | *see* | *dump* | *time* | *time* | *make* | *f**** | *come* |
| | *time* | *squeez* | *money* | ***nok*** | *pump* | *next* | *see* | *see* | *good* | *time* |

or preference toward specific stocks based on subjective factors. The phrase signifies a departure from conventional investment strategies, as investors prioritize personal beliefs, community-driven narratives, and the collective enthusiasm surrounding these stocks. The frequent appearance of the phrase "I just like the stock" across different meme stocks further highlights the emotional and speculative nature of this investing phenomenon. It underscores that investors in meme stocks are driven by personal sentiments and connections rather than traditional valuation metrics or company fundamentals. This aspect of meme stock investing reflects a shift in the investment landscape, with individuals placing greater emphasis on subjective factors and community-driven narratives when making investment decisions.

In addition to textual analysis, the use of emojis in meme stock Twits has revealed intriguing insights, particularly when differentiating between bullish and bearish Twits. Emojis have gained immense popularity as a means of expression, conveying contextual meaning and infusing online conversations with emotional nuances. These visual representations of emotions, attitudes, and reactions play a pivotal role in capturing the essence of meme culture and its profound impact on investor sentiment. By analyzing the usage and frequency of emojis, we can gain deeper insights into the emotional landscape and nuanced dynamics of investor perceptions within meme stock investing on social media platforms.

Our analysis of emojis used in bullish and bearish Twits has revealed notable differences in the frequency of specific emojis, as illustrated in Figure 6. The rocket emoji was the most frequently used emoji in bullish Twits across various stocks, symbolizing enthusiasm and anticipation. Conversely, the crying face emoji appeared most frequently in bearish Twits, indicating disappointment and negativity. These findings underscore the emotive nature of emojis as visual cues and their ability to convey sentiment within the context of meme stock discussions.

Of particular interest is the distinct prevalence of certain emojis in bullish Twits that are specifically associated with meme stocks, such as the ape and diamond hands emojis. The ape emoji has become a symbol of strength and resilience, embodying the tenacity and unwavering determination of retail investors. Redditors on r/WallStreetBets often refer to themselves as "apes", forging an unprecedented bond and fostering a sense of camaraderie. The ape emoji serves as a visual representation of this community and its shared values. Another significant emoji within the meme stock discourse is the diamond hands emoji. This particular symbol carries profound significance, representing the unwavering commitment of retail investors to hold onto their investments despite extreme market volatility. It reflects a steadfast resolve and refusal to succumb to panic selling or short-term fluctuations. The diamond hands emoji has become synonymous with the meme stock community's resilience, emphasizing the importance of maintaining positions and weathering market storms.

Importantly, these emojis, as illustrated in Figure 6, are specific to meme stocks and are noticeably absent in


**FIGURE 6. Distribution of emoji usage in bullish and bearish Twits for each stock. The chart displays the relative frequency of the top 10 most commonly used emojis, represented as horizontal bar charts. The upper section shows the emoji distribution for bullish Twits (green), while the lower section depicts the distribution for bearish Twits (red).**

Twits related to SPY. This further reinforces their strong association with the meme stock phenomenon. The ape and diamond hands emojis have emerged as powerful visual cues within the meme stock community, embodying the unwavering conviction and unity of retail investors. These emojis symbolize a collective determination to challenge established financial norms and hold onto investments for long-term gains. They capture the unique culture and spirit of the meme stock phenomenon, representing the strong bond among retail investors and their collective defiance of traditional market expectations.

Upon analyzing the distribution of Twits and their associated emojis, we observe a striking pattern in the dominance of bullish Twits over bearish Twits. This trend is evident in Figure 7, where we present the ratio of bullish Twits to bearish Twits for each stock. The green bars represent the ratio of bullish Twits, while the red bars depict the ratio of bearish Twits. During the explosive period, there is a significant disparity between the number of bullish and bearish Twits. Notably, stocks such as AMC, BB, and NOK exhibit a clear scarcity of bearish Twits in comparison to the substantial volume of bullish Twits.

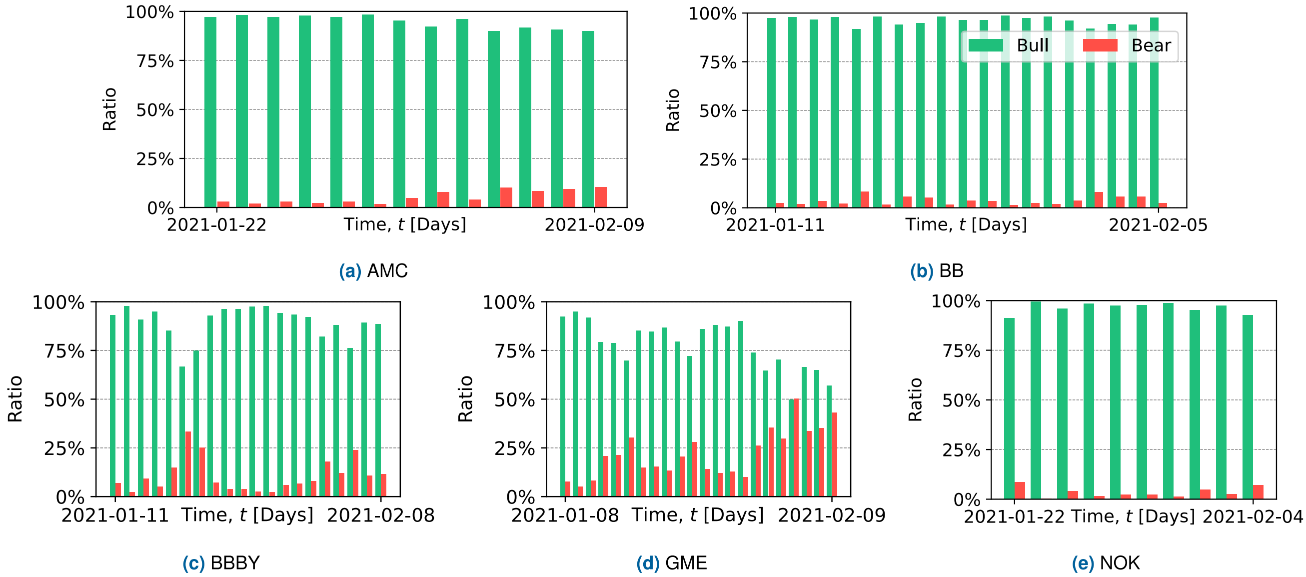



**FIGURE 7. Sentiment analysis during the explosive period. Each plot illustrates the ratio of bullish (green) and bearish (red) Twits. Across all stocks, a high proportion of Twits show bullish sentiment, with only a small number expressing bearish sentiment, especially during the initial phase of the explosive period.**

The dominance of bullish Twits and the scarcity of bearish Twits imply a prevailing sentiment of positivity and enthusiasm among retail investors. This sentiment can be attributed to the collective belief in the potential for significant gains and the conviction to hold onto their investments despite market fluctuations. This phenomenon aligns with the concept of herding behavior in investing. Herding behavior refers to the tendency of investors to follow the actions and sentiments of others [36]. As the price experiences a significant surge at the initial phase of the explosive period, it reflects the collective inclination of investors to join the bandwagon and partake in the momentum. The overwhelming volume of bullish Twits and the relatively low presence of bearish Twits indicate a general consensus and optimism among investors during the explosive period. It underscores the impact of social media discussions on investor sentiment and decision-making processes, exemplifying the influence of online communities in shaping market perceptions and driving the collective behavior of retail investors.

## V. CONCLUSION

The rise of meme stocks, driven by social media engagement, has reshaped market dynamics and introduced new complexities in trading behavior. Our study provides quantitative evidence of the strong interplay between social sentiment and meme stock trading activity. We find a significant correlation between the number of Twits and trading volume, further supported by bidirectional Granger causality, suggesting that social media both reflects and drives market activity. Additionally, we demonstrate that applying the BSADF test to social media data reveals explosive periods that price-based methods alone may miss, underscoring the value of integrating sentiment analysis into speculative asset monitoring. Moreover, our examination of meme stock discourse highlights a distinct network effect, where discussions frequently reference multiple meme stocks, reinforcing coordinated trading patterns. Finally, sentiment analysis reveals a strong bullish bias during explosive periods, reflecting herd-driven optimism.

These findings offer several implications for regulators and investors. For regulators, the strong link between social media and meme stock trading activity suggests that monitoring online sentiment could serve as an early-warning mechanism for excessive speculation and potential market manipulation. Regulatory bodies could integrate sentiment-based indicators into market surveillance systems to identify sudden shifts in investor sentiment that may precede price volatility. Regulators could also leverage the distinct linguistic and structural patterns in online discourse, such as the repeated use of specific linguistic features and emoji-based signals, as early signals of speculative behavior. By leveraging these insights, regulators could develop proactive measures, such as enhanced disclosure requirements or temporary trading restrictions, to mitigate the risks of herd-driven market distortions before they escalate.

For investors, our findings highlight both the potential and risks of using social media sentiment in trading strategies. The market-social media interaction analysis results indicate that social media engagement not only reflects market movements but also drives trading behavior. This suggests that traders can use social sentiment as a complementary tool for identifying shifts in momentum. The BSADF findings further highlight that social media trends capture certain explosive periods overlooked by price-based methods, making them a valuable complementary tool. However, the dominance of bullish sentiment and the prevalence of emotionally charged narratives, particularly during speculative surges, suggest that traders should exercise caution. The interconnected nature of meme stock discussions implies that momentum in one meme stock may spill over into others, reinforcing speculative waves. As such, investors should not rely solely on sentiment indicators but instead combine them with traditional risk management techniques to navigate the volatility of meme stocks more effectively.

Despite its contributions, this study has limitations that future research can address. Our analysis relies on data from StockTwits, which, while relevant, does not capture the full scope of social media activity across platforms such as Reddit and Twitter. Expanding the dataset to include multiple social media sources would provide a more comprehensive view of sentiment trends. Furthermore, while our analysis establishes a strong relationship between social media activity and market behavior, determining the causal mechanisms remains a challenge. Additionally, our analysis does not explicitly account for macroeconomic factors like market volatility or liquidity conditions, which may also shape investor behavior. Future research could refine this by incorporating regression models with control variables to better isolate the effects of social media sentiment. Future studies could also apply machine learning techniques to develop predictive models that incorporate sentiment dynamics, network structures, and price movements to improve forecasting accuracy. Additionally, exploring the role of exogenous shocks and regulatory interventions in shaping meme stock trends could provide valuable insights into market stability and investor behavior.

As social media continues to play a growing role in financial markets, understanding its impact on asset pricing, investor sentiment, and speculative behavior will remain a critical area of research. By integrating social sentiment analysis with established financial models, future studies can further bridge the gap between online discourse and market dynamics, offering deeper insights into the evolving relationship between retail traders and financial markets.

## REFERENCES

[1] T. Renault, "Intraday online investor sentiment and return patterns in the U.S. stock market," *J. Banking Finance*, vol. 84, pp. 25–40, Nov. 2017.

[2] M. Baker and J. Wurgler, "Investor sentiment and the cross-section of stock returns," *J. Finance*, vol. 61, no. 4, pp. 1645–1680, 2006.

[3] M. Münster, F. Reichenbach, and M. Walther, "Robinhood, Reddit, and the news: The impact of traditional and social media on retail investor trading," *J. Financial Markets*, vol. 71, Nov. 2024, Art. no. 100929.

[4] S. Behrendt and A. Schmidt, "The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility," *J. Banking Finance*, vol. 96, pp. 355–367, Nov. 2018.

[5] R. Batra and S. M. Daudpota, "Integrating StockTwits with sentiment analysis for better prediction of stock price movement," in *Proc. Int. Conf. Comput., Math. Eng. Technol. (iCoMET)*, Mar. 2018, pp. 1–5.

[6] F. Audrino, F. Sigrist, and D. Ballinari, "The impact of sentiment and attention measures on stock market volatility," *Int. J. Forecasting*, vol. 36, no. 2, pp. 334–357, Apr. 2020.

[7] E. F. Fama, "Efficient capital markets," *J. Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[8] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," in *Handbook of the Fundamentals of Financial Decision Making: Part I*. Singapore: World Scientific, 2013, pp. 99–127.

[9] M. Costola, M. Iacopini, and C. R. Santagiustina, "On the 'mementum' of meme stocks," *Econ. Lett.*, vol. 207, Oct. 2021, Art. no. 110021.

[10] I. Yousaf, L. Pham, and J. W. Goodell, "The connectedness between meme tokens, meme stocks, and other asset classes: Evidence from a quantile connectedness approach," *J. Int. Financial Markets, Institutions Money*, vol. 82, Jan. 2023, Art. no. 101694.

[11] A. Aloosh, S. Ouzan, and S. J. H. Shahzad, "Bubbles across meme stocks and cryptocurrencies," *Finance Res. Lett.*, vol. 49, Oct. 2022, Art. no. 103155.

[12] A. H. Elsayed, M. E. Hoque, M. Billah, and M. K. Alam, "Connectedness across meme assets and sectoral markets: Determinants and portfolio management," *Int. Rev. Financial Anal.*, vol. 93, May 2024, Art. no. 103177.

[13] E. Vasileiou, "Does the short squeeze lead to market abnormality and antileverage effect? Evidence from the gamestop case," *J. Econ. Stud.*, vol. 49, no. 8, pp. 1360–1373, Oct. 2022.

[14] Z. Umar, I. Yousaf, and A. Zaremba, "Comovements between heavily shorted stocks during a market squeeze: Lessons from the GameStop trading frenzy," *Res. Int. Bus. Finance*, vol. 58, Dec. 2021, Art. no. 101453.

[15] A. Betzer and J. P. Harries, "How online discussion board activity affects stock trading: The case of GameStop," *Financial Markets Portfolio Manage.*, vol. 36, no. 4, pp. 443–472, Dec. 2022.

[16] S. Warkulat and M. Pelster, "Social media attention and retail investor behavior: Evidence from r/wallstreetbets," *Int. Rev. Financial Anal.*, vol. 96, Nov. 2024, Art. no. 103721.

[17] Š. Lyócsa, E. Baumöhl, and T. Výrost, "YOLO trading: Riding with the herd during the GameStop episode," *Finance Res. Lett.*, vol. 46, May 2022, Art. no. 102359.

[18] A. Mancini, A. Desiderio, R. Di Clemente, and G. Cimini, "Self-induced consensus of Reddit users to characterise the GameStop short squeeze," *Sci. Rep.*, vol. 12, no. 1, p. 13780, Aug. 2022.

[19] X. Zheng, H. Tian, Z. Wan, X. Wang, D. D. Zeng, and F.-Y. Wang, "Game starts at GameStop: Characterizing the collective behaviors and social dynamics in the short squeeze episode," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 1, pp. 45–58, Feb. 2022.

[20] Z. Umar, M. Gubareva, I. Yousaf, and S. Ali, "A tale of company fundamentals vs sentiment driven pricing: The case of GameStop," *J. Behav. Exp. Finance*, vol. 30, Jun. 2021, Art. no. 100501.

[21] I. Pandey and M. Guillemette, "Social media, investment knowledge, and meme stock trading," *J. Behav. Finance*, vol. 2024, pp. 1–17, Jun. 2024.

[22] T. Klein, "A note on GameStop, short squeezes, and autodidactic herding: An evolution in financial literacy?" *Finance Res. Lett.*, vol. 46, May 2022, Art. no. 102229.

[23] D. Bradley, J. Hanousek, R. Jame, and Z. Xiao, "Place your bets? The value of investment research on Reddit's wallstreetbets," *Rev. Financial Stud.*, vol. 37, no. 5, pp. 1409–1459, Apr. 2024.

[24] R. G. Chacon, T. G. Morillon, and R. Wang, "Will the Reddit rebellion take you to the moon? Evidence from WallStreetBets," *Financial Markets Portfolio Manage.*, vol. 37, no. 1, pp. 1–25, Mar. 2023.

[25] K. Wang, B. Wong, M. A. Khoshkholghi, P. Shah, R. K. Naha, A. Mahanti, and J.-K. Kim, "Dissecting the hype: A study of WallStreetBets' sentiment and network correlation on financial markets," in *Proc. Int. Conf. Adv. Inf. Netw. Appl.* Cham, Switzerland: Springer, Jan. 2024, pp. 263–273.

[26] S. Agrawal, P. D. Azar, A. W. Lo, and T. Singh, "Momentum, mean-reversion, and social media: Evidence from StockTwits and Twitter," *J. Portfolio Manage.*, vol. 44, no. 7, pp. 85–95, Jul. 2018.

[27] S. Lee, J. Lee, and Y. Lee, "Dissecting the Terra-LUNA crash: Evidence from the spillover effect and information flow," *Finance Res. Lett.*, vol. 53, May 2023, Art. no. 103590.

[28] J. Piñeiro-Chousa, M. Á. López-Cabarcos, A. M. Pérez-Pico, and B. Ribeiro-Navarrete, "Does social network sentiment influence the relationship between the S&P 500 and gold returns?" *Int. Rev. Financial Anal.*, vol. 57, pp. 57–64, May 2018.

[29] R. A. Fisher, "On the probable error of a coefficient of correlation deduced from a small sample," *Metron*, vol. 1, p. 3, Jun. 1921.

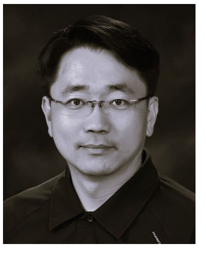
[30] G. Y. Zou, "Toward using confidence intervals to compare correlations," *Psychol. Methods*, vol. 12, no. 4, p. 399, 2007.

[31] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969.

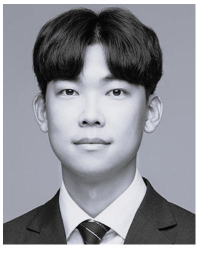
[32] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *J. Amer. Stat. Assoc.*, vol. 74, no. 366, pp. 427–431, Jun. 1979.
[33] P. C. B. Phillips, S. Shi, and J. Yu, "Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500," *Int. Econ. Rev.*, vol. 56, no. 4, pp. 1043–1078, Nov. 2015.
[34] P. C. B. Phillips and P. Perron, "Testing for a unit root in time series regression," *Biometrika*, vol. 75, no. 2, pp. 335–346, Jun. 1988.
[35] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, vol. 57, no. 2, pp. 357–384, Mar. 1989.
[36] D. Sornette, "Why stock markets crash," in *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton, NJ, USA: Princeton Univ. Press, 2009.
[37] V. Monschang and B. Wilfling, "Sup-ADF-style bubble-detection methods under test," *Empirical Econ.*, vol. 61, no. 1, pp. 145–172, Jul. 2021.
[38] C.-Z. Yao and H.-Y. Li, "A study on the bursting point of Bitcoin based on the BSADF and LPPLS methods," *North Amer. J. Econ. Finance*, vol. 55, Jan. 2021, Art. no. 101280.
[39] M. Tomal, "Identification of house price bubbles using robust methodology: Evidence from Polish provincial capitals," *J. Housing Built Environ.*, vol. 37, no. 3, pp. 1461–1488, Sep. 2022.
[40] M. Guo, S. Wang, and Y. Wei, "Bubbles in NFT markets: Correlated with cryptocurrencies or sentiment indexes?" *Appl. Econ. Lett.*, vol. 32, no. 4, pp. 491–497, Feb. 2025.
[41] T. Wolf et al., "Transformers: State-of-the-art natural language processing," in *Proc. Conf. Empirical Methods Natural Lang. Process., Syst. Demonstrations*, 2020, pp. 38–45.
[42] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2019, pp. 4171–4186.
[43] U. Naseem, I. Razzak, K. Musial, and M. Imran, "Transformer based deep intelligent contextual embedding for Twitter sentiment analysis," *Future Gener. Comput. Syst.*, vol. 113, pp. 58–69, Dec. 2020.
[44] J. Ravi and S. Kulkarni, "Text embedding techniques for efficient clustering of Twitter data," *Evol. Intell.*, vol. 16, no. 5, pp. 1667–1677, Oct. 2023.
[45] F. Barbieri, J. Camacho-Collados, L. E. Anke, and L. Neves, "Tweet-Eval: Unified benchmark and comparative evaluation for tweet classification," in *Proc. Findings Assoc. Comput. Linguistics, EMNLP*, Nov. 2020, pp. 1644–1650. [Online]. Available: https://aclanthology.org/2020.findings-emnlp.148
[46] S. Majeed, M. Uzair, U. Qamar, and A. Farooq, "Social network analysis visualization tools: A comparative review," in *Proc. IEEE 23rd Int. Multitopic Conf. (INMIC)*, Nov. 2020, pp. 1–6.
[47] W. Zeng, C.-W. Fu, S. M. Arisona, A. Erath, and H. Qu, "Visualizing mobility of public transportation system," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 1833–1842, Dec. 2014.
[48] Y. Abdelsadek, K. Chelghoum, F. Herrmann, I. Kacem, and B. Otjacques, "Community extraction and visualization in social networks applied to Twitter," *Inf. Sci.*, vol. 424, pp. 204–223, Jan. 2018.
[49] U. Aslak and B. Maier, "Netwulf: Interactive visualization of networks in Python," *J. Open Source Softw.*, vol. 4, no. 42, p. 1425, Oct. 2019.
[50] A. Tengulov et al., "Squeezing shorts through social news platforms," Swiss Finance Inst., Zürich, Switzerland, Tech. Rep. 21-31, 2021.
[51] J. J. Murphy, *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. Baltimore, MD, USA: Penguin, 1999.
[52] D. L. Western, *Booms, Bubbles and Busts in U.S. Stock Markets*. London, U.K.: Psychology Press, 2004.
[53] S. Long, B. Lucey, Y. Xie, and L. Yarovaya, "'I just like the stock': The role of Reddit sentiment in the GameStop share rally," *Financial Rev.*, vol. 58, no. 1, pp. 19–37, Feb. 2023.


**SEUNGJU LEE** received the B.S. and M.S. degrees in industrial engineering from Seoul National University, where he is currently pursuing the Ph.D. degree. His research interests include quantitative finance and the integration of deep learning methodologies into financial market analysis.

**YUNYOUNG LEE** received the B.S., M.S., and Ph.D. degrees in industrial engineering from Seoul National University, in 2018, 2020, and 2023, respectively. He is currently an Assistant Professor with the Department of Artificial Intelligence and Data Science, Sejong University, Seoul, South Korea. His research interests include blockchain, central bank digital currencies, and time series analysis.



**JAEWOOK LEE** received the B.S. degree in mathematics from Seoul National University, Seoul, South Korea, in 1993, and the Ph.D. degree in applied mathematics from Cornell University, in 1999. He is currently a Professor with the Department of Industrial Engineering, Seoul National University. His research interests include machine learning, neural networks, and global optimization and their applications to data mining and financial engineering.



**HOKI KIM** received the B.S. and Ph.D. degrees in industrial engineering from Seoul National University, South Korea. He is currently a Professor with the Department of Industrial Security, Chung-Ang University. His research interests include trustworthy AI, aiming to enhance the robustness, generalization, and explainability of machine learning and deep learning models. He is also interested in addressing these challenges in applications such as industrial manufacturing and finance.