

The Relationship Between Social Media Data and Stock Market Volatility

Zhengxin Yang^{1,a,*}

¹*Beijing Technology and Business University, Beijing 102488, China*

a. zhenglin@bjucd.com

**corresponding author*

Abstract: In recent years, the influence of social media data on the stock market has been deeply concerned and studied by academics and professionals, because the information reflected in these data has become increasingly important in contemporary society. By offering a complete synthesis of current research, evaluating the contributions made by earlier studies, and highlighting knowledge gaps warranting additional inquiry, this literature review seeks to examine the complicated link between social media data and stock market volatility. In order to accomplish this, the paper delves into a wide range of topics, such as where to get social media data, how to gauge stock market volatility, and what kinds of approaches have been taken to examine the possible connections between the two. Some research suggests that social media sentiment may forecast market patterns, while other research suggests a lower effect, as shown by our study, which concludes that there is either no relationship or a positive one between these factors. More study should focus on perfecting sentiment analysis methods, delving into nonlinear correlations, exploring variations in social media platforms, and addressing ethical problems in data utilization. We hope that these efforts will help us better comprehend the relationship between social media data and stock market volatility, allowing us to create more precise forecasting instruments.

Keywords: social media data, stock market volatility, sentiment analysis

1. Introduction

The advent and subsequent meteoric rise in popularity of social media platforms has resulted in the collection and dissemination of enormous quantities of data produced by its users. Financial markets are just one area where this information might be invaluable. Investors and the general public are profoundly affected by stock market volatility since it is a vital barometer of economic health. Thus, studies connecting social media data with market fluctuations have gathered steam. This literature review aims to do just that by providing a synopsis of the relevant studies and highlighting their contributions to the area. We can learn more about the intricate connection between social media data and stock market volatility by analyzing the results of many research. This review will also identify the gaps in our knowledge and provide a roadmap for future research, which will help to further enhance our understanding of this critical area.

This study will combine empirical data with theoretical models to answer the research question. Many scholars have been focusing on the correlation between social media data and market price fluctuations as a subject of empirical research. One such research that looked at the link between

social media mood and stock prices was done by Liu et al. [1]. According to their results, the mood on social media accounts for a significant proportion of the variance in stock prices. To a similar extent, Zhang et al. [2] discovered that data collected from social media platforms may foretell market volatility and enhance the efficacy of volatility forecasting algorithms.

This study will examine the theoretical frameworks that have investigated the link between social media data and market volatility, as well as the existing empirical research in this area. Bollen et al. [3] established a theoretical framework that makes use of data from Twitter to gauge public sentiment and forecast stock prices. They suggested that stock market fluctuations may be explained in part by data collected from social media platforms.

The limitations of past studies and the difficulty in establishing a causal connection between social media data and market volatility will be explored, despite the extensive study on the subject. It's hard to tell, for example, whether market volatility is being caused by social media data or if it's just reflecting it. Ethical issues associated with the use of social media data include users' privacy and the potential for market manipulation.

This study is significant because it sheds light on how social media data might affect international stock markets. The findings of this literature review will be helpful to scholars, policymakers, and funders. For instance, regulators might utilize this study's findings to shape regulations governing the financial sector's use of social media data. In addition, the data collected here may serve as a springboard for research into how social media data and market volatility are connected in the future.

2. Overview of Social Media Data and Stock Market Volatility

2.1. Social Media Data

Social media data is already considered a valuable source of information for financial analysts and investors. It provides real-time data on consumer sentiment, market trends, and breaking news that can impact market outcomes. Social media data can be accessed through various sources, including social media platforms, APIs provided by social media companies, and third-party data providers.

Table 1: Social Media Type.

Social Media Type	Supporting Resource
User-generated content (e.g., text posts, comments, likes)	Bollen et al. (2011) used Twitter sentiment to predict changes in the stock market with a high degree of accuracy.
Multimedia content (e.g., images, videos)	Zhang et al. (2022) used Weibo data to predict market volatility and improve the performance of volatility forecasting models.
Metadata (e.g., user demographics, location information)	Yuan et al. (2023) used Twitter data to predict election outcomes based on political sentiment.

Several studies have looked at the relationship between social media data and economic results. Twitter data has been used to gauge public sentiment and forecast market values [3]; see, for instance, the work of Bollen et al. A high degree of accuracy was discovered when using Twitter sentiment to

forecast stock market movements. To do the same for market volatility, Zhang et al. [2] included Weibo data into their forecasting algorithms.

In many economic studies, researchers have raised important ethical questions about the use of social media data. Protecting social media users' privacy is essential, as is preventing market manipulation. Authorities have begun to respond to these concerns, with the banking industry restricted from using social media data. For instance, the ESMA has developed rules for using social media data into investment research [4].

2.2. Stock Market Volatility

Fluctuations in the value of stocks and other assets traded on the stock market are referred to as volatility. It is a statistical measure of the volatility of a stock's price over a certain time period. Volatility in the stock market may be broken down into two categories: historical volatility and implied volatility [5].

Historical volatility is calculated based on past stock prices and is used to calculate the standard deviation of returns over a specific period. The most common method of calculating historical volatility is the annualized standard deviation of daily returns. The formula for historical volatility is as follows:

$$\text{Historical Volatility} = \sqrt{\frac{\sum(R_i - R)^2}{(n-1)}} \times \sqrt{n} \quad (1)$$

Where R_i is the daily return on day i , R is the average daily return, and n is the amount of observed data (usually 252 days).

Conversely, implied volatility is a measure of how the market anticipates future stock price volatility to behave; it is derived from the pricing of options contracts. The VIX index, which is based on S&P 500 index option prices, is the most extensively used indicator of implied volatility [6]. Complex option pricing models, such as the Black-Scholes model, are needed in order to arrive at the formula for implied volatility.

Factors such as economic indicators, geopolitical events, market sentiment, and changes in monetary policy might contribute to stock market volatility. Economic indicators such as inflation, interest rates and gross domestic product all play a role in stock market volatility. For example, interest rate fluctuations can affect borrowing costs, which can lead to a knock-on effect on stock prices. In addition, geopolitical events such as political elections, wars, and terrorist attacks can also contribute to stock market volatility.

Market sentiment, refers to the collective attitude of investors towards the stock market, which can also play a major role in stock market volatility. Positive market sentiment can lead to increased buying activity and higher stock prices, while negative market sentiment can lead to decreased buying activity and lower stock prices. Additionally, changes in regulations, technological advancements, and market speculation can also lead to uncertainty in the stock market.

Investors, policymakers, and other market players would do well to familiarize themselves with stock market volatility since it sheds light on market patterns and helps in the management of risks related to stock market investments. It also has an important influence in determining the profitability of investments and the health of the economy as a whole.

2.3. The Relationship Between Social Media Data and Stock Market Volatility

Research on the correlation between social media data and market volatility has abounded in recent years. Sentiment analysis has been used to social media data like tweets, postings, and blogs to learn

how investors feel about a company or the market as a whole [3]. There is either a positive [7] or a neutral link between social media data and stock market volatility, according to empirical research.

It has been discovered that the mood on social media greatly affects the volatility of financial markets, especially during extraordinary occurrences like the COVID-19 epidemic, wars, and presidential elections. Some research, for instance [8], have shown the favorable influence of social media mood during the COVID-19 epidemic on stock market volatility. Similar results were seen between social media mood and market volatility during the 2016 US presidential election. Sentiment on social media may enhance the effect of news and information at these times, leading to more market volatility.

Many hypotheses have been advanced to explain the link between social media data and market fluctuations. Theories like as sentiment analysis, news sentiment, and behavioral finance have been advanced to explain the connection. Understanding how the market feels about a stock may be achieved via sentiment analysis, which entails looking at how people feel about it online. Examining the overall tone and substance of news items about a company or the stock market might help you gauge market mood. According to behavioral finance theory, market volatility may be affected by investors' emotions, biases, and attitudes.

Until now, there are a number of variables that might complicate the link between social media data and stock market volatility. Stock market volatility may be more sensitive to the mood on social media during times of high market uncertainty and volatility than during times of low market volatility. Additionally, positive and negative sentiment on social media might have varying effects on stock market volatility, suggesting a nonlinear relationship between the two. Because of the inherent intricacy of the connection between social media data and market volatility, nonlinear models have been developed [7].

2.4. Theoretical Perspectives on Using Social Media and News Sentiment to Predict Stock Market Volatility

According to the concept of sentiment analysis, important information about market sentiment can be revealed by feelings, thoughts and opinions shared on social media, which in turn can affect stock prices and market volatility [3]. According to He et al., sentiment research delves into the tenor and substance of social media data to comprehend how investors feel about a stock or the market as a whole [9]. Eachempati et al. propose that investors' emotions can be used to predict stock market swings [8].

Joshi et al. apply the theory of news sentiment, which states that the tone and substance of news items pertaining to a particular stock or the stock market as a whole can provide valuable information about market mood, which in turn can effect stock prices and stock market volatility [10]. Additionally, Eachempati et al. note that News sentiment analysis is the process of assessing the tone and substance of news items to learn how the market feels about a stock or the market as a whole [8].

Corporate Finance Institute defines behavioral finance theory in the context of cryptocurrency research, arguing that investors' emotions, biases, and sentiments play a role in the stock market and can effect stock market volatility [11]. In addition, Corporate Finance Institute discovered that this theory implies that social media data might provide useful information about investors' emotions, prejudices, and feelings, all of which can affect stock prices and stock market volatility [11].

To predict stock market volatility using social media and news sentiment, various theoretical perspectives have been proposed. These perspectives aim to understand how social media sentiment can reveal important information about market sentiment, which in turn can affect stock prices and market volatility.

One popular approach is sentiment analysis, which delves into the tenor and substance of social media data to comprehend how investors feel about a stock or the market as a whole [12]. This

approach involves using natural language processing techniques to analyze social media data and identify the sentiment expressed in the text. Machine learning algorithms are then used to classify the sentiment as positive or neutral. These sentiment scores can then be used to predict future stock market movements.

Another approach is news sentiment analysis, which involves assessing the tone and substance of news items to learn how the market feels about a stock or the market as a whole [13]. This approach involves using natural language processing techniques to analyze news articles and identify the sentiment expressed in the text. Machine learning algorithms are then used to classify the sentiment as positive or neutral. These sentiment scores can then be used to predict future stock market movements.

The connection between social media sentiments and stock market volatility has also been explained from the standpoint of behavioral finance theory [14]. According to this hypothesis, market volatility is influenced by investors' subjective experiences, including their emotions, prejudices, and sentiments. Information about investors' biases, sentiments, and emotions may be gleaned from social media data and used to the prediction of stock market movements.

Using social media and news sentiment data, machine learning and deep learning methods have also been used to anticipate stock market volatility. The goal of these methods is to anticipate future stock market movements based on patterns found in social media and news sentiment data. For instance, Yuan et al. [15] utilized Twitter sentiment data to inform a machine learning system that predicted stock market movements during the COVID-19 epidemic.

Notwithstanding the challenges, experts and investors are wondering whether there is a correlation between social media data and stock market fluctuations. Further research is needed to fully understand the link and to develop effective techniques of using social media data to forecast stock market volatility.

2.5. Empirical Studies on the Relationship

2.5.1. Positive Relationship

Some empirical studies have found a positive correlation between social media data and stock market volatility. These studies have shown that social media data can provide valuable information about market sentiment, which can impact stock prices and stock market volatility [2,7,8]. For example, Joshi et al. examined that positive sentiment expressed in social media data has been found to be associated with increased stock prices, while negative sentiment expressed in social media data has been associated with decreased stock prices [10]. Zhang et al. found that social media data may cause investor fear and the propagation of false information, which in turn increases market volatility and decreases stock prices [2]. Similar to what we see here, P H and Rishad discovered that negative sentiment in social media data is linked to increased market volatility and falling stock prices [14]. Hence, these studies underscore the necessity of monitoring and interpreting social media data for market participants and imply that inactive social media data might have a negative influence on the stock market.

2.5.2. Neutral Relationship

There appears to be no correlation between social media data and market volatility, according to certain empirical research. Studies such as the one conducted by Reed demonstrate that social media data has little effects on the stock market and that other factors, such as economic indicators, have a more substantial role in predicting stock market volatility.

Nonetheless, there are challenges to be solved in determining whether stock market movements can be predicted using social media data. The literature has identified a number of difficulties,

including concerns about data quality and dependability and methodological flaws [7]. Accuracy and completeness of social media data, as well as the difficulties of confirming the information, were highlighted as Data quality and reliability challenges by Yuan et al. [15]. The difficulty of reliably assessing sentiment in social media data and the selection of a suitable approach for evaluating the data are two examples of methodological challenges [7].

3. Conclusion

In sum, this evaluation of the literature seeks to resolve to the question that follows: "What is the relationship between social media data and stock market volatility?" This review seeks to answer this issue by looking at the theoretical and empirical work done on this topic from a variety of perspectives. Topics include theoretical and empirical looks at how social media and news sentiment may be used to anticipate stock market volatility, as well as social media data, stock market volatility, and the link between the two.

The primary results of this study imply that social media data may have a positive, neutral, or inverse association with stock market volatility. Research has shown that social media sentiment may be used as a credible predictor of market mood, although other studies have shown its impacts to be rather moderate.

Based on the results, researchers may want to further investigate the nonlinear connections between social media data and stock market volatility, or create more sophisticated sentiment analysis algorithms. Ethical issues with utilizing social media data in financial markets should also be addressed, and further research should be conducted to examine the effects of various social media platforms.

To effectively use social media data for predicting stock market volatility, we need to pursue these study lines to better comprehend the complicated interaction between the two.

References

- [1] Liu, Q., Lee, W.-S., Huang, M., Wu, Q.: Synergy between stock prices and investor sentiment in social media. *Borsa Istanbul Review* 23, 76–92 (2023).
- [2] Zhang, H., Chen, Y., Rong, W., Wang, J., Tan, J.: Effect of social media rumors on stock market volatility: A case of data mining in China. *Frontiers in Physics* 10 (2022).
- [3] Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. *Journal of Computational Science* 2, 1–8 (2011).
- [4] ESMA, <http://www.esma.europa.eu> (2020), last accessed 2023/3/25.
- [5] Hull, J.: *Options, Futures, and Other Derivatives, Global Edition*. Pearson (2021).
- [6] COBE, https://www.cboe.com/tradable_products/vix/ (2021), last accessed 2023/3/26
- [7] Li, Q., Chen, Yan, Wang, J., Chen, Yuanzhu, Chen, H.: Web Media and Stock Markets: A Survey and Future Directions from a Big Data Perspective. *IEEE Transactions on Knowledge and Data Engineering* 30, 381–399 (2018).
- [8] Duz Tan, S., Tas, O.: Social Media Sentiment in International Stock Returns and Trading Activity. *Journal of Behavioral Finance* 22, 221–234 (2020).
- [9] Jorion, P.: *Value at Risk, 3rd Ed.: The New Benchmark for Managing Financial Risk*. McGraw Hill Professional (2006).
- [10] Corporate Finance Institute.: *Behavioral Finance*. Corporate Finance Institute (2019).
- [11] Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., Mozetič, I.: The Effects of Twitter Sentiment on Stock Price Returns. *PLOS ONE* 10, e0138441 (2015).
- [12] Agarwal, A.: Sentiment Analysis of Financial News. In: 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN). IEEE (2020).
- [13] P H, H., Rishad, A.: An empirical examination of investor sentiment and stock market volatility: evidence from India. *Financial Innovation* 6 (2020).
- [14] Yuan, C., Ma, X., Wang, H., Zhang, C., Li, X.: COVID19-MLSF: A multi-task learning-based stock market forecasting framework during the COVID-19 pandemic. *Expert Systems with Applications* 217, 119549 (2023).

- [15] Ma, F., Guo, Y., Chevallier, J., Huang, D.: *Macroeconomic attention, economic policy uncertainty, and stock volatility predictability. International Review of Financial Analysis* 84, 102339 (2022).