ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business
Specialization: Financial Economics

Evaluating the Impact of Twitter Sentiment on Stock Market Performance: A Multi-Period Analysis

Author: Aleksander Walenciak
Student number: 610851
Thesis supervisor: Ruben de Bliek
Second reader: Antti Yang
Finish date: 30.06.2024

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second reader, Erasmus School of Economics or Erasmus University Rotterdam.

ABSTRACT

This paper investigates the relationship between the firm-specific Twitter sentiment and stock returns of companies listed on the Dow Jones Industrial Average (DJIA) from January 2017 to December 2023. Furthermore, it examines whether the impact of the Twitter sentiment on stock returns changes across the pre-pandemic, pandemic, and post-pandemic periods. The data used for this study was collected from Bloomberg, and includes daily firm-specific Twitter sentiment and other firm-specific financial and non-financial characteristics. General Least Squares (GLS) regression is used to analyze the predictive power of Twitter sentiment on stock returns. The results indicate that Twitter sentiment significantly influences stock returns, showing the increasing importance of social media information in financial markets. The results also suggest that this relationship does not change significantly over the different periods studied, implying a stable influence of Twitter sentiment across varying market conditions. This study contributes to the behavioral finance literature by finding further evidence supporting the significant role of social media sentiment predicting stock returns. Furthermore, the paper contributes to the understandings of the dynamics of the social media effects on stock returns throughout prolonged periods of time, including the times of extreme economic shocks. Additionally, this research offers practical insights for investors and policymakers regarding the integration of sentiment analysis into investment strategies.

Keywords: Twitter sentiment, Social media, Stock returns

TABLE OF CONTENTS

CHAPTER 1 Introduction

Predicting stock price movements is still an unsolved mystery that not only researchers, but also many individual traders and trading companies try to answer. The empirical research conducted into the topic of stock price movements has found evidence both supporting and rejecting the hypothesis that the stock prices follow a random walk. Although the stock prices may not perfectly follow the random walk, the efficient market hypothesis states that the asset prices reflect all available information, thus it is impossible to consistently beat the market, as market prices should only react to new information. However, due to human imperfections some investors suffer from behavioral biases, while others confuse old information with new information, causing abnormal fluctuations in stock prices. Nowadays, with the growing use of social media, the amount of information available and displayed to people is increasing and can be overwhelming, causing further confusions of old and new information. Furthermore, social media creates a perfect environment for propagating behavioral biases such as herding biases. Thus, extensive research has been conducted into whether the information shared on social media, its topics and sentiment impacts the stock prices. A Substantial number of papers find evidence of the relation between the sentiment of the social media posts, such as Twitter tweets, and the stock prices of the related companies. The social media sentiment importance is underlined by the fact that Bloomberg integrated a sentiment analysis tool into its platform in 2016 to give investors and traders an even better, firm-specific overview. Further examples of how tweets can impact the market are Elon Musk’s tweets and their impact on the stock prices or Carl Icahn’s two tweets that raised Apple value by 17 billion (Pierce, 2013).

The relation between the news and the stock market has been analyzed throughout the years. In his paper, Tetlock (2011) studies the impact of stale news on the investors’ behavior and the stock market’s reaction. He finds that stock returns respond less to stale news, however stale news still impacts the stock price. This suggests that although part of the shared information on Twitter is stale, it can still impact the stock prices. Moreover, Gu and Kurov (2020) find that the daily, firm-specific Twitter sentiment contains useful information for predicting the next day’s stock returns, as the information may include, among others, the analyst recommendations. Tan and Tas (2020) examined the relationships between the volume of the tweets and the stock’s abnormal turnover. Moreover, they investigated the effect of the sentiment of the tweets on the stock returns of the companies listed on the S&P 500, S&P 350 Europe and on Emerging Markets. The authors find Twitter sentiment is correlated with the trading volume and stock returns. Furthermore, to emphasize found correlation, Tan and Tas (2020) discuss their trading algorithm, which takes the advantage of the Twitter sentiment from the previous days to go short and long on certain stock. The annualized returns of such an algorithm were significant, emphasizing the role of Twitter sentiment on the stock prices. Yousaf, Youssef, and

Goodell (2022) find that the relation between Twitter sentiment and financial markets is significantly higher during the periods of extreme sentiment, suggesting that the Twitter sentiment can have significant predictive power during extreme positive or negative sentiment shifts. The literature also discussed the relation between the Twitter sentiment and stock market during the Covid-19 pandemic. Katsafados, Nikoloutsopoulos, and Leledakis (2023) in their study find that during the Covid-19 pandemic, the positive sentiment is positively correlated with the stock market in the short-term, while negative sentiment has a long-term negative impact on the stock returns.

Although there is already an extensive literature describing the relation between the firm-specific sentiment on social media platforms and the related company stock returns, there is little research on whether this impact changes over time. While most of the research focuses on finding the relation in a given period, a few use the panel data to find the differences in impact between across several periods. Furthermore, academic literature discusses the increased trading activity during the Covid-19 pandemic due to increased activity on social media. However, there is a little discussion on the impact of such increased activity on the stock market during Covid-19 pandemic compared to pre or post pandemic periods. Thus, with the increasing social media user base, ever-growing volume of social media posts and the rising importance of social media in people's lives, it is unknown what happens to the relation between the twitter sentiment and stock prices over a prolonged period of time. This raises a question, does the impact of the firm-specific Twitter sentiment on the stock returns of DJIA index companies change across the pre, during and post Covid-19 pandemic times?

The studied period will be from January 2017 till December 2023. The companies that will be examined are listed on the DJIA index during this period. Thus, the analysis will consist of only the companies that were part of the index throughout the whole period, which results in the sample of 25 firms and 43310 total observations. The data used in this research is collected from Bloomberg. It includes company’s financial characteristics, such as opening prices and market capitalization. Furthermore, as Bloomberg incorporated Twitter feed in 2016, the data also contains the daily firm-specific Twitter sentiment and firm-related news sentiment. Bloomberg uses its own algorithm to monitor Twitter and look for cashtags, which is a way of looking for tweets with information about tagged company stock. The relevant tweets are scraped and processed by Bloomberg's NLP algorithm, providing social sentiment analytics. Thus, the provided data about the daily volume of company related tweets, the average firm-specific Twitter sentiment and average firm-related news sentiment is collected by Bloomberg over every 24-hour period. The variables employed in this study are constructed similarly to those used by Tan and Tas (2020). Furthermore, to find the relation between the firm-specific Twitter sentiment and the company’s stock returns I will use hierarchal approach with the General Least Squares (GLS) regressions. Several models will be used, starting from a base model of stock returns and Twitter sentiment to the full model containing several control variables such as firm's size, past volatility, average past illiquidity and others. Moreover, to investigate whether the impact of the firm-specific Twitter sentiment on the company's stock returns changes across the periods before, during and after Covid-19 pandemic, I will also use GLS regression. To capture the significance of such changes I will use the interactions of the Twitter sentiment and period variable in the model. Moreover, I will conduct Wald test for difference in coefficients to estimate the effects.

The academic literature is full of empirical studies with findings that support the hypothesis that Twitter's firm-specific sentiment impacts the stock price of the respective company. Thus, there is an expectation to find that the relation between the firm-specific Twitter sentiment and the company's stock returns is statistically significant. Tetlock (2011) finds that there is an impact of stale news on the stock price, and although it is smaller than the impact of the new information, it is still significant. Moreover, Gu and Kurov (2020) find that the Twitter posts also do contain new information that is relevant for predicting the stock returns. Based on that, and the fact that the volume of daily Twitter posts is growing with the constantly increasing social media user base and its activity, there is an expectation that the impact of the firm-specific Twitter sentiment on the companies' stock returns will be significant throughout the sampled period, as well as increasing with time.

The remainder of this paper is structured as follows. Section 2 discusses the relevant literature and previous research. Sections 3 explains the data collection and variables creation process, as well as provides summary statistics. Section 4 discussed the methodology used to conduct this research, while section 5 shows the results and provides the answers to hypotheses and discusses the findings. Section 6 concludes the paper.

CHAPTER 2 Theoretical Framework

2.1 Efficient Market Hypothesis and Behavioral Finance

Prediction of stock returns has been a topic of ongoing debate for decades. The early economic literature on this topic is based on the Efficient Market Hypothesis (EMH), and argues that if markets are efficient then the stock returns are random, and thus cannot be predicted (Samuelson 1965, Eppen & Fama 1969). One of the implications of that theory is that individual investors cannot consistently outperform the market. However, empirical research found evidence supporting the argument that the markets are inefficient. By applying variance bound tests to stock prices and dividends, the actual variance of stock prices was compared with variance implied by discounted sum of expected future dividends, showing that stock price movements are significantly more volatile than the movements in dividends would predict (Cambell & Shiller 1988). Secondly, trading strategies were proposed that generated statistically significant abnormal results (Rosenberg, Reid, & Lanstein, 1985), further undermining the EMH.

The amount of research undermining the Efficient Market Hypothesis gave rise to behavioral finance, which focused on explaining inefficiencies and mispricing in financial markets. The literature of behavioral finance rejects the assumption of individual rationality. It claims that individual investors suffer from several behavioral biases, such as overconfidence, disposition effect and herding biases (Mushinada & Veluri, 2018; Jaiswal & Kamil, 2012; Jain, Jain, & Jain, 2015). Furthermore, the research suggests that individual investors affect asset prices and that the movements in asset prices may be partly attributed to the changes in investors sentiment (Lee, Shleifer, & Thaler, 1991). Moreover, Lee, Jiang, and Indro (2002) show that investors sentiment is a systematic risk that is priced, while Baker and Wurgler (2007) construct a proxy for the sentiment and find that it can predict future stock returns.

2.2 Effects of Media

The literature has also been studying the effects of traditional media and news on individual investing behavior and stock prices. The information of the news is often quantified by the sentiment of the text, which is approximated by the frequency of the positive and negative words. By quantifying the sentiment of the Wall Street Journal’s “Abreast of the Market" column, Tetlock (2007) finds that high news pessimism predicts downward pressure on market prices. Mass media coverage of securities can disseminate information broadly, reaching a big audience and affecting the prices of such securities (Fang & Peress, 2009). Furthermore, the stocks with slow information diffusion exhibit stronger momentum effects and more pronounced long-term price reversals (Hong & Stein, 1999). By analysing Dow Jones's newswire's news stories Tetlock (2011) finds that individual investors tend to overreact to stale information about publicly traded firms, thus creating stock return anomalies.

While the analysis of traditional media is limited to quantifying the article's sentiment, by analysing social media, researchers can approximate general sentiment by aggregating posts, comments and explicit feedback such as likes and dislikes. Thus, social media allows researchers to analyse individuals' behavior and sentiment at aggregate level. Another characteristic of social media is the fast diffusion of the information and overwhelming amount of information that individuals are exposed to. Furthermore, some posts can "resurface" after some period of time, making it increasingly difficult for the individual investors to differentiate between old and new information. This can amplify the individual investor's overreaction, causing significant stock return anomalies, as argued by Tetlock (2011). This would be in line with the finding that coverage by social media increases subsequent volatility and turnover, while traditional media coverage decreases it (Jiao, Veiga, & Walther, 2020). Li, Wang, Li, Liu, Gong, and Chen (2014) found that the sentiment of the comments on financial discussion boards can interfere with individuals' decision making, impacting stock prices.

With the increasing popularity of social media and its simplicity to extract aggregate data about the company related news sentiment, the empirical literature about the effects of the information sentiment on stock prices is also expanding. By exploiting Facebook's Gross National Happiness Index for daily sentiment, research has found that sentiment has a positive contemporaneous effect on stock returns (Siganos, Vagenas-Nanos, & Verwijmeren, 2014). Moreover, by investigating StockTwits data, Liew and Budavari (2017) find that the derived sentiment has a significant power in explaining daily returns of the sampled stocks, even after controlling for the traditional financial factors, such as those from the Fama-French Five-Factor Model. The analysis of the sentiment of Reddit's comments from the r/WallStreetBets subreddit on the GameStop's intraday returns indicates strong role in influencing stock during up market movements, but weak during downward movements (Long, Lucey, Xie, & Yarovaya, 2023).

The empirical literature extensively covers the effects of the post volume and sentiment extracted from social media platform X, formerly known as Twitter, on the trading volume and stock returns of various securities. By analysing tweets, Gu and Kurov (2020) find that Twitter sentiment predicts stock returns without subsequent reversals, suggesting that tweets also provide information that is not incorporated in stock prices yet. Applying the event study technique to the Twitter data of Dow Jones Industrial Average (DJIA) index companies, where peaks of Twitter volume are identified as events, Ranco et al., (2015) find a significant dependence between the Twitter sentiment and abnormal returns, which is statistically significant for several days after the events. Tan and Tas (2020) analyze the role of Twitter's activity and sentiment on several financial markets indexes. By following the methodology of Tetlock (2011) the authors find that trading volume is associated with Twitter's activity and sentiment, which predicts subsequent-day trading volume. Moreover, the results suggest that daily firm-specific Twitter sentiment holds explanatory power in predicting future stock returns. Splitting the sample into tweets posted by the users with less than sample's median number of followers and more than the median, Sul, Dennis, and Yuan (2017) find that the non-retweeted tweets that were posted by the users with fewer than the median number of followers had a significant impact on the stock's returns the next trading day, next 10 and 20 days. On the other hand, the retweeted posts and tweets from the users with more than a median of followers had no significant impact on future stock returns. Bollen, Mao, and Zeng (2011) investigate the influence of the collective Twitter mood's influence on the DJIA index. Using OpinionFinder to measure mood as positive vs. negative and Google-Profile of Mood States (GPOMS) to assess mood in 6 dimensions, they find the accuracy of their model's predictions of DJIA price movements can be significantly improved by accounting for calmness. Following the above paper, Mittal and Goel (2011) investigate the public's mood, approximated by Twitter mood, impact on the DJIA index prices. By measuring mood in 4 dimensions, the authors find that the calmness and happiness are Granger causative of the DJIA by 3-4 days. Contrary to the standard mean-based connectedness measures, by applying a novel quantile-based connectedness approach it has been found that connectedness between sentiment and financial markets is stronger at upper and lower tails. Thus, indicating that the impact of Twitter sentiment on financial markets is much stronger during extreme sentiment shocks (Yousaf, Youssef, & Goodell, 2022).

2.3 Covid-19 Pandemic

The Covid-19 pandemic has had profound and far-reaching impacts on the society and global economy. The academic literature examining Covid-19's impact and consequences on financial markets is growing. One of such consequences is the collapse of stock prices in March 2020, which can be considered as one of the biggest stock market crashes in history (Mazur, Dang, & Vega, 2021). Zaremba et al., (2020) argues that the governmental non-pharmaceutical interventions significantly increase equity market volatility, while Ashraf (2020) finds that the stock market returns decrease with the increases in the growth of confirmed Covid-19 infections. Furthermore, using transaction data, Ortmann, Pelster, and Wengerek (2020) find that as the Covid-19 pandemic continued, individual investors significantly increased their trading activities. Interestingly, the Twitter activity rose significantly in the period of Covid-19 pandemic (GDELT Blog, 2022). Katsafados, Nikoloutsopoulos, and Leledakis (2023) investigate the relationship between Twitter sentiment, measured by VADER sentiment analyzer, and stock prices indexes of several countries during the

Covid-19 pandemic. The authors find that the positive sentiment is associated with higher returns and lower volatility in the short-run, while negative sentiment is associated with lower returns in the short-run.

2.4 Hypothesis and Expectations

In this study, I test whether the information contained in the firm-specific posts on the social media platform holds explanatory power in predicitng the next day's stock market performance of those companies. Moreover, if that would be true, I will test whether the explanatory power is changing throughout the examined period. To analyze the effect of the firm-specific tweets' sentiment on stock returns I will investigate the following hypotheses:

H1: There is no influence of the firm-specific Twitter sentiment and activity on the company's next trading day stock returns.

Several studies have indicated that individuals are susceptible to suffer from many behavioral biases, such as disposition effect and herding (Jaiswal & Kamil, 2012; Jain, Jain, & Jain, 2015). Furthermore, Tetlock (2011) shows that individual investors tend to overreact to stale information, while other findings suggest that tweeted information is not fully incorporated in stock prices (Gu & Kurow, 2020). Thus, as social media platforms are perfect for sharing information with others, it is also prone to propagate biases, such as herding and overconfidence, by individual investors who react to stale or new information. The academic literature contains many studies indicating the predictive role of social media sentiment on the future asset performance. On the other hand, there are paper in the acedmic literature that find this relation to be insignificant. However, due to increasing role of social media in propagating information, as well as growing user base I believe that it plays a role in impacting stock markets. Therefore, I expect that the hypothesis will be rejected, and that the firm-specific Twitter sentiment will have an explanatory power in predicting the company's next day stock returns.

Furthermore, in this paper, I also examine the changes in the relationship between Twitter sentiment and stock returns. The second hypothesis aims to answer this question.

H2: The impact of firm-specific Twitter sentiment on the next day's company stock returns remains consistent across the periods before, during, and after the Covid-19 pandemic.

Although the academic literature on this topic is limited, there are changes on the investing landscape indicating that the differences might exist. The investing landscape is evolving with more women deciding to invest their savings into assets. The shrinking gap between the number of male and female investors impacts the average behavior of individual investors, as academic literature provides a lot of evidence for the differences in the investing behavior of men and women. Women are more risk averse and process information more carefully compared to men, thus they invest in safer assets with lower risk and more stable growth (Deb & Chavali, 2009). Moreover, Cueva et. al (2019) shows that men trade significantly more than women even after controlling for confidence. Therefore, the growing share of women investors that assess information more carefully, trading less aggressively on new information should indicate that the impact of the Twitter activity and sentiment should be decreasing throughout the periods. On the other hand, during Covid-19 pandemic both the number of investors opening their first broker account and trading volume increased. Thus, with growing Twitter user base and increase in the number of new, inexperienced individual investors susceptible to behavior biases and overreaction to stale information, the impact of firm-specific tweets' sentiment on the next day company's returns might be stronger during the pandemic. As the two phenomena should have conflicting effects on the impact of the Twitter sentiment on companies' stocks performances, the total effect is ambiguous. However, I would expect Twitter's sentiment to have stronger explanatory power in predicting the stock returns during the Covid-19 period, compared to the pre-Covid-19 times.

CHAPTER 3 Data

3.1 Twitter Sentiment Measure

In 2013 Bloomberg integrated Twitter feeds into its platform, offering its subscribers the ability to monitor social media developments. To provide its users with timely and valuable information, Bloomberg developed Bloomberg Social Velocity (BSV) alerts. The algorithm scans Twitter and StockTwits searching for posts that contain the company's cashtag or any mention of the company’s name (Bloomberg, n.d.). By applying the natural language processing algorithms, BSV preprocesses tweets to be analyzed. Then, using a supervising machine learning model, it assigns the numerical scores to the messages, representing estimated financial sentiment. The score takes categorical values (e.g., -1,+1,0), which are then multiplied by the confidence score. Confidence indicates the level of certainty of correct categorization and ranges from 0 to 100%. Thus, the sentiment of a tweet is a confidence-weighted score with values ranging from -1 to +1, with -1, 0, +1 representing very negative, neutral and very positive sentiment respectively. Furthermore, by aggregating and averaging the confidence-weighted sentiment scores from company related tweets during a 24-hours period, Bloomberg computes the firm-specific daily Twitter sentiment. The 24-hour period starts at 9:20 a.m. (10 minutes before NYSE opening time) on the previous day and ends at 9:20 a.m. on the current day.

3.2. Non-financial Variables

The Twitter activity and sentiment data used in this study comes from the daily firm-specific activity and sentiment measured by the BSV. For every firm that was a part of the DJIA index in the years 2017 to 2023 the data contains firm-specific tweets volume and daily average value of twitter sentiment, which are collected for the parent company over a 24-hour period, and published 10 minutes before the start of every trading day. However, as all the studied companies are independent, the Twitter data is related only to those companies. The academic literature finds evidence that the news sentiment is significant in predicting stock returns (Tan & Tas 2020; Tetlock 2007). This predictive power suggests that it might be a good control variable in this study, especially after considering that investors might trade on the sentiment of information from other sources, not necessarily from Twitter. Therefore, the data also contains the values of the average daily news sentiment, which is approximated similarly to the daily average twitter sentiment. It is also collected over a 24-hour period and published 10 minutes before the start of every trading day.

Considering the fact that Covid-19 pandemic had a significant impact on the economy and stock markets, I have segmented the sampled data into three distinct periods: before, during, and after the Covid-19 pandemic. To indicate in which period was the observation recorded, I construct the categorical variable period. This variable takes on values 0,1 or 2, if the observation occurred before, during or after the Covid-19 pandemic, repsectively. As the studied companies are a part of the DJIA index, which contains only US based firms, the pandemic period is determined by the arbitrary pandemic dates in the US. Thus, in the study I consider the period of Covid-19 pandemic to start on 31st January 2020, when the Secretary of Health and Human Services (HHS) declared the U.S. outbreak a public health emergency (The White House, 2020), and end on 18th of September 2022, when president Joe Biden, in his appearance in 60 minutes, declared that in his belief, the COVID-19 pandemic was "over" in the United States (Collinson, 2022).

3.3 Financial Variables

In addition to non-financial variables, I use constructed firm-specific financial variables. In constructing such variables I use financial data such as market capitalization, opening stock prices, highest and lowest daily stock prices, and daily trading volume, which are all collected from Bloomberg. Fama and French (1992), have shown that firm size significantly affects stock returns. Similarly to (Tan & Tas 2020), I use the natural logarithm of the firm's market capitalization to represent the company's size. Moreover, the dataset includes the firm's stocks' highest and lowest prices during a trading day, which are used in the calculation of the Park volatility (Parkinson 1980). Firstly, for each trading day t, I take the logarithm of the high/low ratio, which represent the ratio of highest to lowest stock price and then square the result. Then I take the average of such values of the 5 trading days from t-5 to t-1 and divide the result by a constant 4*ln(2). Lastly, I take a square root of such value, and thus construct Park volatility measure used in this research to account for the past volatility of a stock. The collected trading volume data is used to calculate Amihud's illiquidity measure (Amihud 2002). I take several steps to construct this variable. Firstly, I calculate absolute stock returns on day t. Then, for each stock I divide its absolute return on trading day t by its total turnover on that day. I take average of those values from trading days t-5 to t-1 to create a illiquidity measure variable used in this research. Moreover, academic literature finds that extreme trading volume, in comparison with the usual trading volume of a stock, significantly impacts subsequent stock returns (Gervais et. al, 2001). Therefore, following the methodology of Tetlock (2011), we construct abnormal turnover, as the natural logarithm of trading volume on day t minus its average log trading volume on days t-5 to t-1. Furthermore, as stock returns can be partly explained by the returns on previous days, following the Tan and Tas (2020) methodology, I construct a control variable HRet, which represents cumulative returns from day t-1 to t-5. In their research, Barclay and Hendershott (2003) find that low trading activity after hours can significantly impact stock prices. Thus, using open-to-open changes in stock prices to calculate stock returns is the most suitable choice for this study, as the daily Twitter sentiment data is published 10 minutes before the start of the trading day. This timing means that the sentiment values include data from after-hours trading, which indicate that there is a risk that this information could already be reflected in the opening prices of stocks. Therefore, such overlap between the sentiment measures and the stock returns could potentially bias the results of predictive regressions.

3.4 Variable Adjustments

The study examines a cross-sectional time series data for several companies. This data structure suggests that the time series for some stocks and variables may exhibit non-stationarity. Conducting an analysis on the sample that consists of both stationary and non-stationary time series using traditional methods would lead to spurious results. Therefore, before conducting any analysis I check for the stationarity of the data. To do so, I conduct augmented Dickey-Fuller tests for every time series of a company and variable. The results of those tests suggest that only time series of companies' size variable exhibit non-stationarity at the 5% significance level. Thus, I detrend the size variable by taking the first difference in it, and replacing it with the variable Diff_size. The newly constructed variable Diff_size represents the difference in the size of the company on day t and t-1. The results of the augmented Dickey-Fuller test on the Diff_size imply that the variable is stationary. Hence, after detrending the size variable, all variables used in the analysis exhibit stationarity.

However, some variables take especially small or especially big values. Thus, to gain better insights into the impact and behavior of those variables I adjust their values. The variable representing the firm's illiquidity is constructed by taking the average of the 5 trading days period Amihud's illiquidity measures. The values of that variable are particularly low, thus I multiply its value by a factor of 1,000,000. Furthermore, as the numbers of tweets are relatively high, I multiply this variable by a factor of 1/10000. The scaling of the variables does not have any effects on influence of the independent variables on the stock returns and the results of the regression, except for the fact that the coefficients of the scaled variables should be descaled, while interpreting.

3.5 Summary Statistics

The data sample consists of stocks that were part of the DJIA index in years from 2017 to 2023. The academic literature suggests that joining and leaving the index might impact the company's stock price (Bennett, Stulz, & Wang, 2020; Miller & Ward 2015). Thus, to avoid such bias, I investigate only the stocks of the companies that were part of the DJIA index throughout the whole period. Hence, the dataset consists of 44,000 observations, which represent data about 25 companies from all trading days since the start of 2017 till the end of 2023. However, the dataset contains 43,920 and 43,407 values representing the number of company related tweets and firm-specific daily average Twitter sentiment respectively. This implies that 80 and 593 values for those variables are missing. Furthermore, the dataset does not contain values for the 3rd of January 2017 and 29th of December 2023. The observations with the missing values are not included in the study, thus resulting in the final sample of 43310 observations. The summary statistics of the sample are shown below in the Table 1.

Table 1.

	N	Mean	Median	SD	P10	P90
tweets	43310	323.8	98	872.67	17	753
TwitterS	43310	0.0072	0.0017	0.1212	-0.0534	0.08
News	43310	0.0032	0	0.2111	-0.1066	0.1336
Vola	43310	0.0121	0.0106	0.0067	0.0064	0.0188
Diff_size	43310	0.0003	0.0006	0.0178	-0.0171	0.0172
HRet	43310	0.0029	0.0039	0.0374	-0.0371	0.0418
Illiq	43310	0.0020	0.0013	0.0020	0.0004	0.0044
AbTurn	43310	0.0003	-0.0288	0.3387	-0.3790	0.4116
Ret	43310	0.0006	0.0008	0.0171	-0.0169	0.0176

The presented above mean of the Twitter sentiment is 0.007, implying that the average Twitter sentiment regarding sampled companies is slightly positive. The average news sentiment is somewhat neutral as the mean is 0.003 and the median is 0. The mean of the firm-specific open-to-open returns is 0.001, indicating on average, small daily increases in the stock prices, but the mean is consistent with the general upward trend of the market in the investigated period, except for the Covid-19. Furthermore, the diff_size, volatility, abnormal turnover, illiquidity, cumulative returns and number of firm-specific tweets have the means of 0.001, 0.012, 0.001, 0.002, 0.003, 323.814, respectively. Furthermore, as I investigate the change between the impact of the firm-specific Twitter sentiment on the companies' stock prices the sample was split into three periods. The periods are designated by the Covid-19 pandemic. The pre Covid period has 12990 observations, Covid period has 22588 observations, while post Covid has 7732

CHAPTER 4 Method

4.1 Hypothesis 1

The sampled data consists of 25 company stocks over the period of 6 years, hence the small number of entities and large time period are the characteristic of the data. Furthermore, the dataset is not balanced, thus to get the estimation of the external sentiment indicators on stock performance, I use the General Least Squares (GLS) regression model. This model is efficient when estimating the regression with low number of cross-section units and high number of time periods, as it corrects for heteroskedasticity and autocorrelation, enhancing estimation accuracy. Additionally, GLS effectively manages the unbalanced nature of the dataset in this research, thus eliminating the need to balance the panels — a process that would remove valuable information and lead to less robust results. Thus, GLS is a suitable model to use, considering the characteristics of the sampled dataset. To investigate whether the Twitter sentiment impact the stock returns, I adopt a hierarchical approach. This approach allows for a step-by-step assessment of the incremental contributions of the included variables, while highlighting the changes in the impact of the variables of interest on stock returns. Thus, to measure the relation between the firm-specific Twitter sentiment and company's stock returns, I use GLS regressions with the following equations:

$$Ret_{i,t} = \alpha_i + \beta_1 TwitterS_{i,t-1} + \varepsilon_{it} \tag{1}$$

$$Ret_{i,t} = \alpha_i + \beta_1 TwitterS_{i,t-1} + \beta_2 Ntweet_{i,t-1} + \varepsilon_{it} \tag{2}$$

$$Ret_{i,t} = \alpha_i + \beta_1 TwitterS_{i,t-1} + \beta_2 Ntweet_{i,t-1} + \beta_3 HRet_{i,t} + \varepsilon_{it} \tag{3}$$

$$Ret_{i,t} = \alpha_i + \beta_1 TwitterS_{i,t-1} + \beta_2 Ntweet_{i,t-1} + \beta_3 HRet_{i,t} + \beta_4 Diff\_size_{i,t-1} + \varepsilon_{it} \tag{4}$$

$$Ret_{i,t} = \alpha_i + \beta_1 TwitterS_{i,t-1} + \beta_2 Ntweet_{i,t-1} + \beta_3 HRet_{i,t} + \beta_4 Diff\_size_{i,t-1} + \beta_5 AbTurn_{i,t-1} + \beta_6 Illiq_{i,t} + \beta_7 Vola_{i,t} + \beta_8 News_{i,t-1} + \varepsilon_{it} \tag{5}$$

The dependent variable labelled Ret, is an open-to-open stock return from day t to t+1. The independent variables include TwitterS and Ntweet representing the firm-specific Twitter sentiment and the firm-related number of tweets collected, both collected over the 24-hour period before the market opening on trading day t. HRet, Vola and Illiq represent cumulative returns, Park volatility (Parkinson 1980) measure and average of Amihud's illiquidity measure (Amihud 2002) respectively, over the trading days t-5 to t-1. Additionally, I use Diff_size and News as control variables. Former is calculated by substracting firm's size on trading day t-1 from firm's size on trading day t, where size is the natural logarithm of the firm's market capitalization. The latter is numerical representation of news sentiment over the 24-hour period before market opening on day t, captured by BSV. Control variables representing size, historical returns, volatility measure, illiquidity measure and return on previous days are commonly used in the academic literature to represent stock characteristics (Tetlock 2011; Sprenger et. al 2013). In addition to the variables above, in their research to find the role of Twitter sentiment in predicting the stock returns, Tan and Tas (2020) also control for abnormal turnover and news sentiment, as Tetlock (2011) finds both variables having a significant impact on the stock performance. Hence, not controlling for those variables would lead to omitted variable bias.

4.2 Hypothesis 2

To answer the second hypothesis, whether the impact of Twitter's firm-specific sentiment on the next day company's stock returns is changing between the periods, I use a GLS regression. To ensure the same level of effect of the control variables on dependent variable throughout all periods, I do not perform 3 regressions, one for each time period, but one regression. This approach allows me to capture the impact of the firm specific Twitter sentiment for each period, while controlling for the consistent influence of other variables on the dependent variable throughout the whole sampled period. The GLS regression used in investigating whether the impact of the firm-specific Twitter sentiment on the companies' stock returns changes throughout the time, has the following equation:

$$Ret_{i,t} = \alpha_i + \beta_1 TwitterS_{i,t-1} + \beta_2 TwitterS_{i,t-1} * Period + \beta_3 TwitterS_{i,t-1} * Period + \beta_4 Period + \beta_5 Period + \sum \beta_{i+5} Control_{i,t-1} + \varepsilon_{it}$$

The equation contains interactions of the Twitter sentiment with the period variable. Furthermore, the control variables included in the regression are Ntweets, News, HRet, Vola, Iliiq, Diff_size and AbTurn, which represent the same variables as in equation (5). Upon estimation of the model, the coefficient of the TwitterS variable will represent the average influence of this variable on the stock returns in the pre-Covid-19 period. The coefficients of the interactions of TwitterS with the period variable will indicate how much different is the impact of Twitter sentiment on stock returns in respective period, compared to the pre-Covid-19 period. Thus, the significance of those coefficients implies whether the impact of the variables of interest on stock returns differs from the influence in pre-Covid-19 period. In order to test whether the impact of the Twitter sentiment on stock returns is significantly different between Covid-19 pandemic and post pandemic period, I conduct a Wald test on the coefficients of the interaction terms. I choose the Wald test over the F-test because, unlike the F-test, the GLS model does not require the assumption of normally distributed errors. Therefore, to accommodate potential non-normality of errors, the Wald test is preferred.

CHAPTER 5 Results & Discussion

5.1 Hypothesis 1

This section inspects whether the firm-specific Twitter sentiment have predictive power in estimating company's stock returns. To investigate that hypothesis I use the hierarchal approach, estimating five GLS regressions. The results of that method are shown in Table 2. The first column shows the results of the basic model that estimates firm-specific Twitter sentiment impact on the company's stock returns, without controlling for any characteristics. The coefficient of the Twitter sentiment is positive and significant at 10% level, with the value of 0.0011, indicating that on average the increase in the firm-specific Twitter sentiment by 1 leads to the increase in the next day's firm's stock returns by 0.11%. Furthermore, the results of the chi-squared test on the significance of the model indicate that it provides a statistically significant improvement in predicting the outcome over a baseline model with no predictors. Columns 2 and 3 show the results of the regressions that subsequently add the variables Ntweets and Diff_size, respectively. Controlling for the number of firm-specific tweets does not significantly change the results. However, after accounting for the company's change in size, the influence of the firm-specific Twitter sentiment on company's stock returns changes from positive to negative. Furthermore, its significance increases to the 1% level. Results shown in columns 4 and 5 indicate that while more control variables are added to the model the sign and significance of the Twitter sentiment on stock returns does not change. In the fourth and in the final model, the impact of the Twitter sentiment on stock returns is negative and significant at 1% level. The coefficient of the Twitter sentiment estimated by the full model indicate that an increase in the firm-specific Twitter sentiment by 1, on average results in decrease in next trading day's company's stock returns by 0.14%. Moreover, the results of the final model indicate that other than Twitter sentiment, only variables representing the size, abnormal turnover and historical returns have significant explanatory power in explaining the company's next trading day stock returns. Interestingly, the coefficient of the HRet is negative, implying that the positive historical returns have a negative impact on the next trading day firm's stock price.

In conclusion, the results of the base model that do not include any control variables imply that the influence of the firm-specific Twitter sentiment on the company's next trading day stock returns is positive and significant at 10% level. On the other hand, after controlling for the firm's changes in market capitalization, which approximate the company's size, the effect of the firm-specific Twitter sentiment on the company's stock returns is negative. This negative impact is consistent throughout the models that account for differences in size and for other financial and non-financial firm characteristics. Thus, the results suggest that effect of the Twitter sentiment on the stock returns

Table 2.

Predicting Stock returns using Twitter sentiment and activity

	(1)	(2)	(3)	(4)	(5)
TS	0.0011* (0.0006)	0.0011* (0.0006)	-0.0031*** (0.0005)	-0.0014*** (0.0005)	-0.0014*** (0.0005)
Tweets		-0.0006 (0.0010)	-0.0010 (0.0008)	-0.0005 (0.0008)	-0.0005 (0.0008)
Diff_size			0.5571*** (0.0037)	0.5721*** (0.0037)	0.5710*** (0.0037)
HRet				-0.0513*** (0.0018)	-0.0516*** (0.0018)
Illiq					-0.0276 (0.0333)
Vola					0.0127 (0.0104)
AbTurn					-0.0008*** (0.0002)
News					0.0002 (0.0003)
Intercept	0.0005*** (0.0001)	0.0005*** (0.0001)	0.0004*** (0.0001)	0.0005*** (0.0001)	0.0004*** (0.0001)
N	43310	43310	43310	43310	43310
Prob > chi2	0.0865	0.1882	0.0000	0.0000	0.0000

Note: This table shows the results from the GLS regressions. Dependent variable Ret is open-to-open stock returns from trading day t to t+1. Twitter sentiment and Twitter publication count are firm-specific Twitter sentiment and number of firm-specific tweets. Control variable news sentiment represents the firm-specific news sentiment score, while size is the logarithm of the firm's market capitalization on trading day t-1. Control variables HRet, Illiq, Vola and AbTurn represent cumulative returns in the trading days t-5 to t-1, average of Amihud’s illiquidity measure over days t-5 to t-1, Park’s volatility measure on trading days from t-5 to t-1 and logarithm of turnover on day t minus the average logarithm of turnover on days t-5 to t-1, respectively. Standard errors are Newey-West (Newey and West 1987) adjusted up to 5 lags for heteroskedasticity and autocorrelation, and appear in parenthesis. *, **, *** indicate statistical significance at 10%, 5%, 1% levels, respectively.

in base model may be biased as it does not include several variables that significantly impact company’s stock returns. This indicates that the results of the final model that includes all constructed control variables should be considered when answering the hypothesis. The results of the conducted analysis show that after controlling for several variables the impact of the Twitter sentiment on the stock returns in negative. Therefore, I conclude that I reject the hypothesis of no effect of firm-specific Twitter sentiment on company’s stock returns.

5.2 Hypothesis 2

This sections shows the results of the investigation of hypothesis 2 and provides an answer to it. The results of the GLS regression are shown in the table below.

Table 3
Predicting stock returns using Twitter sentiment and activity for different periods

	(1)
TS	-0.0007 (0.0009)
Covid Period	-0.0001 (0.0001)
Post Covid Period	-0.0000 (0.0002)
TS during Covid	-0.0018 (0.0012)
TS after Covid	-0.0005 (0.0012)
Ntweets	-0.0005 (0.0008)
Diff_size	0.5710*** (0.0037)
HRet	-0.0515*** (0.0018)
Illiq	-0.0250 (0.0334)
Vola	0.0140 (0.0106)
AbTurn	-0.0008*** (0.0002)
News	0.0002 (0.0003)
Intercept	0.0005*** (0.0002)
N	43310
Prob > chi2	0.0000

Note: This table shows the results from the GLS regressions. Dependent variable Ret is open-to-open stock returns from trading day t to t+1. Twitter sentiment and Twitter publication count are firm-specific Twitter sentiment and number of firm-specific tweets. Control variable news sentiment represents the firm-specific news sentiment score, while size is the logarithm of the firm's market capitalization on trading day t-1. Control variables HRet, Illiq, Vola and AbTurn represent cumulative returns in the trading days t-5 to t-1, average of Amihud's illiquidity measure over days t-5 to t-1, Park's volatility measure on trading days from t-5 to t-1 and logarithm of turnover on day t minus the average logarithm of turnover on days t-5 to t-1, respectively. Column 1,2,3 represents results that include Twitter sentiment variable, while coumns 4,5,6 do not include that varibale. Standard errors are Newey-West (Newey and West 1987) adjusted up to 5 lags for heteroskedasticity and autocorrelation, and appear in parenthesis. *, **, *** indicate statistical significance at 10%, 5%, 1% levels, respectively.

The coefficients for the period variable during the Covid-19 pandemic and the post-pandemic period are not statistically significant. This indicates that there has been no significant shift in the general trend across these periods, meaning that the pandemic itself or its subsequent period do not have a significant effect on the stock returns. The coefficient of the Twitter sentiment is negative and insignificant at 10% level. This indicates that before the Covid-19 pandemic the impact of the firm-specific Twitter sentiment on company's next trading day stock returns was insignificant. Furthermore, both coefficients of the interactions between Twitter sentiment and variable period are negative and insignificant at 10% level. The insignificance of those coefficients implies that there is no significant difference in the influence of the firm-specific Twitter sentiment on the company's stock returns between the pre Covid-19 pandemic period and, during and after Covid-19 pandemic periods. Moreover, the results of the conducted Walt test for the difference in the coefficients of the Twitter sentiment interactions with period variable, indicate that there is no significant difference in the impact of the Twitter sentiment on stock returns during and after the Covid-19 pandemic. Thus, the effect of the Twitter sentiment on stock returns does not differ significantly between the three periods. Therefore, I conclude that there is no sufficient evidence to reject the hypothesis of no significant impact of firm-specific Twitter sentiment on the next day's company stock returns remains consistent across the periods before, during, and after the Covid-19 pandemic.

5.3 Discussion

The findings of this paper suggest that for the companies that were part of the DJIA index in years from 2017 to 2023, there is a significant impact of the firm-specific Twitter sentiment and activity on the next trading day company's stock returns at the 1% significance level. Although I find the significant relationship between the Twitter sentiment and stock returns the impact is negative. Thus, the results are contrary to the findings of Tan and Tas (2020), Sul, Dennis, and Yuan (2017) and Gu and Kurov (2020), as they find that the Twitter sentiment has a positive significant impact on the company's stock returns. The results shown in my research further deviate from those described in academic literature by Tetlock (2011) and Tan and Tas (2020), as contrary to those papers, I find that both Park's volatility measure and the average of the Amihud's illiquidity measure over the past 5 trading days and news sentiment, do not have a significant impact on the firm's stock returns. On the other hand, the results suggest that the company's size has a significant influence on the company's stock returns, which is consistent with the existing academic literature that the size of the company has an impact on the firm's stock prices (Tetlock 2011). Furthermore, similarly to the Tetlock (2011) and Tan and Tas (2020) I find that the abnormal turnover is significantly impacting the company's stock returns.

The finding that the firm-specific Twitter sentiment has explanatory power in predicting the next trading day company's stock returns, further undermines the EMH as assumed by Samuelson (1965), which suggests that stock prices reflect all available information. On the other hand, the results support the findings of Gu and Kurov (2020), as they find argue that the Twitter posts contain new information that once posted, is incorporated into the stock prices, and thus is relevant for predicting the stock returns. Furthermore, the findings of this paper give additional evidence of the irrationality of the individual investors, as they suggest that the firm-specific sentiment of a company on social media platform impacts its stock price. Moreover, the results shown in this paper are in line with the research conducted by Lee, Shleifer and Thaler (1991). The authors of that paper find that individual investors affect the asset prices, and that movement in asset prices can be partly attributed to the changes in investors sentiment, which is consistent with the results that Twitter sentiment negatively impact the stock returns, as the investors that make decisions based on the sentiment of the read posts, buy or sell stocks causing fluctuations in their prices.

The negative impact of the firm-specific Twitter sentiment on the next trading day company's stock returns is somewhat counterintuitive, as one could expect that the positive opinions and mood about a particular stock should increase its price and not decrease it. One possible explanation of such a counterintuitive phenomena could strive from the fast that the social media is a perfect environment for the fast diffusion of the information. Fang and Peress (2009) argue that mass media coverage of securities can dissimate information broadly, reaching the big audience and affecting the stock priecs. Thus, once new information is shared with the users, it quickly reaches significant user base. Furthermore, with the easy access to the internet and online trading platforms, individual investors can change their portfolio within minutes. Therefore, it is likely that upon seeing the news, individual investors trade on sentiment within the same day, causing overall overreaction, which is then prone to the reversals. Hence, causing the negative stock returns.

Despite far-reaching impacts of the COVID-19 pandemic on global markets, highlighted by increase in trading activities (Ortmann, Pelster, & Wengerek, 2020), I do not find evidence that the effects of firm-specific Twitter sentiment on the company's next trading day stock returns change across different periods, the pre-pandemic, pandemic, and post-pandemic times. Ashraf (2020) finds that the stock market returns decrease with the increases in the growth of confirmed Covid-19 infections, however I find that the effect of the Covid-19 pandemic, although negative, does not significantly impact the stock returns. This result suggests that other factors should cause the decrease in stock market returns and not the Covid-19 pandemic itself. Katsafados, Nikoloutsopoulos and Leledakis (2023) find a significantly positive relationship between the Twitter's Covid-19 sentiment and the prices of stock indexes of several countries during the Covid-19 pandemic. On the other hand, I find that the effect of the firm-specific Twitter sentiment on the company's stock returns is negative during the Covid-19 pandemic. This implies that while the positive overall sentiment about the Covid-19 pandemic was increasing prices of stock indexes, the positive sentiment about the company was decreasing its stock returns.

CHAPTER 6 Conclusion

6.1 Conclusion

This paper aimed to investigate the persistency of the influence of firm-specific Twitter sentiment on company's next trading day stock returns across three periods, pre, during, and post-COVID-19 pandemic times. Additionally, in this study I tried to examine whether the firm-specific Twitter sentiment significantly impacts company's stock returns. The motivation behind this study comes from an increasing reliance on social media for financial information and ever growing social media user base, which should imply increasing influence of social media on stock prices. Moreover, there exists a very limited academic literature that would provide a comprehensive understanding about whether the impact of social media on stock prices or returns varies significantly during periods of market stress, such as a global pandemic, thus providing further incentive to conduct research in this unexplored area. Two research questions were studies in this paper. Firstly, I studied the question: "What is the influence of the firm-specific Twitter sentiment on the company's next day stock returns". The second question I examined was: "Does the impact of the firm-specific Twitter sentiment on the company's next trading day stock returns of DJIA index companies changes significantly across the periods of pre, during and post Covid-19 pandemic".

To investigate both questions I employed a robust methodology. Firstly, I collected firm-specific financial and non-financial data including the firm-specific Twitter and news sentiment as well as the tweet volume from Bloomberg. Upon the collection of the data I constructed several control variables that are commonly used in the academic literature to represent stock characteristics. Additionally, I addressed non-stationarity in the data by taking the first difference of variables that exhibited non-stationary characteristics. To investigate the first hypothesis I used hierarchal approach, in which I used several General Least Squares (GLS) regressions with subsequently added control variables to the base model, which estimated effects of the firm-specific Twitter sentiment on the company's next trading day stock returns. The results I found indicate that the firm-specific Twitter sentiment has a significantly negative impact on the company's next trading day stock returns. To answer the second research question I also used the GLS regression. However, this time in addition to all control variables, I included the interaction terms of the Twitter sentiment variable with period variable, to estimate the differences in the effects of the Twitter sentiment across the periods. Furthermore, I used Wald test to examine whether there is a significant difference in the impact of the Twitter sentiment on stock returns during and after the Covid-19 pandemic. The analysis showed that the coefficients for the interactions of the Twitter sentiment and period variable were statistically insignificant. This indicates a consistent influence of Twitter sentiment across these times. Hence, no indicators of significant shifts in how Twitter sentiment impacts stock returns across the periods were found.

6.2 Implications

This study shows that despite the fact that the academic literature includes several studies, which indicate that the relationship between the Twitter sentiment and stock returns is positive for particular periods and indexes, the impact of the firm-specific Twitter sentiment on the stock returns of the companies that are part of the DJIA index is negative. Furthermore, this research enriches our understanding of the social media influence on the stock returns, by demonstrating that the relationship between Twitter sentiment and stock returns remains stable over time. Moreover, the results of this research indicate that the impact of the Twitter sentiment on the stock returns is robust to the external shocks such as Covid-19 pandemic. Therefore, this study provides insights for investors and policymakers on the enduring influence of social media sentiment in financial decision-making. Future research could delve deeper into the mechanisms of this persistency and explore if similar patterns hold across different social media platforms or other stock indices.

The study's findings suggest practical implications for investors, particularly in how they can utilize Twitter sentiment analysis as a reliable component of their investment strategies. Investors might benefit from incorporating sentiment analysis tools into their trading algorithms to help predict short-term stock movements. Furthermore, the predictive power of the Twitter sentiment is shown to be consistent throughout the periods, thus becoming a component that could be consistently included in the trading strategy, as the effect of the sentiment are robust to the economic shocks.

6.3 Limitations

There are couple of limitations of this study. One such limitation of this study comes from the inability to fully address autocorrelation in the error terms when using the GLS model, due to the unbalanced nature of the panel data. The GLS model was chosen for its efficiency in dealing with heteroscedasticity and its capability to handle panels with a low number of cross-sectional units and a high number of time periods. However, having unbalanced panels I was not able to address the issue of the autocorrelation in the error terms, which could potentially affect the reliability of the regression results, as this limitation could lead to over or underestimation of the effect of the firm-specific Twitter sentiment on company's stock returns.

Another limitation of this research is the potential for omitted variable bias. While this study controlled for several variables that academic literature finds to influence the stock returns, such as company size, illiquidity, stock volatility and historical returns, there may be other variables that could influence the relationship between the Twitter sentiment and stock returns, but were not included. Not accounting for those variables could distort the estimated effects of Twitter sentiment, leading to biased results.

REFERENCES

1. Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. https://doi.org/10.1016/s1386-4181(01)00024-6
2. Ashraf, B. N. (2020). Stock markets' reaction to COVID-19: Cases or fatalities? *Research in International Business and Finance*, *54*, 101249. https://doi.org/10.1016/j.ribaf.2020.101249
3. Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *□the □Journal of Economic Perspectives/□the □Journal of Economic Perspectives*, *21*(2), 129–151. https://doi.org/10.1257/jep.21.2.129
4. Barclay, M. J., & Hendershott, T. (2003). Price discovery and trading after hours. *Review of Financial Studies/□the □Review of Financial Studies*, *16*(4), 1041–1073. https://doi.org/10.1093/rfs/hhg030
5. Bennett, B., Stulz, R., & Wang, Z. (2020). Does the stock market make firms more productive? *Journal of Financial Economics*, *136*(2), 281–306. https://doi.org/10.1016/j.jfineco.2019.09.006
6. Bloomberg. (n.d.). Trending on Twitter: Social sentiment analytics. Retrieved from https://www.bloomberg.com/company/press/trending-on-twitter-social-sentiment-analytics/
7. Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007
8. Campbell, J. Y., & Shiller, R. J. (1988). The Dividend-Price Ratio and expectations of future dividends and discount factors. *Review of Financial Studies/□the □Review of Financial Studies*, *1*(3), 195–228. https://doi.org/10.1093/rfs/1.3.195
9. Collinson, S. (2022, September 18). Biden declares the pandemic is over on '60 Minutes.' CNN. Retrieved from https://www.cnn.com/2022/09/18/politics/biden-pandemic-60-minutes/index.html
10. Cueva, C., Iturbe-Ormaetxe, I., Ponti, G., & Tomás, J. (2019). An experimental analysis of the disposition effect: Who and when? *Journal of Behavioral and Experimental Economics*, *81*, 207–215. https://doi.org/10.1016/j.socec.2019.06.011
11. Deb, M., & Chavali, K. (2009). A study of gender differences in investment behaviour. *Asian-Pacific Business Review*, *5*(3), 45–55. https://doi.org/10.1177/097324700900500305
12. Eppen, G. D., & Fama, E. F. (1969). Cash Balance and Simple Dynamic Portfolio Problems with Proportional Costs. *International Economic Review*, *10*(2), 119. https://doi.org/10.2307/2525547
13. Fama, E. F., & French, K. R. (1992). The Cross-Section of expected stock returns. *□the □Journal of Finance/□the □Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x
14. Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *□the □Journal of Finance/□the □Journal of Finance*, *64*(5), 2023–2052. https://doi.org/10.1111/j.1540-6261.2009.01493.x
15. Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The High-Volume Return premium. *□the □Journal of Finance/□the □Journal of Finance*, *56*(3), 877–919. https://doi.org/10.1111/0022-1082.00349

16. GDELT Blog. (2022). Visualizing a decade of Twitter's evolution: Jan 2012 - Nov 2022 & Musk's purchase. Retrieved from https://blog.gdeltproject.org/visualizing-a-decade-of-twitters-evolution-jan-2012-nov-2022-musks-purchase/
17. Gu, C., & Kurov, A. (2020). Informational role of social media: Evidence from Twitter sentiment. *Journal of Banking & Finance*, *121*, 105969. https://doi.org/10.1016/j.jbankfin.2020.105969
18. Hong, H., & Stein, J. (1999). *Differences of opinion, rational arbitrage and market crashes*. https://doi.org/10.3386/w7376
19. Jain, R., Jain, P., & Jain, C. (2015). Behavioral biases in the decision making of individual investors. *Journal of Management and Research*, *14*(3), 7. https://www.questia.com/library/journal/1P3-3776384251/behavioral-biases-in-the-decision-making-of-individual
20. Jaiswal, B., & Kamil, N. (2012). Gender, behavioral finance and the investment decision. *Business Review*, *7*(2), 8–22. https://doi.org/10.54784/1990-6587.1201
21. Jiao, P., Veiga, A., & Walther, A. (2020). Social media, news media and the stock market. *Journal of Economic Behavior & Organization*, *176*, 63–90. https://doi.org/10.1016/j.jebo.2020.03.002
22. Katsafados, A. G., Nikoloutsopoulos, S., & Leledakis, G. N. (2023). Twitter sentiment and stock market: a COVID-19 analysis. *Journal of Economic Studies*, *50*(8), 1866–1888. https://doi.org/10.1108/jes-09-2022-0486
23. Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking & Finance*, *26*(12), 2277–2299. https://doi.org/10.1016/s0378-4266(01)00202-3
24. Charles M. C. Lee, Shleifer, A., & Thaler, R. H. (1991). Investor Sentiment and the Closed-End Fund Puzzle. *The Journal of Finance*, *46*(1), 75–109. https://doi.org/10.2307/2328690
25. Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., & Chen, Y. (2014). The effect of news and public mood on stock movements. *Information Sciences*, *278*, 826–840. https://doi.org/10.1016/j.ins.2014.03.096
26. Liew, J., & Budavari, T. (2017). The "Sixth" Factor—A Social Media Factor Derived Directly from Tweet Sentiments. *Journal of Portfolio Management*, *43*(3), 102–111. https://doi.org/10.3905/jpm.2017.43.3.102
27. Long, S., Lucey, B., Xie, Y., & Yarovaya, L. (2022). "I just like the stock": The role of Reddit sentiment in the GameStop share rally. *the Financial Review/the Financial Review*, *58*(1), 19–37. https://doi.org/10.1111/fire.12328
28. Mazur, M., Dang, M., & Vega, M. (2021). COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. *Finance Research Letters*, *38*, 101690. https://doi.org/10.1016/j.frl.2020.101690
29. Miller, C., & Ward, M. (2015). The market impact on shares entering or leaving JSE indices. *the Investment Analysts Journal*, *44*(1), 84–101. https://doi.org/10.1080/10293523.2015.994450
30. Mittal, A., & Goel, A. (2011). Stock prediction using Twitter sentiment analysis. Proceedings of the 2011 Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA 2011).

31. Mushinada, V. N. C., & Veluri, V. S. S. (2018). Investors overconfidence behaviour at Bombay Stock Exchange. *International Journal of Managerial Finance*, *14*(5), 613–632. https://doi.org/10.1108/ijmf-05-2017-0093
32. Ortmann, R., Pelster, M., & Wengerek, S. T. (2020). COVID-19 and investor behavior. *Finance Research Letters*, *37*, 101717. https://doi.org/10.1016/j.frl.2020.101717
33. Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *□the □Journal of Business/□the □Journal of Business*, *53*(1), 61. https://doi.org/10.1086/296071
34. Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PloS One*, *10*(9), e0138441. https://doi.org/10.1371/journal.pone.0138441
35. Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, *11*(3), 9–16. https://doi.org/10.3905/jpm.1985.409007
36. Samuelson, P. A. (1965). A Theory of Induced Innovation along Kennedy-Weisacker Lines. *□the □Review of Economics and Statistics*, *47*(4), 343. https://doi.org/10.2307/1927763
37. Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. *Journal of Economic Behavior & Organization*, *107*, 730–743. https://doi.org/10.1016/j.jebo.2014.06.004
38. Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2013). Tweets and Trades: the Information Content of Stock Microblogs. *European Financial Management*, *20*(5), 926–957. https://doi.org/10.1111/j.1468-036x.2013.12007.x
39. Sul, H. K., Dennis, A. R., & Yuan, L. (2016). Trading on Twitter: Using social media sentiment to predict stock returns. *Decision Sciences*, *48*(3), 454–488. https://doi.org/10.1111/deci.12229
40. Tan, S. D., & Tas, O. (2020). Social media sentiment in international stock returns and trading activity. *□the □Journal of Behavioral Finance*, *22*(2), 221–234. https://doi.org/10.1080/15427560.2020.1772261
41. The White House. (2020). Proclamation on Declaring a National Emergency Concerning the Novel Coronavirus Disease (COVID-19) Outbreak. Retrieved from https://trumpwhitehouse.archives.gov/presidential-actions/proclamation-declaring-national-emergency-concerning-novel-coronavirus-disease-covid-19-outbreak/
42. Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *□the □Journal of Finance/□the □Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x
43. Tetlock, P. C. (2011). All the News That's Fit to Reprint: Do Investors React to Stale Information? *Review of Financial Studies/□the □Review of Financial Studies*, *24*(5), 1481–1512. https://doi.org/10.1093/rfs/hhq141
44. Yousaf, I., Youssef, M., & Goodell, J. W. (2022). Quantile connectedness between sentiment and financial markets: Evidence from the S&P 500 twitter sentiment index. *International Review of Financial Analysis*

(Online)/International Review of Financial Analysis, *83*, 102322. https://doi.org/10.1016/j.irfa.2022.102322

45. Zaremba, A., Kizys, R., Aharon, D. Y., & Demir, E. (2020). Infected Markets: Novel Coronavirus, Government Interventions, and Stock Return Volatility around the Globe. *Finance Research Letters*, *35*, 101597. https://doi.org/10.1016/j.frl.2020.101597