5-3-2024

Twitter Sentiments and Stock Prices: An Event Study on the Role of Influencers

Denys Wincel Lu

PERMISSION TO BORROW

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Denys Wincel Pua Lu
J. Mack Robinson College of Business
Georgia State University
Atlanta, GA 30302-4015

The director of this dissertation is:

Dr. Likoebe M. Maruping
J. Mack Robinson College of Business
Georgia State University
Atlanta, GA 30302-4015

Twitter Sentiments and Stock Prices: An Event Study on the Role of Influencers

by

Denys Wincel Pua Lu

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Executive Doctorate in Business

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY

ROBINSON COLLEGE OF BUSINESS

2024

ACCEPTANCE

This dissertation was prepared under the direction of the *DENYS WINCEL PUA LU* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Dr. Likoebe Maruping

Dr. Po-An Hsieh

Dr. Denish Shah

ACKNOWLEDGEMENTS

I want to thank my family, especially my wife, Sharlene, and our amazing children, Madeline and Anthony, for their unconditional love, enthusiasm, and patience. They have always believed in me and supported me in this significant undertaking. They are my pride and joy, and I dedicate this dissertation to them.

I would also like to express my deepest gratitude to my dissertation committee, my advisor, Dr. Likoebe Maruping, Dr. Po-An Hsieh, and Dr. Denish Shah for their invaluable guidance, feedback, and support throughout this research journey. They have been inspiring mentors, encouraging critics, and influential sources of knowledge and wisdom. I have learned so much from them and am honored to have worked with such distinguished scholars. I want to thank the DBA Program Office which helped provide critical support and guidance to create the environment to make this journey possible.

My old and new friends have been a source of joy, comfort, and motivation during this challenging journey. I am grateful for their friendship, understanding, and companionship. I want to especially acknowledge my friends from the DBA 2024, who have shared this journey with me and have been there through it all. They have enriched my academic, personal, and professional life in countless ways.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Twitter Sentiments and Stock Prices: An Event Study on the Role of Influencers

by

Denys Wincel Pua Lu

May 2024

Chair: Dr. Likoebe Maruping

Major Academic Unit: Department of Computer Information Systems

For many companies, some consumers enthusiastically follow brands and may have insights rivaling those of professional financial analysts based on their knowledge of the companies. Often, these brand-loyal consumers express their thoughts and opinions on social media may be received by others in their communities, driving users to follow them based on the perception of expertise and trustworthiness. In academic literature, such users are referred to as *Social Media Influencers* (SMI).

The present study investigates the relationship between sentiments of Twitter posts and abnormal stock price returns. It further explores if source credibility operationalized as followership affects this relationship. SMIs are perceived to have higher source credibility, and it is expected that the relationship is stronger with SMIs than with non-SMI users. Sentiment analysis categorizes tweets into positive, neutral, and negative based on pre-trained models and machine learning.

Tweets made between 2017 and 2022 for four specific firms are analyzed using an event study approach. Events are identified using an automated algorithm, and abnormal returns are estimated using the Market model. Tweets are split based on percentile rankings of posting authors' follower counts.

The study finds that positive sentiments are generally statistically significant in identifying positive cumulative abnormal returns. Furthermore, the novel approach of gradually including fractiles of follower numbers shows that the significance of abnormal returns is not homogenous across all users. Contrary to expectations, statistical significance is stronger for a longer duration around identified events with tweets posted by users at the bottom 30% of followership. In contrast, tweets made by the users with more followers are not statistically significant until 80%. This finding suggests that the sentiments of tweets from users with a lower number of followers are more strongly related to the abnormal returns of stock prices.

This study shifts the focus of most extant research by using a broader set of tweets from "ordinary" users instead of investment-oriented users in exploring the relationship. The study also contributes to the relatively underexplored effect of SMI by taking an iterative approach to studying tweets through comparative analysis across different numbers of followers.

INDEX WORDS: Sentiment Analysis, Source Credibility, Twitter, Stock Price, Event Study, Social Media Influencers

I INTRODUCTION

I.1 Background

Over the past few decades, social media has risen in popularity, connecting people to engage in conversations and “discuss their opinions, spread information, and let others be part of their thoughts and experiences” (Stieglitz et al., 2020, p. 395). Using social media, individual users can interact with other individuals and collective entities such as companies and organizations. Users can also self-organize into groups where members have a common interest, such as video games, hobbies, or professions. A key feature of social media platforms is allowing users to subscribe or follow each other – typically, someone influential, attractive, or of celebrity status.

The COVID-19 pandemic also affected engagement on social media networks – increasing the use and adoption of social media by new users previously not active due to the increasing feeling of loneliness as they were separated. This loneliness, in turn, further increased individuals' consumption and engagement of social media (Donthu & Gustafsson, 2020; Dubbelink et al., 2021) during the pandemic.

With so many users engaged in social media, it has become a powerful tool across many areas for prediction and understanding social phenomena. To engage and build awareness directly with consumers, companies use social media to build brand awareness, acquire new customers, introduce new products and services, and retain current customers (Moorman, 2018). Often, companies engage with potential and current customers directly through their owned social media accounts (Liadeli et al., 2022). These accounts are not tied to a specific individual but allow a company to interact as a single entity on social media.

Jansen (2009, p. 2184) found that about 19% of tweets mentioned an organization or product brand. Of those, 20% expressed a sentiment or opinion concerning that company, product, or service. Companies and organizations can benefit from understanding the relationship between online social media sentiments and the financial performance of firms. Such understandings can provide insight into how consumers' reactions to different firm actions and events may impact firm performance. Jansen further asserts that the remaining 80% of posts that do not express sentiment are opportunities for firms to provide value to current and potential customers seeking information.

One key shift social media provides has been elevating consumers into advertisers and promoters by enabling individuals to have a voice (Lamberton & Stephen, 2016). With the ease of access to the internet and social media platforms, more individuals can engage on social media platforms. Some users may share content of interest and relevance to others, motivating them to "follow" them. Users who amass many followers become social media influencers (SMI) (Vrontis et al., 2021). SMI with higher numbers of followers can transmit information and sentiment quickly and broadly, *influencing* their followers' perceptions and thereby changing their aggregate sentiment or opinion. Firms can leverage social media by partnering with individual SMI (as advertiser/ promoter) to garner a broader reach to their followers and influence their intents and opinions in favor of the firm. The practice of incorporating SMIs into marketing strategies is referred to as Influencer Marketing (IM) (Vrontis et al., 2021). Firms today consider having a social media presence critical to their marketing strategies, and IM can be a valuable part of their social media engagement. However, partnering with an SMI with a high follower count is not necessarily enough to increase brand adoption. Wies et al. (2022) found that influencers in the middle range of follower counts tend to have higher engagement with sponsored content than those at the lower and higher ends of the range. They determined that influencers in the middle range are more in line with their followers, presenting nuanced content to put more focus on their specific subscriber base. Their finding means that not all posts will have an equal influential effect on followers.

Djafarova and Rushworth (2017) observed that social media users perceived non-traditional celebrities, such as bloggers and SMIs, as more credible than traditional celebrities' endorsements. Participants in their study were aware that some products marketed by traditional celebrities were overpriced "… and were therefore more likely to be influenced by lower-end celebrities who endorsed more affordable brands" (p. 5) – the lower-end celebrities were more relatable. The study also found that influencers were considered more trustworthy and had expertise if they had gone through the "journey." – for example, the influencers themselves had overcome personal challenges or achieved significant weight loss while endorsing a diet product. Ultimately, companies hope social media can be leveraged to encourage and promote sales to increase revenue and enhance shareholder value by broadly spreading influence through electronic word of mouth (Jin & Phua, 2014). With SMI, the concepts of trustworthiness and expertise matter. Together, these compose the construct of source credibility, "a term commonly used to imply a communicator's positive characteristics that affect the receiver's acceptance of a message" (Ohanian, 1990, p. 41).

With the explosive growth of social media use, researchers and practitioners alike have found it a rich data source for analysis and research. Unlike other sources, such as surveys and questionnaires, posts on social media reflect a person's beliefs and feelings at almost any given time (He et al., 2016). With the ease of accessibility and high engagement of its users, the content posted on social media can be viewed "as a form of collective wisdom" (Asur &

Huberman, 2010, p. 492) containing information and insight that, in the aggregate, can rival individual experts and professionals in their accuracy.

The behavioral finance body of knowledge includes research on the influence of social moods on markets (Nofsinger, 2005). Pagolu et al. (2016) used sentiment analysis and machine learning to identify a strong correlation between sentiments of tweets made on Twitter and the direction of a company's stock price movement. Bollen et al. (2011) investigated the relationship between Google-Profile of Mood States (GPOMS) and the closing value of the Dow Jones Industrial Average (DJIA). They found that certain public moods accurately predicted the direction of change for the DJIA index, whereas others did not. Cakra & Trisedya (2015) observed similar results using linear regression of tweet sentiments and prices of stocks in the Indonesian stock market.

The construct of sentiment is operationalized in various studies using different variables: the ratio of positive to negative tweets, average sentiment, average positive sentiment, or average negative sentiment (He et al., 2016). To operationalize the construct of source credibility, this study uses the number of followers of users posting tweets. The number of followers has been used in previous research studies as a variable for measuring source credibility. As reported in a study by Djafarova & Rush, many of their participants “held the common view that their [Instagram celebrities] number of followers was an indication of a person’s credibility” (2017, p. 4). It is expected that the number of followers will affect the relationship between tweet sentiments and abnormal returns of stock prices because influencers can propel information across a broad reach of audiences. Furthermore, their expansive reach is due, in part, to their perceived higher source credibility, making their followers more receptive to their word of mouth.

I.2 Purpose of Study

"The efficient market hypothesis (EMH) states that financial market movements depend on news, current events, and product releases, and all these factors will have a significant impact on a company's stock value" (Pagolu et al., 2016, p. 1345). According to EMH, if markets are efficient, the stock price includes all news and information about its respective company. It "suggests that individual actors cannot outperform the market as the current price of a stock should incorporate all available information" (Eickhoff & Muntermann, 2016). However, due to the idiosyncrasies of human nature, the news is not perceived equally and in the same way by all people. Therefore, the price of a stock is subject to the persistent push and pull of investors' perspectives on the underpricing or overpricing of a given stock. As new technologies and tools emerge, retail investors rely less on professional financial analysts' advice and opinions. Savvy investors begin to make investment decisions with like-minded peers and communicate their analyses on platforms such as social media (Chen et al., 2014). This research explores the influence of users' sentiments on posts made on the social media platform Twitter and the abnormal returns of the stock price of four specific firms. It is proposed that social media users interpret corporate decisions and express these interpretations as opinions and beliefs in social media posts.

While extant research has covered the relationship between social media sentiments and stock market movements, very few have examined the effects of influencers and source credibility. This study extends previous research that has examined the influence of social media sentiments and predicting stock pricing by investigating sentiments' effects on firm financial fundamentals. Firms' financial fundamentals measure firm performance and are a major component of stock pricing. The study further enriches the body of knowledge by incorporating the concept of social media influencers and source credibility.

I.3 Theoretical Framing

Typically, investors rely on professional experts to evaluate firm performance. While the average social media user is not likely to have expertise in a particular firm, an aggregate group of individuals, such as those engaged in social media, may collectively manifest expertise through the wisdom of crowds (Surowiecki, 2005). Bartov (2018) studied the sentiments of tweets posted using Twitter's cashtag feature. Cashtags are stock symbols preceded by a dollar sign used in tweets to reference specific securities such as stocks or cryptocurrency (for example, $DIS, $TGT, or $WMT to reference Disney, Target, and Walmart). By leveraging the wisdom of crowds effect with financial enthusiasts on Twitter, Bartov observed that collective sentiments posted up to ten days before earnings announcements accurately predicted the firm's earnings and returns. Similarly, Sprenger et al. "explored the relationship between aggregate sentiment measures and aggregate stock market indices" (2014, p. 927). "Twitter allows users to tap into the Wisdom of Crowds, where the aggregation of information provided by many (non-expert) individuals often predicts outcomes more precisely than experts" (Bartov et al., 2018, p. 26).

In this study, the crowd comprises Twitter users who are unified in their interests in specific firms. For example, Disney commands a large following of consumers that are colloquially referred to as DisTwitter, "a community of Disney fans who discuss the company's movies, TV shows, and theme parks on Twitter" (*DisTwitter - What Is DisTwitter?*, n.d.). This community enthusiastically discusses firm actions, product releases, and the industry it operates. These fans are often brand-loyal consumers, and even when not, investors in the DIS stock have an intrinsic desire for the company to be successful. Their knowledge and experience can extend beyond financial enthusiasts. Instead of limiting to cashtags that reference stocks specifically, this study broadens the scope of tweets by leveraging more general hashtags. In this approach, tweets are collected that reference the companies instead of solely focusing on the stocks of those companies.

Additionally, this study explores if there is an effect from influencers on the relationship between the sentiment of tweets and stock price. Influencers are Twitter users with many followers–often thousands if not millions. Users choose to follow influencers because they are considered to have high source credibility. Source credibility comprises expertise and trustworthiness (Hovland, 1953). For forecasting purposes, these are desirable traits to consider – the tweets are genuine of the user's perspectives and accurate reflections of reality. Source credibility can be an important factor in users' decisions to follow an influencer (Djafarova & Rushworth, 2017). Extending from the DisTwitter example above, many Twitter users have an interest in Disney. However, these users vary in familiarity and experience with their products and services. This background of users is one reason why the wisdom of crowds theory is of great interest. While a single individual may not have extensive knowledge on a topic, a crowd will provide higher expertise. Can collective expertise be made stronger using the concept of source credibility and the idea that higher numbers of followers (such as those of influencers) are related to higher source credibility?

The theories of the *wisdom of crowds* and *source credibility* provide framing in this study to explore how a collection of users linked to each other in the common interest of a firm can form a crowd with wisdom and insight into how a company is performing. Moreover, can a crowd of influencers defined by their many followers be considered more expert than the average user, further sharpening that wisdom? Can crowds of Twitter users with high source credibility bring about even more wisdom? Generally speaking, is there more wisdom to gain from more expert crowds?

I.4 Research Questions and Hypotheses

The proposed research investigates the relationship between sentiments of social media posts and firms' financial performance. Sentiment analysis of tweets has been done in previous research to explore its relationship with the movement of stock prices (Agarwal et al., 2011; Bae & Lee, 2012; Cakra & Distiawan Trisedya, 2015; Pagolu et al., 2016; Wu He et al., 2016). This research extends that body of knowledge further to examine the effect of the source credibility of the user generating the post as measured by their number of followers. The extant literature does not cover how follower count may affect stock price. Given the growing interest in social media influencers (SMI), this research can be a valuable contribution in two ways. First, it uses a broader user base for social media's predictive power as a data source for predicting securities pricing. Secondly, it extends the body of literature around SMI, which is becoming an increasingly important phenomenon in today's society with the growing number of influencers and content creators on social media platforms. This study is motivated by two related research questions:

I.4.1 *Research Question 1 (RQ1)*

Do the sentiments of tweets made about firms relate to abnormal returns of those firms' stock prices?

Null Hypothesis 1 ($H1_0$)

The sentiments of identified events have no relation to abnormal returns of firms' share prices.

Alternative Hypothesis 1 ($H1_A$)

The sentiments of identified events are directly related to abnormal returns of firms' stock prices. Positive events will have positive cumulative abnormal returns, whereas negative events have negative cumulative abnormal returns.

I.4.2 *Research Question 2 (RQ2)*

Does the number of followers a posting user has affect the significance of the relationship between the sentiments of their tweets and the abnormal returns of firms' stock prices?

Null Hypothesis 2 ($H2_0$)

The number of followers a posting user has does not affect the relationship between the sentiments of their tweets and abnormal returns of stock prices.

Alternative Hypothesis 2 ($H2_A$)

Tweets made by users with more followers will have a stronger relationship between the sentiments of their tweets and the abnormal returns of firms' stock prices. In contrast, the opposite will be true with sentiments of tweets posted by users with fewer followers. As suggested by Wies et al. (2022) and Djafarova and Rushworth (2017), it is also expected that tweets made by users with the highest number of followers will not have a significant relationship because they are less relatable and less credible regarding the specific firms they are tweeting about.

Tweets posted by users with higher follower counts are expected to strengthen the positive relationship between their tweets' sentiments and the firm's financial performance. Users with higher follower counts have higher source credibility and are perceived (by the users that follow them) to be more trustworthy and to have greater expertise. The sentiments of their tweets are expected to be more "accurate" in their opinions and sentiments about firms. Conversely, users with low numbers of followers are likely to be less accurate, thus weakening the relationship.

II LITERATURE REVIEW

II.1 Social media and the wisdom of crowds

With such large user bases, social media attracts and engages users from all backgrounds and competencies. There are bound to be experts on a wide variety of topics, including those with the ability to predict stock price movements (Bar-Haim et al., 2011). Chen et al. (2014) found predictive power in the sentiment of words expressed in articles and comments posted on the website Seeking Alpha and future stock returns and earnings surprises. While not social media in the traditional sense, the comments sections of websites allow individuals to easily share their opinions and beliefs, often in discourse with each other. This research leverages social media data to explore the influence of tweet sentiments on the abnormal returns of stock prices of certain firms based on the collective sentiment of individual posts using a "wisdom of crowds" approach where the aggregate sentiment approximates the real outcome (Schoen et al., 2013, p. 530).

Not all crowds are the same. "Large independent and heterogeneous groups can outperform smaller groups in their assessments even if the smaller group consists of subject matter experts" (Eickhoff & Muntermann, 2016, p. 836). Communities on social media tend to be made up of many users. Furthermore, social media platforms like Twitter are easy to access and post new content, allowing a diverse group to participate. The literature on the wisdom of crowds has identified three key conditions for crowd wisdom to manifest: diversity, independence, and decentralization, all strengths of social media platforms. Hong et al. (2020) add crowd size as a fourth dimension that significantly moderates the influence of diversity, independence, and decentralization on crowd performance. Similarly, crowd size can be a strength of social media's ease of access. In other words, social media platforms are ideal for forming wisdom-generating crowds.

Social media is effective at generating crowd wisdom and expertise because of its unregulated nature. Information disseminated on social media tends to flow in and throughout networks without friction, allowing users to uncover information effectively from inside and outside sources (Ray, 2006). Eierle et al. (2022) examined whether social media sentiment contained information worth investors' consideration regarding pricing markets and stocks. They found that "social media sentiment may be relevant to investors because it may contain either 'better and/or new information'" (Eierle et al., 2022, p. 11). They also found that this information had a "permanent effect on stock returns."

II.2 Social media sentiment analysis to predict stock prices

Social media and content posted on it can be valuable for research. Wu He et al. (2016) outline four advantages to leveraging social media for predicting the performance of companies:

> First, because of its natural occurrence, social media data can reflect direct and immediate market reactions. Second, since social media data are mostly generated by individual consumers rather than marketers or companies, consumers often consider social media content to be more trustworthy (Levy, Duan, & Boo, 2013). Third, social media data have the ability to simultaneously capture a limitless variety of events and topics occurring in the market. Fourth, social media data provide a continuous stream of consumers' and investors' thoughts, feelings, and behaviors over time. (p.75)

The real-time nature of social media posts is immensely beneficial to practitioners and researchers, an "idiosyncrasy of social media-based forecasting the fact that sometimes it is not forecasting but nowcasting" (Schoen et al., 2013, p. 529). Furthermore, the high engagement and participation of such a vast number of users on social media closely represent public sentiment and opinion of current events (Pagolu et al., 2016, p. 1345). Currently, the literature includes some research exploring using sentiment analysis to leverage social media content to investigate relationships with different outcomes. Such research includes analyzing social media posts to forecast box-office revenues (Asur & Huberman, 2010), political opinions (O'Connor et al., 2010; Tumasjan et al., 2010), and stock prices (Huang & Liu, 2020; Stieglitz et al., 2020; X. Zhang et al., 2011). Karamptatsas et al. (2023) found that investor sentiment on specific firms significantly influenced stock returns and earnings surprises.

Reed (2016) takes a different approach to sentiment analysis and stock prices by creating a measure of consumer confidence using Twitter data and observing its influence on stock market indices, such as the S&P 500 and the Dow Jones Industrial Average (DJIA). Prantl and Mičík (2019) also looked at the sentiment of tweets on stock price movements through an electronic word-of-mouth lens. They found that relationships were statistically significant for B2C companies but not for B2B companies. By its very nature, social media will surface users' reactions to new topics and trends very quickly – as an illustrative example, new research has started investigating the predictive power of social media on the returns of cryptocurrencies such as Ethereum (Rousidis et al., 2020, p. 6281).

Furthermore, research such as that done by Bae & Lee (2012) found that Twitter allowed deeper exploration of influence and how it impacts "real-world audience sentiments," valuable insight for how online social media is a valid and useful data source for investigating relationships to different outcomes (such as job approvals in politics). The high diversity in all dimensions potentially improves the quality of predictions (Schoen et al., 2013, p. 532). This diversity can manifest in different interpretations of the same news regarding stocks and companies, leading to different sentiments (Dong & Gil-Bazo, 2020, p. 105).

Sentiment analysis of social media posts, such as tweets made on Twitter, has been explored in the current literature. Typically, quantitative methods such as regression analysis are performed using some metrics of Twitter posts, such as volume and sentiment (X. Zhang et al., 2011). Operationalizing the construct of sentiment can also differ between studies. Pal et al. (2020) use a Python module called NLTK (Natural Language Toolkit), which returns the polarity of text sentiments between a range of -1 to 1. In this case, the sentiments of tweets could be described as more or less negative or positive with each other. Pal et al. further weight the sentiments based on whether the user is verified or the tweet has been retweeted (shared) to provide some incorporation of source credibility. In the work done by Agarwal et al. (2011), they build and examine different models to categorize sentiment into three possible classes: positive, negative, and neutral.

In some cases, the sentiment of posts is operationalized in a very different manner based on the researchers' goals. Bouadjenek et al. (2023) initially explored both StockTwits and Twitter social media platforms in their study. In their case, the sentiment construct is not defined by positivity or negativity but by whether the user (likely an investor) is bullish or bearish on specific stocks.

The techniques used to perform sentiment analysis are also varied. Feldman (2013, pp. 82–83) notes that sentences consumed in sentiment analysis fit one of two principal classes: “objective sentences that contain factual information and subjective sentences that contain explicit opinions, beliefs, and views about specific entities.” Research such as the current one using sentiment analysis will typically focus on subjective sentences. Feldman provides many options for sentiment analysis, including support vector machines (SVM), Naïve Bayes, and

Logistic Regression. Statistical methods tend to focus on regression and correlation analyses, but there is also a growing number of research leveraging event studies from econometrics.

Twitter also presents some challenges to sentiment analysis, including the interpretation of emoticons, slang, and other tokens (such as *gr8* being shorthand for "great" – typically, an expression of positive sentiment) (Agarwal et al., 2011, p. 32).

II.3 Social media influencers and source credibility

Brown et al. (2007) note that word of mouth (WOM) is considered by traditional communication theory as having a powerful influence on behavior and the exchanges that occur between individuals to have "informational value." Their studies observed that individuals evaluated the opinions and reviews shared on websites and the individuals themselves. This finding supports the notion that the information content is more easily spread via WOM when the source is considered more credible.

The source credibility model consists of expertness and trustworthiness (Hovland, 1953). As depicted in **Error! Reference source not found.**, Hovland defines *expertise* as "the extent to which a communicator is perceived to be a source of valid assertions," and *trustworthiness* as "the degree of confidence in the communicator's intent to communicate the assertions he considers most valid." In other words, "a source should be perceived as more credible when it (1) possesses greater expertise and (2) is less prone to bias." (J. Brown et al., 2007, p. 6).

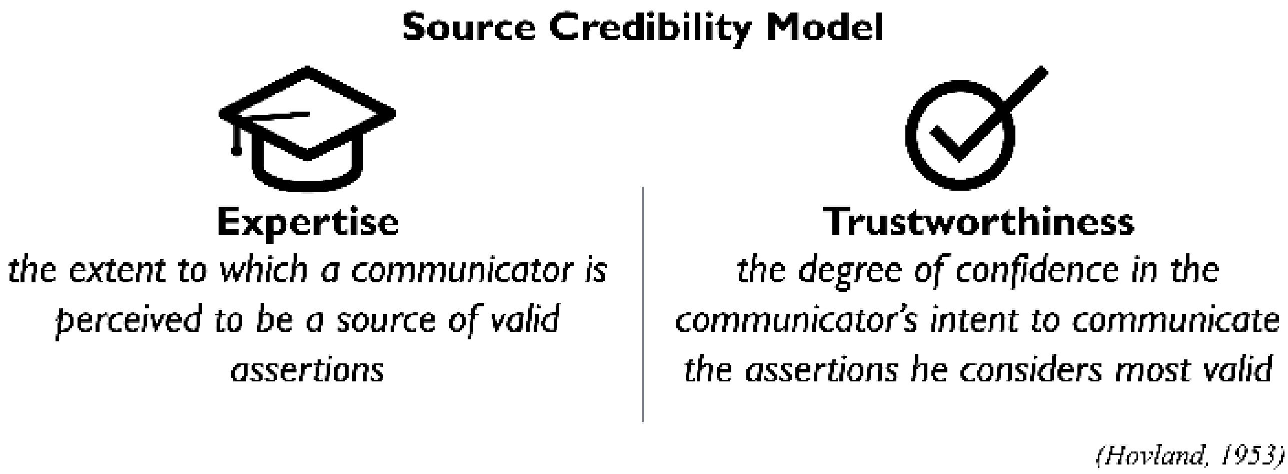


Figure 1 Hovland's Source Credibility Model

Bar-haim et al. (2011) studied tweets posted on Twitter. They found that distinguishing between experts from non-experts made predictive models of stock performance beneficial and created an expert-finding framework specific to users' ability to predict stock movements. This distinction is helpful because motivations for posting tweets differ between individuals. As a result, their relevance and accuracy will impact their usefulness for forecasting specific outcomes. Furthermore, these characteristics are considered valuable to prospective followers. Therefore, users who want to increase their followership will make efforts to increase their source credibility: "Microbloggers have a strong incentive to publish valuable information to maintain or increase mentions, the rate of retweets, and their followership" (Sprenger et al., 2014, p. 930). In a social network like Twitter, users are motivated to be credible to increase their numbers of followers for reputational prestige: "The size of the followership and the rate of retweets may represent the Twittersphere's 'currency' and provide it with its own kind of a pricing mechanism" (Tumasjan et al., 2010, p. 184).

Social media influencers (SMI) have grown exponentially over the past few years. A key feature of social media is that individuals can follow other users who provide engaging and relevant content. This connection between users can be quantitatively used for practical purposes. Companies can seek out individuals with high follower counts because they have expansive reach to other users who are potential customers for a company's products. This influencer marketing (IM) has become prevalent in marketing strategies (Vrontis et al., 2021). The construct of source credibility should be considered when trying to understand the influence and trustworthiness of users. Source credibility can be defined as "a communicator's positive characteristics that affect the receiver's acceptance of a message" (Ohanian, 1990, p. 41). In other words, is the user trustworthy, and are their posts and actions aligned with reality?

Notably, "consumers perceive individuals with a large number of subscribers [followers] as more attractive and trustworthy" (Djafarova & Rushworth, 2017, p. 1). Jin and Phua (2014) concur that source credibility is associated with celebrities with more followers. Not only were they perceived to be more credible, but they were also perceived to be more physically attractive, trustworthy, and competent. Since source credibility has been studied in the context of social media influencers, most studies are in marketing – influencing consumers to start or continue using products and services.

Source credibility and the phenomenon around social media influencers have applications outside of marketing. Users follow SMIs because of their perceived expertise in domains relevant to their followers – they "would prefer to interact with their ingroup members and are influenced by [their] opinions" (Liu et al., 2015, p. 51). This "congruence between the characteristics of the actors in a social network" is called homophily. While physical attributes such as gender and age are dimensions for homophily in the physical world, in an online context, research suggests the dimensions for homophily tend to be around similar interests and mindsets (J. Brown et al., 2007). In effect, users will seek out and follow other users with similar interests – for example, interests in a certain company or particular goods and services.

II.4 Gaps in research and contribution of current research

While there has been previous research on the relationship between online social media sentiments and stock prices (Ajjoub et al., 2021; Bollen et al., 2011; Reboredo & Ugolini, 2018; Sul et al., 2017; W. Zhang et al., 2018; X. Zhang et al., 2011), tweets were typically specific to investors' sentiments or that of professional financial analysts – their views and beliefs on the stocks themselves and not necessarily the companies. This research focuses on the tweets about companies underlying the stock to capture more opinions expressed by consumers who "are important stakeholders of a firm because the firm's ability to generate cash flows depends in large part on the value created for its [consumers]" (Huang, 2018, p. 164). Furthermore, very little research explores the effects of source credibility of individuals and the relationship between tweet sentiments and stock price. The source credibility of individuals is an important dimension to consider because users do not always have altruistic intentions when making their posts on online social networks – they may "seek to mislead or influence others for self-gain or other unknown motives" (Bouadjenek et al., 2023, p. 9:16).

In identifying a research gap, I performed two related searches using Web Of Science's database of research articles. In the first search, I used the query:

("Sentiment") AND ("Social Med*" OR "Social Network*") AND ("Stock Price*")*

The search query intends to identify research involving sentiment (such as sentiment analysis) and stock prices in social media or social networks. Web of Science identified 195 articles between 2011 and 2024 as of the time of this writing (March 2024).

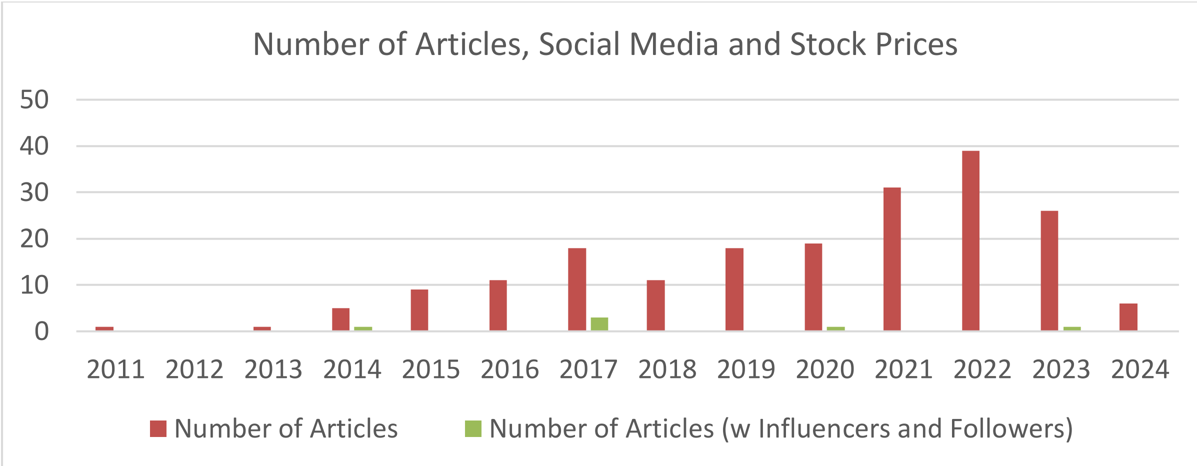


Figure 2 Number of Articles returned by Web of Science (March 16, 2024)

To compare, I added search terms for consideration of followers or influencers using the query:

("Sentiment") AND ("Social Med*" OR "Social Network*") AND ("Stock Price*")* ***AND ("Influencer*" OR "Follower*")***

For this search, Web of Science identified six articles published between 2014 and 2023 showing that while there is research interest in using sentiments of posts on online social media relating to stock prices, there is very little done with consideration of the effects of followers and Social Media Influencers (SMI). A chart presenting article counts for comparison over the years 2011 and 2024 is presented in Figure 2.

Table 1 Articles with coverage of SMI

1	Bouadjenek, M. R., Sanner, S., & Wu, G. (2023). A User-Centric Analysis of Social Media for Stock Market Prediction. *ACM TRANSACTIONS ON THE WEB*, *17*(2), 9. https://doi.org/10.1145/3532856
2	Coyne, S., Madiraju, P., & Coelho, J. (2017). Forecasting Stock Prices using Social Media Analysis. *2017 IEEE 15TH INTL CONF ON DEPENDABLE, AUTONOMIC AND SECURE COMPUTING, 15TH INTL CONF ON PERVASIVE INTELLIGENCE AND COMPUTING, 3RD INTL CONF ON BIG DATA INTELLIGENCE AND COMPUTING AND CYBER SCIENCE AND TECHNOLOGY CONGRESS(DASC/PICOM/DATACOM/CYBERSCI*, 1031–1038. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.169
3	Pal, R., Pawar, U., Zambare, K., & Hole, V. (2020). Predicting Stock Market Movement Based on Twitter Data and News Articles Using Sentiment Analysis and Fuzzy Logic. In S. Smys, T. Senjyu, & P. Lafata (Eds.), *SECOND INTERNATIONAL CONFERENCE ON COMPUTER NETWORKS AND COMMUNICATION TECHNOLOGIES, ICCNCT 2019* (Vol. 44, pp. 561–571). Springer International Publishing Ag. https://doi.org/10.1007/978-3-030-37051-0_63
4	Sul, H. K., Dennis, A. R., & Yuan, L. (2014). Trading on Twitter: The Financial Information Content of Emotion in Social Media. In R. H. Sprague (Ed.), *2014 47TH HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS)* (pp. 806–815). IEEE. https://doi.org/10.1109/HICSS.2014.107
5	Sul, H. K., Dennis, A. R., & Yuan, L. (Ivy). (2017). Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns. *Decision Sciences*, *48*(3), 454–488. https://doi.org/10.1111/deci.12229
6	Zhang, Y., An, Y., Feng, X., & Jin, X. (2017). Celebrities and ordinaries in social networks: Who knows more information? *Finance Research Letters*, *20*, 153–161. https://doi.org/10.1016/j.frl.2016.09.021

The six articles identified are focused on investor users. As seen in **Error! Reference source not found.**, four of the six articles use datasets where tweets are extracted from Twitter using queries based on cashtags. Article 3 (Pal et al., 2020) is unclear on how tweets are gathered, though the authors mention using the *tweepy* Python module to extract them through the Twitter API. The remaining article by Zhang et al. (2017) uses Weibo as its platform and focuses on postings made by "celebrities." They define "celebrities as users who are industry professionals and are influential in social networks" (2017, p. 155).

Moreover, they focus on verified users likely to be active in the financial-industry. Therefore, while these six articles push the boundaries into SMI in social network sentiment and stocks, they are limited to the sentiments of active investors, which limits their practical contribution to marketers and the firms themselves in providing value to their customers. In contrast, this study uses hashtags to broaden the insights made to more general users more likely to be consumers or potential consumers. Twitter also allows @ mentions to call out other users or Twitter handles. For example, a user can make a tweet followed by @disney so that the mentioned user is notified of the post and the tweet appears in searches. In this study, I do not use mentions because I want to capture the sentiments of tweets made about the company, even if the tweet author does not necessarily intend to notify or respond to the firm.

Table 2 Summary of articles with coverage of SMI

Paper	**Method**	**Platform**	**Cashtags or Hashtags**
	Followers: How follower counts were incorporated into the study		
1 - Bouadjenek, M. R., Sanner, S., & Wu, G. (2023)	Logistic regression	Twitter and StockTwits	Cashtag
	Followers: N/A		
2 - Coyne, S., Madiraju, P., & Coelho, J. (2017)	ML-based algorithms	StockTwits	Cashtag
	Followers: Used to identify “Smart Users” to strengthen the predictive power of their model		
3 - Pal, R., Pawar, U., Zambare, K., & Hole, V. (2020)	Regression	Twitter	Unclear
	Followers: Discarded tweets when users with less than 100 followers posted them		
4 - Sul, H. K., Dennis, A. R., & Yuan, L. (2014)	Regression	Twitter	Cashtag
	Followers: Explore the speed at which information spreads based on the number of followers		
5 - Sul, H. K., Dennis, A. R., & Yuan, L. (Ivy). (2017)	Regression	Twitter	Cashtag
	Followers: Sentiments of tweets from users and the effect of the number of followers on stock returns		
6 - Zhang, Y., An, Y., Feng, X., & Jin, X. (2017)	Regression and event study	Weibo	References to stock codes
	Followers: Compare “ordinaries” against "celebrities"		

This research extends the currently sparse literature by exploring social media influence at more granular levels. Specifically, I perform an event study for different groupings of tweets posted by users split by follower count. The approach allows a more granular exploration of the relationship between the abnormal return of stock prices and sentiments of posted tweets. Because social media platforms like Twitter are generally open and accessible to the world's population (assuming internet connectivity), individual accounts can be opened and used for different purposes, in some cases with malicious intent.

By incorporating follower count as a proxy for source credibility, it is believed that there is a difference in the strength of the relationship. More specifically, users with more followers will have stronger relationships because they are perceived to be more trustworthy and have more expertise. In contrast, users with fewer followers are perceived to have less source credibility and are expected to have less accuracy in their tweet sentiments.

III RESEARCH DESIGN

III.1 Social media as a data source for research

Social media is an invaluable dataset for researchers and practitioners to tap into for insight into the sentiments and reactions of many people. This social media dataset will only grow in size and detail as more and more users engage in social networks. With broader and easier access to high-speed internet, social media has become important for individuals and organizations to share information (Gu & Kurov, 2020). "Compared with other information sources such as survey and archival data, social media has several advantages in unveiling individual's thinking and feelings" (Wu He et al., 2016, p. 75). Schoen et al. (2013, p. 530) categorize the process of collecting and analyzing data in three broad ways: (1) researchers can use historical data, or (2) data is collected on-demand, for example, surveys or polling. Social media provides a third approach by allowing researchers to observe users' behaviors unobtrusively in real-time or historically. Moreover, "information from social media is timelier than information from traditional media and is more likely to be value relevant" (Gu & Kurov, 2020, p. 3).

One successful social network is Twitter, a microblogging service launched on July 13, 2006. With Twitter, users post status updates, called tweets, which may include their musings, beliefs, and values or information they want to put out to the public. Users can follow users, increasing their influence and the likelihood that their posts will appear on their followers' Twitter page or mobile app. "Due to its huge reach, news organizations increasingly use Twitter to filter news updates through the community… Companies also use it to advertise products and provide outreach to consumers and other stakeholders (Asur & Huberman, 2010). With a large number of users actively engaged on social media, the Securities and Exchange Commission has approved the use of social media sites such as Facebook and Twitter to be just as permissible as company websites and press releases when communicating to investors (Holzer & Bensinger, 2013) further legitimizing these channels for quality information.

With social media a mainstay in contemporary culture and day-to-day life, researchers have found different ways to leverage it as a dataset for exploration. Common approaches include platform aggregated quantitative metrics such as volume of posts, number of users being followed, or number of followers for a given user. Such metrics can also be analyzed according to different time aggregations per day, month, quarter, or other relevant duration.

Before March 2022, researchers with academic affiliations could request academic access to Twitter. Academic access allowed researchers to retrieve 10,000,000 tweets monthly through Twitter's official API at no cost. From the current body of literature, Twitter was a very popular data source used in a wide variety of academic research, such as exploring trends in public perception about COVID-19 vaccines (Saleh et al., 2023), understanding crisis communications (Acar & Muraki, 2011), and media communications and information flow (Wu et al., 2011). On February 8, 2023, Twitter announced they were dramatically changing their access tiers, and Academic Access was effectively discontinued (Developers [@XDevelopers], 2023b). As reported in the mainstream media, removing easy access to Twitter has dramatically affected the academic research community (Calma, 2023; Jingnan, 2023).

III.2 Variable Operationalization

Data used for this study was collected from two different sources: Twitter (now known as X) and Yahoo Finance.

III.2.1 *Sentiment*

The primary data for this study are tweets – posts made on Twitter (now named X but still referred to as Twitter) by users. The collection was done primarily through open-source

Python and Python modules, such as *snscrape* (for tweets) and *tweepy* (for user profile information).

Tweets are collected programmatically, leveraging the Python module *snscrape*,[1] a popular module for scraping supported social networks (such as Twitter, Facebook, and Instagram). At the time of the extraction for this study (December 2022 - January 2023), *snscrape* was not limited to API thresholds because it did not authenticate to Twitter using a login. Instead, it used Twitter's native search functionality to extract tweets. Using a developer account to access the Twitter API is subject to certain limitations, such as throttling (about one call to the API per second) to ensure stability and reliability. With these advantages, *snscrape* has been effectively used in other academic research (Ağrali & Aydin, 2021; Blair et al., 2021; Sarkar & Rajadhyaksha, 2021; Verma et al., 2023) to extract tweets without needing developer credentials and Twitter API access. However, a drawback of using *snscrape* is that it is unsupported by the Twitter company and is subject to sudden breakages. Indeed, *snscrape* is no longer a feasible tool to collect tweets from Twitter as of Summer 2023, after significant restrictions were placed on the formerly free Twitter API and access modes.

The users of Twitter are highly engaged and active on the platform - collecting the entire population of tweets is quite difficult. In a recent study by Pfeffer et al. (2023), an entire day's worth of tweets was collected, resulting in about 374,937,971 tweets posted by 40,199,195 accounts during the 24 hours starting September 20, 2022, 1500 UTC. On average, there were 4,340 tweets per second. Therefore, for this study, only a subset of tweets could be feasibly collected within resource constraints using *snscrape* constrained by two key dimensions: 1)

---

[1] https://github.com/JustAnotherArchivist/snscrape

timing and 2) content relevance. I prioritized having complete data for each target firm; therefore, tweets were extracted for each firm before moving on to the next. Unfortunately, Twitter's Academic API access was discontinued during the data extraction phase, and only four companies were fully extracted and usable for this study. These four firms (The Walt Disney Company, Nike, Target, and Tesla) were selected for their B2C business models, relatively well-known brands, and active engagement on social media, including brand-loyal followings of consumers. Other firms initially targeted but not extracted for were competitors identified within their respective North American Industry Classification Systems (NAICS). For example, the Walt Disney Company is listed in NAICS 71311 – Amusement and Theme Parks. Other major companies in this classification include NBCUniversal Media, SeaWorld Entertainment, Cedar Fair, and Six Flags Entertainment.

Table 3 Number of tweets per year per firm

Number of tweets per target firm per year

Between Jan 01, 2017 thru Dec 31, 2022

	2017	2018	2019	2020	2021	2022	TOTAL
DIS	684,026	571,254	603,632	470,571	295,826	394,098	3,019,407
NKE	287,223	384,650	341,436	306,075	183,957	278,194	1,781,535
TGT	93,105	64,808	60,194	56,117	39,021	38,306	351,551
TSLA	175,038	205,107	216,859	244,342	378,568	407,665	1,627,579
TOTALS	1,239,392	1,225,819	1,222,121	1,077,105	897,372	1,118,263	6,780,072

Number of tweets collected from Twitter (X) using *snscrape* between December 2022 and January 2023

The timing of the tweets was constrained to the calendar years 2017-2022 and limited to English-language tweets to take advantage of existing Natural Language Processing (NLP) modules. Finally, Twitter's concept of hashtags was used to filter further and focus on tweets about specific companies. The search queries were constructed using hashtags used most commonly for the four targeted firms (#disney, #target, #nike, and #tesla) to identify the relevant subset of tweets to be collected. 6,780,072 tweets posted between January 1, 2017, and December 31, 2022, were collected using this approach. The breakdown of the tweets by firm and year of the post is shown in **Error! Reference source not found.**. Appendix B – Chart of Tweets Per Firm Per Day (page 69) shows a chart of posting activities per firm per day.

III.2.2 *Source Credibility*

This study also explores the moderation effect of source credibility on the relationship between the sentiment of tweets and stock prices. This construct is operationalized using the number of followers of users posting tweets. The number of followers is available through a user's profile, accessible through the Twitter API's *users* endpoint. Similar to tweets, the population of users on Twitter is not feasible to collect.

Table 4 Distinct authors and descriptive statistics of followers per firm

Distinct authors (and followers)

Between Jan 01, 2017 thru Dec 31, 2022

		Followers						
	Distinct Authors	Minimum	Maximum	Average	Std Dev	Median	Skew	Kurtosis
DIS	601,875	0	18,788,527	4,225	105,230	209	96	12,216
NKE	364,935	0	107,608,618	5,117	235,728	231	303	125,074
TGT	135,040	0	12,960,339	4,240	82,988	238	91	11,256
TSLA	367,418	0	76,641,724	5,107	171,376	196	276	112,544

Based on tweets collected from Twitter (X) using *snscrape* between December 2022 and January 2023

After collecting relevant tweets, a list of unique Twitter users was consolidated. For each user, a query was made against the Twitter API to retrieve the profile information. The Twitter API's *users* endpoint cannot be used anonymously, so a different approach must be used than *snscrape*. In this case, a different Python module was used - *tweepy*[2] - to work with the API authenticating with academic research credentials. *tweepy* is "an open source python package which makes it convenient to use Twitter API with its classes and methods" (Chaudhary & Niveditha, 2021, p. 4513). Many researchers use *tweepy* to make it easier to access Twitter as a data source (Kaur & Sharma, 2020; Manguri et al., 2020; Shelar & Huang, 2018).

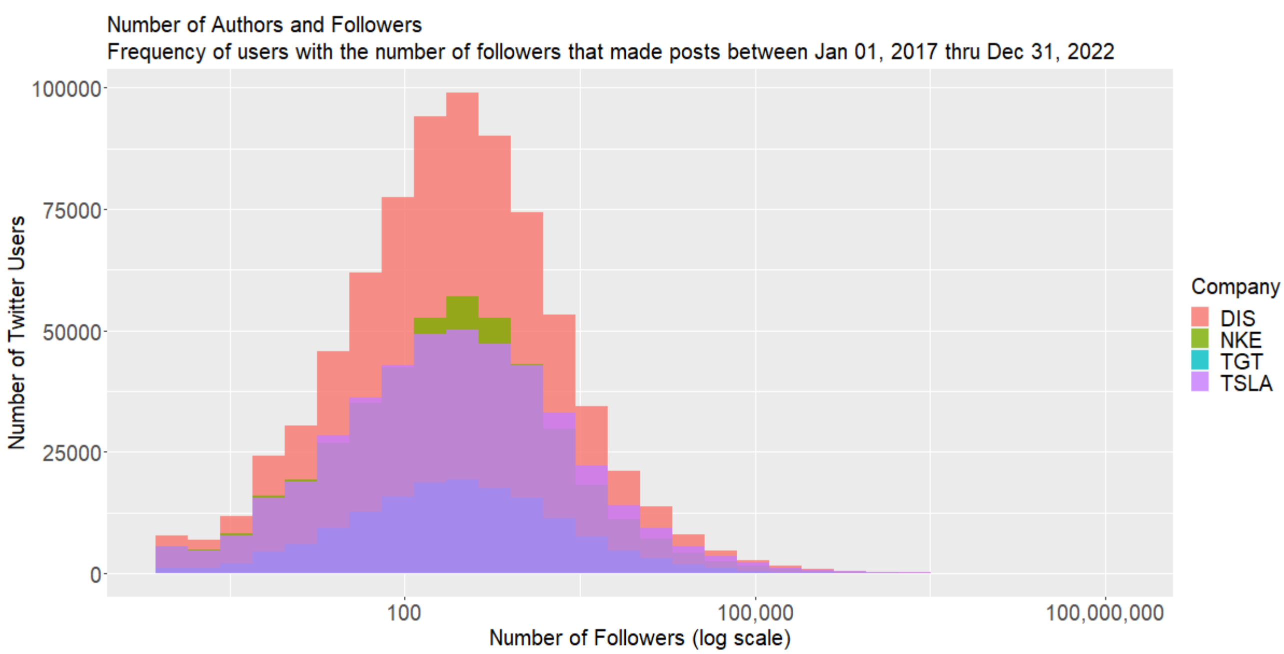


Figure 3 Chart of number of users with 99% winsorized number of followers per firm

[2] https://www.tweepy.org/

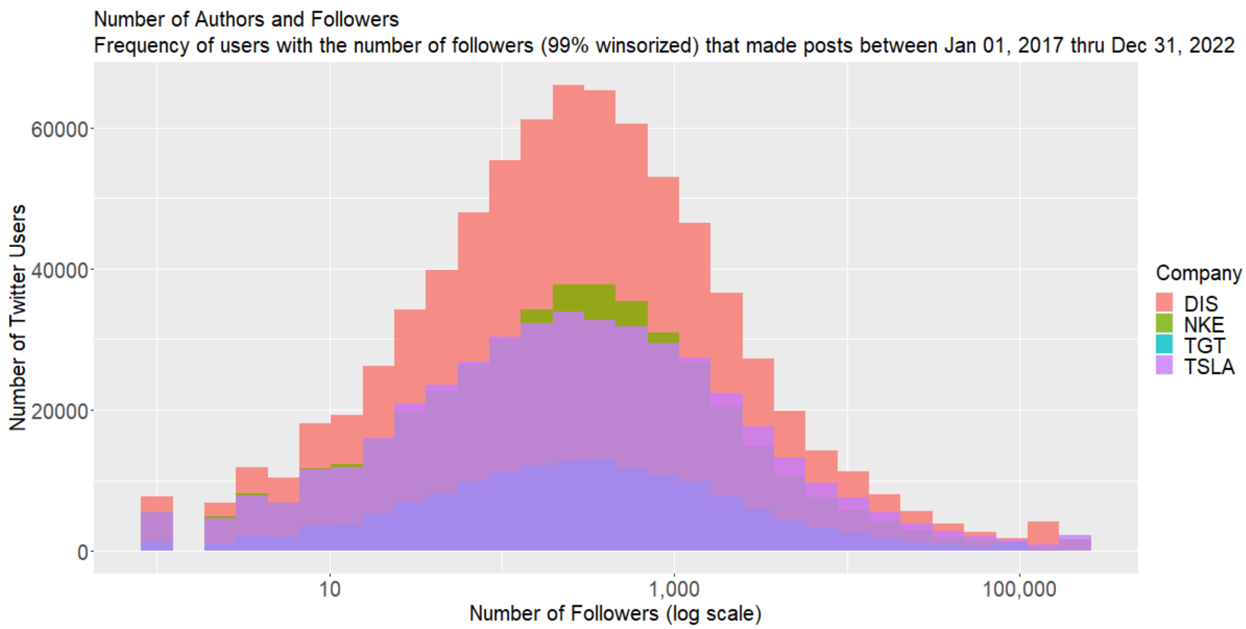


Figure 4 Chart of number of users with number of followers per firm

Before 2023, Twitter was open and easily accessible as a data source for academic research. However, in a tweet posted on Twitter on February 2nd, 2023, the company announced that it would "no longer support free access to the Twitter API" (Developers [@XDevelopers], 2023a). It was unclear the extent of the impact that this change would have on the academic research community (Lukpat, 2023). Ultimately, the impact on academic research has been extensive, and free academic access to the Twitter API was also discontinued and additional tweets are no longer feasible to extract. Table 4 shows the number of distinct authors that posted at least one tweet about the given firm between 2017 - 2022. Descriptive statistics are also presented. Notably, the number of followers is highly varied, especially at the higher end.

Error! Reference source not found. shows the distribution of the number of followers. The magnitude of skew varies by firm, with Nike having the highest (positive) skew of all four.

The skew of the distribution of the number of followers should be considered because there is no specific number of followers to qualify an individual as an influencer. The handles with the ten highest number of followers for each firm are presented in Appendix C – Top Followers for Each Firm (page 70). There are large differences in the number of followers across and within firms. For example, for Nike, *Cristiano,* has the largest followers, with 107,608,618. *Cristiano* also has the largest number of followers across all four firms, and the handle with the second largest number is *premierleague,* with 39,597,974 followers. The difference of 68 million is larger than every other handle in the top 10 lists of the four firms, except for *EllenDeGeneres,* with 76,641,724 followers.

Figure 3 shows the same data as **Error! Reference source not found.**, but with 99% winsorization. Winsorization is an approach to treating datasets for extreme values. In winsorization, a researcher-selected cutoff is applied to a variable in a dataset. All observations with values greater than the cutoff are equal to the cutoff. An alternate approach is truncation, where the observations with values greater than the cutoff are removed from the dataset (Leone et al., 2019). In **Error! Reference source not found.**, there may only be 1 or 2 users at the higher numbers of followers, making it difficult to visualize the number of users that make up the long tail. Using a 99% winsorization for the number of followers, the users with more than 99 percentile followers are equal to the 99 percentile, making the number of users more apparent to the reader in Figure 3. The winsorization procedure is only used in this study for visual and descriptive purposes, not statistical testing.

III.2.3 *Stock Price*

This study looks at relationships between the aggregate sentiments of tweets and abnormal returns of stock prices for each targeted company. For this study, I extracted daily stock prices between 2016 – 2022 for the four target companies (Disney, Nike, Target, and Tesla) using the Python module *yfinance*,[3] a wrapper around the Yahoo! Finance API. The same data can be extracted from the Center for Research in Security Prices (CRSP) at the University of Chicago daily stock price database, accessible through Wharton Research Data Services (WRDS).

III.3 Sentiment Analysis

With the advent of cheaper computing resources and cloud computing, a common approach is analyzing posts' content, typically through sentiment analysis and labeled as positive, negative, or neutral, to capture whether the posting user is expressing positivity or negativity about a firm at the given time. Typically studies in online social media will aggregate posts to describe sentiments at a higher level. For example, the number of positive or negative tweets can be counted and compared for a given day. The day is labeled as positive if there are more positive than negative tweets. Alternatively, the number of tweets for each sentiment can also be used directly in further analyses – for example, correlation, significance testing, and regression analysis. Sentiments of online social media are useful and provide additional information when predicting stock returns. In a study by Gu and Korov (2020, p. 2), they used Twitter sentiment data and found that it was not reflected in the price of stocks:

> "This finding suggests that Twitter sentiment does not simply reflect sentiment of uninformed traders. Instead, Twitter sentiment contains relevant information incorporated into stock prices with a one-day delay. We also find that Twitter sentiment contains more information about firms with limited analyst coverage. This finding

[3] https://pypi.org/project/yfinance/

further supports the conclusion that firm-level Twitter sentiment plays an informational role."

Using tweets requires additional considerations, including what period and how to collect and prepare the Twitter data for analysis (Schoen et al., 2013, p. 533). Employing sentiment analysis to explore tweets is a fairly typical approach to analyzing social media posts and their relationship with other variables. Sentiment analysis aims to categorize opinions to summarize the underlying text better.

III.3.1 *Natural Language Processing (NLP) packages (for R and Python)*

In the context of Natural Language Processing (NLP), several packages are available open-source for sentiment analysis. Utilizing a pre-built package has numerous benefits. Firstly, R packages are typically distributed open-source via the Comprehensive R Archive Network (CRAN)[4]. Likewise, Python packages are typically installed using the *pip* tool that accesses the Python Package Index (PyPI)[5]. Centralized repositories mean that packages are easily discoverable and usable by virtually any R or Python developer. Secondly, users – many academic researchers and data practitioners – will use and validate the packages through continued and increasing use. Thirdly, by using a package that has been used and tested by others, a custom solution does not need to be created at the expense of time and resources while possibly including errors typical of immature software. Therefore, to work within this study's time and budget constraints, an available R or Python package will be used in place of training a bespoke model for sentiment analysis.

---

[4] https://cran.r-project.org/

[5] https://pypi.org/

Naldi (2019) compares four packages that provide a numeric score conveying the direction and valence of inputted text. Their comparison of *syuzhet*, *Rsentiment*, *SentimentR*, and *SentimentAnalysis* yielded a recommendation for *SentimentR*. Their recommendation for *SentimentR* is primarily because it is the sole package of the four to account for negators and valence. Naldi provides a helpful example:

> the scores assigned to the sentences ”This device is good”, ”This device is very good”, ”This device is very very good”, and ”This device is good but bad” (sorry for the contradiction) are respectively 1,2, 3, and 0. (p.5)

The last update was in 2021, with the latest release, version 2.6.1, in 2018 (*Sentimentr/Inst/CITATION at Master · Trinker/Sentimentr · GitHub*, n.d.).

TweetNLP is a Python library that provides a set of Natural Language Processing (NLP) tasks, including sentiment analysis (Camacho-collados et al., 2022). The library is based on Transformer-based language models similar to the large language models popular today, such as ChatGPT (Gupta et al., 2023). TweetNLP is trained specifically on social media text, particularly Twitter. Therefore, TweetNLP was selected for the sentiment analysis of tweets for this study.

TweetNLP has many built-in applications for applying NLP on tweets, including Named Entity Recognition (NER), Topic Classification, and Sentiment Analysis. The model output consists of 4 specific data points for sentiment analysis: a label (positive, negative, or neutral), positive probability, neutral probability, and negative probability. The label value is based on which of the other three probability values is the greatest. This study uses the probability scores with a threshold of .75 to assign a sentiment label. In other words, a tweet is considered positive if the positive probability is greater than or equal to .75, negative if the negative probability is greater than or equal to .75, and neutral in all other cases. Because sentiment analysis is an imperfect science, having less uncertainty in assigning a positive or negative sentiment label is desirable. This approach is also specified by Prantl (2019, p. 6): "The evaluation of the posts' sentiment will be done based on words appearing in them. Based on this analysis, we will divide the posts into positive and negative categories. Posts containing ambiguous sentiments will be eliminated from the analysis."

III.3.2 *Stopwords*

Tweets are first cleansed of non-sensical artifacts such as emojis, emoticons, retweets, and hashtags. Stopwords are also removed. Stopwords are "…words that do not express any emotion… like 'a,' 'is,' 'the,' 'with,' etc." (Pagolu et al., 2016, p. 1346). In this study, I used a stopword dictionary, as suggested by Agarwal (2011, p. 32). The dictionary[6] "is a comprehensive list of words ignored by search engines" (*Stop Words*, n.d.). Overall, the sentiment categorizations did not change significantly before and after removing the stopwords. For example, on December 31, 2022, Twitter user *warnilla5* tweeted,

> Watching CNN earlier, Tesla stock has now fully collapsed, could this be the end for Tesla? Seemingly they need to stop production of cars soon and have already given refunds to customers? Any truth in this?. #tesla https://t.co/RZXFMxf9yj

TweetNLP categorized this tweet as 84.65% probability of being negative sentiment (14.58% neutral and .76% positive). After preprocessing to remove stop words, the resultant tweet was worded as:

> Watching CNN earlier, Tesla stock fully collapsed, Tesla? Seemingly stop production cars refunds customers? Any truth this?. #tesla

[6] http://www.webconfs.com/stop-words.php

The probability of the tweet being negative went up to 88.27% (11.12% neutral and .60 % positive). Based on the study's threshold of .75, this tweet would be classified as negative in both cases.

III.3.3 *Daily aggregations*

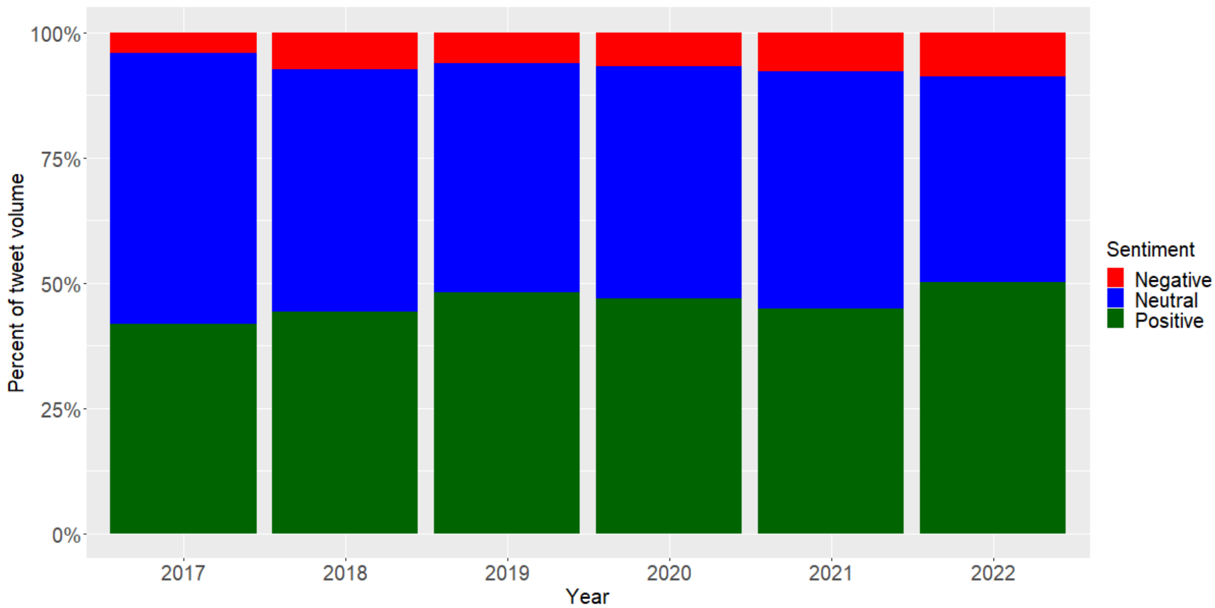


Figure 5 Percentage of tweet volume per sentiment per year (*yearly for clarity*)

As noted above in section III.3.1, positive tweets are tweets with a positive probability greater than or equal to .75. A percentage of sentiment sentiments of tweet volume is presented in Figure 5. While aggregations for testing and analysis were done at a daily resolution, the chart shows a summary of the aggregations at an annual resolution for visual clarity. Similar to other studies, there is a larger percentage of positive tweets than negative ones.

III.4 Analytical Approach

With the data collected as described in the preceding sections, steps were taken to transform and integrate the data for preparation for analysis. This study utilizes an event study approach described in Ranco et al. (2015). Ranco et al. performed an event study to identify non-earnings announcement events for the 30 companies that form the Dow Jones Industrial Average (DJIA). This common index represents the larger market of stock securities. Their principal data are tweets containing cash tags denoting the specific company's stocks. Cash tags are similar to hashtags, linking tweets to specific securities. The scope of their study is the 15 months, from June 1, 2013, to September 18, 2024 (Ranco et al., 2015, p. 5).

III.4.1 *Event Study*

Many event studies "provide a direct test of market efficiency" (S. J. Brown & Warner, 1980, p. 205). Therefore, event studies are a good approach to examine if additional information is present in the sentiments of tweets made about a company on the abnormal return of its stock price. MacKinlay (1997) describes the general procedure for an event study while noting that there is no single definitive approach to an event study. While James Dolley first published the event study in 1993 (MacKinlay, 1997, p. 13) to study the effects of stock splits on stock price, their applications extend far beyond the areas of finance and economics, resulting in many different approaches to how an event study is conducted.

III.4.2 *Sub-datasets, Influencers, and Number of Followers*

To test Hypothesis 2, I create different subsets of tweets based on the number of followers posting users have – this allows each subset of tweets to be tested in isolation, and resulting outcomes can be compared across different subsets for analysis. Primarily, I create two different dataset *variants* that differ in their approach to subsetting the population of tweets. Within each dataset variant are *iterations* that gradually build up from a small range of authors based on their numbers of followers to including the entire population of collected tweets.

Influencers are users that have a "high" number of followers. However, there is no specific number or threshold between influencers and non-influencers. Therefore, I iteratively test subsets of the population of collected tweets around different ranges of the number of followers from the top and the bottom. In this way, I can see how the relationship may differ if influencers are defined as the Top 10% of users based on the number of followers or the Top 30%. Using a split by quantiles also avoids additional definitions that may further delineate between users with micro-influencers with thousands of followers and mega-influencers with more than a million followers (Li et al., 2024; Tian et al., 2023). To split each subset (iteration) of data, I first extracted the list of distinct authors that had made at least one tweet about each firm. From these lists, I rank the authors based on their number of followers and determine the range of followers for a given percentile. To balance granularity against processing time, I opted for each iteration to be an incremental 2%, meaning there are 50 iterations (building up to 100% of authors and tweets). This data-splitting approach is similar to an approach by Sul et al. (2014). In their study, using Twitter as the study platform, they perform regression analysis with the cumulative abnormal return as a dependent variable. In their model incorporating the classification of tweets based on the number of followers, they split the tweets into three subsets using the thresholds of 177 (the median number of followers in their sample), 1,000, and 100,000. The factor added to their regression model is the sentiment valence (positive or negative) of tweets made by the split groups of tweet authors.

The first variant, *topDown* (seen abridged in Table 5 below), is generated top-down, starting with tweets from authors with the highest number of followers. Then, over 50 iterations gradually including tweets posted by the next two percentile of authors ranked by number of followers. The full population of collected tweets is included in the 50th iteration (100%).

Structuring the dataset this way allows the analysis to be programmatically automated to identify events and perform statistical analyses for each iteration. The second variant, *bottomUp* (see Table 6), is generated in the opposite manner, starting with tweets from authors with the lowest number of followers – including those with none and, with each successive iteration, including the next two percentile higher of authors ranked by tweets. Note that the 50th iteration of the *bottomUp* variant is equivalent to the 50th iteration of the *topDown* variant, which is simply the population of collected tweets.

Notably, each iteration is, in effect, an isolated study of just the tweets made by authors within the specified percentile range. In an event study, this will affect the detection of events because they are based on outliers of daily tweet volume and the number of tweets per sentiment classification. These impacts are desirable for this study because they allow the exploration of how the relationships between tweet sentiments and abnormal returns of stock prices may change based on the tweet authors' numbers of followers.

Table 5 Range of followers for each iteration of the TopDown dataset variant

Range of followers for the TopDown dataset variant

The range of followers used to filter for the given percentile per firm.
Showing multiples of 10 percentiles for brevity.

		DIS		NKE		TGT		TSLA	
Iteration	Percentile	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
5	10	3,617	18,788,527	3,716	107,608,618	4,383	12,960,339	4,369	76,641,724
10	20	1,278	18,788,527	1,334	107,608,618	1,533	12,960,339	1,470	76,641,724
15	30	657	18,788,527	702	107,608,618	782	12,960,339	720	76,641,724
20	40	380	18,788,527	414	107,608,618	440	12,960,339	387	76,641,724
25	50	230	18,788,527	254	107,608,618	262	12,960,339	218	76,641,724
30	60	138	18,788,527	153	107,608,618	156	12,960,339	123	76,641,724
35	70	77	18,788,527	84	107,608,618	89	12,960,339	66	76,641,724
40	80	38	18,788,527	39	107,608,618	45	12,960,339	31	76,641,724
45	90	13	18,788,527	12	107,608,618	17	12,960,339	10	76,641,724
50	100	0	18,788,527	0	107,608,618	1	12,960,339	0	76,641,724

Based on tweets collected from Twitter (X) using *snscrape* between December 2022 and January 2023

Table 6 Range of followers for each iteration of the BottomUp dataset variant

Range of followers for the BottomUp dataset variant

The range of followers used to filter for the given percentile per firm.
Showing multiples of 10 percentiles for brevity.

		DIS		NKE		TGT		TSLA	
Iteration	Percentile	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
5	10	0	11	0	10	0	14	0	9
10	20	0	33	0	33	0	39	0	27
15	30	0	69	0	74	0	79	0	58
20	40	0	125	0	138	0	141	0	110
25	50	0	209	0	231	0	238	0	196
30	60	0	344	0	377	0	397	0	345
35	70	0	585	0	629	0	694	0	633
40	80	0	1,106	0	1,159	0	1,328	0	1,262
45	90	0	2,744	0	2,814	0	3,328	0	3,275
50	100	0	18,788,527	0	107,608,618	0	12,960,339	0	76,641,724

Based on tweets collected from Twitter (X) using *snscrape* between December 2022 and January 2023

III.4.3 *Detection of Events*

The first step toward an event study is to define an event (Campbell et al., 1997) – this may be defined by the purpose of the study, for example, earnings announcements, breaking news, or stock splits. Or the events may be automatically identified through analysis, as Ranco et al. (2015) have explored using tweet volumes. This study extends the work of Ranco et al. by focusing on a different subpopulation of tweets and incorporating the effects of the number of followers for the tweets' authors.

I established the baselines of tweet volumes for each calendar day spanning the study's scope of 2017 – 2022. Baselines are set by aggregating the number of tweets per day for each firm. Then, for each day, a window of five days before and after is used to establish a set of tweet volumes around the given day. The median of the set is the baseline for that day, TW_b.

To identify whether a given day's volume is an event, Ranco et al. calculated an outlier fraction $\phi(t_0)$ as shown in Equation (1). and then set a minimum of $\phi(t_0)$=2 as an event where the volume of tweets for a given day TW_d is significantly higher than the baseline.

$$\phi(d_0) = \frac{TW_d - TW_b}{max(TW_b, n_{min})} \quad (1)$$

They also include the term n_{min} to set a minimum baseline. The term is necessary for Ranco et al.'s study because cash tags are not used as frequently used hashtags, resulting in lower volumes of tweets. Since I am using hashtags for this study, the volume of tweets per day is much larger, and a term for minimum volume is unnecessary. Likewise, the variance of tweet volumes in this study is significantly greater. I set the threshold of the outlier fraction to $\phi(t_0)$=1 to maximize the identification of events. Examples of timelines with tweet volumes and identified events for each firm using the *TopDown* (Figure 9) and *BottomUp* (Figure 10) datasets for Iteration 15 (top

30% and bottom 30% of followers, respectively) are provided in 0 Appendix D – Example of Timeline of Events and Tweet Volumes (p. 71).

Events are identified for each firm and iteration within a dataset variant, thus essentially allowing an isolated event study for each subset based on the number of followers at differing levels (based on percentiles). Sometimes, the events are identified when the stock market is closed (for example, on Saturdays, Sundays, or certain holidays such as Thanksgiving and Christmas). If an event is identified on a non-trading day, the event is assigned to the next market open day. Keeping the event for study allows the high tweet volume event to be considered and analyzed rather than ignored and removed from the analysis. To mitigate against the overlaps and clustering, I remove events that occur five days or less from a previously identified event, in line with the length of the event window of 5 days before and after an event.

III.4.4 *Abnormal Returns*

In this study, I use the Market model similar to Ranco et al. (2015) to calculate the expected, and consequently abnormal, returns. The market model is a one-factor ordinary least squares (OLS) regression equation (see Equation (4) below) that relates the expected returns of a share of stock with a market portfolio, typically approximated with an equity index such as the Dow Jones Industrial Average (DJIA) or S&P 500. While there are many ways to model share price returns, the market model is as good or better than other models, such as the index or average return models (Armitage, 1995). MacKinlay describes the second step as determining the measure of the abnormal return. The return of a stock is the percent difference of the price of a stock for a given day (P_t) from the price the day before (P_{t-1}):

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \qquad (2)$$

"The abnormal return is the actual ex post return of the security over the event window minus the normal return of the firm over the event window. The normal return is defined as the expected return without conditioning on the event taking place." (MacKinlay, 1997, p. 15). Mathematically, this is described in Equation (3) for firm *i* and event date *t* as :

$$AR_{i,t} = R_{i,t} - E\left[R_{i,t}\right] \tag{3}$$

where $AR_{i,t}$ is the abnormal return of the stock of a given company on a given event date, $R_{i,t}$ is the actual return, and $E\left[R_{i,t}\right]$ is the expected return for the stock. The expected return is estimated using the asset pricing model:

$$E\left[R_{i,t}\right] = \alpha_i + \beta_i R_{M,t} \tag{4}$$

To estimate the parameters, α_i and β_i, a regression is performed using the observed daily returns of the stock for firm *i* as the dependent variable and the market portfolio *M*. This study uses the S&P 500 index (Yahoo! Finance ticker ^GSPC) as the market portfolio. The estimation window is 120 trading days with an event window of 5 days before and after the event day. As described by MacKinlay, events dealing with daily stock prices are unlikely to be constrained to a single day. Therefore, event windows aggregate abnormal returns, referred to as the cumulative abnormal return (CAR). Like Ranco et al., this study assumes a five-day window both leading to the event and after it for 11 days for each event.

Given the time constraints for this particular research project, I used a pre-built package to assist in the calculations for abnormal returns. The approach used for this event study is similar to that of Sun and Liao (Sun & Liao, 2011) and described in *Empirical Research in Economics: Growing up with R* (Sun, 2015). The first author (Sun) developed a package in R, *erer* that is helpful to "conduct an event analysis and estimate abnormal returns over time and across firms" (*evReturn Function - RDocumentation*, n.d.). The *evReturn* function returns the abnormal returns for each firm and day within the event window.

To assign a sentiment classification for each event, they are divided into three equal groups based on the rank order of their Positive Ratio. The division of events into equally sized groups is similar to the approach described by Prantl and Mičík (2019), who divided their sample of companies in half based on the rank order of the ratio of positive to negative posts (they discarded neutral posts) and Ranco et al. (2015) who divided the events in their study into thirds by rank order of positive sentiment value. In this study, the events with the highest ratio of positive tweets are categorized as *positive* events, the second highest as *neutral* events, and the lowest ratio of positive tweets as *not positive* events.

In event studies, "the abnormal return observations must be aggregated in order to draw overall inferences for the event of interest. The aggregation is along two dimensions—through time and across securities." (Campbell et al., 1997, p. 160). Therefore, for each event sentiment classification (positive, negative, and neutral), the average CAR is calculated for each day of the event window across all events and stocks. The CAR for each day of the event window is then charted to visualize the accumulation of abnormal returns over each day for each sentiment category.

III.4.5 *Statistical Validation*

"Event studies provide a direct test of market efficiency" (S. J. Brown & Warner, 1980, p. 205). To that end, they rely on identifying the expected return and, ultimately, the excess return of a security (abnormal return). However, it is impossible to know the normal return for security, so statistical validation is necessary to test and provide confidence that the abnormal returns of identified events are in excess due to additional information not reflected in the price. In this study, I used two different tests: the Wilcoxon signed rank test and the t-test.

The t-test is a parametric test assuming abnormal returns are normally distributed. However, there is evidence that abnormal returns tend to skew positively, so this assumption may not hold for event studies such as this one. Therefore, I perform a Wilcoxon signed rank test for statistical significance, a nonparametric test that does not assume a normal distribution. The Wilcoxon signed rank test accounts for the signs and the magnitude of abnormal performance. The null hypothesis is that the proportion of positive measures is about .5 in a given sample. However, Brown and Warner found that “Wilcoxon tests do not appear to reject the null hypothesis often enough.” They may “themselves suffer from such a problem as misspecification” (S. J. Brown & Warner, 1980, p. 218).

In evaluating t-tests, Brown and Warner find that “the differences between the empirical frequency distribution of the test statistics and the t-distribution are generally not large” (1980, p. 248). In their follow-up study using daily returns, they further confirm the non-normality of excess returns but still state, “standard parametric tests for significance of the mean excess return are well-specified…the tests typically have the appropriate probability of Type 1 error” (S. J. Brown & Warner, 1985, p. 25). Given these findings, this study presents results for the Wilcoxon test and the t-test.

III.5 Research Model and Hypotheses

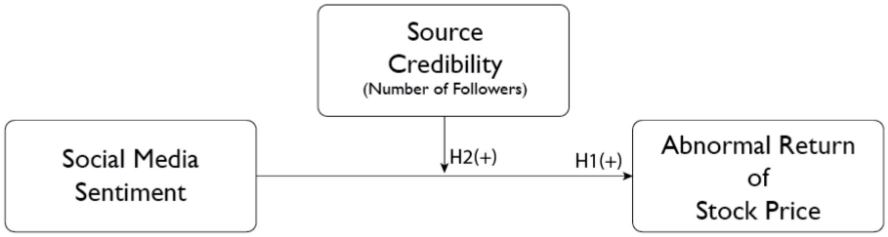


Figure 6. Research model

The research model for this study is presented in Figure 6. Hypothesis 1 (H1) is that I expect that social media sentiment is directly related to the abnormal stock price return. In the event studies, positive events will, on average, have a positive cumulative abnormal return. Furthermore, my Hypothesis 2 (H2) is that source credibility, operationalized with the metric of number of followers, will have a moderating effect on this relationship between social media sentiment and the abnormal return of stock price. Authors with higher numbers of followers will strengthen the relationship between the sentiment of their postings and the abnormal return of the price of a stock.

IV RESULTS AND FINDINGS

The motivation for this study is around the two research questions that explore sentiments of tweets made about specific firms, the number of followers of the users making those tweets, and the abnormal return of stock prices of those firms. To that end, there are two related hypotheses.

Hypothesis 1:

Null Hypothesis 1 ($H1_0$):	The sentiments of identified events have no relation to abnormal returns of firms' share prices.
Alternative Hypothesis 1 ($H1_A$):	The sentiments of identified events are directly related to abnormal returns of firms' stock prices. Positive events will have positive cumulative abnormal returns, whereas negative events have negative cumulative abnormal returns.

Hypothesis 2:

Null Hypothesis 2 ($H2_0$):	The number of followers a posting user has does not affect the relationship between the sentiments of their tweets and abnormal returns of stock prices.
Alternative Hypothesis 2 ($H2_A$):	Tweets made by users with more followers will have a stronger relationship between the sentiments of their tweets and the abnormal returns of firms' stock prices. In contrast, the opposite will be true with sentiments of tweets posted by users with fewer followers.

IV.1 Identifying Events

Events were identified using the automatic detection approach described in Section III.5.C - Detection of Events (p. 38). The baseline was established using tweet volumes for each firm, and, with a sliding window, each day's volume was compared against the median of volumes across a 10-day window (before and after the event day).

Table 7 Number of events identified for top 30% and bottom 30% of followers

Number of events identified

Eventsper firm between posted volume by authors with top 30% and bottom 30% of followers

	Number of Events	
	Top 30%	Bottom 30%
DIS	6	5
NKE	6	6
TGT	29	22
TSLA	33	36
Total	74	69

Firms are DIS = Disney, NKE = Nike, TGT = Target, and TSLA = Tesla

I identified 74 events based on the volume of tweets made by authors in the top 30% of followers and 69 events based on the volume of tweets made by authors in the bottom 30% of the number of followers, as summarized in Table 7. A more detailed listing of the events and their characteristics can be found in 0 Appendix E – Example details of identified events (p. 74). The table in the appendix also details tweets' sentiments, the ratio of positive tweets, the volume of tweets, and the baseline. The last column, Outlier Fraction, is the value of $\phi(d_0)$ from Equation (1) (p. 40). For this study, events are those days that the Outlier Fraction is greater than 1.

Tweets posted between January 1, 2017, through December 31, 2022, were used for this study, resulting in 2,191 possible event days.

IV.2 Calculating Abnormal Returns and Cumulative Abnormal Returns

For each day, abnormal returns were calculated using the market model and performing regressions with the *erer* package's *evReturn* function. The abnormal returns this function returns are averaged across all events and stock securities for each day of an event to calculate the average abnormal return (AAR).

Along with the AARs for each day of an event, the running sum is calculated for each passing day of the event, referred to as the cumulative abnormal return (CAR).

For each iteration, events are categorized into one of three possible sentiments: Positive, Neutral, and Non-Positive. Non-positive is used instead of Negative because the ranking is based on the PositiveRatio, which only takes into account the number of positive sentiment tweets (that *tweetNLP* has assigned a positive probability of >= .75) over the total number of tweets, which includes a mix of neutral and negative tweets. The PositiveRatio metric is preferred over the NegativeRatio because tweets are biased toward the positive sentiment, as shown in Figure 5 (p. 35).

I assigned each event into one of the three sentiment categories by fractiling the set of events into three equally sized groups based on their PositiveRatio ranking – the highest third are categorized as *Positive* events, the middle third as *Neutral*, and the bottom third as Non-Positive.

The CARs are then aggregated based on the average abnormal returns of events in each sentiment category and plotted in charts for visual analysis. Generally, cumulative abnormal returns tended to trend positive for events categorized as positive and negative for events categorized as non-positive. Neutral events tended to have CARs that trended between the positive and non-positive and were closer to 0 CAR throughout events. The trend is generalizable across different dataset variants and iterations. This trend aligns with previous research that found similar effects based on event sentiments on Twitter and social media postings (Ranco et al., 2015). Still, it extends the finding for the broader population of tweets using hashtags instead of limiting them to cash tags to incorporate regular users and not limiting them to investors, analysts, and others who may be focused on stocks. Examples of the CARs plotted for each sentiment category are presented in Figure 7 (top 30%) and Figure 8 (bottom 30%).

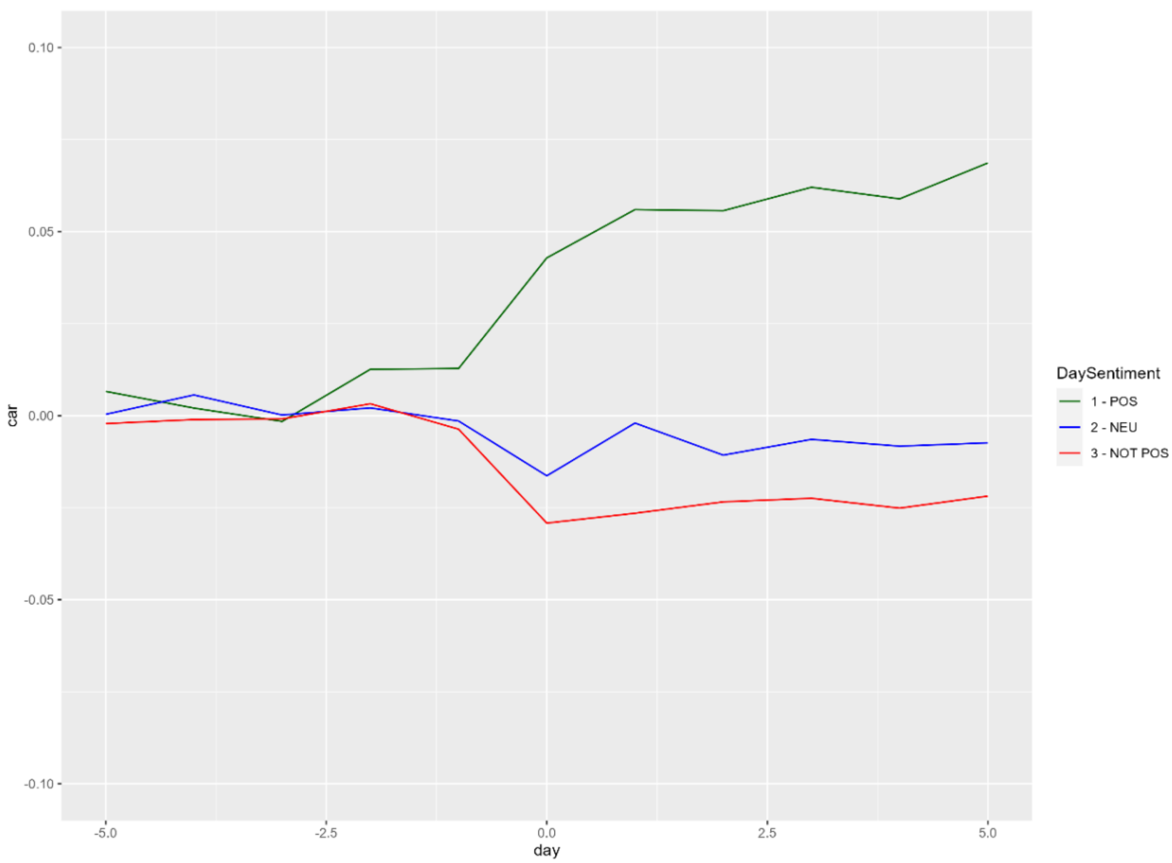


Figure 7 CARs for events identified in the top 30% of authors

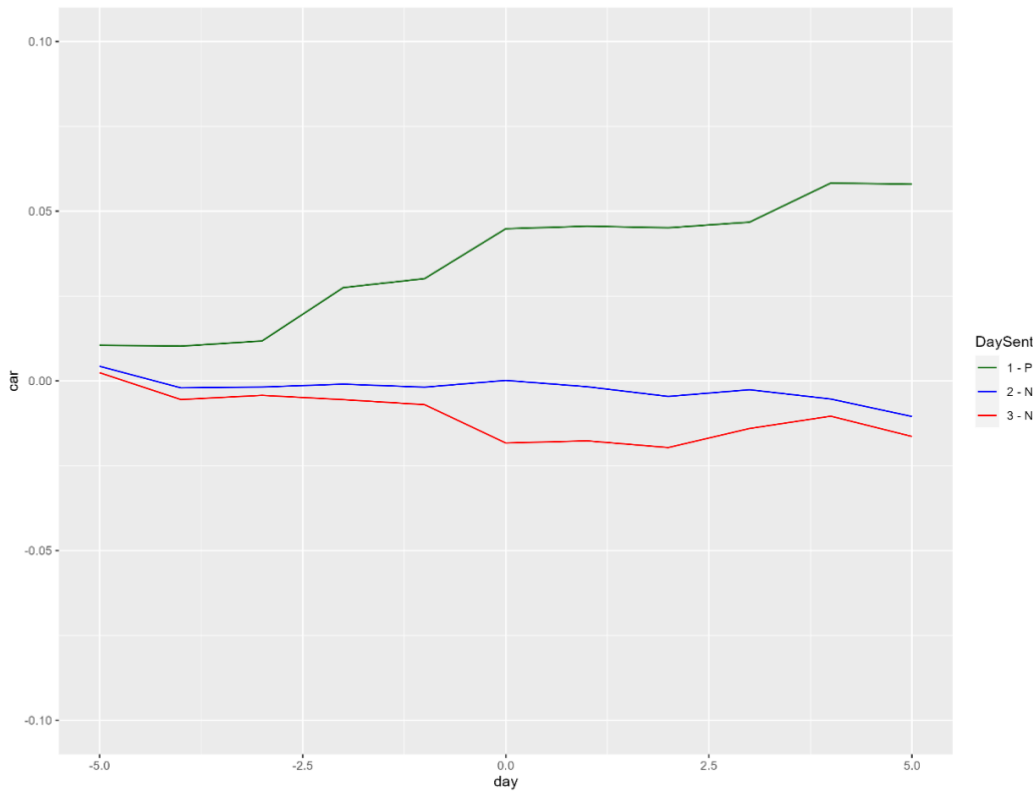


Figure 8 CARs for events identified in the bottom 30% of authors

In line with other research around the stock price, signals show information leakage before the event day, as shown by inflection points preceding the event day. These tend to be more pronounced in the previous day or two.

IV.3 Statistical Tests

Generally, when using the *PositiveRatio* metric for sentiment categorization, there were differences between sentiments of tweets made by users with more followers and sentiments of tweets made by those with fewer followers. More specifically, using the Wilcoxon signed rank test and t-test, there was little to no significance in abnormal returns identified by events when only analyzing authors with a higher rank by number of followers (*TopDown*). In contrast, when analyzing the subsets of tweets posted by lower-ranked authors, there seemed to be more statistical significance (*BottomUp*).

Statistical Tests - TopDown						
Iteration 15 - Top 30 %						
	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0066	0.3254	0.0845	-0.0022	0.7915	0.6870
-4	0.0021	0.8532	0.6605	-0.0011	0.4578	0.8722
-3	-0.0016	0.9578	0.7757	-9e-04	0.4578	0.9136
-2	0.0126	0.1485	0.1585	0.0032	0.8119	0.7189
-1	0.0128	0.2635	0.2605	-0.0037	0.4742	0.6972
0	0.0429 *	0.0588	0.0494	-0.0292	0.0203	0.0536
1	0.0560 *	0.0667	0.0429	-0.0265	0.0755	0.1065
2	0.0557	0.1199	0.0739	-0.0234	0.0802	0.1844
3	0.0620 *	0.1073	0.0477	-0.0224	0.0851	0.2338
4	0.0589 *	0.0755	0.0367	-0.0251	0.1409	0.2007
5	0.0686 *	0.0551	0.0278	-0.0218	0.2200	0.2449

[1] *p<0.05; **p<0.01; ***p<0.001

Table 8. CARs and p-values from the Top 30% of authors by event sentiment

Statistical Tests - BottomUp						
Iteration 15 - Bottom 30 %						
	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0105 *	0.0211	0.0190	0.0024	0.6010	0.5433
-4	0.0102	0.0646	0.0541	-0.0055	0.5399	0.3873
-3	0.0118 *	0.0211	0.0328	-0.0043	0.2467	0.5675
-2	0.0275 **	0.0014	0.0019	-0.0055	0.4274	0.4651
-1	0.0301 **	0.0014	0.0031	-0.0070	0.2226	0.4436
0	0.0448 **	0.0022	0.0042	-0.0183	0.1511	0.3077
1	0.0456 *	0.0315	0.0246	-0.0177	0.3001	0.3521
2	0.0451 *	0.0179	0.0287	-0.0196	0.2726	0.3065
3	0.0468 *	0.0164	0.0186	-0.0140	0.3765	0.5275
4	0.0583 **	0.0079	0.0071	-0.0104	0.6010	0.6589
5	0.0580 *	0.0138	0.0132	-0.0164	0.4274	0.4583

[1] *p<0.05; **p<0.01; ***p<0.001

Table 9. CARs and p-values from the Bottom 30% of authors by event sentiment

Table 8 shows the statistical test results on the 15th iteration of the TopDown approach. This iteration includes the subset of tweets posted by users with the top 30% of the number of followers for each firm. The average CARs for the Positive and Non-Positive events (in addition to the Neutral events) are also plotted in Figure 7. As previously referenced in III.4.5 Statistical Validation (p. 43), there appear to be fewer rejections of the null hypothesis when using the Wilcoxon signed rank test in contrast with the t-test (S. J. Brown & Warner, 1980). The results of the Wilcoxon and t-test are shown alongside the Average CAR for each day of the event window. For the TopDown dataset, the mean CARs are not statistically different from 0 until the day of the event. Then, except for the second day after the event, the p-value of the mean CARs remains under the critical value of .05.

In comparison, the 15th iteration of the BottomUp shows a similar trend for the mean CAR, as shown in Figure 8. This iteration comprises the subset of tweets users post with the Bottom 30% of followers. The mean CARs are significantly statistically different than 0 five days before through five days after the event (except for four days before when the p-value is slightly outside the critical value of .05). The p-values are also much smaller than the analogous days of the event window in the TopDown variant, signifying a stronger relationship between the tweet sentiments and abnormal returns of the stock prices. Furthermore, the critical value is surpassed for both the Wilcoxon test and the t-test strengthening support of this finding. Of note, the Non-Positive events are not strongly significant. This result may be due to the categorization of events using the rank ordering of the PositiveRatio, which combines the Neutral and Negative events into the Non-Positive classification.

I test the statistical significance of the abnormal returns using a combination of t-test for parametric testing and Wilcoxon rank test (for non-parametric testing). I look for trends in the statistical significance of CARs across the different iterations on the day of the event and the day before and after it. For clarity and brevity, Table 10 shows the statistical significance of the event day and one day before and after it (three days total) when including the Top 30% and Bottom 30% of authors based on their numbers of followers. A more detailed listing can be found in Appendix F – Detailed indicators of significance for iterations (p. 78).

Table 8 and Table 9 show the p-values for the positive CARs that are depicted in Figure 7 and Figure 8, respectively (p. 28). The evidence supports rejecting the null hypothesis for Hypothesis 1 ($H1_0$: The sentiments of identified events have no relation to abnormal returns of firms' share prices.). Cumulative Abnormal Returns tend to be greater for Positive events, followed by Neutral and Non-Positive events. In many cases, the differences in CARs are statistically significantly different than 0.

Table 10 Comparison of statistical significance across bottom-up and top-down dataset variants by deciles

Significant iterations - TopDown

Only deciles are presented for clarity and brevity

Sentiment	Top 10%	Top 20%	Top 30%	Top 40%	Top 50%	Top 60%	Top 70%	Top 80%	Top 90%	Top 100%
Positive								*	*	*
Non-Positive										

*p<0.05

Significant iterations - BottomUp

Only deciles are presented for clarity and brevity

Sentiment	Bottom 100%	Bottom 90%	Bottom 80%	Bottom 70%	Bottom 60%	Bottom 50%	Bottom 40%	Bottom 30%	Bottom 20%	Bottom 10%
Positive	*	*	*	*	*	*	*	*		
Non-Positive					*					

*p<0.05

Table 10 shows differences in the relationships between the tweets made by authors depending on their numbers of followers. From the TopDown approach, the relationship is not statistically significantly different from 0 until authors from near the Top 80% are included. This dynamic is also reflected in the BottomUp approach. The relationship is not statistically significant at the very low end of the rankings. However, they are found to be significant around the Bottom 30%. Notably, these two boundaries are relatively similar for both datasets when approaching from either direction.

This finding supports the rejection of the second null hypothesis ($H2_0$: The number of followers a posting user has does not affect the relationship between the sentiments of their tweets and abnormal returns of stock prices.). However, the second hypothesis is only partially supported because, while I expected the strength of the relationship to be stronger with sentiments of tweets made by users with a high number of followers, the results of my analysis are that the relationships are stronger with those made by users with a lower number of followers. As will be discussed later, I do not believe this means source credibility (as measured by number of followers) is not important to fully understanding the relationship between tweet sentiments and abnormal returns of stocks. Still, it may be mediated through additional constructs important to eWOM.

Alternative Hypothesis 2 ($H2_A$):	Tweets made by users with more followers will have a stronger relationship between the sentiments of their tweets and the abnormal returns of firms' stock prices. In contrast, the opposite will be true with sentiments of tweets posted by users with fewer followers.

Table 10 observes that the significance is stronger on the lower-end, which contrasts with $H2_A$, supported by the Source Credibility Model. Therefore, Hypothesis 2 should be considered to be only partially supported.

V DISCUSSION

V.1 Social Media Influencers and Source Credibility

Social media is a valuable data source for academics and practitioners. It allows researchers to access the opinions and sentiments of many users across a diverse range of topics. This study uses empirical evidence through the collection of tweets to explore the relationship between their sentiments and the abnormal return of stock prices. It extends the current body of knowledge by leveraging a larger dataset that includes users in “normal conversation,” whereas extant research has typically focused on investors using cashtags. Furthermore, little research delves into the effect of influencers and follower counts. In this study, I leverage Twitter (now known as X) to explore the sentiments of opinions through users' tweets and a novel approach, considering their number of followers as a variable to the concept of source credibility. With the user’s follower counts, we can get a sense of their credibility as a source to explore if there is an effect of credibility related to the abnormal price movements of certain stocks. My findings are that source credibility is not sufficient as there appears to be a stronger relationship at the lower end of rankings of Twitter (X) users based on their follower count. From the source credibility perspective, these users would be considered “less credible.” However, I do not believe this to be the full explanation.

Extant literature supports that there is a relationship between sentiments of postings made on online social media and the price of stocks. Previous studies have used regression analysis of tweet sentiments using different models on stock indices or specific industries such as finance. Very few have explored the effect of followers on such relationships. One such study was done by Sul et al. (2017), where they studied how followers may affect the speed of spreading sentiment as a social contagion and being priced into certain stocks: “Tweets spread positive or negative sentiment about a stock through the market and can influence prices, and thus the returns from trading those stocks. Sentiment can spread quickly; for example, network hubs" (Sul et al., 2017, p. 475). Network hubs, in their case, would be influencers with a high number of followers. In this study, we look at followers through the lens of source credibility. The idea is that followers would be more receptive to considering and internalizing the sentiments and opinions of those they follow. Sprenger et al. (2014) also found that, in stock-related tweets (using cash tags), there is evidence that users are more source credible and receive "greater attention in microblogging forums through higher levels of retweets as well as a larger followership."

When related to all tweets using the more popular hashtags, how a user's followership affects the relationship between sentiments and abnormal stock returns may be indirect. An article by Coyne et al (2017) supports that source credibility should not be dismissed. Their study compares the accuracy of different prediction models based on posts and users on the StockTwits platform. In their third model, they incorporate three additional factors: the number of likes, the user's follower count, and how often the user is correct. This third model, which they labeled "Smart Users," was more accurate than other models tested. The three factors are source credibility, expertise (correctness), and trustworthiness (likes and follower count). In a study by Pal et al. (2020), the authors only incorporate tweets they consider authentic through a process that considers the number of followers, whether the user is a verified user, or whether the retweet count is greater than or equal to 50. Notably, the latter two criteria do not matter if a user's follower count is not greater than 100 and the tweet is immediately discarded.

V.2 Relationships and effects of tweet sentiments and abnormal returns

Based on the findings and results, there may be a mediation effect between source credibility and abnormal stock returns. Visualizations of the cumulative abnormal returns across event windows and categorized by sentiments depict a picture of a consistent trend – positive events trend upwards over an event window to a CAR higher than that of neutral and non-positive CARs. In support, neutral events trend around a stable return, and non-positive events negatively affect CARs, supporting Hypothesis 1. The data shows these are consistent regardless of how the data was subsetted, though the CARs' magnitude differs.

Statistical significance is much more varied based on how the number of followers is incorporated into the analysis. Generally, there are stronger statistical validities with fewer followers than those with more - partially supporting Hypothesis 2. Higher source credibility will have stronger relationships with abnormal price returns, so I would not discount that there is an indirect effect. The datasets filtered for the lowest numbers of followers (including no followers) were also statistically non-significant. Low-to-moderate number of followers tended to have statistically significant abnormal returns.

From the data, one reason is that more users fall into lower ranges – it becomes increasingly difficult to gather more followers and rise in rankings. In other words, there are fewer individuals with more followers, as shown in **Error! Reference source not found.**, where a long tail skews the distribution right. From the collected tweets, the top 30% in terms of number of followers consists of 346,095 distinct authors who posted 2,880,609 tweets in 2017-2022. The bottom 30% consisted of 405,924 users who posted 1,509,194 tweets. On average, the influencers posted 8.32 tweets per user vs 3.72 for the lower-ranked users. Notably, this difference persists in the 25th iteration. The top 50% comprises 607,610 distinct authors posting 4,130,520 tweets, whereas the bottom 50% comprises 671,934 authors posting 2,508,346 tweets. The average tweets posted per user are 6.80 and 3.73, respectively. As a baseline for comparison, the full dataset of collected tweets is 6,735,873 tweets[7] posted by 1,302,380 distinct authors (about 5.17 tweets per author).

Based on the comparative numbers above, fewer individuals are considered at a higher range of followers. On the other hand, the influencers, on average, tend to be more expressive, posting more tweets (twice as many) than the lower-end users. This study does not attempt to study how information and sentiment are dispersed through electronic Word of Mouth. However, trust and source credibility is "the basic pre-requisite for successful WOM" (Prantl & Mičík, 2019, p. 2). Therefore, while source credibility may be important, the relationships may be more directly related by word-of-mouth. Source credibility, in actuality, may be mediated through this phenomenon.

The topic of discussion may also moderate source credibility. In some cases, influencers have a high followership because they are celebrities or experts in unrelated topics. Take, for example, Ellen DeGeneres, a well-known celebrity and the author with the highest number of followers in 0 Appendix C – Top Followers for Each Firm (p. 70). On December 4, 2017, Ellen posted a tweet[8] with the hashtag Tesla:

> .@TiffanyHaddish had a dream, and I made that dream come true. #Tesla #EllenShowMeMore https://t.co/YLh2GxcbTT

The tweet includes a link to a YouTube video clip of her show posted on the same day. In the video, she is gifting a Tesla to another celebrity, Tiffany Hadish, who expresses her desire for a Tesla. Source credibility applies in this context in two different ways. First, does Ellen have

---

[7] Note this number differs from the total 6,780,072 in **Error! Reference source not found.** (p.27) because 44,199 (about 0.65%) tweets were posted by users for which profile data was not available or accessible. Consequently, these tweets were removed from this study. Details of the distribution of missing profile data can be found in 0 Appendix G – Table of Completeness (p. 69)

[8] https://twitter.com/EllenDeGeneres/status/937793051782606848

expertise in the Tesla firm or their products and services? Secondly, does the tweet affect others through word of mouth, not necessarily through her source credibility as a Tesla expert, but as a person who is willing to acquire and use Tesla? Therefore, is she also motivating others (who are likely not to have as many followers) to spread sentiments and opinions through their social media? Wies et al. (2022) also find that indegree (the number of followers) is a part of the evaluation of the effectiveness of influencers in marketing campaigns. On their own, they do not provide a comprehensive measure of source credibility and eWOM. It is also important to consider the connections between users. This study aligns with their findings that there is no statistically significant information relating to abnormal stock price returns at the extreme ends of the rankings. Recent studies on SMI further delineate influencers into micro-influencers and mega-influencers and have found that the two groupings may have different motivations for generating WOM, which can have an impact on followers' (or potential followers) perceptions of their trustworthiness (Li et al., 2024).

Users with fewer followers may consider information in the form of opinions and sentiments from multiple influencers that are equally credible. If they differ, they may further aggregate with other influencers and perhaps with other peer users with similar or higher followers. One other consideration is the concept of relevance. While a user has many followers and is generally credible, users may moderate their perception of credibility based on the specific expertise of the influencers they are following. A network or graph approach in future research could provide valuable insights into considering how relevance may be measured and incorporated into such analyses.

V.3 Contributions

This study makes contributions to research and practice. There is interest in Social Media Influencers (SMI), and market research has called for more research in this space (Vrontis et al., 2021). In practice, SMI is becoming a critical component of marketing campaigns, especially because they are perceived and internalized differently from traditional celebrities (Djafarova & Rushworth, 2017; Jin & Phua, 2014; Ohanian, 1990; Wies et al., 2022). However, there is still a dearth of extant literature covering SMI.

First, this study shifts the focus from sentiments of tweets made by investors and analysts to those by ordinary users. By using hashtags instead of cashtags (Bartov et al., 2018; Bouadjenek et al., 2023; Cookson et al., 2023; Hong et al., 2020; Hossain et al., 2022), this shift captures a broader set of tweets about companies and may provide a stronger means of capturing consumer sentiment than the current approach. The benefit of cashtags is that they can "increase confidence that the tweets relate to the firm financial performance and value, thereby increasing the relevance of [their] measures" (Bartov et al., 2018, p. 31). However, they limit the type of users and largely ignore the consumers. While investors are important firms' stakeholders, this study can provide more practical insight for marketing departments to understand their (potential) customers better.

Second, this study considers social media influencers and their effect on the relationship between tweet sentiment and stock price, an effect largely overlooked by extant research. Supported by this research, the followership of users impacts the relationship's strength. Furthermore, the strongest effect is not necessarily found with mega-influencers. While this supports previous studies in consumer behavior and social media (Li et al., 2024; Tian et al., 2023; Wies et al., 2022), research has not explored relationships between online social media sentiment and stock prices.

V.4 Limitations and Constraints

This study was performed within certain constraints and, therefore, has limitations. First, the Twitter (now known as X) platform has closed its (free) access to academic research. Therefore, the tweets collected for this study were done before the sudden closure and are limited. Tweets were collected for four separate firms: The Walt Disney Company, Nike Inc, Target Corp, and Tesla Inc. The tweets collected are longitudinally robust (covering the years 2010 – 2022), and the firms serve B2C, which Prantl & Mičík (2019) found to be more appropriate for studies involving online social media data. Even so, there is a limitation to using only these four firms in this study. Second, in order to focus on additional information provided by sentiments of tweets posted by more ordinary consumers, the search queries used for extracting tweets from the Twitter API used hashtags exclusively. However, since cashtags have been studied (as shown in Table 1, p. 24), there may be correlations between the sentiments of ordinary users and investors, creating an endogeneity concern in the form of omitted variable bias. To mitigate this, incorporating tweets from investors and financial analysts (including tweets using cashtags) as a control could help separate effects specifically associated with ordinary users.

With the new leadership installed Twitter (X), there are new features and perceptions of how to continue platform development. These changes, or the unknown expectations of future changes, may affect the population, culture, and usage of Twitter and may also have inherent changes simply because of the fears around possible future changes that may motivate users to leave the Twitter platform and move to a different one. This change could further limit the findings and approaches of this study.

This event study provides a novel approach to including the Social Media Influencers (SMI) concept using a subset method based on users' follower counts. A technical limitation to obtaining user profile data from the Twitter API is that the follower counts are accurate as of the extraction time, not the posting time. Therefore, there may be some inaccuracies if users' follower counts significantly changed between the post and the extraction time (in early 2023). If the change is significant enough, in either direction, the tweet may fall and become included in a different subset than it would have during the original post date. There is also evidence that followers may be false. Zhang et al. point out that while Twitter studies use followers as a proxy of influence, it was inappropriate for their study on Weibo because "some Weibo users manipulate influence by purchasing fake followers" (2017, p. 155). This manipulation may merit additional research for Twitter (X). Using the number of followers alone as a variable for source credibility may not provide measures of expertise. Previous research has suggested that "retweets are driven by the content value of a tweet, while mentions are driven by the name value of the user" (Cha et al., 2010, p. 17) – including these metrics should be considered to strengthen the operationalization of source credibility, which in turn would facilitate more robust analysis.

While all of the firms in this study serve consumers (B2C), there may be differences in how they engage with SMIs and what kinds of influencers they partner with. For example, firms like Nike work with traditional celebrities often but may not have the same levels of engagement with their followers (Jin & Phua, 2014; Wies et al., 2022; Y. Zhang et al., 2017). In contrast, Disney's marketing group provides special access and events for the influencers they partner with. Those influencers engage closely with their followers – through live vlogs (video blogging), posts, and different online social networks. In many cases, the message may be more strongly delivered on YouTube than on Twitter (X).

This study, like many others, explores users when aggregated as groups. The work by Bouadjenek et al. (2023) describes the importance that users' behaviors and predictive accuracies may differ significantly. They classify users as consistently correct (or incorrect) over time for their stock price predictions (whether they are bullish or bearish on a stock). Based on their analysis, users can fall into three different temporal horizons regarding their predictive accuracy (within 20 days, 60-100 days, and more than 100 days). Integrating their approach into a study such as this could provide an approach to considering topical relevance and expertise supporting source credibility.

Quantitative analyses require researchers to make certain decisions, in some cases arbitrarily or based on a different study, concerning calculations and modeling. For example, to estimate the abnormal returns for each event, numerous decisions were made: which model to use (the market model), which market portfolio to use as the independent variable (S&P 500), and the number of days for estimation (120 days). The factors going into event studies are an area of research interest (Armitage, 1995; S. J. Brown & Warner, 1980, 1985; Campbell et al., 1997). Another notable decision is the specification of the Outlier Fraction threshold. Much of this study followed the approach specified in the work done by Ranco et al. (2015). One notable diversion from their approach is that they used an Outlier Fraction threshold of 2.0, whereas, in this study, I used 1.0. Where their study used cashtags for extracting tweets, this one uses the more commonly used hashtags, yielding a much greater volume of tweets with higher volatility. Using a higher threshold would yield a much smaller and arguably less robust population of events. Utilizing an Outlier Fraction of 2.0 with the approach specified in this study yielded a very small number of events per iteration and resulted in little statistical significance and insight.

Likewise, there are different approaches to sentiment analysis. In some studies, it is a core component of the research design, including efforts for building machine learning algorithms (Agarwal et al., 2011; Naresh et al., 2022; Pagolu et al., 2016) for a more purpose-built approach to sentiment classification. I use the tweetNlp python module in this study due to resource constraints. The benefit is that using a prepackaged library allows for easier reproducibility. While the package has its classification algorithms, I restrict the classifications of positive and negative sentiments more tightly in this study, requiring at least a .75 probability for assigning one of the two labels. Otherwise, the tweet is considered neutral. Some sensitivity testing was done, adjusting the threshold between .6 and .9 with no major changes in findings.

V.5 Future Research Opportunities

Many research studies are made feasible (or infeasible) by access to data. Twitter (X) as a platform for research has become more constrained since Spring 2023. A critical need in research is to find a substitute for this new void and test the findings made in this study across a larger group of firms and users. If access to Twitter is available, it would also be prudent to evaluate if the demographic of users has changed and if those changes affect the applicability of previously established findings and tests. While this study was constrained to the four specific firms, future studies may focus on firms individually to provide some mitigation against selection bias to compare analyses and findings across industries, consumer types, or firm size.

The relationships between users and their tweets would also be of great interest. A key finding in this study is that there is support for relationship effects from the number of followers. However, the strongest relationships tended to be toward the lower end. Future research should explore how source credibility and the number of followers may indirectly affect abnormal returns of stock prices through word-of-mouth transmission to larger groups of users at lower ranks of followership. This transmission would be important not only in securities and finance but also in how users may follow influencers for purely hedonistic value – perhaps not giving much credibility to those high follower influencers but finding their tweets and opinions entertaining. Twitter's retweet and Mention feature – where users can notify other users of their tweet using the @ symbol followed by the firm's Twitter handle, could provide a useful way of following a conversation between users and firms or between influencers and followers. Including additional factors has been supported by Coyne et al. (2017). They improved the accuracy of their predictive models by incorporating additional factors such as likes and users' follower counts.

In the near term, additional datasets will be made available and research published using data collected from Twitter before the discontinuation of API access (this study, for example). These future studies could present additional opportunities to test and perform meta-analysis using historical data, especially those that previously had access to the Academic API. For example, Pfeffer et al. (2023) collected a "complete" dataset over 24 hours starting September 21, 2022. Through a collective and collaborative effort, they collected 374,937,971 tweets posted by 40,199,195 accounts during the period.

Other platforms, such as Meta (the parent company of Facebook, WhatsApp, and Instagram) may also provide additional access for academic research. While these social networks may serve different purposes and serve different types of users, they may include a different representation of the population of people and provide different insights into their predictive power for financial securities. While this study is specifically concerned with users that are not necessarily investors (I use hashtags instead of cashtags), the same questions regarding SMI apply to groups focused on investments and financial securities. Research supports that there may be more relevance to different use cases (such as stock price prediction) found in other platforms, such as StockTwits. StockTwits is very similar to Twitter but is primarily focused on communications and sharing information regarding financial securities: "StockTwits is a social media platform that is more likely to be used by experts in the stock market as it is intended for that purpose" (Bouadjenek et al., 2023, p. 9:11).

This study focused on daily stock price returns subject to many other factors not in individual consumers' control or influence. Future research should also explore how online social media users relate to a firm's fundamentals since they directly affect revenue and contribution by their actions.

Finally, the approach taken in this study to consider the number of followers is novel. Future research can look for other methods to study the phenomenon of SMIs to develop methods further and incorporate followers as a control variable in univariate and multivariate regressions, which are more typical in current research. While using the TopDown and BottomUp approaches solely focus on the number of followers and use statistical methods (such as quantiles) to delineate between the follower rankings of authors. A valuable research opportunity could look at integrating further concepts such as relevance or topical source credibility that can provide meaningful interpretations of relationships between users in online social networks.

VI CONCLUSION

Social media platforms have become essential for global users to express their views and learn from others. These platforms' ease of access and engagement allows information to spread rapidly. The credibility of the sources users choose to follow significantly influences the type of information conveyed and received. Current research explores how social media sentiments can predict outcomes like abnormal stock price returns. This study shifts the focus from investors and financial analysts to a broader user base, revealing that additional information may not be reflected in a stock's price. Interestingly, a higher number of followers doesn't necessarily strengthen this relationship.

The study also proposes incorporating source credibility into event studies, as represented by the number of followers. This approach can be further refined to include additional variables that represent the expertise construct, providing a better measure of source credibility. Considering these additional measures is an area for future research consideration.

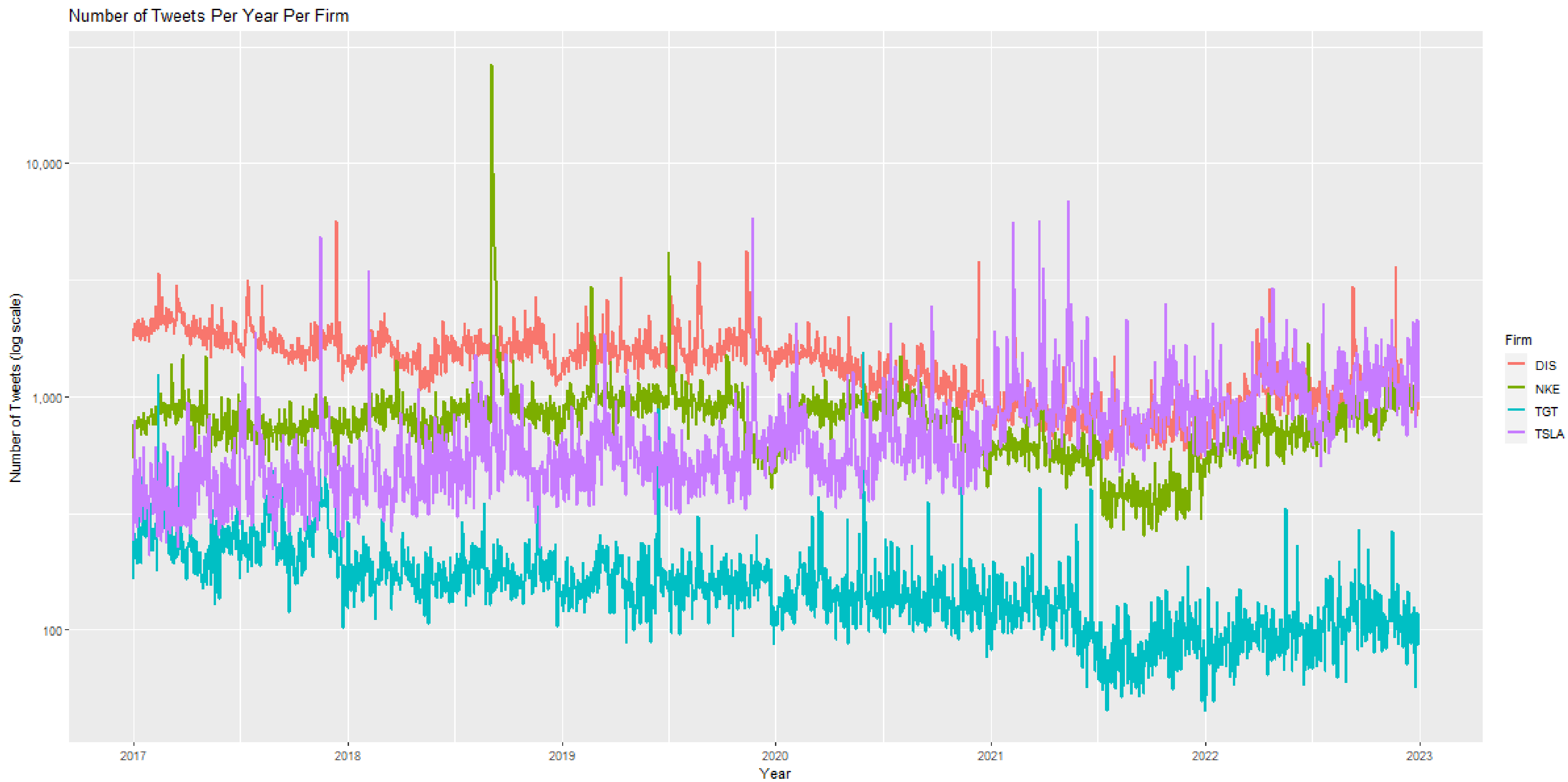
Social media influencers, who serve as hubs for their followers, will continue to spark interest in academic research and marketing practices. As social media platforms evolve and new ones emerge, they reflect society's dynamic tastes. While these changes pose challenges for research leveraging these data sources, they also offer significant opportunities to understand human interactions in diverse social settings better.

VII APPENDICES

Appendix A – Table 2 from Mathiassen (2017, p. 21)

Twitter Sentiments and Stock Prices: An Event Study on the Role of Influencers	
Component	**Specification**
Target Journals	Journal of Marketing, Journal of Marketing Research, Journal of Consumer Research
Problem setting (P)	Social Media Influencers (SMI) have become an area of academic research and practice interest. News and information tend to disperse – widely in many cases – through users following other users on social media, referred to as electronic word of mouth (eWOM). This dispersion is affected by many factors, including source credibility. It is expected that users with a high number of followers will have a greater quality of eWOM in terms of its distance and rate of dispersion. While the relationship between tweet sentiments and stock prices has been studied and mostly supported, there is currently no extant research on how influencers may affect that relationship.
Area of concern (A)	Relationship between tweet sentiments and abnormal stock returns
Conceptual framing (F)	Framing related to A (F_A): social media, social media influencers, source credibility, wisdom of crowds, electronic word of mouth Framing related to M: Sentiment analysis; social networks, panel (fixed effects) regression analysis; moderation analysis
Research method (M)	A quantitative approach using an event study to analyze sentiments of tweets and their relationship with abnormal stock returns and if there is a difference in effect depending on the number of followers. Two types and sources of data – 1) sentiment analysis and 2) stock price data
RQ	Does the sentiment of tweets relate to abnormal returns of stocks during events of high tweet volumes? Is there a difference in the relationship in (1) depending on the number of followers of the tweet authors?
Contribution (C)	Contribution to A (CA): (1) This extends the body of knowledge in SMI and explores the effect of number of followers in a quantitative study of tweet sentiments and stock prices, which responds to the call for more research regarding social media influencers (Vrontis et al., 2021). Contribution to P (CP): Understand the relationship between consumers' sentiments as shared on online social media and firms' abnormal stock price returns. Extant research has focused on investors using cashtags whereas here, I use more general hashtags from all users

Appendix B – Chart of Tweets Per Firm Per Day



Appendix C – Top Followers for Each Firm

Top Followers - DIS

User	Followers
Forbes	18,788,527
MTV	17,765,155
ndtv	17,718,301
TheRock	17,065,140
9GAG	16,849,200
timesofindia	14,684,832
billboard	14,113,981
ABPNews	13,246,147
CGTNOfficial	13,128,907
WWE	12,979,787

Top Followers - NKE

User	Followers
Cristiano	107,608,618
premierleague	39,597,974
KevinHart4real	37,618,162
Reuters	25,746,422
LFC	23,333,837
10Ronaldinho	21,552,763
ImRaina	20,860,622
9GAG	16,849,200
RafaelNadal	15,782,316
timesofindia	14,684,832

Top Followers - TGT

User	Followers
chrissyteigen	12,960,339
enews	11,975,871
thalia	9,811,797
IGN	9,624,984
DCComics	5,485,188
ARYNEWSOFFICIAL	5,253,549
USATODAY	4,946,173
2chainz	4,515,032
EconomicTimes	4,324,069
InStyle	4,284,960

Top Followers - TSLA

User	Followers
EllenDeGeneres	76,641,724
Reuters	25,746,422
Forbes	18,788,527
ndtv	17,718,301
timesofindia	14,684,832
ABPNews	13,246,147
CGTNOfficial	13,128,907
XHNews	12,100,680
MeekMill	11,514,624
Riteishd	11,403,412

Appendix D – Example of Timeline of Events and Tweet Volumes

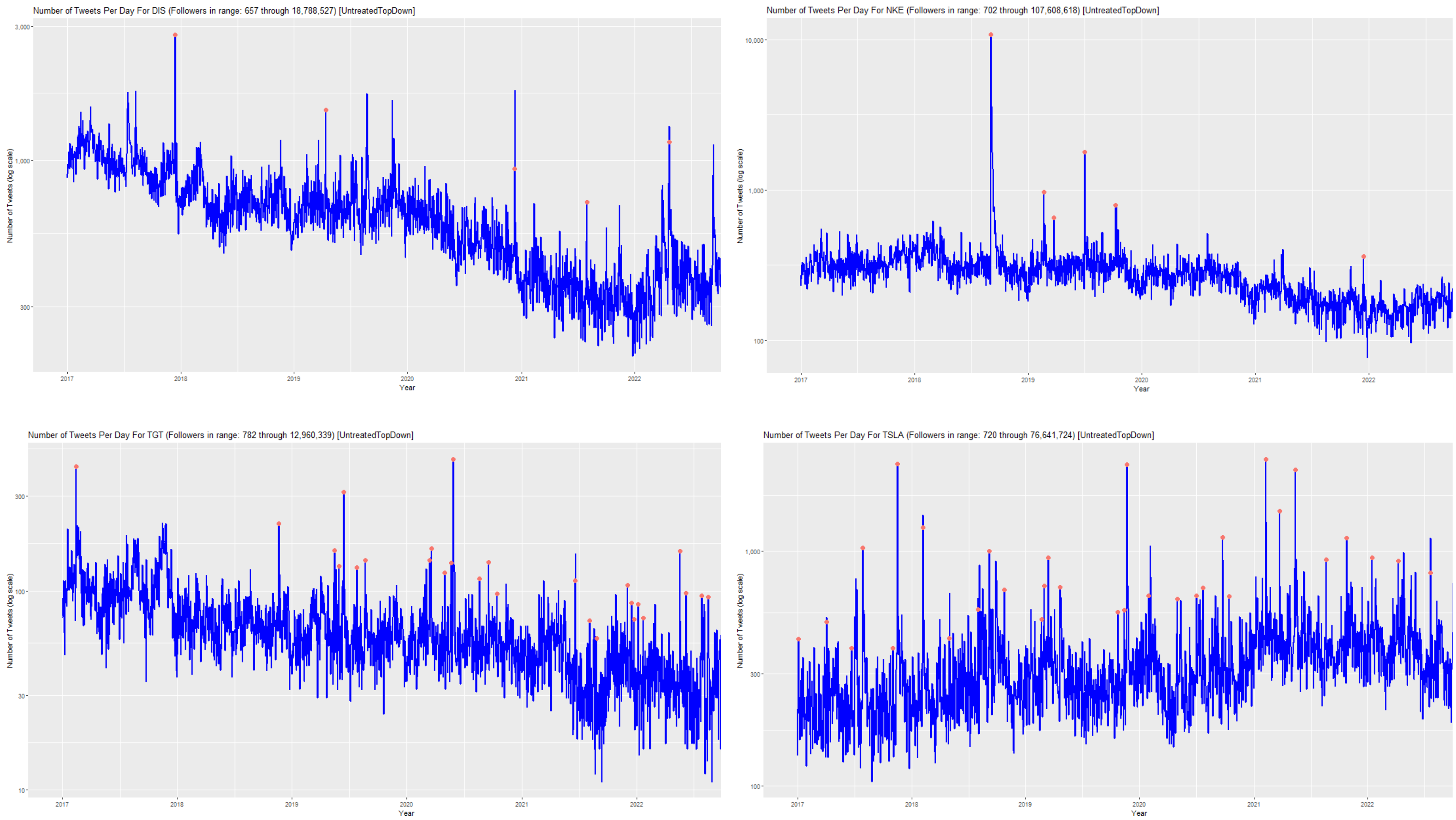
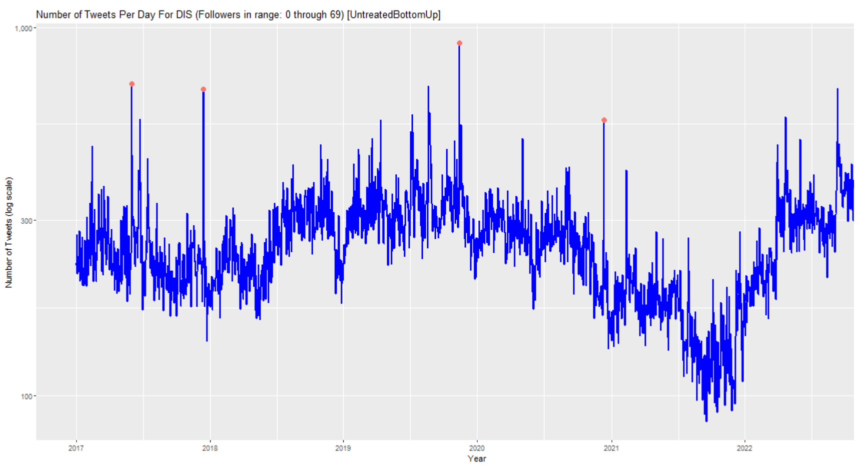
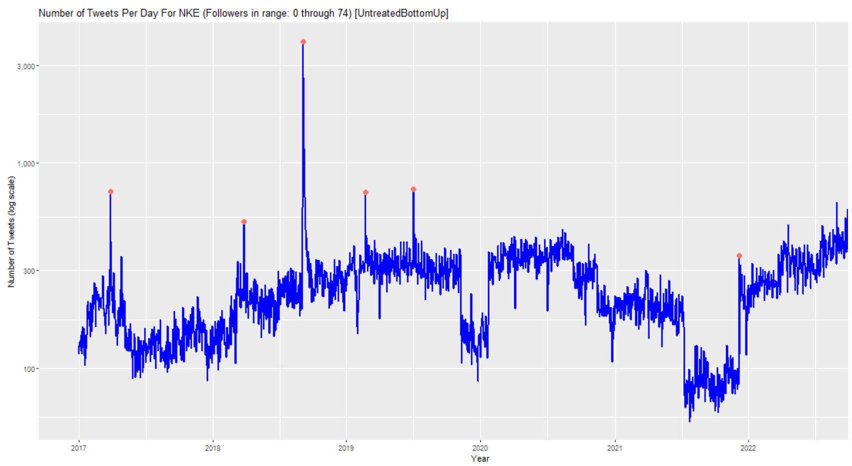
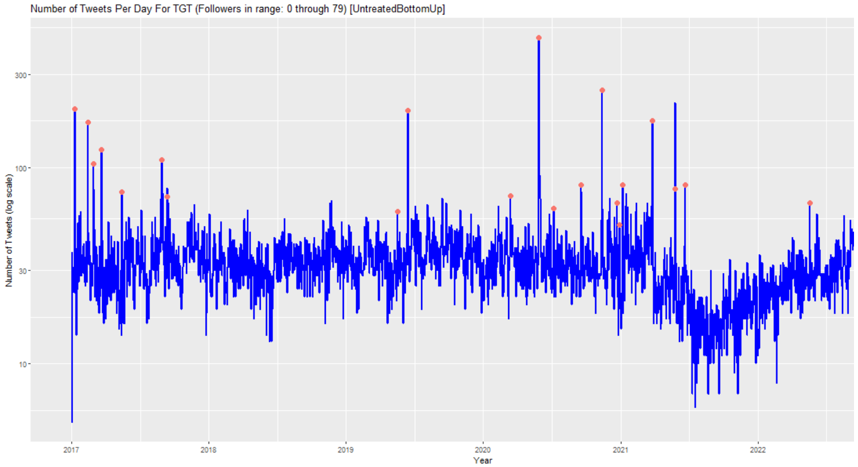


Figure 9. Iteration 15 of the TopDown dataset.

Events are identified with red diamonds (♦) for each firm (clockwise from top left: Disney, Nike, Tesla, and Target). The y-axis is the log volume of tweets from users with the top 30% of followers.







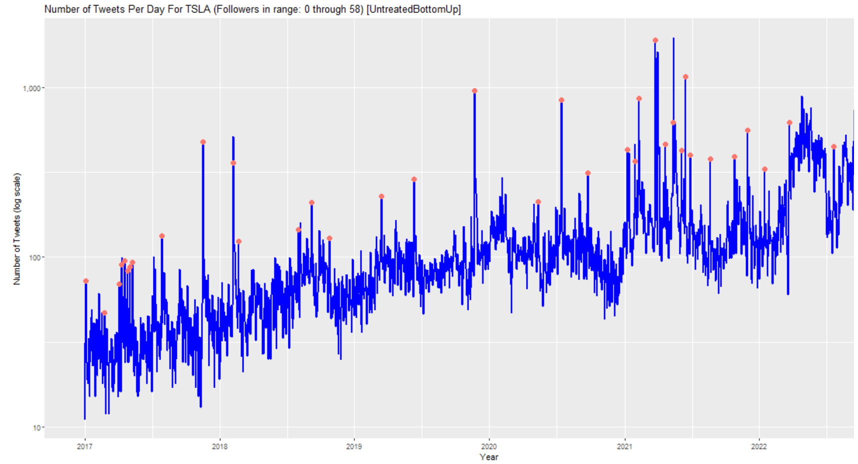


Figure 10. Iteration 15 of the BottomUp dataset.

Events are identified with red diamonds (♦) for each firm (clockwise from top left: Disney, Nike, Tesla, and Target). The y-axis is the log volumes of tweets from users with the bottom 30% of followers.

Appendix E – Example details of identified events

Table 11 Events for tweets from users with the top 30% of the number of followers

Using the TopDown dataset, iteration 15 (74 events identified)

		Counts					Ratios		Event Identification	
Firm	Date	Posting Users	Positive	Neutral	Negative	Total	Positive	Negative	Baseline	Outlier Fraction
DIS	12/14/2017	2,048	697	2,017	79	2,793	0.25	0.028	1,006	1.776
DIS	4/12/2019	970	398	1,086	26	1,510	0.264	0.017	725	1.083
DIS	12/10/2020	494	298	600	33	931	0.32	0.035	461	1.02
DIS	7/30/2021	535	151	479	78	708	0.213	0.11	350	1.023
DIS	4/21/2022	916	150	731	280	1,161	0.129	0.241	564	1.059
DIS	11/21/2022	1,022	397	958	88	1,443	0.275	0.061	442	2.265
NKE	9/4/2018	8,366	1,381	6,800	2,660	10,841	0.127	0.245	1,677	5.465
NKE	2/21/2019	769	76	751	141	968	0.079	0.146	357	1.711
NKE	3/25/2019	468	61	532	62	655	0.093	0.095	314	1.086
NKE	7/2/2019	1,415	158	1,080	557	1,795	0.088	0.31	378	3.749
NKE	10/10/2019	227	74	699	19	792	0.093	0.024	350	1.263
NKE	12/14/2021	307	97	259	7	363	0.267	0.019	176	1.062
TGT	2/13/2017	385	107	249	65	421	0.254	0.154	146	1.884
TGT	11/20/2018	196	123	81	15	219	0.562	0.068	65	2.369
TGT	5/16/2019	51	15	144	1	160	0.094	0.006	64	1.5
TGT	5/30/2019	59	13	116	4	133	0.098	0.03	50	1.66
TGT	6/15/2019	249	30	171	113	314	0.096	0.36	103	2.049
TGT	7/26/2019	48	17	111	3	131	0.13	0.023	47	1.787
TGT	8/21/2019	110	43	96	4	143	0.301	0.028	66	1.167
TGT	3/13/2020	123	17	103	23	143	0.119	0.161	68	1.103
TGT	3/19/2020	134	17	105	41	163	0.104	0.252	72	1.264
TGT	5/1/2020	107	17	97	10	124	0.137	0.081	53	1.34
TGT	5/22/2020	67	13	90	35	138	0.094	0.254	59	1.339
TGT	5/28/2020	340	20	250	189	459	0.044	0.412	74	5.203
TGT	8/19/2020	81	36	75	4	115	0.313	0.035	48	1.396
TGT	9/16/2020	107	18	82	40	140	0.129	0.286	53	1.642
TGT	10/13/2020	56	19	75	3	97	0.196	0.031	48	1.021
TGT	6/20/2021	85	24	89	0	113	0.212	0	38	1.974
TGT	8/5/2021	51	16	54	1	71	0.225	0.014	28	1.536
TGT	8/26/2021	37	17	38	3	58	0.293	0.052	24	1.417
TGT	12/2/2021	38	16	88	3	107	0.15	0.028	43	1.488
TGT	12/16/2021	35	10	74	3	87	0.115	0.034	39	1.231
TGT	12/23/2021	35	16	55	1	72	0.222	0.014	33	1.182
TGT	1/6/2022	33	8	78	0	86	0.093	0	35	1.457
TGT	1/21/2022	35	7	63	3	73	0.096	0.041	35	1.086
TGT	5/18/2022	128	10	109	40	159	0.063	0.252	32	3.969
TGT	6/7/2022	81	8	81	9	98	0.082	0.092	25	2.92
TGT	7/27/2022	29	4	91	0	95	0.042	0	47	1.021
TGT	8/17/2022	75	12	68	13	93	0.129	0.14	36	1.583
TGT	10/6/2022	42	18	78	2	98	0.184	0.02	38	1.579
TGT	11/16/2022	91	9	86	27	122	0.074	0.221	43	1.837

TSLA	1/4/2017	329	90	310	22	422	0.213	0.052	209	1.019
TSLA	4/3/2017	380	191	304	5	500	0.382	0.01	241	1.075
TSLA	6/23/2017	287	45	330	10	385	0.117	0.026	192	1.005
TSLA	7/29/2017	579	283	731	19	1,033	0.274	0.018	375	1.755
TSLA	11/2/2017	300	38	274	75	387	0.098	0.194	183	1.115
TSLA	11/17/2017	1,561	736	1,581	40	2,357	0.312	0.017	369	5.388
TSLA	2/6/2018	1,024	511	727	22	1,260	0.406	0.017	315	3
TSLA	5/2/2018	289	80	312	34	426	0.188	0.08	202	1.109
TSLA	8/1/2018	315	118	405	42	565	0.209	0.074	252	1.242
TSLA	9/7/2018	687	140	735	129	1,004	0.139	0.128	390	1.574
TSLA	10/25/2018	388	195	478	14	687	0.284	0.02	329	1.088
TSLA	2/21/2019	236	166	314	34	514	0.323	0.066	226	1.274
TSLA	3/1/2019	525	146	539	26	711	0.205	0.037	337	1.11
TSLA	3/15/2019	544	257	652	32	941	0.273	0.034	375	1.509
TSLA	4/22/2019	376	131	524	47	702	0.187	0.067	312	1.25
TSLA	10/24/2019	365	165	349	37	551	0.299	0.067	275	1.004
TSLA	11/13/2019	423	111	422	27	560	0.198	0.048	268	1.09
TSLA	11/22/2019	1,689	493	1,713	134	2,340	0.211	0.057	483	3.845
TSLA	1/30/2020	437	199	429	19	647	0.308	0.029	300	1.157
TSLA	5/1/2020	400	84	473	67	624	0.135	0.107	279	1.237
TSLA	7/2/2020	441	233	400	12	645	0.361	0.019	308	1.094
TSLA	7/22/2020	466	197	482	18	697	0.283	0.026	285	1.446
TSLA	9/22/2020	611	332	779	33	1,144	0.29	0.029	351	2.259
TSLA	10/13/2020	212	382	248	12	642	0.595	0.019	274	1.343
TSLA	2/8/2021	1,850	471	1,958	36	2,465	0.191	0.015	438	4.628
TSLA	3/24/2021	1,140	347	1,114	23	1,484	0.234	0.015	437	2.396
TSLA	5/13/2021	1,606	261	1,734	229	2,224	0.117	0.103	707	2.146
TSLA	8/20/2021	593	166	712	43	921	0.18	0.047	381	1.417
TSLA	10/25/2021	759	355	756	29	1,140	0.311	0.025	499	1.285
TSLA	1/14/2022	610	212	709	20	941	0.225	0.021	440	1.139
TSLA	4/8/2022	535	212	668	31	911	0.233	0.034	448	1.033
TSLA	7/20/2022	568	107	653	47	807	0.133	0.058	335	1.409
TSLA	10/20/2022	406	89	525	28	642	0.139	0.044	314	1.045
Firms are DIS = Disney, NKE = Nike, TGT = Target, and TSLA = Tesla										

Table 12 Events for tweets from users with the bottom 30% of the number of followers

Using the BottomUp dataset, iteration 15 (69 events identified)

		Counts					Ratios		Event Identification	
Firm	Date	Posting Users	Positive	Neutral	Negative	Total	Positive	Negative	Baseline	Outlier Fraction
DIS	6/1/2017	483	233	446	27	706	0.33	0.038	289	1.443
DIS	12/14/2017	567	196	453	34	683	0.287	0.05	226	2.022
DIS	11/12/2019	718	347	509	54	910	0.381	0.059	369	1.466
DIS	12/11/2020	465	237	295	30	562	0.422	0.053	203	1.768
DIS	11/21/2022	707	327	418	64	809	0.404	0.079	342	1.365
NKE	3/28/2017	603	24	702	4	730	0.033	0.005	235	2.106
NKE	3/27/2018	297	215	303	1	519	0.414	0.002	236	1.199
NKE	9/4/2018	3,362	713	2,192	1,011	3,916	0.182	0.258	688	4.692
NKE	2/21/2019	583	209	421	91	721	0.29	0.126	329	1.191
NKE	7/2/2019	665	192	319	238	749	0.256	0.318	335	1.236
NKE	12/5/2021	339	75	276	4	355	0.211	0.011	117	2.034
TGT	1/10/2017	92	5	192	2	199	0.025	0.01	29	5.862
TGT	2/13/2017	155	40	102	29	171	0.234	0.17	40	3.275
TGT	2/28/2017	37	16	73	16	105	0.152	0.152	34	2.088
TGT	3/21/2017	96	8	115	1	124	0.065	0.008	40	2.1
TGT	5/15/2017	74	60	15	0	75	0.8	0	29	1.586
TGT	8/29/2017	38	24	86	0	110	0.218	0	49	1.245
TGT	9/12/2017	55	10	38	23	71	0.141	0.324	32	1.219
TGT	5/18/2019	51	23	30	7	60	0.383	0.117	29	1.069
TGT	6/15/2019	181	38	97	61	196	0.194	0.311	37	4.297
TGT	3/13/2020	70	17	39	16	72	0.236	0.222	35	1.057
TGT	5/28/2020	351	23	280	157	460	0.05	0.341	47	8.787
TGT	7/6/2020	57	13	34	15	62	0.21	0.242	30	1.067
TGT	9/16/2020	65	9	57	16	82	0.11	0.195	30	1.733
TGT	11/12/2020	192	45	148	55	248	0.181	0.222	31	7
TGT	12/21/2020	49	21	35	10	66	0.318	0.152	30	1.2
TGT	12/29/2020	45	23	24	4	51	0.451	0.078	22	1.318
TGT	1/6/2021	37	6	68	8	82	0.073	0.098	22	2.727
TGT	3/26/2021	146	38	87	48	173	0.22	0.277	32	4.406
TGT	5/25/2021	59	46	27	5	78	0.59	0.064	25	2.12
TGT	6/21/2021	69	33	49	0	82	0.402	0	26	2.154
TGT	5/18/2022	53	12	44	10	66	0.182	0.152	29	1.276
TGT	9/20/2022	105	23	95	4	122	0.189	0.033	40	2.05
TSLA	1/4/2017	63	22	35	15	72	0.306	0.208	24	2
TSLA	2/23/2017	41	5	41	1	47	0.106	0.021	21	1.238
TSLA	4/4/2017	62	19	50	0	69	0.275	0	32	1.156
TSLA	4/11/2017	64	25	65	0	90	0.278	0	37	1.432
TSLA	4/19/2017	40	15	78	2	95	0.158	0.021	36	1.639
TSLA	4/26/2017	31	11	67	5	83	0.133	0.06	37	1.243
TSLA	5/3/2017	34	14	69	4	87	0.161	0.046	36	1.417
TSLA	5/10/2017	41	20	70	3	93	0.215	0.032	33	1.818
TSLA	7/29/2017	111	47	84	2	133	0.353	0.015	49	1.714
TSLA	11/17/2017	407	169	287	18	474	0.357	0.038	63	6.524
TSLA	2/6/2018	327	168	186	3	357	0.471	0.008	108	2.306

TSLA	2/21/2018	121	13	105	6	124	0.105	0.048	51	1.431
TSLA	8/2/2018	79	27	111	7	145	0.186	0.048	72	1.014
TSLA	9/7/2018	150	40	154	16	210	0.19	0.076	92	1.283
TSLA	10/25/2018	92	49	76	4	129	0.38	0.031	53	1.434
TSLA	3/15/2019	198	79	139	10	228	0.346	0.044	95	1.4
TSLA	6/11/2019	71	23	260	5	288	0.08	0.017	119	1.42
TSLA	11/22/2019	821	244	644	71	959	0.254	0.074	188	4.101
TSLA	5/12/2020	185	33	160	19	212	0.156	0.09	102	1.078
TSLA	7/14/2020	748	357	478	8	843	0.423	0.009	142	4.937
TSLA	9/22/2020	232	115	184	15	314	0.366	0.048	123	1.553
TSLA	1/7/2021	382	255	173	2	430	0.593	0.005	162	1.654
TSLA	1/28/2021	298	109	241	18	368	0.296	0.049	180	1.044
TSLA	2/8/2021	763	220	613	25	858	0.256	0.029	231	2.714
TSLA	3/24/2021	1,748	1,417	459	11	1,887	0.751	0.006	165	10.44
TSLA	4/20/2021	411	116	328	19	463	0.251	0.041	169	1.74
TSLA	5/11/2021	533	188	417	16	621	0.303	0.026	270	1.3
TSLA	6/4/2021	360	50	282	91	423	0.118	0.215	178	1.376
TSLA	6/14/2021	166	559	586	3	1,148	0.487	0.003	153	6.503
TSLA	6/28/2021	354	277	112	10	399	0.694	0.025	156	1.558
TSLA	8/20/2021	305	112	252	15	379	0.296	0.04	121	2.132
TSLA	10/25/2021	326	128	248	13	389	0.329	0.033	165	1.358
TSLA	11/29/2021	516	254	295	8	557	0.456	0.014	203	1.744
TSLA	1/14/2022	281	81	237	12	330	0.245	0.036	116	1.845
TSLA	3/22/2022	133	406	209	7	622	0.653	0.011	221	1.814
TSLA	7/21/2022	341	90	336	21	447	0.201	0.047	186	1.403
Firms are DIS = Disney, NKE = Nike, TGT = Target, and TSLA = Tesla										

Appendix F – Detailed indicators of significance for iterations

Significant iterations - UntreatedTopDown

		p-value	
Iteration	Percentile	Positive	Non-Positive
2	Top 4%		
3	Top 6%		
4	Top 8%		
5	Top 10%		
6	Top 12%		
7	Top 14%		
8	Top 16%		
9	Top 18%		
10	Top 20%		
11	Top 22%		
12	Top 24%		
13	Top 26%		
14	Top 28%		
15	Top 30%		
16	Top 32%		
17	Top 34%		
18	Top 36%		
19	Top 38%		
20	Top 40%		
21	Top 42%		
22	Top 44%		
23	Top 46%		
24	Top 48%		
25	Top 50%		
26	Top 52%		
27	Top 54%		
28	Top 56%		
29	Top 58%		
30	Top 60%		
31	Top 62%		
32	Top 64%		
33	Top 66%		
34	Top 68%		
35	Top 70%		
36	Top 72%		
37	Top 74%		
38	Top 76%		
39	Top 78%		
40	Top 80%	*	
41	Top 82%		
42	Top 84%	*	
43	Top 86%	*	
44	Top 88%	*	
45	Top 90%	*	
46	Top 92%	*	
47	Top 94%		
48	Top 96%	*	
49	Top 98%	*	
50	Top 100%	*	

Significant iterations - UntreatedBottomUp			
		p-value	
Iteration	Percentile	Positive	Non-Positive
1	Bottom 2%		
2	Bottom 4%		
3	Bottom 6%		
4	Bottom 8%		
5	Bottom 10%		
6	Bottom 12%		
7	Bottom 14%		
8	Bottom 16%		
9	Bottom 18%		
10	Bottom 20%		
11	Bottom 22%		
12	Bottom 24%		
13	Bottom 26%		
14	Bottom 28%	*	
15	Bottom 30%	*	
16	Bottom 32%	*	
17	Bottom 34%		
18	Bottom 36%	*	
19	Bottom 38%	*	
20	Bottom 40%	*	
21	Bottom 42%	*	
22	Bottom 44%	*	
23	Bottom 46%	*	
24	Bottom 48%		
25	Bottom 50%	*	
26	Bottom 52%	*	
27	Bottom 54%	*	*
28	Bottom 56%	*	*
29	Bottom 58%	*	*
30	Bottom 60%	*	*
31	Bottom 62%		
32	Bottom 64%	*	
33	Bottom 66%	*	
34	Bottom 68%	*	
35	Bottom 70%	*	
36	Bottom 72%	*	
37	Bottom 74%	*	
38	Bottom 76%	*	
39	Bottom 78%	*	
40	Bottom 80%	*	
41	Bottom 82%	*	
42	Bottom 84%	*	
43	Bottom 86%	*	
44	Bottom 88%	*	
45	Bottom 90%	*	
46	Bottom 92%	*	
47	Bottom 94%	*	
48	Bottom 96%	*	
49	Bottom 98%	*	
50	Bottom 100%	*	

Appendix G – Table of Completeness

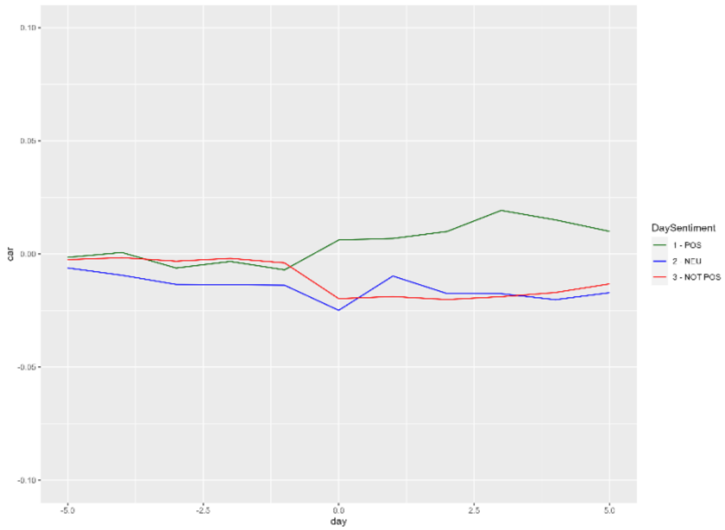
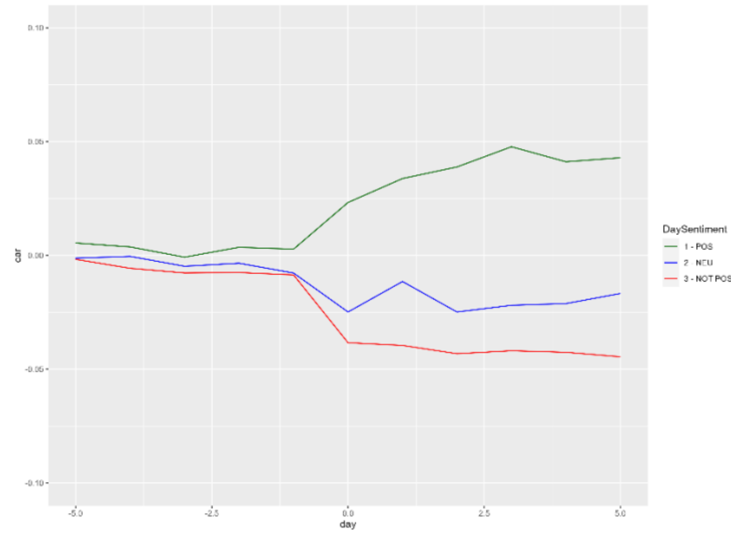
Completeness of Tweets with User Profiles

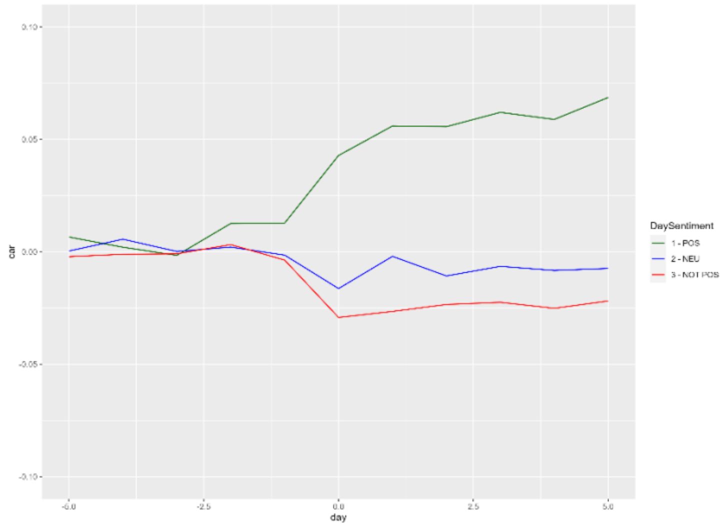
6,780,072 total tweets collected between Jan 01, 2017 thru Dec 31, 2022

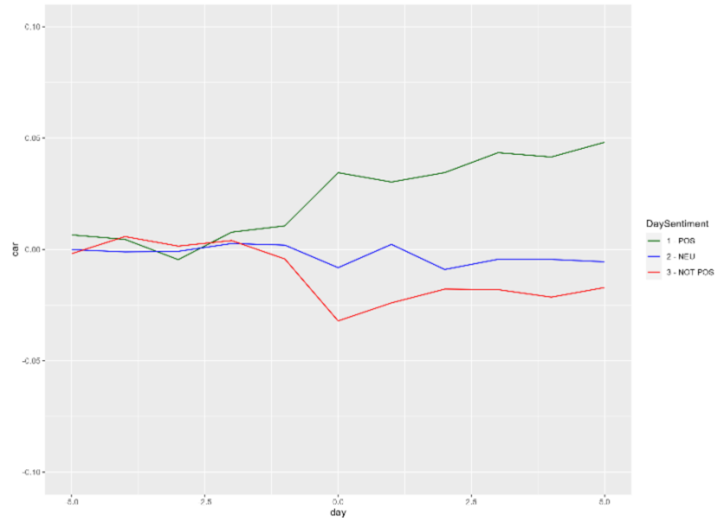
Year	Total	Complete	Incomplete	% Incomplete
2017	1,239,392	1,230,350	9,042	99.27%
2018	1,225,819	1,218,757	7,062	99.42%
2019	1,222,121	1,214,849	7,272	99.40%
2020	1,077,105	1,070,266	6,839	99.37%
2021	897,372	891,441	5,931	99.34%
2022	1,118,263	1,110,210	8,053	99.28%

Number of tweets collected from Twitter (X) API between December 2022 and January 2023

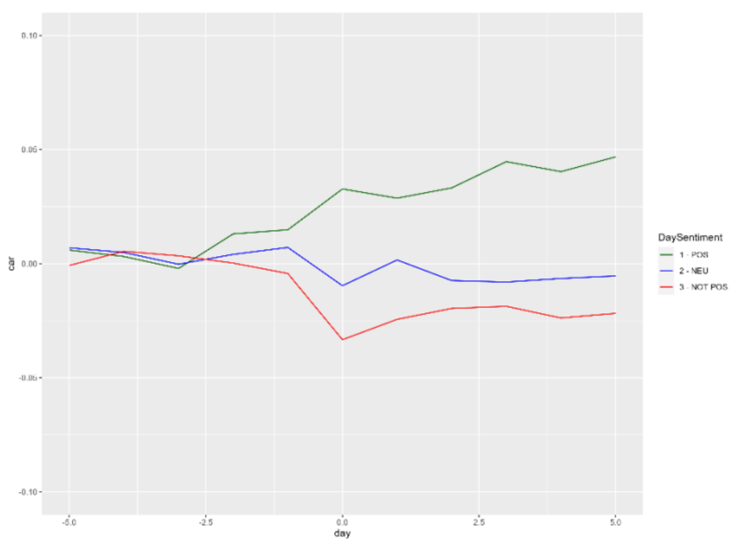
Appendix H – Select CAR Plots (every five iterations)

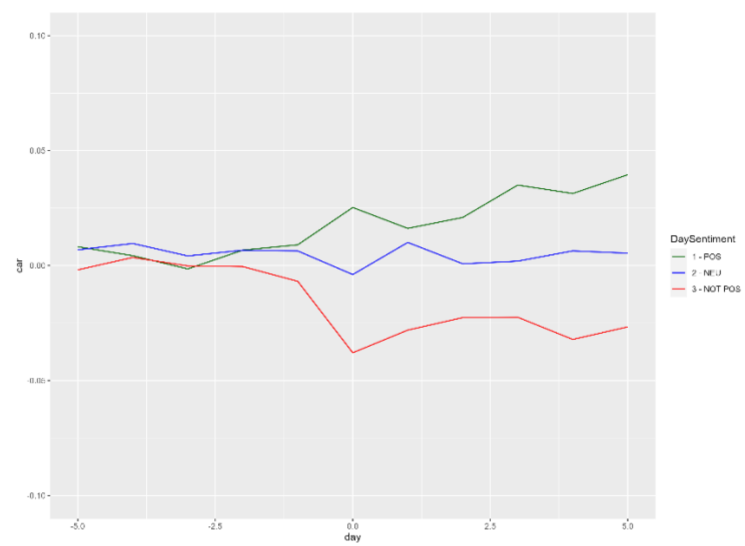
TopDown – Iteration 5 – Top 10%	TopDown – Iteration 10 – Top 20%
	
TopDown – Iteration 15 – Top 30%	**TopDown – Iteration 20 – Top 40%**

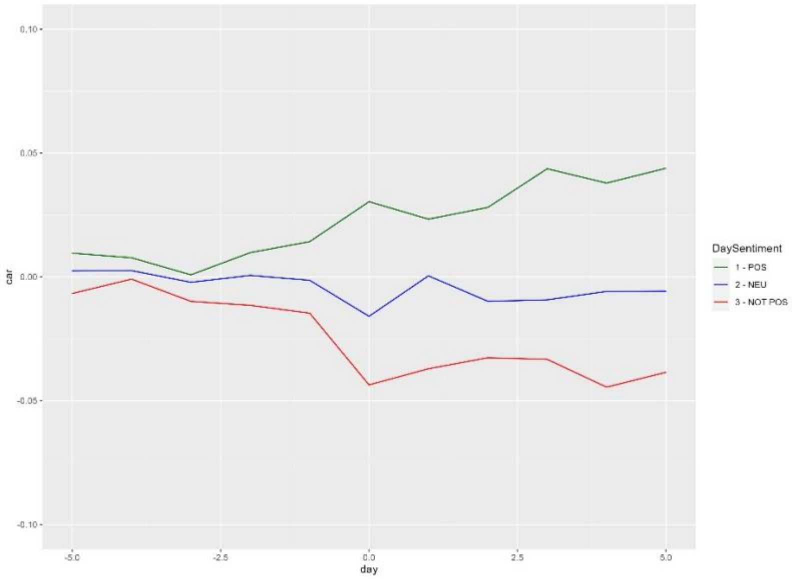
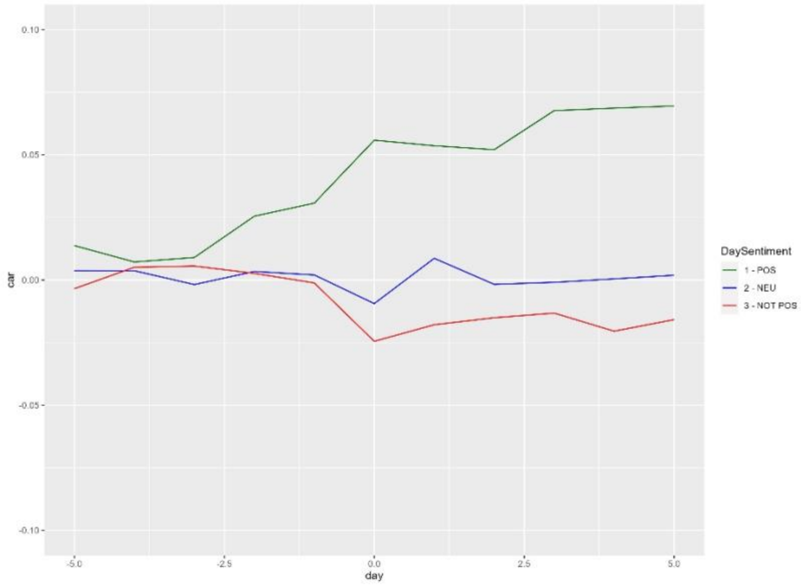


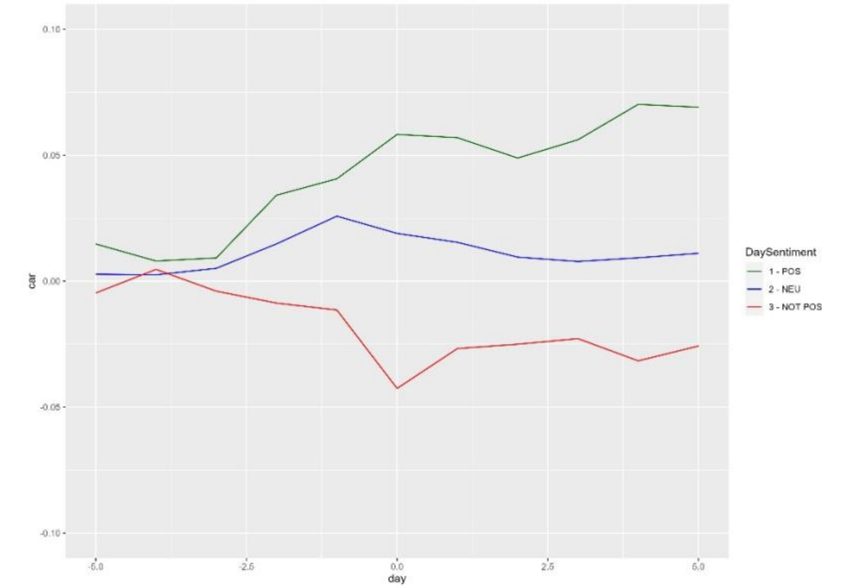


TopDown – Iteration 25 – Top 50%	**TopDown – Iteration 30 – Top 60%**

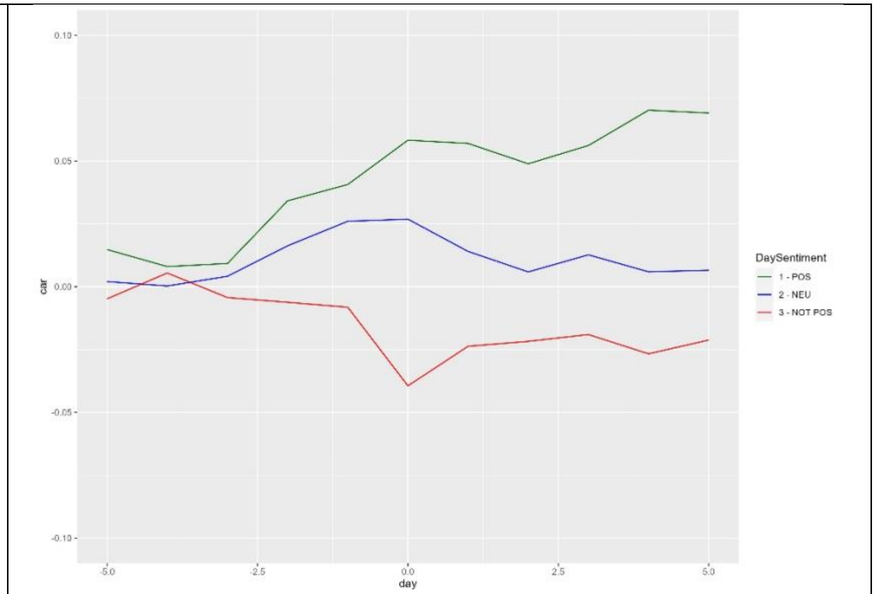




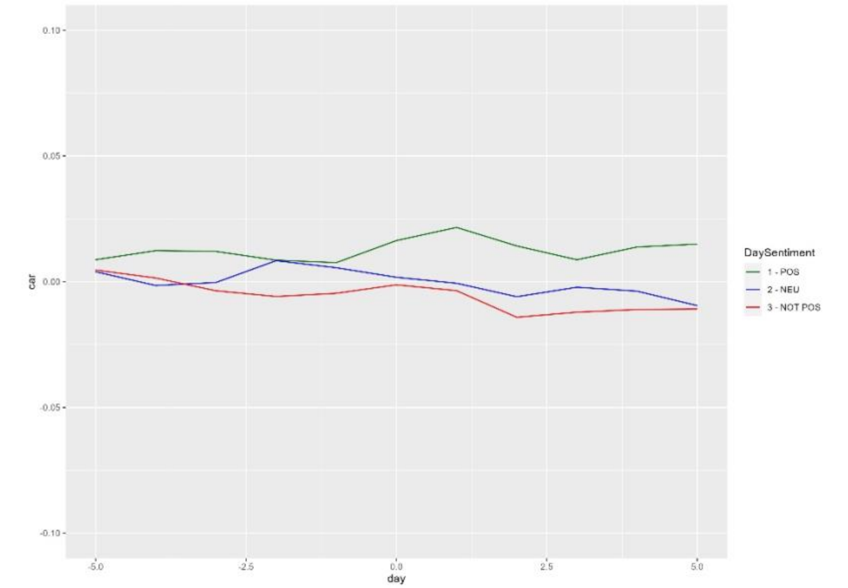
TopDown – Iteration 35 – Top 70%	TopDown – Iteration 40 – Top 80%
	
TopDown – Iteration 45 – Top 90%	**TopDown – Iteration 50 – Top 100%**



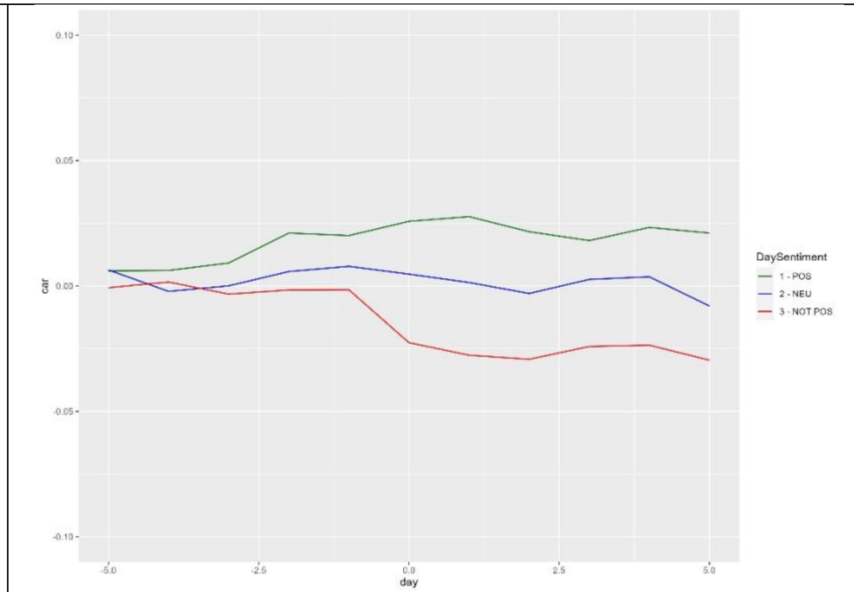
BottomUp – Iteration 5 – Bottom 10%



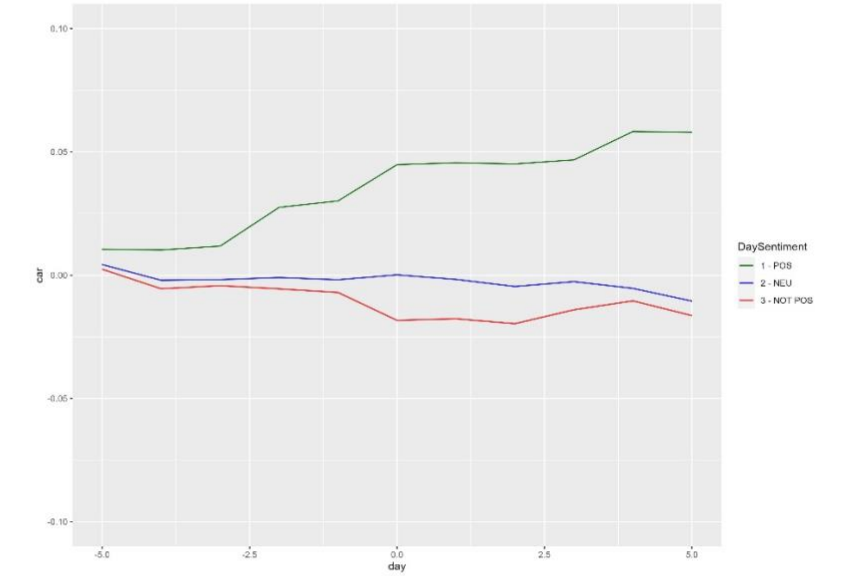
BottomUp – Iteration 10 – Bottom 20%



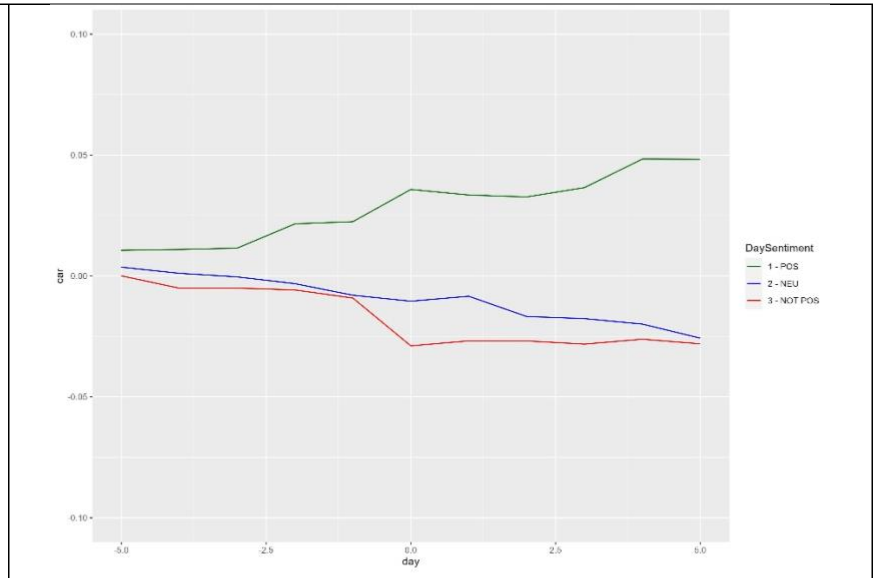
BottomUp – Iteration 15 – Bottom 30%



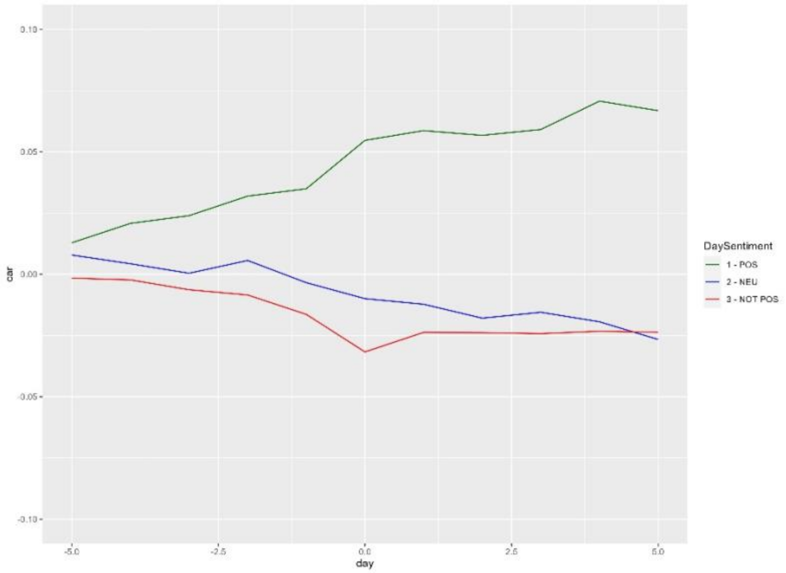
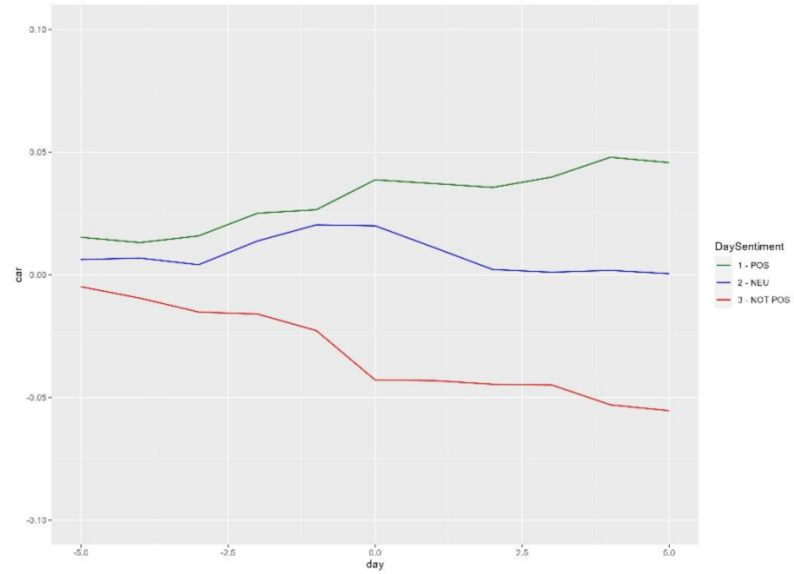
BottomUp – Iteration 20 – Bottom 40%

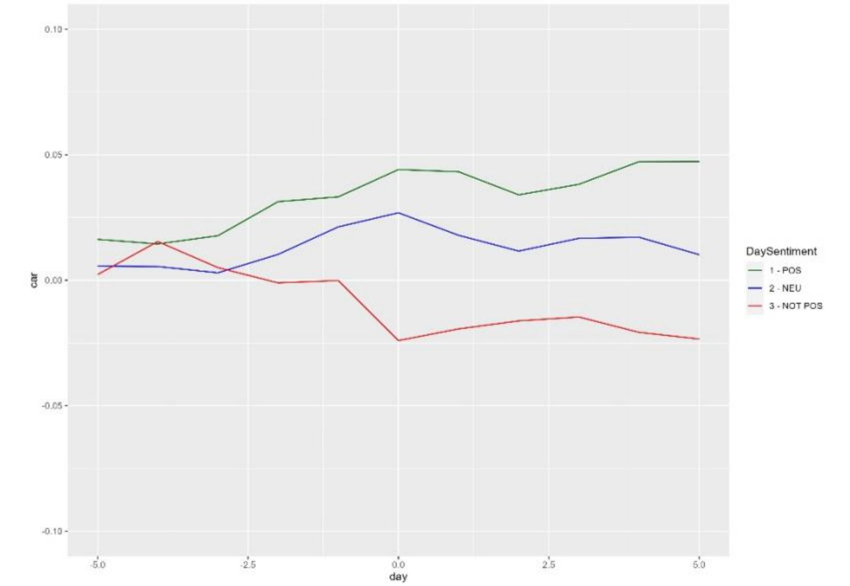


BottomUp – Iteration 25 – Bottom 50%

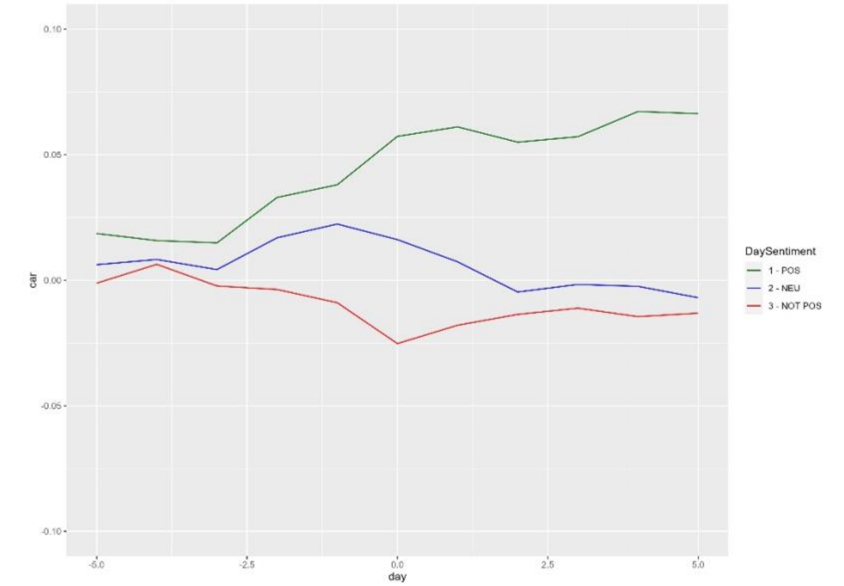


BottomUp – Iteration 30 – Bottom 60%

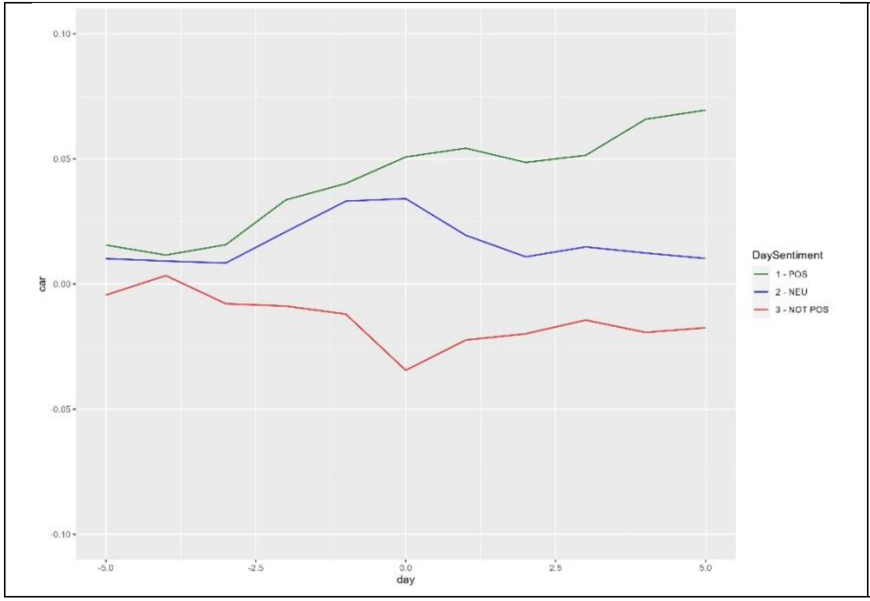
	
BottomUp – Iteration 35 – Bottom 70%	**BottomUp – Iteration 40 – Bottom 80%**

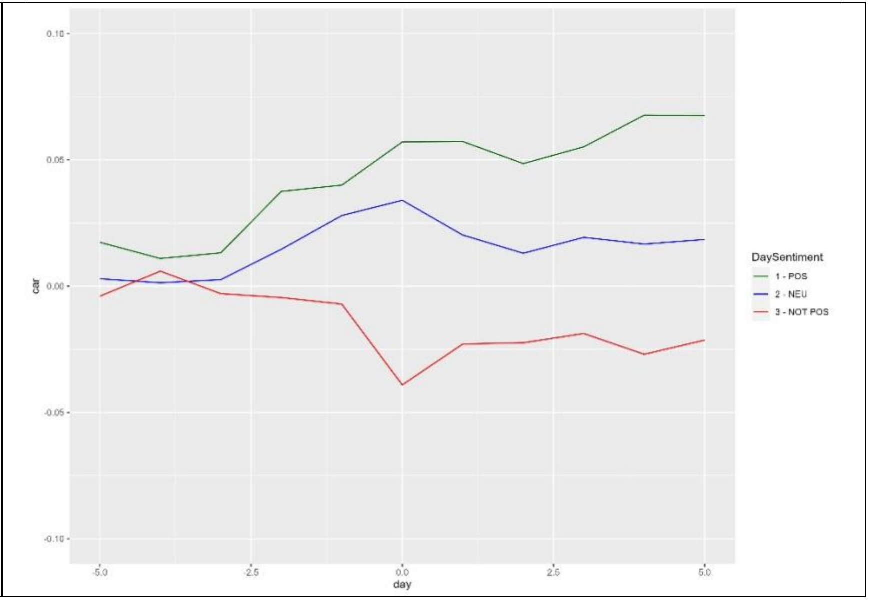


BottomUp – Iteration 45 – Bottom 90%



BottomUp – Iteration 50 – Bottom 100%


0.10
0.05
0.00
-0.05
-0.10
car
-5.0
-2.5
0.0
2.5
5.0
day
DaySentiment
1 - POS
2 - NEU
3 - NOT POS


0.10
0.05
0.00
-0.05
-0.10
car
-5.0
-2.5
0.0
2.5
5.0
day
DaySentiment
1 - POS
2 - NEU
3 - NOT POS

Appendix I – Select Statistical test results (every five iterations)

Statistical Tests - TopDown

Iteration 5 - Top 10 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	-0.0015	0.0498	0.6375	-0.0025	0.5682	0.4842
-4	7e-04	0.8754	0.8593	-0.0016	0.9001	0.6907
-3	-0.0062	0.4000	0.2550	-0.0032	0.8544	0.5908
-2	-0.0033	0.7468	0.6532	-0.0020	0.9154	0.7720
-1	-0.0070	0.6247	0.3711	-0.0039	0.7942	0.6119
0	0.0062	0.9339	0.6527	-0.0197	0.2023	0.0914
1	0.0069	0.9632	0.6764	-0.0188	0.2167	0.1408
2	0.0100	0.8320	0.6171	-0.0201	0.1302	0.1076
3	0.0193	0.6117	0.3625	-0.0189	0.2724	0.1473
4	0.0151	0.6117	0.4199	-0.0170	0.3992	0.2371
5	0.0100	0.8034	0.6242	-0.0132	0.5948	0.3352

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 10 - Top 20 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0055	0.6431	0.1563	-0.0017	0.8996	0.6399
-4	0.0037	0.5088	0.4428	-0.0056	0.2768	0.2878
-3	-8e-04	0.9218	0.8880	-0.0077	0.2405	0.2463
-2	0.0036	0.3902	0.6055	-0.0075	0.3165	0.3078
-1	0.0028	0.5088	0.7409	-0.0085	0.3305	0.3275
0	0.0233	0.3165	0.2223	-0.0383 *	0.0164	0.0156
1	0.0338	0.1515	0.1325	-0.0395 *	0.0211	0.0202
2	0.0389	0.1875	0.1494	-0.0432 *	0.0079	0.0130
3	0.0478	0.1780	0.1029	-0.0419 *	0.0138	0.0193
4	0.0412	0.1355	0.0996	-0.0425 *	0.0229	0.0270
5	0.0429	0.1688	0.1197	-0.0445 *	0.0115	0.0154

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 15 - Top 30 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0066	0.3254	0.0845	-0.0022	0.7915	0.6870
-4	0.0021	0.8532	0.6605	-0.0011	0.4578	0.8722
-3	-0.0016	0.9578	0.7757	-9e-04	0.4578	0.9136
-2	0.0126	0.1485	0.1585	0.0032	0.8119	0.7189
-1	0.0128	0.2635	0.2605	-0.0037	0.4742	0.6972
0	0.0429 *	0.0588	0.0494	-0.0292	0.0203	0.0536
1	0.0560 *	0.0667	0.0429	-0.0265	0.0755	0.1065
2	0.0557	0.1199	0.0739	-0.0234	0.0802	0.1844
3	0.0620 *	0.1073	0.0477	-0.0224	0.0851	0.2338
4	0.0589 *	0.0755	0.0367	-0.0251	0.1409	0.2007
5	0.0686 *	0.0551	0.0278	-0.0218	0.2200	0.2449

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 20 - Top 40 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0065	0.4100	0.1083	-0.0019	0.9406	0.7433
-4	0.0045	0.3765	0.3612	0.0058	0.8932	0.4648
-3	-0.0046	0.6221	0.3982	0.0015	0.7540	0.8650
-2	0.0077	0.4100	0.3887	0.0040	0.7768	0.6818
-1	0.0106	0.4634	0.3809	-0.0043	0.4820	0.6719
0	0.0345	0.1793	0.1340	-0.0321 *	0.0123	0.0445
1	0.0303	0.3604	0.2477	-0.0240	0.1511	0.1740
2	0.0345	0.4100	0.2631	-0.0178	0.2002	0.3477
3	0.0434	0.3765	0.1738	-0.0180	0.2345	0.3746
4	0.0415	0.2726	0.1413	-0.0214	0.2002	0.3048
5	0.0481	0.1793	0.1079	-0.0171	0.3604	0.3900

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 25 - Top 50 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0059	0.4826	0.1869	-8e-04	0.8917	0.8948
-4	0.0031	0.6327	0.5393	0.0054	0.7593	0.5691
-3	-0.0021	1.0000	0.6902	0.0034	0.7593	0.7486
-2	0.0130	0.1207	0.1519	3e-04	0.9457	0.9826
-1	0.0148	0.2902	0.2280	-0.0043	0.5621	0.7174
0	0.0327	0.0794	0.1375	-0.0333	0.0263	0.0577
1	0.0287	0.2756	0.2622	-0.0244	0.3038	0.2187
2	0.0332	0.3535	0.2807	-0.0196	0.3205	0.3655
3	0.0447	0.2348	0.1572	-0.0187	0.4120	0.4247
4	0.0403	0.2348	0.1450	-0.0238	0.3554	0.3337
5	0.0468	0.1762	0.1090	-0.0218	0.4733	0.3519

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 30 - Top 60 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0082	0.2455	0.0747	-0.0019	0.6507	0.7695
-4	0.0043	0.4524	0.4396	0.0035	0.9530	0.7382
-3	-0.0015	0.8695	0.7995	-2e-04	0.7983	0.9902
-2	0.0067	0.2611	0.3605	-4e-04	0.8906	0.9736
-1	0.0090	0.3118	0.3700	-0.0068	0.3955	0.5954
0	0.0253	0.1140	0.2383	-0.0379 *	0.0204	0.0485
1	0.0162	0.4091	0.4876	-0.0281	0.2935	0.2017
2	0.0209	0.5217	0.4739	-0.0226	0.3124	0.3466
3	0.0350	0.3683	0.2683	-0.0225	0.3321	0.3866
4	0.0313	0.2943	0.2377	-0.0321	0.2413	0.2410
5	0.0395	0.2024	0.1632	-0.0267	0.3321	0.3017

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 35 - Top 70 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0096	0.1564	0.0549	-0.0067	0.4171	0.1944
-4	0.0077	0.2101	0.1827	-9e-04	0.8650	0.9201
-3	8e-04	0.3525	0.8793	-0.0099	0.6397	0.3855
-2	0.0098	0.0799	0.1577	-0.0115	0.5509	0.3796
-1	0.0142	0.0955	0.1486	-0.0146	0.2462	0.2143
0	0.0304	0.0446	0.1392	-0.0436 *	0.0077	0.0156
1	0.0233	0.2579	0.3081	-0.0371	0.1187	0.0822
2	0.0281	0.3736	0.3427	-0.0326	0.1964	0.1542
3	0.0437	0.1956	0.1751	-0.0333	0.1540	0.1783
4	0.0379	0.2101	0.1592	-0.0445	0.1297	0.0983
5	0.0439	0.1447	0.1350	-0.0385	0.1674	0.1251

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 40 - Top 80 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0137 **	0.0104	0.0062	-0.0035	0.8317	0.4829
-4	0.0072	0.2645	0.2027	0.0051	0.7019	0.5194
-3	0.0090	0.1540	0.1958	0.0056	0.7987	0.7086
-2	0.0255 **	0.0016	0.0011	0.0027	0.8986	0.8692
-1	0.0307 **	0.0023	0.0017	-0.0012	0.7019	0.9349
0	0.0558 **	0.0010	0.0026	-0.0245	0.0737	0.3071
1	0.0536 *	0.0237	0.0205	-0.0179	0.4171	0.4809
2	0.0521	0.0737	0.0685	-0.0151	0.5798	0.5475
3	0.0676 *	0.0385	0.0351	-0.0132	0.4951	0.6240
4	0.0686 *	0.0237	0.0129	-0.0204	0.4423	0.4709
5	0.0695 *	0.0208	0.0195	-0.0158	0.5226	0.5309

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 45 - Top 90 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0148 **	0.0034	0.0030	-0.0047	0.6095	0.3254
-4	0.0080	0.2288	0.1598	0.0047	0.6705	0.5418
-3	0.0092	0.0987	0.1787	-0.0040	0.8986	0.6954
-2	0.0341 ***	0.0003	0.0010	-0.0087	0.7987	0.4953
-1	0.0406 **	0.0016	0.0024	-0.0114	0.3927	0.3052
0	0.0583 **	0.0019	0.0054	-0.0426 *	0.0159	0.0216
1	0.0570 *	0.0432	0.0369	-0.0267	0.3465	0.1925
2	0.0489	0.0987	0.0936	-0.0250	0.3465	0.2536
3	0.0561 *	0.0539	0.0361	-0.0228	0.3927	0.3275
4	0.0702 *	0.0268	0.0159	-0.0316	0.2462	0.1928
5	0.0691 *	0.0304	0.0253	-0.0257	0.3465	0.2456

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - TopDown

Iteration 50 - Top 100 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0148 **	0.0034	0.0030	-0.0048	0.5798	0.3105
-4	0.0080	0.2288	0.1598	0.0055	0.5798	0.4716
-3	0.0092	0.0987	0.1787	-0.0043	0.9323	0.6692
-2	0.0341 ***	0.0003	0.0010	-0.0062	0.9661	0.6182
-1	0.0406 **	0.0016	0.0024	-0.0082	0.5798	0.4446
0	0.0583 **	0.0019	0.0054	-0.0394 *	0.0304	0.0327
1	0.0570 *	0.0432	0.0369	-0.0237	0.5226	0.2467
2	0.0489	0.0987	0.0936	-0.0217	0.4951	0.3193
3	0.0561 *	0.0539	0.0361	-0.0190	0.5226	0.4115
4	0.0702 *	0.0268	0.0159	-0.0267	0.3927	0.2666
5	0.0691 *	0.0304	0.0253	-0.0212	0.5226	0.3328

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 5 - Bottom 10 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0087 *	0.0774	0.0464	0.0046	0.2516	0.3028
-4	0.0124 *	0.0439	0.0338	0.0015	0.5876	0.8341
-3	0.0121 *	0.0336	0.0264	-0.0036	0.2071	0.6430
-2	0.0086	0.2486	0.1886	-0.0059	0.2516	0.3972
-1	0.0076	0.3160	0.3264	-0.0046	0.2947	0.6060
0	0.0164	0.1866	0.0863	-0.0012	0.5988	0.9026
1	0.0216	0.1005	0.0507	-0.0036	0.6799	0.7129
2	0.0142	0.1815	0.2468	-0.0141	0.1310	0.1795
3	0.0087	0.3617	0.5007	-0.0121	0.3180	0.3093
4	0.0138	0.3160	0.3207	-0.0111	0.3341	0.4037
5	0.0150	0.2364	0.3015	-0.0109	0.3945	0.3933

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 10 - Bottom 20 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0060	0.1252	0.0691	-7e-04	0.9923	0.8629
-4	0.0062	0.4102	0.1627	0.0016	0.7646	0.8410
-3	0.0092	0.0942	0.0640	-0.0032	0.2639	0.7114
-2	0.0211 **	0.0048	0.0046	-0.0016	0.7352	0.8503
-1	0.0201 *	0.0187	0.0265	-0.0015	0.4102	0.8998
0	0.0258 *	0.0233	0.0288	-0.0226	0.0456	0.1474
1	0.0277	0.0793	0.0701	-0.0276	0.0414	0.0879
2	0.0216	0.1302	0.1745	-0.0292	0.0374	0.0899
3	0.0181	0.2395	0.2393	-0.0241	0.1023	0.1786
4	0.0233	0.1203	0.1409	-0.0236	0.1757	0.2197
5	0.0211	0.2167	0.1830	-0.0296	0.1066	0.1217

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 15 - Bottom 30 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0105 *	0.0211	0.0190	0.0024	0.6010	0.5433
-4	0.0102	0.0646	0.0541	-0.0055	0.5399	0.3873
-3	0.0118 *	0.0211	0.0328	-0.0043	0.2467	0.5675
-2	0.0275 **	0.0014	0.0019	-0.0055	0.4274	0.4651
-1	0.0301 **	0.0014	0.0031	-0.0070	0.2226	0.4436
0	0.0448 **	0.0022	0.0042	-0.0183	0.1511	0.3077
1	0.0456 *	0.0315	0.0246	-0.0177	0.3001	0.3521
2	0.0451 *	0.0179	0.0287	-0.0196	0.2726	0.3065
3	0.0468 *	0.0164	0.0186	-0.0140	0.3765	0.5275
4	0.0583 **	0.0079	0.0071	-0.0104	0.6010	0.6589
5	0.0580 *	0.0138	0.0132	-0.0164	0.4274	0.4583

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 20 - Bottom 40 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0106 *	0.0384	0.0238	1e-04	0.9168	0.9883
-4	0.0110 *	0.0522	0.0469	-0.0051	0.3931	0.4327
-3	0.0115 *	0.0415	0.0477	-0.0050	0.2226	0.5372
-2	0.0216 **	0.0060	0.0043	-0.0058	0.3447	0.4968
-1	0.0225 **	0.0083	0.0043	-0.0091	0.1424	0.3369
0	0.0358 **	0.0054	0.0074	-0.0289	0.0918	0.0994
1	0.0335 *	0.0563	0.0415	-0.0268	0.2345	0.1664
2	0.0327 *	0.0254	0.0392	-0.0269	0.1511	0.1827
3	0.0365 *	0.0214	0.0267	-0.0282	0.1262	0.1901
4	0.0483 *	0.0135	0.0105	-0.0262	0.2861	0.2746
5	0.0482 *	0.0233	0.0212	-0.0281	0.2226	0.2169

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 25 - Bottom 50 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0129 *	0.0532	0.0396	-0.0016	0.9217	0.6990
-4	0.0208	0.0441	0.0797	-0.0023	0.9843	0.7258
-3	0.0239	0.0153	0.0565	-0.0063	0.4653	0.3791
-2	0.0319 *	0.0073	0.0134	-0.0084	0.3321	0.3172
-1	0.0349 *	0.0073	0.0196	-0.0164 *	0.0446	0.0389
0	0.0547 **	0.0017	0.0063	-0.0317	0.0204	0.0503
1	0.0587 *	0.0121	0.0107	-0.0237	0.2253	0.2174
2	0.0567 *	0.0083	0.0135	-0.0238	0.1956	0.2447
3	0.0592 **	0.0064	0.0095	-0.0242	0.2413	0.2707
4	0.0707 **	0.0037	0.0046	-0.0233	0.5412	0.3032
5	0.0668 **	0.0083	0.0100	-0.0237	0.3955	0.2673

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 30 - Bottom 60 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0153 **	0.0101	0.0081	-0.0049	0.5621	0.2522
-4	0.0131 *	0.0351	0.0373	-0.0095	0.3377	0.1909
-3	0.0159 *	0.0239	0.0274	-0.0152	0.1571	0.1085
-2	0.0251 **	0.0043	0.0026	-0.0160	0.1470	0.1139
-1	0.0265 **	0.0043	0.0022	-0.0227 *	0.0195	0.0120
0	0.0387 **	0.0071	0.0083	-0.0429 **	0.0033	0.0081
1	0.0372 *	0.0547	0.0388	-0.0431 *	0.0142	0.0219
2	0.0356 *	0.0319	0.0408	-0.0446 *	0.0127	0.0287
3	0.0398 *	0.0263	0.0281	-0.0449 *	0.0239	0.0396
4	0.0479 *	0.0263	0.0236	-0.0530 *	0.0290	0.0255
5	0.0457 *	0.0547	0.0485	-0.0554 *	0.0175	0.0155

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 35 - Bottom 70 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0163 **	0.0043	0.0044	0.0023	0.4524	0.6844
-4	0.0145 *	0.0290	0.0218	0.0154	0.2774	0.1641
-3	0.0177 **	0.0049	0.0098	0.0050	0.8124	0.6860
-2	0.0313 ***	0.0001	0.0001	-0.0011	0.8983	0.9329
-1	0.0332 ***	0.0002	0.0001	-1e-04	0.4749	0.9937
0	0.0441 **	0.0010	0.0022	-0.0240	0.0897	0.2416
1	0.0432 *	0.0175	0.0160	-0.0194	0.4304	0.3703
2	0.0340	0.0502	0.0554	-0.0162	0.4304	0.4777
3	0.0382 *	0.0460	0.0369	-0.0146	0.5459	0.5462
4	0.0473 *	0.0319	0.0210	-0.0208	0.5217	0.3908
5	0.0473 *	0.0351	0.0349	-0.0234	0.4304	0.3067

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 40 - Bottom 80 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0186 ***	0.0010	0.0008	-0.0012	0.9530	0.8102
-4	0.0158 *	0.0215	0.0112	0.0063	0.5678	0.3941
-3	0.0149 *	0.0215	0.0274	-0.0023	0.9843	0.8131
-2	0.0330 ***	0.0003	0.0009	-0.0037	1.0000	0.7458
-1	0.0380 **	0.0004	0.0013	-0.0089	0.4413	0.3935
0	0.0573 **	0.0006	0.0020	-0.0252	0.1447	0.1570
1	0.0611 *	0.0083	0.0109	-0.0179	0.7086	0.3721
2	0.0550 *	0.0172	0.0260	-0.0136	0.7381	0.5267
3	0.0572 *	0.0136	0.0157	-0.0111	0.9217	0.6261
4	0.0672 **	0.0107	0.0083	-0.0145	0.8596	0.5439
5	0.0663 *	0.0136	0.0151	-0.0131	0.8288	0.5535

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 45 - Bottom 90 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0156 **	0.0160	0.0089	-0.0044	0.5678	0.2852
-4	0.0115	0.1232	0.0789	0.0034	0.9530	0.6361
-3	0.0157 *	0.0446	0.0377	-0.0079	0.4900	0.4035
-2	0.0336 **	0.0017	0.0020	-0.0088	0.6794	0.4415
-1	0.0401 **	0.0014	0.0020	-0.0120	0.2579	0.2019
0	0.0507 *	0.0062	0.0141	-0.0345 *	0.0446	0.0478
1	0.0542 *	0.0494	0.0392	-0.0224	0.5153	0.2482
2	0.0486	0.0663	0.0680	-0.0199	0.5153	0.3410
3	0.0514 *	0.0546	0.0430	-0.0144	0.7381	0.5172
4	0.0658 *	0.0141	0.0124	-0.0193	0.6507	0.4068
5	0.0695 *	0.0141	0.0159	-0.0175	0.6794	0.4110

[1] *p<0.05; **p<0.01; ***p<0.001

Statistical Tests - BottomUp

Iteration 50 - Bottom 100 %

	Positive Events			Non-Positive Events		
		p-value			p-value	
Days	Mean CAR[1]	Wilcoxon	t-test	Mean CAR[1]	Wilcoxon	t-test
-5	0.0174 **	0.0034	0.0031	-0.0040	0.7660	0.4113
-4	0.0110	0.1674	0.1025	0.0059	0.4683	0.4327
-3	0.0132	0.0898	0.1058	-0.0030	0.7337	0.7699
-2	0.0375 ***	0.0002	0.0005	-0.0045	0.8650	0.7177
-1	0.0400 **	0.0016	0.0025	-0.0071	0.6705	0.5089
0	0.0571 **	0.0019	0.0059	-0.0391 *	0.0385	0.0344
1	0.0573 *	0.0385	0.0362	-0.0229	0.5509	0.2639
2	0.0485	0.0987	0.0956	-0.0224	0.4171	0.3038
3	0.0552 *	0.0539	0.0385	-0.0188	0.5798	0.4182
4	0.0677 *	0.0304	0.0185	-0.0270	0.3692	0.2612
5	0.0676 *	0.0304	0.0276	-0.0213	0.5226	0.3296

[1] *p<0.05; **p<0.01; ***p<0.001

REFERENCES CITED

Acar, A., & Muraki, Y. (2011). Twitter for crisis communication: Lessons learned from Japan's tsunami disaster. *International Journal of Web Based Communities*, *7*(3), 392–402. https://doi.org/10.1504/IJWBC.2011.041206

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. J. (2011). Sentiment analysis of twitter data. *Proceedings of the Workshop on Language in Social Media (LSM 2011)*, 30–38.

Ağrali, Ö., & Aydin, Ö. (2021). Tweet Classification and Sentiment Analysis on Metaverse Related Messages. *Journal of Metaverse*, *1*(1), Article 1.

Ajjoub, C., Walker, T., & Zhao, Y. (2021). Social media posts and stock returns: The Trump factor. *International Journal of Managerial Finance*, *17*(2), 185–213. https://doi.org/10.1108/IJMF-02-2020-0068

Armitage, S. (1995). Event Study Methods and Evidence on Their Performance. *Journal of Economic Surveys*, *9*(1), 25–52. https://doi.org/10.1111/j.1467-6419.1995.tb00109.x

Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, *1*, 492–499.

Bae, Y., & Lee, H. (2012). Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers. *Journal of the American Society for Information Science and Technology*, *63*(12), 2521–2535. https://doi.org/10.1002/asi.22768

Bar-Haim, R., Dinur, E., Feldman, R., Fresko, M., & Goldstein, G. (2011). Identifying and following expert investors in stock microblogs. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 1310–1319.

Bartov, E., Faurel, L., & Mohanram, P. S. (2018). Can Twitter Help Predict Firm-Level Earnings and Stock Returns? *The Accounting Review*, *93*(3), 25–57. https://doi.org/10.2308/accr-51865

Blair, J., Hsu, C.-Y., Qiu, L., Huang, S.-H., Huang, T.-H. K., & Abdullah, S. (2021). Using Tweets to Assess Mental Well-being of Essential Workers During the COVID-19 Pandemic. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3411763.3451612

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

Bouadjenek, M. R., Sanner, S., & Wu, G. (2023). A User-Centric Analysis of Social Media for Stock Market Prediction. *ACM TRANSACTIONS ON THE WEB*, *17*(2), 9. https://doi.org/10.1145/3532856

Brown, J., Broderick, A. J., & Lee, N. (2007). Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of Interactive Marketing*, *21*(3), 2–20.

Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of Financial Economics*, *8*(3), 205–258. https://doi.org/10.1016/0304-405X(80)90002-1

Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics*, *14*(1), 3–31. https://doi.org/10.1016/0304-405X(85)90042-X

Cakra, Y. E., & Distiawan Trisedya, B. (2015). Stock price prediction using linear regression based on sentiment analysis. *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 147–154. https://doi.org/10.1109/ICACSIS.2015.7415179

Calma, J. (2023, May 31). *Scientists say they can't rely on Twitter anymore*. The Verge. https://www.theverge.com/2023/5/31/23739084/twitter-elon-musk-api-policy-chilling-academic-research

Camacho-collados, J., Rezaee, K., Riahi, T., Ushio, A., Loureiro, D., Antypas, D., Boisson, J., Espinosa Anke, L., Liu, F., & Martínez Cámara, E. (2022). TweetNLP: Cutting-Edge Natural Language Processing for Social Media. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–49. https://doi.org/10.18653/v1/2022.emnlp-demos.5

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press. https://doi.org/10.2307/j.ctt7skm5

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. (2010). Measuring User Influence in Twitter: The Million Follower Fallacy. *Proceedings of the International AAAI Conference on Web and Social Media*, *4*(1), Article 1. https://doi.org/10.1609/icwsm.v4i1.14033

Chaudhary, J., & Niveditha, S. (2021). Twitter sentiment analysis using tweepy. *Int Res J EngTech*, *8*, 4512–4516.

Chen, H., De, P., Hu, Y. (Jeffrey), & Hwang, B.-H. (2014). Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media. *Review of Financial Studies*, *27*(5), 1367–1403. https://doi.org/10.1093/rfs/hhu001

Cookson, J. A., Fox, C., Gil-Bazo, J., Imbet, J. F., & Schiller, C. (2023). Social media as a bank run catalyst. *Available at SSRN 4422754*.

Coyne, S., Madiraju, P., & Coelho, J. (2017). Forecasting Stock Prices using Social Media Analysis. *2017 IEEE 15TH INTL CONF ON DEPENDABLE, AUTONOMIC AND SECURE COMPUTING, 15TH INTL CONF ON PERVASIVE INTELLIGENCE AND COMPUTING, 3RD INTL CONF ON BIG DATA INTELLIGENCE AND COMPUTING AND CYBER SCIENCE AND TECHNOLOGY CONGRESS(DASC/PICOM/DATACOM/CYBERSCI*, 1031–1038. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.169

Developers [@XDevelopers]. (2023a, February 2). *Starting February 9, we will no longer support free access to the Twitter API, both v2 and v1.1. A paid basic tier will be available instead* 🧵 [Tweet]. Twitter. https://twitter.com/XDevelopers/status/1621026986784337922

Developers [@XDevelopers]. (2023b, February 8). *Paid basic access that offers low level of API usage, and access to Ads API for a $100 monthly fee.* [Tweet]. Twitter. https://twitter.com/XDevelopers/status/1623467617037221889

DisTwitter—What is DisTwitter? (n.d.). [DisTwitter]. Slang.Net. Retrieved June 5, 2023, from https://slang.net/meaning/distwitter

Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, *68*, 1–7. https://doi.org/10.1016/j.chb.2016.11.009

Dong, H., & Gil-Bazo, J. (2020). Sentiment stocks. *International Review of Financial Analysis*, *72*, 101573. https://doi.org/10.1016/j.irfa.2020.101573

Donthu, N., & Gustafsson, A. (2020). Effects of COVID-19 on business and research. *Journal of Business Research*, *117*, 284–289. https://doi.org/10.1016/j.jbusres.2020.06.008

Dubbelink, S. I., Herrando, C., & Constantinides, E. (2021). Social Media Marketing as a Branding Strategy in Extraordinary Times: Lessons from the COVID-19 Pandemic. *Sustainability*, *13*(18), Article 18. https://doi.org/10.3390/su131810310

Eickhoff, M., & Muntermann, J. (2016). Stock analysts vs. the crowd: Mutual prediction and the drivers of crowd wisdom. *Information & Management*, *53*(7), 835–845. https://doi.org/10.1016/j.im.2016.03.008

Eierle, B., Klamer, S., & Muck, M. (2022). Does it really pay off for investors to consider information from social media? *International Review of Financial Analysis*, *81*, 102074. https://doi.org/10.1016/j.irfa.2022.102074

evReturn function—RDocumentation. (n.d.). Retrieved February 11, 2024, from https://www.rdocumentation.org/packages/erer/versions/3.1/topics/evReturn

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82–89.

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

Gupta, B., Mufti, T., Sohail, S. S., & Madsen, D. Ø. (2023). ChatGPT: A brief narrative review. *Cogent Business & Management*, *10*(3), 1–17. Business Source Complete. https://doi.org/10.1080/23311975.2023.2275851

He, W., Guo, L., Shen, J., & Akula, V. (2016). Social Media-Based Forecasting: A Case Study of Tweets and Stock Prices in the Financial Services Industry. *Journal of Organizational and End User Computing*, *28*(2), 74–91. https://doi.org/10.4018/JOEUC.2016040105

Holzer, J., & Bensinger, G. (2013, April 2). SEC Embraces Social Media. *Dow Jones Institutional News*. https://www.proquest.com/abicomplete/docview/2097913609/citation/EE3472CECAE840B1PQ/2

Hong, H., Ye, Q., Du, Q., Wang, G. A., & Fan, W. (2020). Crowd characteristics and crowd wisdom: Evidence from an online investment community. *Journal of the Association for Information Science and Technology*, *71*(4), 423–435. https://doi.org/10.1002/asi.24255

Hossain, M. M., Mammadov, B., & Vakilzadeh, H. (2022). Wisdom of the crowd and stock price crash risk: Evidence from social media. *Review of Quantitative Finance and Accounting*, *58*(2), 709–742. https://doi.org/10.1007/s11156-021-01007-x

Hovland, C. I. (1953). *Communication and persuasion: Psychological studies of opinion change*. Yale university press.

Huang, J. (2018). The customer knows best: The investment value of consumer opinions. *Journal of Financial Economics*, *128*(1), 164–182. https://doi.org/10.1016/j.jfineco.2018.02.001

Huang, J., & Liu, J. (2020). Using social media mining technology to improve stock price forecast accuracy. *Journal of Forecasting*, *39*(1), 104–116. https://doi.org/10.1002/for.2616

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science & Technology*, *60*(11), 2169–2188. https://doi.org/10.1002/asi.21149

Jin, S.-A. A., & Phua, J. (2014). Following Celebrities' Tweets About Brands: The Impact of Twitter-Based Electronic Word-of-Mouth on Consumers' Source Credibility Perception, Buying Intention, and Social Identification With Celebrities. *Journal of Advertising*, *43*(2), 181–195. https://doi.org/10.1080/00913367.2013.827606

Jingnan, H. (2023, February 9). Twitter's new data access rules will make social media research harder. *NPR*. https://www.npr.org/2023/02/09/1155543369/twitters-new-data-access-rules-will-make-social-media-research-harder

Karampatsas, N., Malekpour, S., Mason, A., & Mavis, C. P. (2023). Twitter investor sentiment and corporate earnings announcements. *European Financial Management*, *29*(3), 953–986. https://doi.org/10.1111/eufm.12384

Kaur, C., & Sharma, A. (2020). Twitter sentiment analysis on coronavirus using textblob. *EasyChair2516-2314*.

Lamberton, C., & Stephen, A. T. (2016). A Thematic Exploration of Digital, Social Media, and Mobile Marketing: Research Evolution from 2000 to 2015 and an Agenda for Future Inquiry. *Journal of Marketing*, *80*(6), 146–172. https://doi.org/10.1509/jm.15.0415

Leone, A. J., Minutti-Meza, M., & Wasley, C. E. (2019). Influential Observations and Inference in Accounting Research: Accounting Review. *Accounting Review*, *94*(6), 337–364. https://doi.org/10.2308/accr-52396

Li, W., Zhao, F., Lee, J. M., Park, J., Septianto, F., & Seo, Y. (2024). How micro- (vs. mega-) influencers generate word of mouth in the digital economy age: The moderating role of mindset. *Journal of Business Research*, *171*, 114387. https://doi.org/10.1016/j.jbusres.2023.114387

Liadeli, G., Sotgiu, F., & Verlegh, P. W. J. (2022). A Meta-Analysis of the Effects of Brands' Owned Social Media on Social Media Engagement and Sales. *Journal of Marketing*, 1. Business Source Complete.

Liu, S., Jiang, C., Lin, Z., Ding, Y., Duan, R., & Xu, Z. (2015). Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach. *Information Sciences*, *306*, 34–52.

Lukpat, A. (2023, February 2). Twitter to Remove Free API Access in Elon Musk's Latest Revenue Push. *Wall Street Journal*. https://www.wsj.com/articles/twitter-to-charge-developers-for-api-access-in-elon-musks-latest-revenue-push-11675353876

MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, *35*(1), 13–39.

Manguri, K. H., Ramadhan, R. N., & Amin, P. R. M. (2020). Twitter sentiment analysis on worldwide COVID-19 outbreaks. *Kurdistan Journal of Applied Research*, 54–65.

Mathiassen, L. (2017). Designing Engaged Scholarship: From Real-World Problems to Research Publications. *Engaged Management ReView*, *1*(1). https://doi.org/10.28953/2375-8643.1000

Moorman, C. (2018, March 21). *Social Media Spending, Use And Distraction?* Forbes. https://www.forbes.com/sites/christinemoorman/2018/03/21/social-media-spending-use-and-distraction/

Naldi, M. (2019). *A review of sentiment computation methods with R packages* (arXiv:1901.08319). arXiv. http://arxiv.org/abs/1901.08319

Naresh, E., Ananda, B. J., Keerthi, K. S., & Tejonidhi, M. R. (2022). Predicting the Stock Price Using Natural Language Processing and Random Forest Regressor. *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, 1–5. https://doi.org/10.1109/ICDSIS55133.2022.9915940

Nofsinger, J. R. (2005). Social Mood and Financial Economics. *Journal of Behavioral Finance*, *6*(3), 144–160. https://doi.org/10.1207/s15427579jpfm0603_4

O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *Proceedings of the International AAAI Conference on Web and Social Media*, *4*(1), Article 1. https://doi.org/10.1609/icwsm.v4i1.14031

Ohanian, R. (1990). Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Journal of Advertising*, *19*(3), 39–52. https://doi.org/10.1080/00913367.1990.10673191

Pagolu, V. S., Reddy, K. N., Panda, G., & Majhi, B. (2016). Sentiment analysis of Twitter data for predicting stock market movements. *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, 1345–1350.

Pal, R., Pawar, U., Zambare, K., & Hole, V. (2020). Predicting Stock Market Movement Based on Twitter Data and News Articles Using Sentiment Analysis and Fuzzy Logic. In S. Smys, T. Senjyu, & P. Lafata (Eds.), *SECOND INTERNATIONAL CONFERENCE ON COMPUTER NETWORKS AND COMMUNICATION TECHNOLOGIES, ICCNCT 2019* (Vol. 44, pp. 561–571). Springer International Publishing Ag. https://doi.org/10.1007/978-3-030-37051-0_63

Pfeffer, J., Matter, D., Jaidka, K., Varol, O., Mashhadi, A., Lasser, J., Assenmacher, D., Wu, S., Yang, D., Brantner, C., Romero, D. M., Otterbacher, J., Schwemmer, C., Joseph, K., Garcia, D., & Morstatter, F. (2023). Just Another Day on Twitter: A Complete 24 Hours of Twitter Data. *Proceedings of the International AAAI Conference on Web and Social Media*, *17*, 1073–1081. https://doi.org/10.1609/icwsm.v17i1.22215

Prantl, D., & Mičík, M. (2019). Analysis of the Significance of eWOM on Social Media for Companies. *E+M Ekonomie a Management*, *22*(4), 182–194. ABI/INFORM Collection. https://doi.org/10.15240/tulI/001/2019-4-012

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

Ray, R. (2006). Prediction Markets and the Financial "Wisdom of Crowds." *Journal of Behavioral Finance*, *7*(1), 2–4. https://doi.org/10.1207/s15427579jpfm0701_1

Reboredo, J. C., & Ugolini, A. (2018). The impact of Twitter sentiment on renewable energy stocks. *Energy Economics*, *76*, 153–169. https://doi.org/10.1016/j.eneco.2018.10.014

Reed, M. (2016). A Study of Social Network Effects on the Stock Market. *Journal of Behavioral Finance*, *17*(4), 342–351. https://doi.org/10.1080/15427560.2016.1238371

Rousidis, D., Koukaras, P., & Tjortjis, C. (2020). Social media prediction: A literature review. *Multimedia Tools and Applications*, *79*(9), 6279–6311. https://doi.org/10.1007/s11042-019-08291-9

Saleh, S. N., McDonald, S. A., Basit, M. A., Kumar, S., Arasaratnam, R. J., Perl, T. M., Lehmann, C. U., & Medford, R. J. (2023). Public perception of COVID-19 vaccines through analysis of Twitter content and users. *Vaccine*, *41*(33), 4844–4853. https://doi.org/10.1016/j.vaccine.2023.06.058

Sarkar, T., & Rajadhyaksha, N. (2021). *TLA: Twitter Linguistic Analysis* (arXiv:2107.09710). arXiv. https://doi.org/10.48550/arXiv.2107.09710

Schoen, H., Gayo-Avello, D., Metaxas, P. T., Mustafaraj, E., Strohmaier, M., & Gloor, P. (2013). The power of prediction with social media. *Internet Research*, *23*(5), 528–543. https://doi.org/10.1108/IntR-06-2013-0115

Sentimentr/inst/CITATION at master · trinker/sentimentr · GitHub. (n.d.). Retrieved January 29, 2024, from https://github.com/trinker/sentimentr/blob/master/inst/CITATION

Shelar, A., & Huang, C.-Y. (2018). Sentiment analysis of twitter data. *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 1301–1302.

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

Stieglitz, S., Meske, C., Ross, B., & Mirbabaie, M. (2020). Going Back in Time to Predict the Future—The Complex Role of the Data Collection Period in Social Media Analytics. *Information Systems Frontiers*, *22*(2), 395–409. https://doi.org/10.1007/s10796-018-9867-2

Stop Words. (n.d.). Retrieved February 11, 2024, from https://www.webconfs.com/stop-words.php

Sul, H. K., Dennis, A. R., & Yuan, L. (2014). Trading on Twitter: The Financial Information Content of Emotion in Social Media. In R. H. Sprague (Ed.), *2014 47TH HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS)* (pp. 806–815). IEEE. https://doi.org/10.1109/HICSS.2014.107

Sul, H. K., Dennis, A. R., & Yuan, L. (Ivy). (2017). Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns. *Decision Sciences*, *48*(3), 454–488. https://doi.org/10.1111/deci.12229

Sun, C. (2015). *Empirical research in economics: Growing up with R*. Pine Square Starkville, Mississippi.

Sun, C., & Liao, X. (2011). Effects of litigation under the Endangered Species Act on forest firm values. *Journal of Forest Economics*, *17*(4), 388–398. https://doi.org/10.1016/j.jfe.2011.03.004

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.

Tian, Z., Dew, R., & Iyengar, R. (2023). Mega or Micro? Influencer Selection Using Follower Elasticity. *Journal of Marketing Research*, 00222437231210267. https://doi.org/10.1177/00222437231210267

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the International AAAI Conference on Web and Social Media*, *4*(1), Article 1. https://doi.org/10.1609/icwsm.v4i1.14009

Verma, M., Moudgil, N., Goel, G., Pardeshi, P., Joseph, J., Kumar, N., Singh, K., Singh, H., & Kodali, P. B. (2023). People's perceptions on COVID-19 vaccination: An analysis of twitter discourse from four countries. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-41478-7

Vrontis, D., Makrides, A., Christofi, M., & Thrassou, A. (2021). Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies*, *45*(4), 617–644. https://doi.org/10.1111/ijcs.12647

Wies, S., Bleier, A., & Edeling, A. (2022). Finding Goldilocks Influencers: How Follower Count Drives Social Media Engagement. *Journal of Marketing*, 002224292211251. https://doi.org/10.1177/00222429221125131

Wu He, Lin Guo, Jiancheng Shen, & Akula, V. (2016). Social Media-Based Forecasting: A Case Study of Tweets and Stock Prices in the Financial Services Industry. *Journal of Organizational & End User Computing*, *28*(2), 74–91. https://doi.org/10.4018/JOEUC.2016040105

Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who says what to whom on twitter. *Proceedings of the 20th International Conference on World Wide Web*, 705–714. https://doi.org/10.1145/1963405.1963504

Zhang, W., Wang, P., Li, X., & Shen, D. (2018). Twitter's daily happiness sentiment and international stock returns: Evidence from linear and nonlinear causality tests. *Journal of Behavioral and Experimental Finance*, *18*, 50–53. https://doi.org/10.1016/j.jbef.2018.01.005

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear." *Procedia - Social and Behavioral Sciences*, *26*, 55–62. https://doi.org/10.1016/j.sbspro.2011.10.562

Zhang, Y., An, Y., Feng, X., & Jin, X. (2017). Celebrities and ordinaries in social networks: Who knows more information? *Finance Research Letters*, *20*, 153–161. https://doi.org/10.1016/j.frl.2016.09.021

VITA

Denys Lu is an accomplished IT Executive with over two decades of experience in digital and IT audit, analytics, and IT service delivery. His expertise encompasses customer service, enterprise systems administration, software engineering, analytics, and business intelligence. Denys is recognized for his extensive technical knowledge and ability to communicate with technical and non-technical stakeholders effectively.

Denys has diverse academic interests, with a bachelor's degree in Biology from Emory College, an MBA in Finance and Consulting from Goizueta Business School, and a Doctor of Business Administration from Georgia State University. His academic pursuits reflect a strong commitment to continuous learning and applying advanced knowledge to practical challenges in Information Systems, Finance, and Marketing. He enjoys teaching and is invigorated by academic research.

Outside of his professional and academic endeavors, he enjoys cooking, woodworking, and spending time with his wife and their two children.