Norwegian School of Economics
Bergen, Spring 2021

WallStreetBets on Wall Street

An Empirical Analysis of the Market Power of WallStreetBets

Truls Jacobsen and Tobias Fosser Pedersen

Supervisor: Walter Pohl

Master Thesis, MSc in Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

We would like to thank our supervisor Walter Pohl for his enthusiasm for the thesis. We appreciate his invaluable guidance, insights, and constructive feedback, as well as his support during the writing process. We would also like to thank Thomas Eitzen for great topic discussions, whilst sharing his valuable insights and thoughts with us throughout the process. Finally, a thanks should be directed to August Rolfsen who has helped with LaTeX formatting as well as invaluable feedback over the semester.

This thesis marks the end of our Master of Science in Economics and Business Administration at the Norwegian School of Economics. Our studies have required patience, dedication, and support from those who stand closest to us. We therefore want to extend our heartfelt gratitude to our loved ones for their support throughout our studies.

Abstract

In this thesis, we apply sentiment analysis techniques to test whether sentiment on WallStreetBets has had an impact on stock returns, trading volume, option volume, and implied volatility from January 01, 2020, to March 15, 2021. We analyze each submission and comment posted on WallStreetBets during this time interval that can be linked to discussion of a selected sample of stocks, and apply sentiment analysis techniques to identify whether each post displays positive, neutral, or negative sentiment. We then analyze stocks on an individual and aggregated basis to test the following hypotheses: whether sentiment on WallStreetBets has had an impact on (i) stock returns; (ii) stock volume; and (iii) option volumes and implied volatility. First, there are large variations in the results for sentiment's impact on return on an individual basis, and while reverse causality can be attributed to explain much of the results we observe for some individual stocks, we find indicative evidence of WallStreetBets sentiment having had a statistically significant impact on the return of other stocks. On an aggregated basis, sentiment is shown to explain returns better the day after sentiment is recorded, suggesting an ability to influence future stock returns. Second, by looking at sentiment against volume we find a statistically significant relationship on most stocks in our sample, suggesting forum sentiment drives stock activity. This relationship on an aggregated basis is stronger without lagged effects, meaning same-day sentiment drives stock volumes. Finally, we find the strongest relationship in our study when looking at option-related metrics, showing a clear effect on both call and put volume as well as implied volatility both on an individual and aggregated basis. The results from our minute-by-minute model during the January 2021 rallies suggest that forum activity was a statistically significant driving force behind volume in the affected stocks. However, on the same data we could not find a statistical relationship on return, suggesting there were other influences behind the price increases than comments on the forum alone. We also develop trading strategies based on sentiment on WallStreetBets, and find that these would have yielded remarkable returns in the time interval we explore.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background and motivation

In November 2020, the stock price of the American software company Palantir Technologies Inc. (PLTR) rose significantly. The company had not reported any changes to its fundamentals, and there were no other news that could rationally explain such a sudden and impactful upsurge in the price of the stock. From its close on November 2, 2020, at $10.54, PLTR's stock price had nearly tripled on December 8, 2020, when it reached its 2020 intraday high at $31.24. The Reddit subforum WallStreetBets was quickly identified as the likely actor driving the rally. The upsurge followed a period of heavy discussion of PLTR at this particular forum, and as the price surge showed no apparent sign of stopping, the WallStreetBets forum became increasingly flooded with hyping submissions and comments revolving around PLTR (Whiteman, 2021). Eventually, the PLTR rally came to an end, but WallStreetBets was soon to be world-famous for even more remarkable rallies.

Being occasional observers of the forum, entertained by the forum's fascination for high-risk options trading strategies, the sheer magnitude of comments and submissions revolving around the stock intrigued us. We quickly identified that stock mentions had sprung over to other social media outlets and wanted to explore how the forum interacted with the market. This event was the main motivation behind looking into the topic at hand in our thesis.

2020 was a year for the history books. The COVID-19 pandemic meant that people all across the globe suddenly were forced to change their lives, habits, and behaviors to adapt to a precautionary approach to the pandemic as well as governmentally implemented restrictions (Van Bavel et al., 2021). Consequently, trading volumes in equity and derivative markets have been unprecedented (Chiah & Zhong, 2020), and the number of people entering the stock markets for the first time has been record-breaking (Rooney, 2020, AksjeNorge, 2021). Stimulus checks in the US, intended to be recycled back into

the economy, have for many Americans been an entry point into the stock market, where middle-class Americans traded stocks 90% more the week they received the stimulus checks than the week prior (Fitzgerald, 2020). At the same time, stock market indices have reached all-time highs (Jain & Singh, 2021), the federal funds rate has been just above 0.0% (Board of Governors of The Federal Reserve System, 2020) and yield curves have reached record-low levels (Brettell et al., 2020; Franck & Li, 2020).

Just as what happened with the stock price of PLTR in November 2020, in January/February 2021 financial markets started to witness extreme upsurges in the price of several stocks which had not experienced any noteworthy changes in their respective fundamentals. Many of these stocks had displayed flat, or even negative, stock price developments, but suddenly their prices rose many-fold. The inflicted stocks were many and from a vast range of different sectors, but what they all had in common was that they were heavily discussed at the Reddit subforum WallStreetBets. At the WallStreetBets forum, the inflicted stocks were subject to intense debate, and the community actively encouraged its members to buy the targeted stocks directly or trading in related options with the intention to create an upward price spiral. The goal of these efforts was to push the stocks' prices significantly upwards both as a result of increased demand for the stocks and options, as well as initiating short and gamma squeezes in the particularly targeted stocks. By pursuing a strategy of targeting stocks displaying characteristics such as high short interest, i.e. a characteristic that makes stocks especially prone to short squeezes, the community sought to achieve substantial gains. It quickly became evident that when specific stocks became subject to intense discussion on the WallStreetBets forum, their prices often started to surge.

Although the phenomenon started to catch the attention of the financial media in November 2020 with the PLTR rally, in January 2021 international media's interest in the phenomenon exploded after the gaming and consumer electronics retailer GameStop Corp. (GME) became the new target at WallStreetBets. On January 4, 2021, GME shares closed at $17.25. Roughly three weeks later, on January 28, 2021, the shares reached their peak price of $483.00, i.e. 28 times higher than what they traded at in

the beginning of the same month. At the same time, other stocks displayed similar tendencies, whereas AMC Entertainment (AMC), Blackberry (BB), and Nokia (NOK) serve as some of the most-known companies that suddenly became victims for these targeted attacks. To illustrate the magnitude of some of these parallel rallies, compared to their price levels before their respective price peaks AMC experienced a price surge of 480% (Keshner, 2021), BB a price surge of 112% (Aliaj & Fletcher, 2021) and NOK a price surge of 133% (Reuters Staff, 2021).

During these rallying weeks the stocks that displayed the rally-like tendencies were subject to intense discussion on WallStreetBets, and the forum itself also experienced a significant increase in the number of subscribers, posts and comments in the same time period. These coordinated efforts to try to move stock prices, facilitated and communicated at the WallStreetBets forum, have attracted massive media coverage, extensive investigation by legal and regulatory authorities and both profound criticism and enthusiasm. Some argue that actions must be taken to prevent such rallies in the future, whereas others argue that this is the beginning of a battle between retail investors and institutional investors. Nevertheless, the recent events will likely continue to be subject to extensive scrutiny from regulatory authorities, lawmakers and researchers in the months and years to come.

In this thesis we will analyze the market power of the WallStreetBets forum by examining whether sentiment on this forum has had an impact on stock returns, trading volume, option volumes and implied volatility. We will apply textual and sentiment analysis to analyze each submission and comment posted on the WallStreetBets forum from January 01, 2020, until March 15, 2021, and combine this with statistical tests and financial modelling to analyze whether the forum really can be claimed to exercise market power. Several studies have previously examined different aspects of the WallStreetBets forum, but none have yet to formally study how sentiment on the forum actually impacts the market. Our study of this contributes to the developing span of studies analyzing the forum's market power and investment strategies based on activity on the forum.

1.2 Research question

In this thesis, we combine sentiment variables constructed based on textual data from WallStreetBets with financial data from January 02, 2020, to March 15, 2021 to test;

(i) whether sentiment on WallStreetBets has had an impact on stock returns

(ii) whether sentiment on WallStreetBets has had an impact on stock volume

(iii) whether sentiment on WallStreetBets has had an impact on option volumes and implied volatility

During this time period, a massive number of news articles claimed that WallStreetBets had driven several stocks' prices to unprecedented levels. We test this relationship for 18 stocks individually and aggregated to analyze whether sentiment on WallStreetBets actually can explain these variations. WallStreetBets is characterized by being a forum dedicated to high-risk trading strategies utilizing options to maximize potential returns, and by testing these hypotheses we are also able to analyze whether this strategy has had an impact on option volumes and implied volatility.

1.3 Relevance of this research

As one of the first papers to formally analyze the WallStreetBets forum's market power, this paper is relevant both for further research, policy makers, regulatory authorities and market participants. The recent rallies have led to huge gains for some investors, but also losses amounting to billions of dollars for hedge funds and other short-sellers, in addition to many retail investors who did not liquidate their positions in time. The monumental stock price movements shook up equity markets all around the world. Many have called the rallies detrimental for market efficiency, whereas others have called it market manipulation. Price manipulation is a breach of the efficient market hypothesis, and although there is still debate as to whether the latest rallies were legal price manipulations or not, the rallies carry important implications for further regulation and trading strategies. Therefore, although we do not seek to argue whether these actions

were either legally or morally acceptable, we believe that this research will contribute to bring knowledge and comprehension into the debate that will follow these rallies in the months and years to come.

1.4 Organization of the thesis

The first chapter of this thesis gives an introduction to the background and motivation for the thesis, as well as an introduction to the thesis' research question. The second chapter will provide a contextual overview of some of the most important features surrounding and enabling the events we seek to explore in this thesis. In chapter 3, relevant literature and research is explored. Chapter 4 presents the data and methods we have utilized in our analyses. Chapter 5 presents the results and discussion of our analyses. In chapter 6, we conclude our findings and discuss this research's limitations as well as suggestions for further research.

Throughout the thesis, although the forum is commonly referred to by many acronyms, we will generally refer to the forum as WallStreetBets. We will also use stock tickers to refer to the stocks we discuss (see table 4.2 for an overview of the thesis' selected stock sample and their tickers). Finally, activity will be used as a term describing overall stock mentions, positive sentiment to describe the percentage positive sentiment a stock experiences, whilst sentiment will be used as a broad term incorporating both.

2 Context

2.1 The coronavirus' impact on financial markets

During the ongoing COVID-19 pandemic, at least 37 international equity markets have experienced a large spike in aggregated trading volume (Chiah & Zhong, 2020). Ordered to stay at home, many Americans have flocked to zero-commission trading platforms such as Robinhood, which experienced a three-fold increase in trading volume in March 2020 from its 2019 level (Rooney, 2020). Investors have opened new trading accounts, and on average added funds to their accounts and established more new positions in stocks (Ortmann et al., 2020). Panigirtzoglou et al. (2021) use small traders' equity option flows as a proxy for retail investors and find that the retail impulse in the financial markets has been strong during the COVID-19 pandemic. They also find that the share of retail-driven stock market trading volumes in the US reached unprecedented heights in January 2021, in addition to finding that retail investors have preferred large tech stocks as well as stocks with low market capitalization during 2020 as well as in January and February 2021 (Panigirtzoglou et al., 2021).

Given the extremely low interest rates during the pandemic, where many central banks have operated with interest rates very close to or precisely 0%, investors have had few other alternatives than to tilt their portfolios towards stock markets to achieve notable returns. At the same time, due to cancellation and postponement of sports events, the European sports betting industry saw a massive decrease in their gambling revenues during the start of the COVID-19 pandemic (Auer et al., 2020). Several studies have shown that trading in stock markets may function as a substitute for gambling (see e.g. Gao & Lin, 2015; Kumar, 2009; Li, 2012), meaning that the increased trading volumes experienced world-wide may be partially explained by more time spent at home, the substitution effect of gambling and the introduction and adaptation of several zero-commission trading platforms.

These observations may explain how such price rallies we seek to explore in this thesis

could erupt in the first place. Being heavily discussed on the WallStreetBets forum, the stocks subject to these rallies have been remarkably popular among retail investors during the time period of the rallies. The COVID-19 pandemic can arguably be attributed to explain much of the heightened retail impulse in the financial markets over the last year. We would also argue that the pandemic has contributed to increased activity on WallStreetBets, and thereby having facilitated the foundation for the massive discussion and subsequent stock rallies we explore in this thesis.

2.2 WallStreetBets

Reddit is an American social network platform founded in 2005, offering its members access to submit and interact with various content (Reddit.com, 2020). As of February 2021, Reddit.com was the 18th most visited website in the world and the seventh most visited website in the US (Alexa.com, 2021). The Reddit platform is organized along different subforums, or subreddits, whereas each subreddit is dedicated to a specific category of content. WallStreetBets, alternatively r/wallstreetbets or WSB, is a subreddit created by Jaime Rogozinski in 2012 that is dedicated to discussion and communication revolving around high-risk/reward short-term trading operations (Anthony, 2020).

Starting from January 2020, continuing throughout 2020, WallStreetBets experienced a steady increase in the number of its subscribers, posts and comments per day (Subreddit Stats, 2021). Eventually the subforum became subject to speculations as to whether coordinated trading efforts on the subforum, made possible by pooling the users' financial power, had the potential to move stock prices (see e.g. Kawa, 2020b; Sheetz, 2020; Zweig, 2020). These speculations came specifically after a time of seeing multiple stocks, e.g. PLTR, Plug Power, Inc. (PLUG) and Virgin Galactic Holdings, Inc. (SPCE), with seemingly no news experienced soaring stock prices after having been subject for discussion on WallStreetBets (see e.g. Kawa, 2020a; Lipschultz, 2020; Powell & Stafford, 2020; Wang & Hajric, 2020).

In January 2021, WallStreetBets experienced a surge in its popularity, measured by the

number of subscribers, comments and submissions (Subreddit Stats, 2021). The surge was fueled by the immense scrutiny and attention from the media as the equity and option markets observed significant price movements in particular securities following intense discussions on WallStreetBets. WallStreetBets was as of February 2021 among the four most active subforums in terms of comments per day on the Reddit platform, and has over nine million subscribers (Subreddit Stats, 2021). Below, three charts showing the development in terms of subscribers, posts per day and comments per day on the WallStreetBets subreddit are displayed.

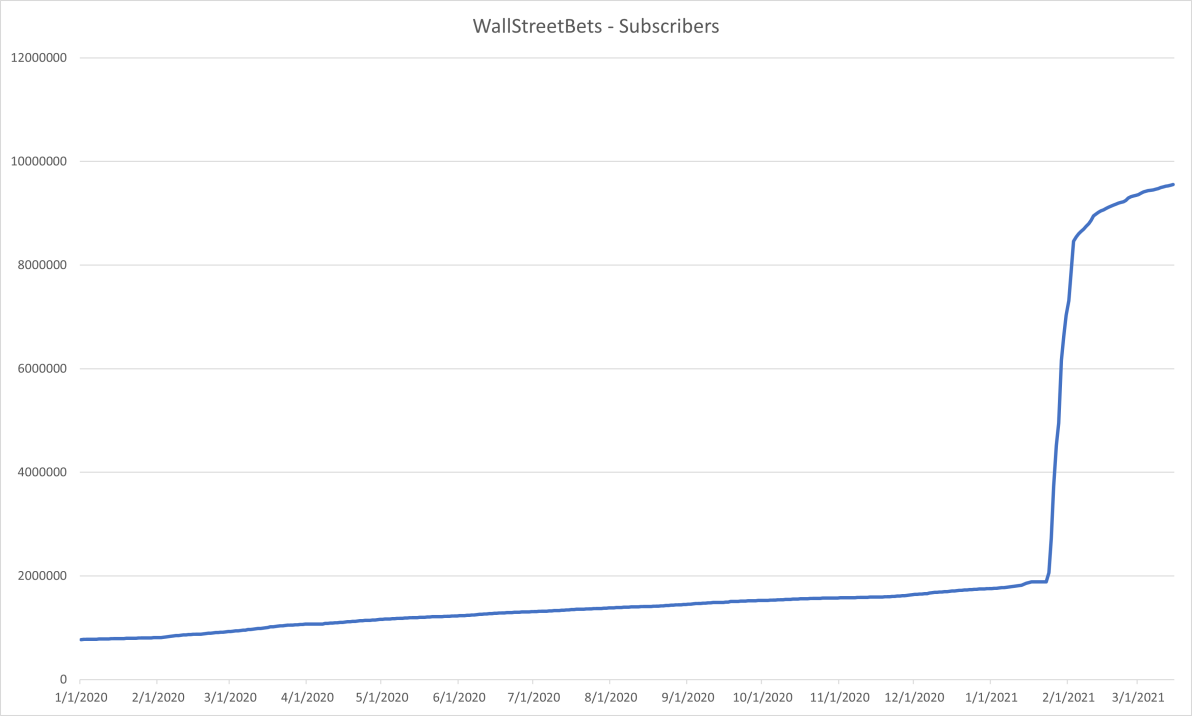

Figure 2.1: WallStreetBets subscribers

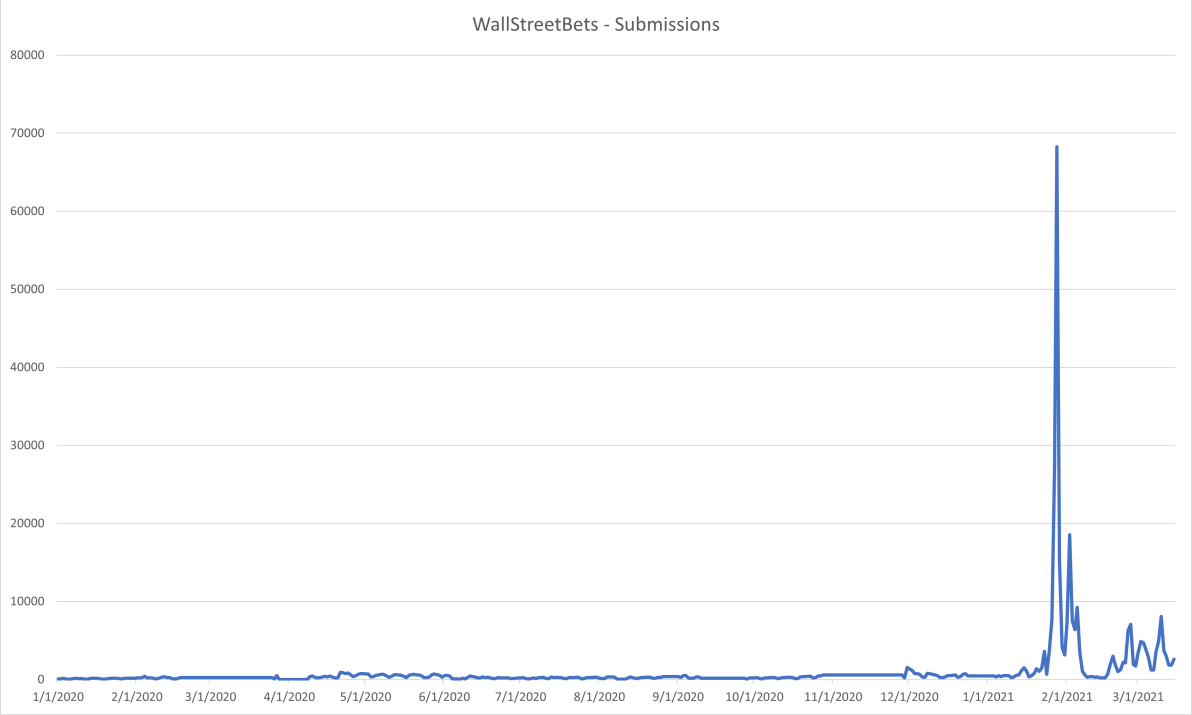

Figure 2.2: WallStreetBets submissions

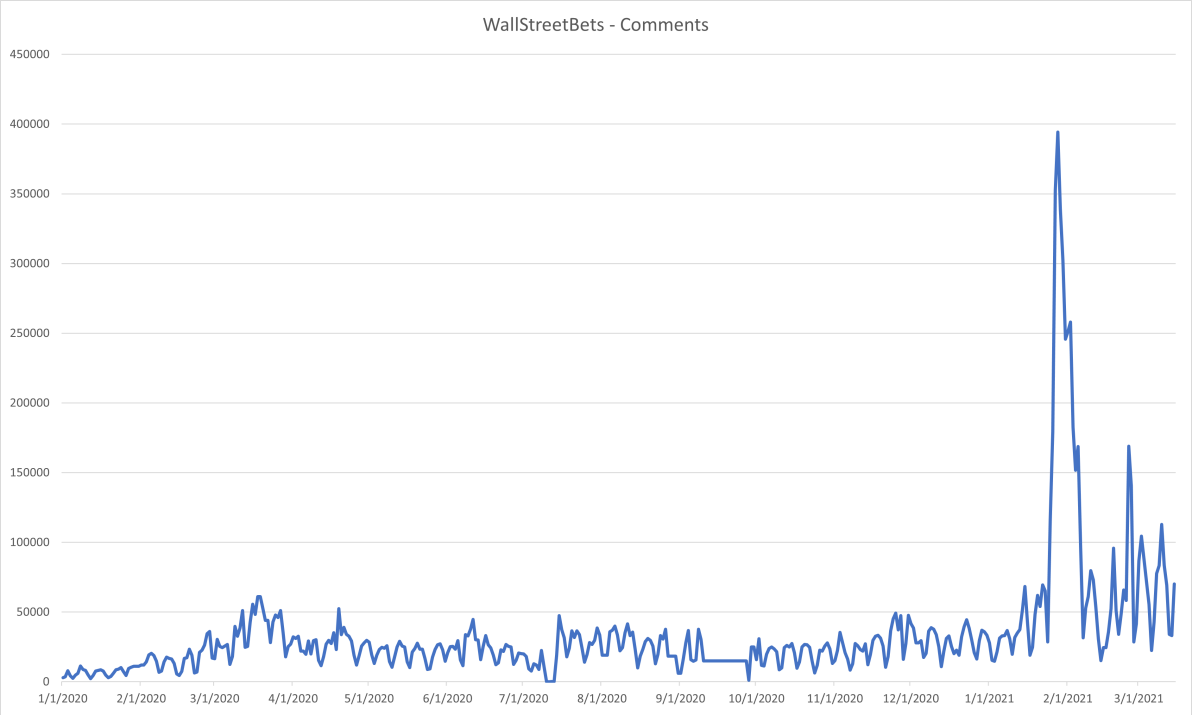

Figure 2.3: WallStreetBets comments

2.3 Robinhood

Robinhood Markets Inc., hereby Robinhood, is an American financial services company headquartered in Menlo Park, California (Craft.co, 2021) that offers trading in stocks, funds, ETFs, options, golds and cryptocurrencies, as well as other products (Robinhood, n.d.c.). Robinhood offers commission-free trading to its customers, and had 13 million users as of 2020 (Rega, 2020). Their revenue streams come from their premium subscription program Robinhood Gold, stock loan income, interest on interest-bearing bank accounts, cash management services, proxy service revenue and fees (Robinhood, n.d.b). In addition, Robinhood earns rebates through payment for order flow by selling its customers' orders to market makers who then can execute these trades (Robinhood, n.d.b.).

The company has announced a vision of democratizing finance for all, thereby offering products aimed at their retail investor-based target group (Robinhood, n.d.a). Retail investors wanting to trade in options and other derivatives would previously have a challenge of doing so, as the financial systems did not allow retail investors easy access to such "advanced" financial products. With the democratisation of finance in recent years, companies such as Robinhood have enabled easy access for retail investors to options and derivative trading. In chapter 5, we find that sentiment on WallStreetBets can be claimed to have had a statistically significant impact on call and put option volumes. We would argue that this is made possible and enabled by the democratisation of finance that has evolved over the recent years, where Robinhood and similar retail investor trading platforms have facilitated easy access for retail investors to products they previously would have troubles of trading in.

While US retail investors use many different trading platforms in addition to Robinhood, such as Etrade, TD Ameritrade and Schwab, Robinhood is often considered as the most popular stock broker among WallStreetBets users (see e.g. Alfonso III, 2021; Sarlin, 2021). Most often when WallStreetBets users post screenshots of their trade balances, gains or losses, we have observed that Robinhood is the depicted trading platform. Also,

using textual analysis, we find that Robinhood is the most frequently mentioned trading platform on the forum, being the 143th most frequently mentioned word of all our comments. We would therefore argue that Robinhood is clearly the most important trading platform for WallStreetBets users. To fully understand the enabling factors for the recent rallies, we argue that it is essential to understand the role of zero-commission platforms such as Robinhood. An introduction of Robinhood therefore seeks to contribute to a comprehensive understanding of enabling factors.

2.4 Stock rallies

2.4.1 Stocks subject to rallies

In the time interval we are exploring in this thesis, i.e. from January 01, 2020, to March 15, 2021, many stocks displayed rally-like tendencies, and many became subject to short and gamma squeeze mechanics. Not all the rallies that erupted in our selected time interval are likely to have erupted as a result of short or gamma squeezes, but most of the stocks that displayed rallying tendencies were heavily discussed at WallStreetBets. Some rallies were more prominent and publicly debated than others, but the mechanics were often just as distinct for those who did not receive as much public attention as e.g. GME and AMC. While the number of stocks that displayed rally-like tendencies in this time interval was high, we have limited the focus to 16 of the stocks that displayed such tendencies between January 01, 2020, and March 15, 2021. All stocks except Apple Inc. (AAPL) and Alphabet Inc. (GOOG) in table 4.2 experienced rallies in this time period, and thereby serve as examples of the vast range of stocks that became subject to these mechanics in the time interval we explore.

2.4.2 The GME rally

The most prominent and discussed stock rally of all the stock rallies that erupted in the time interval we are exploring in this thesis is arguably the GME rally which occurred from January to February 2021. This specific episode sparked unprecedented attention both

from the WallStreetBets forum, media, market participants and regulatory authorities, and GME was on everyone's lips. Just as the other rallies that took place in this time interval, GME had been specifically targeted on the WallStreetBets forum because of its high short ratio which made the stock especially prone to a joint sentiment-fueled targeted attack by investors. Being the most debated and controversial of the rallies, this episode serves as a great illustration of the magnitude and mechanics underlying this thesis, and we will therefore now provide an introduction to this particular episode.

On January 4, 2021, i.e. on the first trading day in 2021, GME's closing price was \$17.25. One year earlier, on January 2, 2020, i.e. the first trading day in 2020, GME's closing price was \$6.31, meaning that during 2020 GME's stock price had nearly tripled. As remarkable as this may seem, this was nothing compared to what was about to evolve over the next month. On January 20, 2021, GME closed at \$39.36, and only two days later, i.e. on January 22, GME closed at \$65.01. Over the next days the rally continued even further, and on January 27 GME closed at \$347.51. On the next day, i.e. January 28, GME reached its peak at \$483.00, before it later collapsed and closed on \$193.60 the same day. This collapse came as a direct result of trading restrictions imposed by several trading venues on multiple stocks, including GME (see chapter 2.4.3 for an introduction to these trading restrictions). The following days after the trading restrictions were imposed, GME's stock price continued to fall steadily, and on February 22, the stock closed at \$46.00. The GME short squeeze had come to a swift end.

In 2020 overall, the daily average trading volume in the GME stock was 6.68 million shares. In January 2021, the daily average trading volume in the GME stock was 66.43 million shares. On January 22 alone, over 197 million GME shares were traded, a daily record of over 120 million more shares than the daily trading volume record for the stock in 2020. Figure 2.4 shows the daily trading volume in the GME stock from January 02, 2020 to February 22, 2021, and shows that the trading volume spiked during the period of the rally in January 2021.

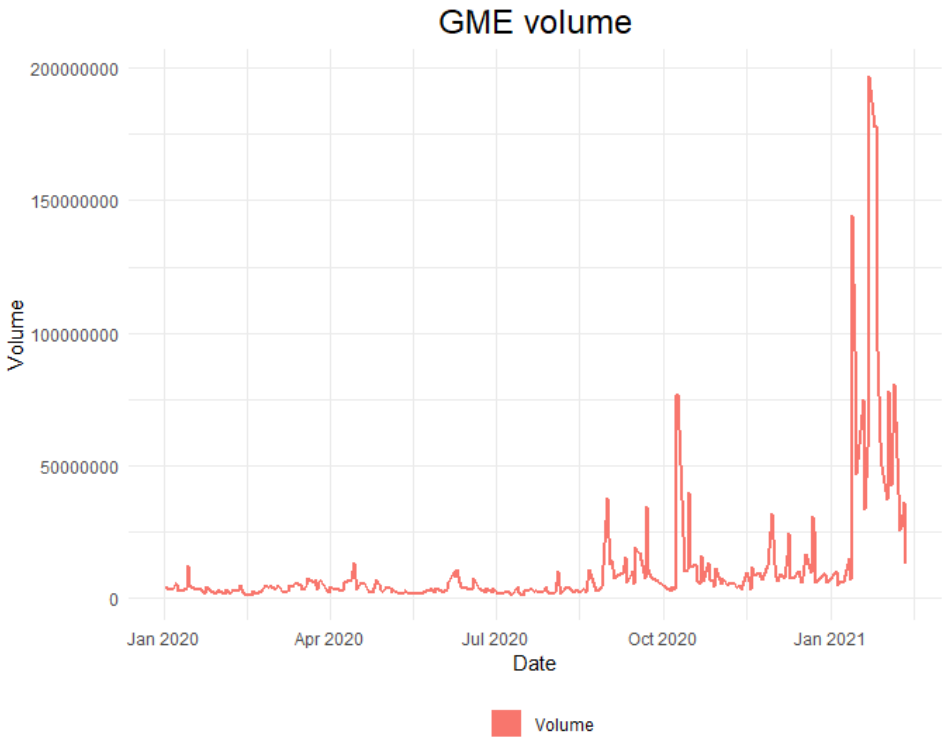

Figure 2.4: GME volume

Having, according to their perspective, successfully launched previous attacks on stocks, the WallStreetBets community started showing explicit interest for the GME stock in November 2020. The stock became increasingly praised on the forum throughout the end of 2020, and in January 2021 the forum's interest for the GME stock exploded (Lyons, 2021). The forum's interest escalated particularly after the investment research company Citron Research announced that they considered GME as overvalued and therefore had started taking short positions in the stock (Ghosh, 2021a). The forum's interest for GME was further fueled as other short-sellers eventually became identified, amongst others the investment management firm Melvin Capital being one of the most prominent of those identified (Aliaj et al., 2021).

When the short-sellers were identified, these were commonly referred to as "enemies" by the forum. This enemification of the forum's counterparties is believed to have helped to fuel the rally further, and many users argued that the GME rally was the beginning of a battle between retail and institutional investors. Another specific event believed to have

contributed to the squeeze's momentum happened on January 26, 2021, when the CEO of Tesla, Inc., Elon Musk published a one-word tweet in the jargon of WallStreetBets saying "Gamestonk!!"(Musk, 2021b). Elon Musk is often believed to move stock prices when he expresses his opinions about certain stocks, and there are many examples of episodes where tweets by Musk are followed by significant stock price movements. For example, on January 7, 2021, following a tweet by Musk saying "Use Signal" (Musk, 2021a) where Musk referred to a chat app, the stock price of the unrelated biotech company Signal Advance Inc. rose manyfold (Gambrell, 2021; Google Finance, 2021). On January 26, 2021, Musk tweeted "I kinda love Etsy" (Musk, 2021c), following up that he had bought a hand knit wool Marvin the Martian helm for his dog on the Etsy e-commerce platform (Musk, 2021c). After publishing the tweet, the value of Etsy Inc. increased by more than $2 billion (Gambrell, 2021). When Musk tweeted "On Clubhouse tonight at 10pm LA time" (Musk, 2021d) the stock price of the unrelated Clubhouse Media Group rose significantly (Gambrell, 2021). Also when Musk announced that Tesla had bought $1.5 billion worth of Bitcoin, the Bitcoin price rose significantly (Dawson & Popina, 2021). The "Gamestonk!!" tweet by Musk is therefore widely believed to have contributed to fuel the GME rally further (see e.g. Bursztynsky, 2021; Gambrell, 2021).

In January 2021, GME was the most shorted stock on Wall Street (Ponciano, 2021), and had an accumulated short interest of 144.34%. Through most of 2020 the short interest in GME had fluctuated around 100%, but starting from September 2020 the short interest rose rapidly towards 144.34%. As the WallStreetBets-fueled targeted attack on short-sellers gradually gained momentum, the GME stock price rose significantly. To cover their positions, both as a result of limiting losses and margin calls, short sellers started buying GME stock, which contributed to the stock price being pushed further upwards. Having to unwind their short positions, the short interest in GME became significantly reduced from over 144.34% on January 14 to 42.61% on January 29, before it was further reduced to 32.78% on February 12, 2021. The following graph shows both the GME stock price and the short interest in the GME stock, and effectively shows how the short interest was reduced as the stock price rose in January 2021.

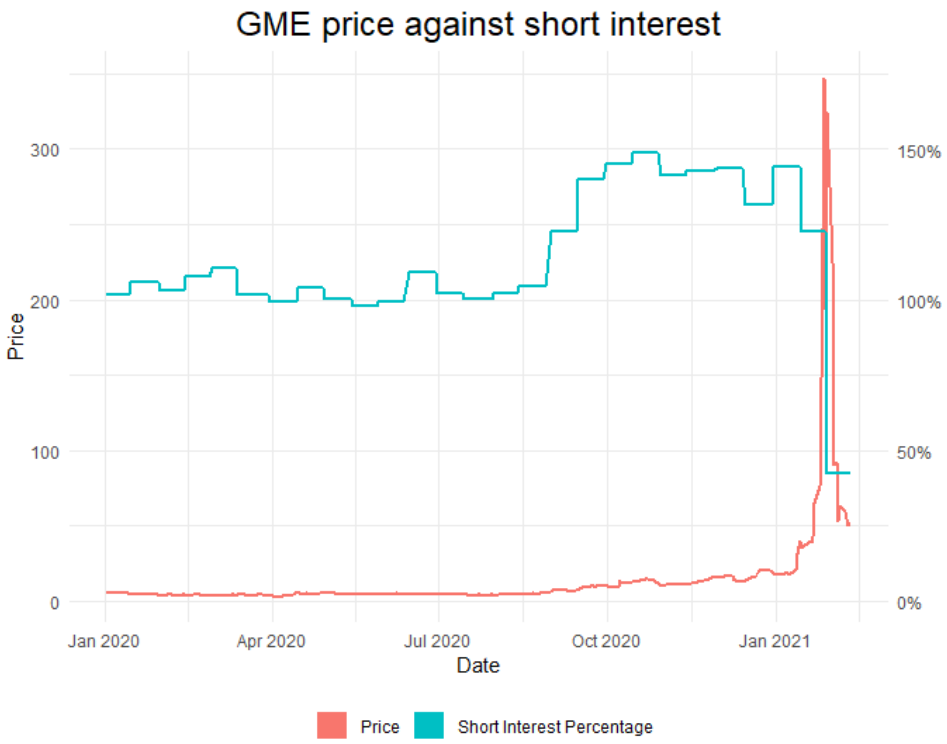

Figure 2.5: GME price and short interest

In conjunction with the short squeeze, a gamma squeeze also contributed to the upsurge in GME's stock price. On the WallStreetBets forum many users advocated buying call options in GME to initiate a gamma squeeze, with the intent of pushing the GME stock price to unprecedented levels. Consequently, as is illustrated in figure 2.6, in the midst of January 2021 call option trading volumes in the GME stock rose significantly. To hedge their exposure, market makers having written the options started buying GME shares, which increased both the trading volume in the GME stock as well as the overall price pressure (Niu, 2021). The following graph shows GME's option volume from January 2020 to February 2021.

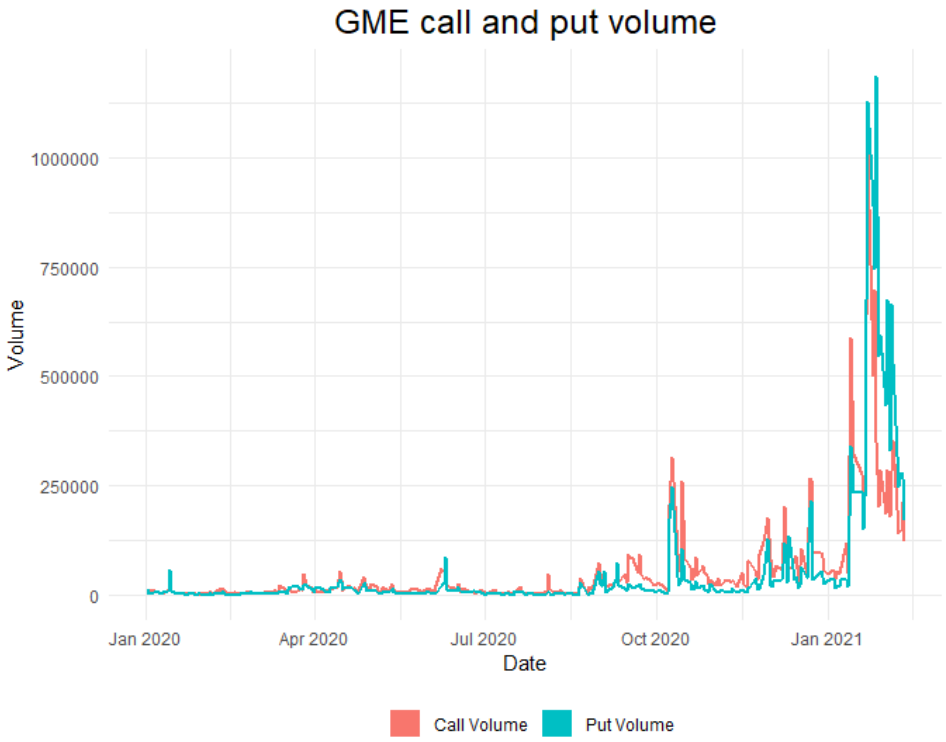

Figure 2.6: GME call and put volume

The combination of the short squeeze and the gamma squeeze made the attack on the GME stock powerful and unparalleled. The WallStreetBets community received extensive coverage in the international media, and investors all across the globe started taking positions in the GME stock to join in on the rally (see e.g. He & Wang, 2021; Hopland, 2021; Kowsmann, 2021). As the trading volume in the GME stock, as well as other stocks also subject to short squeezes in this time period, skyrocketed, several trading platforms imposed trading restrictions on many of these inflicted stocks, including GME, on January 28 (Li & Pound, 2021). The restrictions hampered investors' ability to buy GME shares, and investors using these platforms could only choose to sell their shares in the inflicted stocks. Not being able to trade freely, the demand-driven price pressure on GME dropped immediately, and the stock price fell significantly in only a few days. Although the restrictions were gradually eased over the next few days, the overall price pressure had lost its momentum. Combined with the short ratio having dropped significantly, the GME short squeeze was considered as having come to an end. Remark,

that GME is still as of May 2021 the most frequently discussed stock on WallStreetBets, and in the beginning of March the stock experienced yet another price surge. This time GME had a much lower short ratio, and the temporary surge was also relatively short-lived. Still, the stock has yet to revert back to its low levels prior to the rallies and hovers around USD 170 on May 20, 2021, almost 37 times higher than the year prior.

2.4.3 Trading restrictions

On January 28, 2021, Robinhood, Schwab and several other providers of commission-free trading platforms imposed trading restrictions on certain stocks, including GME (Li & Pound, 2021). This followed the period of significant upsurges in several stocks' prices and option volumes. Robinhood claimed that they had experienced a ten-fold increase in their clearinghouse-mandated deposit requirements as a result of the large increase in trading volumes, and therefore eventually had to impose trading restrictions on certain stocks (Robinhood, 2021). On January 28, these restrictions on the Robinhood platform meant that users could not buy more GME stock or options, but could only choose to sell. Robinhood did not publish an exhaustive list over the trading-limited stocks on January 28, but published a statement saying that due to recent volatility they were "restricting transactions for certain securities to position closing only, including $AAL, $AMC, $BB, $BBBY, $CTRM, $EXPR, $GME, $KOSS, $NAKD, $NOK, $SNDL, $TR and $TRVG" in addition to saying that they raised certain margin requirements (Fitzgerald, 2021).

Following the imposition of the restrictions, Interactive Brokers chairman Thomas Peterffy stated that the extreme increases in trading and option volume nearly made the entire financial system collapse, since short-sellers and market makers were near a point where they could not fulfill their obligations (Stankiewicz, 2021). This could, according to Peterffy, have created a scenario where "brokers default on the clearinghouses, so you end up with a complete mess that is practically impossible to sort out" (Stankiewicz, 2021). The next few days the trading restrictions were gradually eased, and on February 5, 2021, the trading restrictions were completely lifted on Robinhood, meaning that Robinhood-based investors again could freely engage in trades in the previously inflicted

stocks. Schwab has still as of May 4, 2021, special margin requirements for AMC and GME (Charles Schwab, 2021), meaning that although investors can trade relatively freely, the special margin requirements still hamper completely free trading activity on their platform. As of May 4, 2021, TD Ameritrade has still put a 100% margin requirement for long positions in the GME stock, in addition to having temporarily banned short-selling of GME (TD Ameritrade, 2021). In sum, these restrictions made the GME rally, as well as the other rallies occurring in the same time period, come to an abrupt end, and they still limit free trading. The restrictions have been controversial, and many have criticized the platforms for imposing them in the first place.

Figure 2.7: AMC, BB, GME and NOK January 2021 movements

Figure 2.7, which is gathered from Bloomberg, illustrates how certain stocks suddenly experienced significant upturns between January 25 and January 27 before they later collapsed, at least partially, as a result of the trading restrictions. The fact that several stocks' short ratios had fallen dramatically as a result of short squeezes may also have contributed to the collapse. In the figure, we observe how the stock prices of AMC, BB, GME and NOK all rose significantly in this period, before they all fell dramatically on January 28, 2021, when the trading restrictions were imposed. After the dramatic collapse on January 28, we observe how all these stocks then rose again on January 29, before they fell further in February. The figure also shows how correlated the stocks'

prices were during this period, where upturns and downturns of different stocks parallel each other quite remarkably.

2.4.4 Aftermath of the short squeezes

The trading restrictions were met with fury and debate amongst WallStreetBets users who now were unable to trade freely (Ghosh, 2021b), and several of the affected stocks quickly plunged (Davies, 2021). The restrictions also drew wider attention. Two prominent US lawmakers Alexandria Ocasio-Cortez and Ted Cruz both called Robinhood's controversial practice into question (Jones and Carissimo, 2021). The SEC released a statement saying that they were monitoring the situation carefully and that they "will act to protect retail investors when the facts demonstrate abusive or manipulative trading activity that is prohibited by the federal securities laws and several court appeals have since been raised" (SEC, 2021). A Robinhood user filed a class-action lawsuit against Robinhood on January 28, 2021, on the basis of Robinhood's decision to restrict trading in several stocks, calling it manipulation of the stock markets (Duffy, 2021). As of February 1, 2021, at least 25 lawsuits in 11 US states had been filed the previous week against Robinhood (Cridlin, 2021).

On February 18, 2021, a remote hearing titled "Game Stopped? Who Wins and Loses When Short Sellers, Social Media, and Retail Investors Collide" was held in the House Financial Services Committee, which has launched an investigation on the recent activities in stocks such as GME and the trading restrictions imposed by several trading platforms (U.S. House Committee on Financial Services, 2021; Warmbrodt, 2021). The House Financial Services Chair Maxine Waters, in addition to other politicians, have announced their intentions to specifically dig into whether Wall Street hedge funds have had a role in imposing the trading restrictions. In the first hearing, both Robinhood CEO Vladimir Tenev, Reddit CEO Steve Huffman, Melvin Capital CEO Gabriel Plotkin, Citadel CEO Kenneth Griffin and WallStreetBets user and investor Keith Gill, more commonly known amongst WallStreetBets users as DeepFuckingValue (DFV), attended (U.S. House Committee on Financial Services, 2021). On March 17 a second congressional

hearing was held, where the committee focused particularly on the controversial practice of payment for order flow (Phillips, 2021) and gamification of investment and investment apps (Avis, 2021). The ultimate outcome of the hearings and regulatory proceedings are not yet clear as of May 16, 2021, but the short squeeze episodes will likely be subject to immense regulatory scrutiny in the months and years to come.

Ponczek et al. (2021) estimate that 50 of the so-called "meme stocks", i.e. stocks that are praised on the WallStreetBets forum, gained approximately $276 billion in value from the end of 2020 to the peak of the rallies, before they lost approximately $167 billion in total when the rallies came to an abrupt end. Rao (2021) finds that the hedge funds that had taken short-positions in GME lost at least $12.5 billion as a result of the short squeeze. Prominent GME short-sellers, such as Citron Research and Melvin Capital, received extensive negative coverage in international financial media and became known for having lost tremendous amounts of money as a result of the short squeeze (see e.g. Canny, 2021; Kumar, 2021; Winck, 2021). Investors holding long positions in the GME stock during the squeeze are also expected to have incurred large losses when the GME rally swiftly ended (Brown, 2021). Overall, therefore, while some managed to get rich as a result of the short squeezes, others took on large losses and were subject to negative exposure in international financial media.

3 Literature review

3.1 Price manipulation and retail investor affection

Joint coordinated efforts to influence a security's price without releasing falsified information or taking actions meant to change the security's value, e.g. as recently observed in the GameStop stock, is a form of trade-based price manipulation (Allen & Gale, 1992). In a study of stock market manipulation in the United States, Aggarwal and Wu (2003) found that illiquid stocks are more susceptible to manipulation than liquid stocks. They also found that manipulation increases stock volatility, and that over 50% of manipulated stocks in the US stock markets are "penny stocks" with low market capitalization and low average trading volume (Aggarwal & Wu, 2003). Using data from the Istanbul Stock Exchange, Imisiker and Tas (2013) found that firms with low market capitalization, as well as firms with lower free float rate are more likely to be manipulated than larger firms. Higher leverage ratio is also found to have a significant and positive effect on the probability of being manipulated (Imisiker & Tas, 2013).

Han and Kumar (2013) show that high idiosyncratic volatility, skewness and lower prices are stock-specific attributes attracting a higher proportion of retail trading in a stock. Han and Kumar (2013) also find that retail investors exhibiting a greater propensity of speculation and gambling are more likely to engage with trading in stocks with a higher proportion of retail trading. Kumar and Lee (2006) show that lower priced firms, small firms, firms with lower institutional ownership and firms with high book-to-market ratios are displaying higher retail investor concentration and trading activity. Gao and Lin (2015) find that retail investors prefer stocks displaying high past returns, low market capitalization, low earnings per share, high market-to-book ratio and high retail trading concentration.

In a study using trading data from Taiwan, Gao and Lin (2015) find that stock market trading may function as a substitute for gambling activity. Defining "lottery-like" stocks as those with high skewness, they find that these stocks are more prone for trading as a

means for fun and excitement among individual investors. Kumar (2009) and Li (2012) also argue that retail investors use stock markets as a substitute for gambling. Using data from Taiwan, Barber et al. (2009b) also find that retail investors' gambling desire encourages investors to enter the stock market. They propose that a combination of the desire to gamble and the level of overconfidence may explain much of the active trading activity of individual investors (Barber et al., 2009b). Grinblatt and Keloharju (2009) also find that overconfidence and sensation seeking are personal traits increasing stock market trading volumes.

Guiso et al. (2008) define trust as "the subjective probability individuals attribute to the possibility of being cheated". They find that individuals displaying high levels of trust to the financial system are more likely to buy stocks and other risky assets, and that, conditional on investing in stocks, such individuals are more prone to invest a larger share of their wealth in the stock market (Guiso et al., 2008). Chiah and Zhong (2020) find that in societies with higher levels of trust and individualism, investors tend to trade more heavily. They also find that investors living in countries with stronger protection of legal rights, sounder governance systems, higher national wealth and greater gambling opportunities are more willing to trade than others (Chiah & Zhong, 2020). Further, they find that investors use stock markets for reasons such as fear-of-missing-out, quick-to-get-rich-schemes, gambling and bargain hunting, and connect this with the significant increase in investors using the Robinhood platform as their trading platform of choice in the United States (Chiah & Zhong, 2020).

3.2 Retail investors' ability to move stock prices

If, and to what extent, retail investors have the ability to move stock prices has for long been a hot topic among researchers. Many have speculated whether the recent stock rallies actually were driven by retail investors, where many have argued that retail investors cannot account for such large movements in stock prices as have been observed over the last couple of months, but that institutional investors have been monitoring sentiment at the WallStreetBets forum and imposed trading strategies thereof. Nevertheless, several

studies show that retail investors indeed can move stock prices. Barber et al. (2009a) find that retail investors can move stock prices, and that the movement's direction is determined by the direction of retail investors' net trades. Kumar and Lee (2006) also find that collective trading action of retail investors can move stock prices. They also show positive systematic correlation between retail investors' trades, i.e. that retail investor trading displays a tendency of aggregating across individual investors, and that this may move stock prices (Kumar & Lee, 2006). Another study finds that in an environment with heavy aggregated selling by retail investors, this may move stock prices even oppositely to institutional investors' net trading direction (Burch et al., 2016). Other studies also find strong systematic correlation in the aggregate retail investors' trading activity (Dorn et al., 2008; Jackson, 2003).

Foucault et al. (2013) argue that there are especially three explanations of why retail investors' trades may be correlated. First, they argue that this may reflect a correlation in retail investors' liquidity needs. Such needs may be urged by macroeconomic shocks such as economic recessions where massive layoffs force households to liquidate their asset holdings simultaneously (Foucault et al., 2013). Second, they argue that retail investors may display "herd behavior", meaning that investors are imitating one another's trading strategies (Foucault et al., 2013). Third, they argue that retail investors may follow momentum strategies, which in general is buying at the same time when an asset's price is rising and equivalently selling at the same time when an asset's price is falling (Foucault et al., 2013). De Long et al. (1990) coined investors adhering to such strategies "positive feedback traders".

Without having access to stock market order flows we have not been able to identify which orders are coming from retail and institutional investors. By using the insights from these studies, we would nonetheless claim that retail investors can actually be claimed to have contributed to movements in stock returns during the recent rallies. Although we do not seek to distinguish between who actually have been driving the recent significant stock movements in the stocks we concentrate on in this thesis, we would argue that it is important to understand that retail investors in fact do carry the

potential to create such price movements as the stock markets have recently witnessed.

3.3 Investor sentiment

Although there is no single commonly accepted definition of sentiment, most definitions grasp that sentiment is related to individuals' feelings, attitudes and thoughts considering something. By sentiment we therefore adhere to Cambridge Dictionary's definition of sentiment, which captures many of these different aspects of sentiment, whereby they define sentiment as "a thought, opinion, or idea based on a feeling about a situation, or a way of thinking about something" (Cambridge Dictionary, n.d.).

Whether investor sentiment can explain stock price movements has for long been subject to extensive academic research. Shleifer and Summers (1990) argue that investor sentiment is indeed an important determinant for changes in stock prices, meaning that changes in sentiment both can rationalize stock market rallies and market setbacks, such as the rallies and following setbacks we seek to explore in this thesis. Baker and Wurgler (2006) also find that stock prices and trading volume are heavily influenced by sentiment. Furthermore, they find that stocks of firms being unprofitable, younger, having low market capitalization, being non-dividend paying, having high volatility, being growth companies or firms being in financial distress are more sensitive to investor sentiment than others (Baker and Wurgler 2006). Many of these characteristics can be observed in the stocks we are focusing on in this thesis. We would therefore argue that the intense hyping of several stocks at WallStreetBets have created positive sentiment for the stocks subject to the hyping, which is the core feature we test whether has had a statistically significant effect on stock returns. Hong and Stein (2007) also find that stock prices and trading volumes are influenced by investor sentiment, showing that stocks subject to positive sentiment are more likely to display higher trading volume and are more likely to be overvalued by investors, just as what we believe was the case for the targeted stocks on WallStreetBets.

3.4 Short-selling, short squeezes and gamma squeezes

What gave rise to the short squeezes we discuss in this thesis was naturally that the stocks subject to short squeezes were shorted in the first place. Many studies have been devoted to finding why stocks are shorted, whereas Dechow et al. (2001) find that stocks are shorted as a result of low fundamental ratios. Asquith et al. (2005) find that individual stocks have a high short interest either because some investors find the stock to be overvalued, or because some investors find that convertible bonds issued by the company are undervalued, giving rise to arbitrage. Brent et al. (1990) argue that investors may hold short positions on stocks they hold long to defer taxes on capital gains. They also find that stocks with high betas and connected options and convertible securities tend to display higher short interest ratios (Brent et al., 1990). Further, Nagel (2004) finds that short interest is significantly correlated with a stock's ratio of institutional ownership.

A stock's short interest ratio can exceed 100%, which was observed in e.g. the GME stock prior to the burst of the GME bubble in January/February 2021, which can happen for particularly two reasons. First, it may happen if shares sold short are borrowed by another party and sold in the market again (Asquith et al., 2005). Second, it may happen as a result of naked short selling, which is to sell shares short without having borrowed or arranged to borrow shares (Boulton & Braga-Alves, 2009). In extreme cases, naked short sellers may sell shares short that do not even exist (Angel & McCabe, 2009). Naked short selling can be used as a means to manipulate a stock's price (Christian et al., 2006) since aggressive naked short selling may lead to abnormal selling pressure on the shorted stock, pushing the stock's price downwards (Finnerty, 2005). Abusive short sale practices are considered illegal, meaning that using short-selling strategies to manipulate the price of a stock is considered illegal (SEC, 2015).

Short squeeze is a term used to describe a situation where an initial price increase of a stock forces investors with short positions in the stock to cover their short positions, either as a result of covering losses or margin calls (Lamont, 2012, p. 21). Forced to cover their short positions, the short-sellers who initially betted on a price decrease are "squeezed"

out of their short positions, meaning they have to start buying the stock which further fuels the price increase (Xu & Zheng, 2016, p. 1). The short-seller becomes increasingly vulnerable to a short squeeze the larger the total short position the seller takes (Xu & Zheng, 2016, pp. 1-2). A short squeeze often implies that the underlying stock's price is displaying a pattern of rapid and significant increase, which then is followed by a rapid and significant decrease.

Instead of trading in stocks directly, investors can trade indirectly in stocks by trading in options with the specific stock as the underlying asset (Black, 1975). This was frequently observed during the multiple short squeezes we explore in this thesis, as investors entered these stock rallies both by buying both stock and call options. Call options are usually specified for 100 shares of the underlying stock (Corrado & Su, 1997, p. 79), meaning that by buying such call options, investors engage in bets that enable them to increase their potential gains and losses manyfold compared to investing the same dollar amount in the stock directly (Pedersen, 2021, p. 32).

Market makers selling call options to investors will buy some number of the underlying stock to hedge their positions (Armstrong, 2021). As investors joined in on the recent rallies, our data shows that call option volumes spiked, meaning that market makers had to significantly increase the number of shares bought meant for hedging their risk exposure. When the price of an option's underlying stock increases, market makers will start buying more shares of the underlying stock, as a higher price on the underlying stock translates into higher risk for the market makers having sold the options. As the price of the underlying stocks in the recent rallies we explore in this thesis eventually increased significantly, e.g. as happened with the GME stock, market makers had to start buying even more shares, which further increased the already significantly heightened demand price pressure on the inflicted stocks. This contributed to pushing the price of the underlying stock even further upwards, meaning that they had to start buying even more shares. This type of an upward price spiral leading to market makers having sold call options have to buy more and more shares as the stock price increases, which further pushes the stock price upwards, is called a gamma squeeze (Pedersen, 2021, p.

32), which we now will dig further into.

Option traders are concerned with so-called option Greeks, i.e. factors that affect the price of option contracts (Frederick, 2020), where gamma is one such factor. Option Greeks are calculated using option pricing models and are important determinants for what trading operations option traders would undertake. To understand gamma, one must first understand another option Greek, i.e. delta. Delta is commonly referred to as the hedge ratio. The delta factor shows how much the price of an option contract is expected to change if the price of the underlying stock changes by USD 1 (Frederick, 2020). If the delta of an option is 0.20, the price of the option is expected to change by USD 0.20 per USD 1 the price of the underlying stock changes. The delta factor can also be utilized for hedging purposes. If you have a short position in a call option on a stock with a delta of 0.20, you could hedge your position by buying 20 shares of the underlying stock (since option contracts normally are specified for 100 shares (Corrado & Su, 1997, p. 79)). Such a hedging strategy is called a delta hedging strategy.

Gamma is a factor measuring the rate of change in the delta factor, i.e. it explains how delta is expected to change if the price of the underlying stock changes (Summa, 2021). The properties of gamma can therefore be used by option traders to calculate expected price movements in the future. Delta is constantly changing as the price of the underlying stock changes, whereas gamma is constant (Frederick, 2020). Gamma hedging is a strategy which seeks to reduce the risk for significant stock price movements, where the trader tries to create a delta-neutral position (Scott, 2021). Such strategies are particularly in the trader's interest to pursue in the last days before the option expires, since the time value of the option contract by then is almost completely eroded (Scott, 2021).

To illustrate how gamma squeezes such as those we explore in this thesis can erupt, let us assume that WallStreetBets users start buying significant amounts of far out-of-the-money call options in addition to the underlying stock. Even though the call options are far out-of-the-money, the market makers having sold the options need to hedge themselves by buying the underlying stock, i.e. they utilize delta and gamma hedging strategies. If

the price of the underlying stock starts increasing, so does the accumulated amount of shares bought for hedging purposes by the market makers. When using the word gamma squeeze, what now should be clear is that as market makers sell option contracts and the price of the underlying stock increases, they will start buying more and more shares to hedge their positions whereby delta and gamma are two factors explaining how many shares to buy for hedging purposes and the anticipated change in these strategies over time.

As became evident during the recent squeezes, a gamma squeeze has the potential to contribute significantly to increases in a stock's price if the call option volume is sufficiently high. Gamma squeezes may both create short squeezes themselves or coexist with traditional short squeezes (Armstrong, 2021). Short squeezes and gamma squeezes have previously been considered as infrequently observed phenomena (Liu & Xu, 2016), but, as we explore in this thesis, short squeezes and gamma squeezes became more prevalent as several stocks became subject to these mechanics in 2020 and 2021.

3.5 Short squeezes and gamma squeezes in a financial bubble perspective

Most academic definitions of financial bubbles share several distinct characteristics, such as substantial price deviations from aggregate fundamentals (see e.g. DeMarzo et al., 2008, p. 25), significant upward price movements that eventually implodes (see e.g. Kindleberger, 1978, p. 16) and that financial bubbles are characterised as a period of unsustainable accelerating growth (see e.g. Sornette and Cauwels, 2014, p. 1). These characteristics are in our opinion similar to those observed in short squeezes and gamma squeezes. We would therefore argue that short squeezes and gamma squeezes are types of financial bubbles.

Financial bubbles come in many forms, where some serve as classic examples where the underlying asset's value is subject to significant value appraisal, whereas in other bubbles, such as those we are exploring in this thesis, the upward price movement comes

as a result of intentional trade-based price manipulation. Economic history is full of episodes being characterized as financial bubbles, even though economists disagree to some extent as to which historical financial bubbles actually were bubbles (see e.g. Day, 2004, pp. 151-152; Garber, 1990, p. 35). The most well-known classic example of a financial bubble is arguably the Dutch tulip mania (Garber, 1990, p. 35). During the Dutch tulip mania from 1634 to 1637 tulips became the hottest craze, and the prices of tulip bulbs skyrocketed (Garber, 1990). At the peak of the mania in 1635, one single tulip bulb could sell for up to 5,500 florins, which implies a present price of as much as USD 76,000 (Hirschey, 1998). In February 1637 the prices of tulips suddenly collapsed, and it was hard to sell tulips for even a tenth of what they were worth at the peak of the mania (Garber, 1990, p. 37). Following a lengthy period of significant appraisal in the price of the tulips, the tulip mania had come to an abrupt end.

Unlike classic bubbles, the recent short squeezes and gamma squeezes we are seeking to explore in this thesis were, at least partially, intentionally created. Investors specifically targeted stocks seeking to inflate the stocks' prices by trade-based price manipulation. This trade-based price manipulation took the shape of coordinated trading efforts by investors aimed at stocks displaying certain specific characteristics such as a high short ratio. The intent of these coordinated efforts has been to push the price of the subject securities upwards with the intent of trying to squeeze out investors with short positions in the inflicted stocks, thereby further inflating the securities' prices.

Neither this kind of a financial bubble is a new phenomenon. One of the largest short squeezes in economic history to the time of writing is the Volkswagen (VW) short squeeze in 2008 (Allen et al., 2019). On September 25, 2005, Porsche announced their intention to acquire a stake of almost 20 percent in VW's voting capital (Porsche, 2005). Later, on October 26, 2008, Porsche publicly announced that they had acquired control of as much as approximately 74.1 percent of VW's voting capital through both common stock and synthetic derivative positions (Porsche, 2008a). At the same time, Lower Saxony controlled 20 percent of the voting capital (Allen et al., 2019, p. 9), meaning that only 5.9 percent of the VW voting stock was floating freely in the market. Also, at the same

time, a short position of approximately twelve percent had been built up in VW (Allen et al., 2019, p. 9).

What followed was a significant price increase in VW's stock. On October 27, 2008 the share price opened at EUR 350, a 66 percent increase compared to the price at the previous close, and the next day it rose almost 150 percent further to EUR 520 at close (Allen et al., 2019, p. 10). This massive price increase meant that the holders of the amassed short positions of around twelve percent suddenly found themselves caught in a short squeeze and had to start covering their short positions by buying VW's stock (Allen et al., 2019, p. 10). This fueled the price increase even further, and eventually led to that on the next day, i.e. October 28, 2008, VW suddenly had become the company with the biggest market capitalization in the world (CNBC.com, 2008).

The short squeeze became devastating for the short sellers, and the hedge funds that had shorted the VW stock lost approximately $30 billion as a result of the short squeeze (Ramey, 2021). The next four days, VW's share price fell by around 58 percent, and one month later the price had fallen by around 70 percent from its peak during the short squeeze (Li, 2021). This came both as a result of having squeezed out the short sellers (Allen et al., 2019, pp. 9-10) and that Porsche had started unwinding their synthetic ownership positions (Porsche, 2008b). The magnitude of this episode illustrates the forces at play in short squeezes, and therefore serves as an example of how dramatic short squeezes potentially can be.

4 Methodology

In this chapter, the models, methods, and processes used in this thesis are illustrated. The methods used are what has been done regarding text mining, text processing and sentiment scoring, while the models explain how our results are generated.

4.1 Data collection

In this thesis, we have combined a variety of datasets to perform our analyses. The textual dataset contains all posted submissions and comments on the WallStreetBets forum from January 1, 2020, to March 15, 2021. Since Reddit's API does not allow direct access to this data, we have used a python script to extract this data from pushshift.io, an open-source Reddit database that collects daily data from most subreddits. We then ran the script through the Windows command prompt, scraping all submissions from WallStreetBets over our given time period. With pushshift.io being an open-source database run by Jason Baumgartner, there were limitations to how many requests per minute the database could handle, thus the process was time consuming. An unfortunate consequence of the database was that some days of WallStreetBets data was missing, but overall we are grateful to Baumgartner for singlehandedly creating the database, since without it, this thesis would not have been achievable. The missing dates are shown in table 4.1.

Table 4.1: Missing dates

11/5/2020
1/25/2021
1/26/2021
2/4/2021
2/5/2021
2/6/2021
2/28/2021
3/1/2021
3/6/2021

The dataset contains 1 203 556 unique submissions and 27 612 806 unique comments.

Given the relatively large number of submissions and comments and related attributes, the downloading process of this data alone took over two months. The submissions and comments are given a vast range of attributes in this dataset, where most attributes are purely technical attributes used for database and website purposes. We have therefore chosen to include only those we deem useful for textual analysis purposes.

The attributes we have chosen to use in the submissions part of the dataset are structured as follows: id; selftext; title; created_utc. Each submission is given a unique ID, which can be used for identifying purposes, as well as linking comments to the submission at which they were posted. selftext is the textual content underlining the title if the submission is a text post rather than one containing a picture or a link. title is the title of the submission. created_utc shows the specific time the submission was posted, and is on UTC-format.

The attributes we have chosen to use in the comment part of the dataset are structured as follows: body; created_utc; parent_id. body is the actual textual comment. created_utc shows the specific time the comment was posted and is on UTC-format. parent_id is the ID of the submission to which the comment was posted. To be able to use sentiment analysis techniques on comments as well as submissions, we have added an attribute to each comment, Com_ID, which is an unique ID that enables identification of each comment, just as each submission is given a unique ID.

Daily stock prices and trading volumes from January 02, 2020, to December 30, 2020, are collected from the CRSP database. Daily stock prices and trading volumes from January 04, 2021, to March 15, 2021, are collected from Yahoo! Finance.

Intraday prices and trading volumes, implied volatility, call option volumes, put option volumes and short ratios for each stock from January 02, 2020, to March 15, 2021, have been collected from Bloomberg.

Data for the Barclays and Goldman Sachs long and short baskets are collected from Bloomberg. Data from Russell 3000 is collected from Bloomberg.

The excess return on the market portfolio, i.e. Mkt.RF, as well as the Fama-French

factors SMB, HML, RMW, CMA and Mom have been collected from the French data library (French, 2021).

4.2 Data filtering

Due to the large variety of stocks discussed on WallStreetBets, decisions had to be made regarding which stocks to include in our thesis. A selection process was induced where the stocks most frequently mentioned on the forum as well as a random sample of stocks experiencing surges during our research period were chosen. There were also two control stocks introduced with AAPL and GOOG being large tech stocks, the sector with the most attention on the forum, whom were also frequently discussed. On the basis of this selection process, this thesis will focus on the 18 stocks listed in table 4.2 from January 01, 2020, and March 15, 2021.

Table 4.2: Sample of stocks

Ticker	**Stock Name**
AAPL	Apple, Inc.
AMC	AMC Entertainment Holdings, Inc.
AMD	Advanced Micro Devices, Inc.
APHA	Aphria, Inc.
BB	BlackBerry Limited
GME	GameStop Corp.
GOOG	Alphabet, Inc.
NIO	NIO, Inc.
NKLA	Nikola Corporation
NOK	Nokia Corporation
NVDA	NVIDIA Corporation
PLTR	Palantir Technologies, Inc.
PLUG	Plug Power, Inc.
RKT	Rocket Companies, Inc.
SPCE	Virgin Galactic Holdings, Inc.
TLRY	Tilray, Inc.
TSLA	Tesla, Inc.
TSM	Taiwan Semiconductor Manufacturing Company Limited

The text process then begins with the filtering process. In the first step all posts' Submission Title is scanned for the company name, company ticker and other company pseudonyms. For example, for Alphabet, Inc. this would be Alphabet, Inc., Alphabet, Google and GOOG. The process was then repeated for each submission's text, and finally each individual comment was scanned for company related keywords. Due to the fact that each title, submission text or comment could contain discussion about more than one stock, the ID of the comment or submission was stored alongside the Ticker identified in a wide format. This was done to prevent duplicate entries in our dataset.

The dataset was then filtered out by ID to contain only the comments and submissions containing our stocks of interest, merged to a single data frame consisting of all comments and the matched submissions with the associated tickers, reducing the overall size of the data to 6 740 104 rows.

The final step before sentiment analysis could be performed is known as preprocessing and data cleaning. First, all html links are removed from the dataset as they only provide noise, then stemming and lemmatization is performed to reduce words to their root form. An example of this would be to transform playing, played and plays to play. This practice reduces the amount of word alternatives, but keeps the actual words intact. Next, all common stop words in English are removed as these do not contain any information relevant for the sentiment analysis. Finally, all words are transformed to lowercase and all excess whitespace between words are removed. The main benefit of this process is to reduce the actual size of the data without reducing the information value of the content, thus reducing the computational strain when working with big data. With less noise in the textual data, the sentiment analysis will also perform better than if no preprocessing was performed.

4.3 Sentiment analysis

Sentiment analysis is a term used to describe analytical techniques that identify and measure the sentiment of what is subject to analysis. To analyze whether sentiment on

the WallStreetBets forum has had an impact on various stock and option metrics, we have applied sentiment analysis techniques on all submissions and comments posted on WallStreetBets from January 1, 2020, to March 15, 2021. Our goal of applying sentiment analysis is to gauge the sentiment around specific stocks on WallStreetBets, and see whether positive, neutral or negative sentiment or fluctuations in these have any impact on stock prices. The sentiment analysis process in this thesis is shown in figure 4.1.

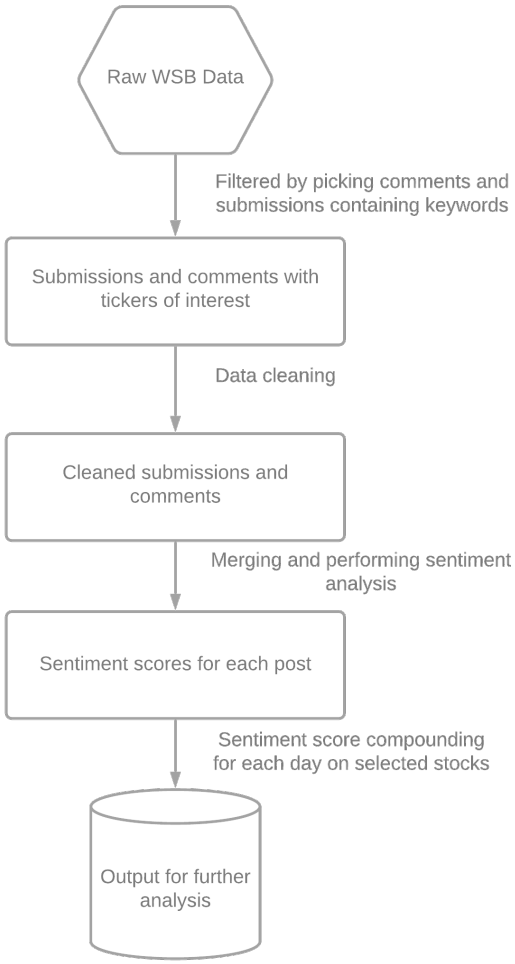

Figure 4.1: Sentiment analysis process

To clarify the different distinctions between the terms submissions, submission title, submission text and comments, below follows a screenshot of a submission on WallStreetBets by the user SuperShortSqueezer (2021), which graphically indicates

what we categorize as submission, submission title and submission text:

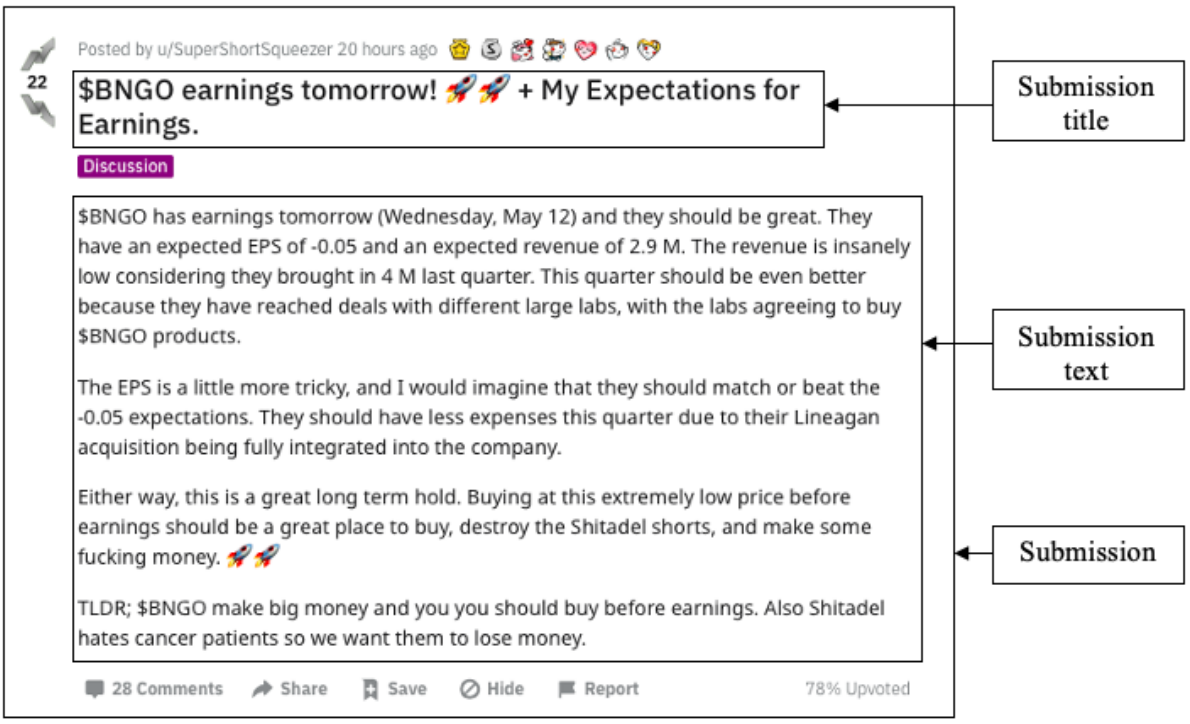

Figure 4.2: Submission example

To illustrate how comments to submissions on WallStreetBets look like, below follows a screenshot of some of the comments to the submission in figure 4.2 above. As is evident in the screenshot, while the submission seems to display a positive sentiment for BNGO, the following comments show that individuals' sentiment around BNGO differ.

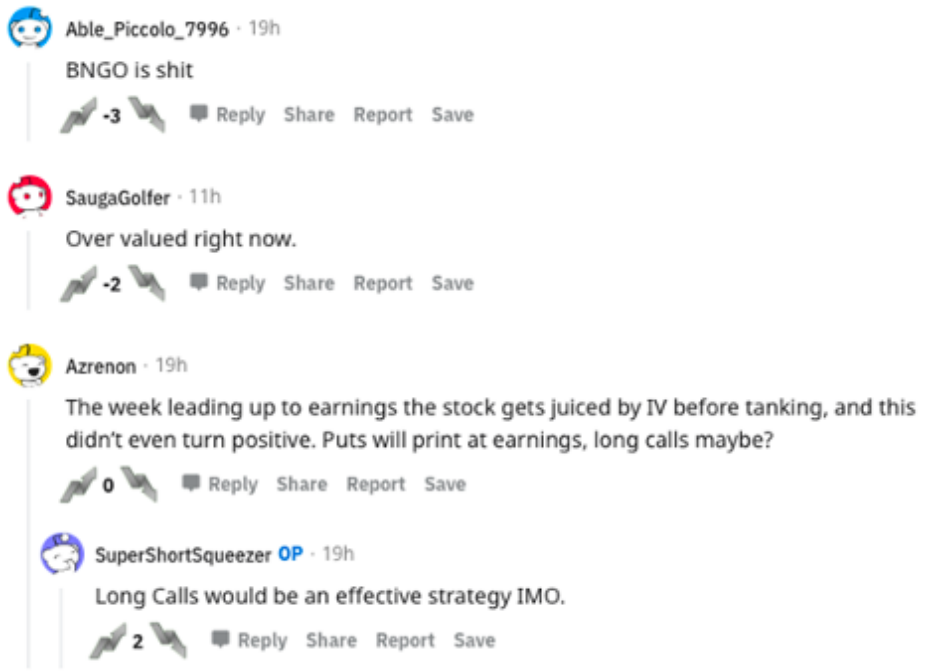

Figure 4.3: Comment example

When defining our sentiment variables, we wanted to create a measure of sentiment of each of the stocks subject to analysis in this thesis based on the textual content in each submission and comment posted on WallStreetBets in this time interval.

One alternative to analyze sentiment could be to use an attribute called "score" in the original dataset that we retrieved from pushshift.io. Score is based on a function on Reddit, where each post can be given an upvote or a downvote by users. Each upvote counts as +1, whereas each downvote counts as -1. Score shows the net sum of upvotes and downvotes and is therefore indicative of sentiment revolving around posts on WallStreetBets. However, due to inconsistencies based on when pushshift.io saves comments and submissions at various periods during the day, comments that could receive thousands of upvotes can potentially be stored in the database as having a score of only 1, meaning that the reported score does not reflect the true score for each post, we have chosen not to incorporate or utilize the score attribute further into our dataset or analysis.

4.3.1 Choice of sentiment method

In R there are many packages available for sentiment analysis. The sentimentanalysis package contains dictionaries like Harvard IV and Loughran-McDonald's "Master Dictionary". Despite this, due to the informal and sometimes rather vulgar language on the forum, both dictionaries gave significantly negative scores on all aggregated posts. The VADER sentiment package, an open-source rule-based sentiment analysis package and lexicon invented specifically for text data on social media, was also an alternative. Unfortunately, most sentences were again classified as negative and all existing sentiment analysis packages based upon dictionaries were thus decided against.

The next alternative was to investigate machine learning models, specifically supervised ML algorithms trained on the comments. However, due to the low signal-to-noise ratio in the comments, as well as linguistic changes that occurred during the peak of WallStreetBets activity around the end of January 2021, a different approach was adopted.

In order to have as much control as possible on the output from the sentiment analysis, a WallStreetBets dictionary was created. From our original dataset containing 27 612 806 unique comments spanning from January 1, 2020, to March 15, 2021, the data was cleaned and the 15 000 most frequently used words were extracted. Scores were assigned from -1 to +1 on words that were either clearly positively or negatively charged, ignoring words that could be used in both contexts. Due to the language used on the forum, words that are usually negatively charged could be used in both a positive and negative fashion such as a WallStreetBets favourite; "ape", a word used both as a term to signal a kinship with the forum and as an expression for a user doing an apelike trade, were subsequently not used in our dictionary.

A sample from the dictionary output is shown in table 4.3.

Emojis were encoded in Unicode and had to be transformed to an interpretable word, where for example "ðŸš€" were converted to "rocketemoji". Emojis were frequently used

Table 4.3: Dictionary output

Word	Score
short	- 1
long	+1
loss	- 1
rocketemoji	+1
tendi	+1
bad	- 1
yolo	+1
crash	- 1
miss	- 1

during the rallies where "paperhands" were used as a negative term for people who sold, whilst the "rocketemoji" and "diamond hands" emojis were used to signal either a stock ready to soar or particularly for the GME-related short squeezes having diamond hands and holding the stock through price descents.

4.3.2 Sentiment scoring

From the filtered dataset, a ticker was assigned to the comment if the comment ticker was missing, but the associated submission had a ticker either in its title or submission text. This was performed since a comment, such as the second comment in figure 4.3, could talk about a stock discussed in a submission without explicitly mentioning the ticker in the comment. The filtered dataset was then converted to a long format where each comment, submission text or title would only have one ticker associated with it. The next step was to create a column with which days the stocks were talked about. Due to American stock exchanges closing at 21:00 UTC, all comments and submissions posted after 21:00 UTC in the dataset were transformed to the next day as it would not impact current day stock activity.

With the dataset ready for the sentiment analysis, each stock was extracted individually, merged with the dictionary tallying each occurrence of positive or negative words in the comments, and given a sentiment value. For example, a comment including two

words defined as displaying positive sentiment, i.e. “buy” and “undervalued”, e.g. saying “buy GameStop, it is undervalued” would be given a score of +2 and indicate positive sentiment for GME, whereas a comment saying “sell GME, it is overvalued” would have a score of -2 and indicate negative sentiment for GME. Neutral comments are comments that either do not indicate any emotionally charged language, or comments that are 50% positive and 50% negative to a stock, which would have a score of 0. For example, a neutral comment could say “On one side I want to buy GameStop, but on the other side I think it is overvalued” or simply “I neither want to buy or sell GameStop”.

Aiming to have a normalized measure of sentiment that did not depend on the frequency of mentions, a decision was made to look at sentiment direction instead of numerical sentiment scores. Thus, sentiment score for the comment, title and submission text on day t is given by:

$$Percentage\ Positive\ Sentiment_t = \frac{\sum Positive\ Sentiment_t}{\sum Negative\ Sentiment_t + \sum Positive\ Sentiment_t} * 100 \quad (4.1)$$

where positive sentiment is a comment, submission title or submission text with a sentiment score above 0, and negative sentiment is a comment, submission title or submission text having a sentiment score below 0. As neutral comments with a score of 0 does not indicate positive or negative sentiment, they were not included in the calculations. Subsequently, percentage positive sentiment for comments a day with 80 positive comments and 20 negative comments would be 80% for the day. Daily occurrences with neither positive or negative comments were given a neutral sentiment score of 50% as these were rare, but a score of 0 or NA would either alter our findings or remove entire rows of data from the regressions.

Sentiment was also calculated with a similar process on a minute-by-minute basis on selected stocks, where instead of aggregating sentiment per day, it was calculated per minute.

4.3.3 Final dataset

The final step in preparation for the regressions is to create datasets for individual stocks. Based on the stock prices gathered from CRSP and Yahoo! Finance we have calculated daily returns for each stock. We have chosen to use logarithmic returns because logarithmic returns better capture aspects such as log-normality, time-additivity and numerical stability than arithmetic returns. We have then subtracted the daily logarithmic returns we have calculated by the risk free rate. Daily logarithmic returns are calculated as follows, where p_t is the price of the stock at day t:

$$logarithmic\ daily\ return = r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \tag{4.2}$$

We wanted to include a measure of liquidity in our analyses. Based on the daily logarithmic returns, we have added Amihud's illiquidity factor for each stock daily from January 02, 2020, to March 15, 2021, in our dataset. This is calculated as follows, where p_t is the price of the stock at day t and vol_t is the trading volume for the stock at day t:

$$AMIHUD = \frac{r_t}{p_t * vol_t} \tag{4.3}$$

Further, our dataset contains implied volatility, call and put volume, daily volume and price data as well as the Fama-French factors SMB, HML, RMW, CMA and Mom. This daily data was then merged with our constructed sentiment variables; Comments per day, Positive sentiment comments, Title per day, Positive sentiment title, Title text per day and Positive sentiment title text.

We decided against utilizing WallStreetBets data gathered from the weekend, as forum activity significantly dropped off during this period, and as we wanted to assess how activity translated to stock movements, looking only at weekday interactions was opted for instead of aggregating weekend activity towards the next week.

Finally, to get a better assessment of whether stock movements were driving forum activity or if forum activity was driving the stock movements, lag and lead variables were

created for both returns and volume.

4.4 Variable definitions

In chapter 5, as well as in chapter A1 and A2 in the appendix, several regression tables are shown. The tables contain a set of different variables, whereby the following lists explain the variables.

Dependent variables

Return	The daily logarithmic return of the stock minus the risk-free rate.
Volume	The daily number of shares traded in the stock.
Implied Volatility	Implied volatility. Calculated as the weighted average of the volatilities of the two call options closest to the at-the-money strike price.
Call Volume	The daily number of traded call options with the specific stock as the underlying asset.
Put Volume	The daily number of traded put options with the specific stock as the underlying asset.
AMIHUD	Amihud's illiquidity ratio.
$\mathbf{Return_{t-1}}$	The stock's logarithmic return minus risk-free rate on day t-1.
$\mathbf{Return_{t+1}}$	The stock's logarithmic return minus risk-free rate on day t+1.
$\mathbf{Return_{t+2}}$	The stock's logarithmic return minus risk-free rate on day t+2.
$\mathbf{Return_{t+3}}$	The stock's logarithmic return minus risk-free rate on day t+3.
$\mathbf{Return_{t+4}}$	The stock's logarithmic return minus risk-free rate on day t+4.
$\mathbf{Volume_{t-1}}$	The stock's trading volume in number of shares on day t-1.
$\mathbf{Volume_{t+1}}$	The stock's trading volume in number of shares on day t+1.
$\mathbf{Volume_{t+2}}$	The stock's trading volume in number of shares on day t+2.
$\mathbf{Volume_{t+3}}$	The stock's trading volume in number of shares on day t+3.
$\mathbf{Volume_{t+4}}$	The stock's trading volume in number of shares on day t+4.

Independent variables

Comments per day	Total number of comments per day considering the stock subject for regression.
Positive sentiment comments	Percentage share of the comments considering the the specific stock categorized as displaying positive sentiment.
Title per day	Total number of submissions titles per day containing the stock's name or stock ticker.
Positive sentiment title	Percentage share of the submission titles containing the specific stock's name or stock ticker categorized as displaying positive sentiment.
Title text per day	Total number of submission texts per day containing the stock's name or stock ticker.
Positive sentiment title text	Percentage share of the submission texts containing the specific stock's name or stock ticker categorized as displaying positive sentiment.
Mkt.RF	Daily excess logarithmic return on the market portfolio. Constructed by calculating logarithmic returns using daily return on the Fama-French Mkt.RF factor. Mkt.RF is the "value-weighted returns of all NYSE, AMEX and NASDAQ firms that have a CRSP share code of 10 or 11 at the beginning of month t, good shares and price data at the beginning of t, and good return data for t minus the one-month Treasury bill rate (from Ibbotson Associates)" French (n.d.).
SMB	Small-minus-big Fama-French factor. Calculated as the average return on small stock portfolios minus the average return on big stock portfolios.
HML	High-minus-low Fama-French factor. Calculated as the average return on high book-to-market portfolios, i.e. value firms, minus the average return on low book-to-market portfolios, i.e. growth firms.
RMW	Robust-minus-weak Fama-French factor. Calculated as the average return of robust operating profitability portfolios minus weak operating profitability portfolios.
CMA	Conservative-minus-aggressive Fama-French factor. Calculated as the average return on conservative stock portfolios minus aggressive stock portfolios.
Mom	Momentum Fama-French factor. Calculated as the average return on high prior return stock portfolios minus the average return on low prior return stock portfolios.

4.5 Regression models

In order to assess the impact of WallStreetBets on various stock metrics, we utilize regressions in order to capture the statistical relationship between the variables. We use the Fama-French Mkt.RF factor as the control variable for excess return on the market portfolio as well as the benchmark index in the thesis. The Nasdaq Composite index is often used as a control variable and/or benchmark index, but this index is very technology-heavy. Given the relative sector-spread in the stocks in the sample subject for analysis in this thesis, we have chosen to use the Fama-French benchmark factor Mkt.RF as the benchmark index. We have therefore used this benchmark factor as a control variable in the regression tables in chapter 5, as well as the benchmark index where this is indicated in chapter 5.

To ensure heteroscedastic-consistent estimations in the regressions, we use vcovHC as our standard error estimator to ensure heteroskedasticity robust standard errors in all our regressions.

We have chosen to perform analyses on both the time interval overall as well as dividing the time period into two subsections, i.e. 2020 and 2021. The foremost reason for this division is that we want to analyze whether the January 2021 rallies changed the mechanics driving the different rallies. Given the magnitude of the January 2021 rallies, by dividing the period into two subsections this makes us able to analyze potential differences in the rallies and thereby analyze whether the rallies in 2021 were structurally different to the rallies in 2020. Another reason is that several of the stocks that caught WallStreetBets' attention in 2020 did not catch WallStreetBets' attention during the multiple rallies in 2021. By separating the two periods, we manage to perform more in-depth analyses of the different sets of stocks in the two sub-periods. Further, the magnitude and implications of the two sub-periods differ greatly. Even though the stocks that caught WallStreetBets' attention during the first sub-period, i.e. 2020, were subject to heavy discussion, the total level of activity was greatly surpassed by the total level of activity in the second period, i.e. 2021. To better be able to analyze the first rally

episodes that WallStreetBets was accused of having caused, a separation of the two periods facilitates a more comprehensive analysis of the first of these episodes.

Another reason for the split is that the activity in the two different sub-periods took different forms. While the activity on WallStreetBets in the first sub-period centered around discussions where a select few stocks were the flavor of the month, the second period was characterized by discussion revolving almost solely around those subject to a potential short squeeze, and with a much more aggressive approach. The jargon on the forum thereby changed from the first period to the second, and a division of the two periods then facilitates a better ground for comparison and analysis of the change in magnitude and forces that evolved from the first period to the second.

The regressions created for 2020 and 2021 are set up the exact same way as the stock regressions outlined below, but with only data from 2020 and 2021 respectively. The data for 2021 alone naturally contains fewer observations than the data for 2020, since we analyze data for 2020 from January 02 to December 31, while the data for 2021 contains a larger portion of missing data, and only ranges from January 04 to March 15, i.e. 44 observations. This generally lower number of observations in 2021 can carry potential distorting effects into the regressions, so we advise to bear in mind the differing number of observations when comparing the different analyses for 2020 and 2021.

4.5.1 Individual regressions

For the stock-individual regressions, we estimate the effect of the sentiment variables, the return on the market portfolio and the Fama-French factors on stock returns, trading volume, implied volatility, call option volume, put option volume and the Amihud illiquidity factor on individual stocks. The models we seek to estimate can be formulated as follows:

$$\begin{aligned} Return_t = \beta_0 + \beta_1 Comments\,per\,day_t + \beta_2 Positive\,sentiment\,comments_t \\ +\beta_3 Title\,per\,day_t + \beta_4 Positive\,sentiment\,title_t + \beta_5 Title\,text\,per\,day_t \\ +\beta_6\,Positive\,sentiment\,title\,text_t + \beta_7 Mkt.RF_t + \beta_8 SMB_t + \beta_9 HML_t + \beta_{10} RMW_t \\ +\beta_{11} CMA_t + \beta_{12} Mom_t \end{aligned} \tag{4.4}$$

$$\begin{aligned} Volume_t = \beta_0 + \beta_1 Comments\,per\,day_t + \beta_2 Positive\,sentiment\,comments_t \\ +\beta_3 Title\,per\,day_t + \beta_4 Positive\,sentiment\,title_t + \beta_5 Title\,text\,per\,day_t \\ +\beta_6 Positive\,sentiment\,title\,text_t + \beta_7 Mkt.RF_t \end{aligned} \tag{4.5}$$

$$\begin{aligned} Implied\,Volatility_t = \beta_0 + \beta_1 Comments\,per\,day_t + \beta_2 Positive\,sentiment\,comments_t \\ +\beta_3 Title\,per\,day_t + \beta_4 Positive\,sentiment\,title_t + \beta_5 Title\,text\,per\,day_t \\ +\beta_6 Positive\,sentiment\,title\,text_t + \beta_7 Mkt.RF_t \end{aligned} \tag{4.6}$$

$$\begin{aligned} Call\,Volume_t = \beta_0 + \beta_1 Comments\,per\,day_t + \beta_2 Positive\,sentiment\,comments_t \\ +\beta_3 Title\,per\,day_t + \beta_4 Positive\,sentiment\,title_t + \beta_5 Title\,text\,per\,day_t \\ +\beta_6 Positive\,sentiment\,title\,text_t + \beta_7 Mkt.RF_t \end{aligned} \tag{4.7}$$

$$\begin{aligned} Put\,Volume_t = \beta_0 + \beta_1 Comments\,per\,day_t + \beta_2 Positive\,sentiment\,comments_t \\ +\beta_3 Title\,per\,day_t + \beta_4 Positive\,sentiment\,title_t + \beta_5 Title\,text\,per\,day_t \\ +\beta_6 Positive\,sentiment\,title\,text_t + \beta_7 Mkt.RF_t \end{aligned} \tag{4.8}$$

$$\begin{aligned} AMIHUD_t = \beta_0 + \beta_1 Comments\,per\,day_t + \beta_2 Positive\,sentiment\,comments_t \\ +\beta_3 Title\,per\,day_t + \beta_4 Positive\,sentiment\,title_t + \beta_5 Title\,text\,per\,day_t \\ +\beta_6 Positive\,sentiment\,title\,text_t + \beta_7 Mkt.RF_t \end{aligned} \tag{4.9}$$

Further, we also estimate the effect of the sentiment variables, the return on the market portfolio and the Fama-French factors on lagged return variables. It is important to note that the control variables are not lagged in the regressions in order to better capture the effect of the lagged sentiment variables. These models can be formulated as follows:

$$\begin{aligned} Return_{t-1} = \beta_0 + \beta_1 Comments\,per\,day_{t+1} + \beta_2 Positive\,sentiment\,comments_{t+1} \\ +\beta_3 Title\,per\,day_{t+1} + \beta_4 Positive\,sentiment\,title_{t+1} + \beta_5 Title\,text\,per\,day_{t+1} \\ +\beta_6 Positive\,sentiment\,title\,text_{t+1} + \beta_7 Mkt.RF_{t+1} + \beta_8 SMB_{t+1} + \beta_9 HML_{t+1} \\ +\beta_{10} RMW_{t+1} + \beta_{11} CMA_{t+1} + \beta_{12} Mom_{t+1} \end{aligned} \tag{4.10}$$

$$\begin{aligned} Return_{t+1} = \beta_0 + \beta_1 Comments\,per\,day_{t-1} + \beta_2 Positive\,sentiment\,comments_{t-1} \\ +\beta_3 Title\,per\,day_{t-1} + \beta_4 Positive\,sentiment\,title_{t-1} + \beta_5 Title\,text\,per\,day_{t-1} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-1} + \beta_7 Mkt.RF_{t-1} + \beta_8 SMB_{t-1} + \beta_9 HML_{t-1} \\ +\beta_{10} RMW_{t-1} + \beta_{11} CMA_{t-1} + \beta_{12} Mom_{t-1} \end{aligned} \tag{4.11}$$

$$\begin{aligned} Return_{t+2} = \beta_0 + \beta_1 Comments\,per\,day_{t-2} + \beta_2 Positive\,sentiment\,comments_{t-2} \\ +\beta_3 Title\,per\,day_{t-2} + \beta_4 Positive\,sentiment\,title_{t-2} + \beta_5 Title\,text\,per\,day_{t-2} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-2} + \beta_7 Mkt.RF_{t-2} + \beta_8 SMB_{t-2} + \beta_9 HML_{t-2} \\ +\beta_{10} RMW_{t-2} + \beta_{11} CMA_{t-2} + \beta_{12} Mom_{t-2} \end{aligned} \tag{4.12}$$

$$\begin{aligned} Return_{t+3} = \beta_0 + \beta_1 Comments\,per\,day_{t-3} + \beta_2 Positive\,sentiment\,comments_{t-3} \\ +\beta_3 Title\,per\,day_{t-3} + \beta_4 Positive\,sentiment\,title_{t-3} + \beta_5 Title\,text\,per\,day_{t-3} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-3} + \beta_7 Mkt.RF_{t-3} + \beta_8 SMB_{t-3} + \beta_9 HML_{t-3} \\ +\beta_{10} RMW_{t-3} + \beta_{11} CMA_{t-3} + \beta_{12} Mom_{t-3} \end{aligned} \tag{4.13}$$

$$\begin{aligned} Return_{t+4} = \beta_0 + \beta_1 Comments\,per\,day_{t-4} + \beta_2 Positive\,sentiment\,comments_{t-4} \\ +\beta_3 Title\,per\,day_{t-4} + \beta_4 Positive\,sentiment\,title_{t-4} + \beta_5 Title\,text\,per\,day_{t-4} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-4} + \beta_7 Mkt.RF_{t-4} + \beta_8 SMB_{t-4} + \beta_9 HML_{t-4} \\ +\beta_{10} RMW_{t-4} + \beta_{11} CMA_{t-4} + \beta_{12} Mom_{t-4} \end{aligned} \tag{4.14}$$

To assess whether activity is driven by the forum, we estimate the effect of sentiment on lagged volume variables on stocks individually. In essence, the models we seek to estimate can be formulated as follows:

$$\begin{aligned} Volume_{t-1} = \beta_0 + \beta_1 Comments\,per\,day_{t+1} + \beta_2 Positive\,sentiment\,comments_{t+1} \\ +\beta_3 Title\,per\,day_{t+1} + \beta_4 Positive\,sentiment\,title_{t+1} + \beta_5 Title\,text\,per\,day_{t+1} \\ +\beta_6 Positive\,sentiment\,title\,text_{t+1} + \beta_7 Mkt.RF_{t+1} \end{aligned} \tag{4.15}$$

$$\begin{aligned} Volume_{t+1} = \beta_0 + \beta_1 Comments\,per\,day_{t-1} + \beta_2 Positive\,sentiment\,comments_{t-1} \\ +\beta_3 Title\,per\,day_{t-1} + \beta_4 Positive\,sentiment\,title_{t-1} + \beta_5 Title\,text\,per\,day_{t-1} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-1} + \beta_7 Mkt.RF_{t-1} \end{aligned} \tag{4.16}$$

$$\begin{aligned} Volume_{t+2} = \beta_0 + \beta_1 Comments\,per\,day_{t-2} + \beta_2 Positive\,sentiment\,comments_{t-2} \\ +\beta_3 Title\,per\,day_{t-2} + \beta_4 Positive\,sentiment\,title_{t-2} + \beta_5 Title\,text\,per\,day_{t-2} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-2} + \beta_7 Mkt.RF_{t-2} \end{aligned} \tag{4.17}$$

$$\begin{aligned} Volume_{t+3} = \beta_0 + \beta_1 Comments\,per\,day_{t-3} + \beta_2 Positive\,sentiment\,comments_{t-3} \\ +\beta_3 Title\,per\,day_{t-3} + \beta_4 Positive\,sentiment\,title_{t-3} + \beta_5 Title\,text\,per\,day_{t-3} \\ +\beta_6 Positive\,sentiment\,title\,text_{t-3} + \beta_7 Mkt.RF_{t-3} \end{aligned} \tag{4.18}$$

$$\begin{aligned} Volume_{t+4} = \beta_0 + \beta_1 Comments\ per\ day_{t-4} + \beta_2 Positive\ sentiment\ comments_{t-4} \\ +\beta_3 Title\ per\ day_{t-4} + \beta_4 Positive\ sentiment\ title_{t-4} + \beta_5 Title\ text\ per\ day_{t-4} \\ +\beta_6 Positive\ sentiment\ title\ text_{t-4} + \beta_7 Mkt.RF_{t-4} \end{aligned} \tag{4.19}$$

4.5.2 Aggregated regressions

For the regressions with categorical variables, AAPL has arbitrarily been chosen as the baseline regressor, utilized as the constant in the regressions. The remaining tickers are to be interpreted as dummy variables where the coefficient is multiplied with 1 if the ticker of interest is investigated, all other dummies being 0. Below, the model we estimate for return is formalized. The rest of the models in the regressions with categorical variables, such as lagged return, lagged volume and the year specific regressions, are identical to the regressions outlined above, only with all the tickers included for an aggregated result over the entire dataset.

$$\begin{aligned} Return_t = \beta_0 + \beta_1 Comments\ per\ day_t + \beta_2 Positive\ sentiment\ comments_t \\ +\beta_3 Title\ per\ day_t + \beta_4 Positive\ sentiment\ title_t + \beta_5 Title\ text\ per\ day_t \\ +\beta_6\ Positive\ sentiment\ title\ text_t + \beta_7 Mkt.RF_t + \beta_8 SMB_t + \beta_9 HML_t + \beta_{10} RMW_t \\ +\beta_{11} CMA_t + \beta_{12} Mom_t + \beta_{13} TickerAMC_t + \beta_{14} TickerAMD_t + \beta_{15} TickerAPHA_t \\ +\beta_{16} TickerBB_t + \beta_{17} TickerGME_t + \beta_{18} TickerGOOG_t + \beta_{19} TickerNIO_t \\ +\beta_{20} TickerNKLA_t + \beta_{21} TickerNOK_t + \beta_{22} TickerNVDA_t + \beta_{23} TickerPLTR_t \\ +\beta_{24} TickerPLUG_t + \beta_{25} TickerRKT_t + \beta_{26} TickerSPCE_t + \beta_{27} TickerTLRY_t \\ +\beta_{28} TickerTSLA_t + \beta_{29} TickerTSM_t \end{aligned} \tag{4.20}$$

4.5.3 Minute-by-minute regressions

For the minute-by-minute regressions we have only included comments, not submissions, to assess what impact sentiment had on stock fluctuations in real-time. Submissions are generally posted less frequently than comments, meaning that by excluding submissions we take into account the general variation in the relative posting frequency of the two. One submission can only be written by one author, which implies that the actual discussion of the stocks occurs in the comment sections on the submissions or, alternatively, in so-called

"Daily Threads" or "Megathreads" on WallStreetBets. Comments thereby function as a better proxy for a minute-by-minute analysis of WallStreetBets sentiment on the stock metrics we perform the following regressions upon.

As only price and volume are metrics we managed to attain on a minute-by-minute basis, our regressions attempt to capture the relationship between returns against sentiment as well as volume against sentiment. The models we estimate are both constructed on an individual and aggregated stock basis. The individual minute-by-minute models can be formulated as follows, where $\text{Return}_{\text{k stock i}}$ is the return on stock i at minute k and $\text{Volume}_{\text{k stock i}}$ is the trading volume on stock i at minute k:

$$Return_{k\ stock\ i} = \beta_0 + \beta_1 Comments\ per\ minute_k + \beta_2 Positive\ sentiment\ comments_k \tag{4.21}$$

$$Volume_{k\ stock\ i} = \beta_0 + \beta_1 Comments\ per\ minute_k + \beta_2 Positive\ sentiment\ comments_k \tag{4.22}$$

In the aggregated minute-by-minute models below, Return_{k} is the estimated return on minute k and Volume_{k} is the estimated volume on minute k. AMC has been arbitrarily chosen as the baseline regressor, while the remaining tickers function as dummy variables. The aggregated minute-by-minute models can be formalized as follows:

$$\begin{aligned} Return_k = \beta_0 + \beta_1 Comments\ per\ minute_k + \beta_2 Positive\ sentiment\ comments_k \\ +\beta_3 TickerBB + \beta_4 TickerGME + \beta_5 TickerNOK \end{aligned} \tag{4.23}$$

$$\begin{aligned} Volume_k = \beta_0 + \beta_1 Comments\ per\ minute_k + \beta_2 Positive\ sentiment\ comments_k \\ +\beta_3 TickerBB + \beta_4 TickerGME + \beta_5 TickerNOK \end{aligned} \tag{4.24}$$

In the models we construct, the number of observations for return and volume can differ on the same stock. The reason for this phenomenon lies in how return is calculated. Return is calculated on the basis of the return in the previous observation. At some minutes in the sample no trades were done. The reason is that during the period we analyze in this section trading was frequently halted due to significant increased volatility

and volumes. No trading at some minutes means that when calculating return the next minute a trade is done the return will be calculated as NA, as the previous price is missing. Rows with NA variables are not included in the regressions. Volume is an absolute number, meaning that all minutes that trades are done, volume will be treated as a non-NA variable. This explains the differences in the number of observations in the regression outputs in chapter 5.3.

4.6 Mention-weighted portfolios

To analyze what cumulative returns could be achieved by utilizing the number of stock mentions on WallStreetBets as a proxy for sentiment and implementing a trading strategy thereof, we constructed mention-weighted portfolios. To do so, we first summed the daily number of stock mentions for each individual stock in both submissions and comments to find the total number of stock mentions on WallStreetBets each day throughout the time interval of this thesis. Then we divided the daily number of mentions of each specific stock by the daily total of mentions of our selected stocks to mention-weigh the portfolios. The portfolios are all rebalanced daily, i.e. each day the relative stock weights adjust to changes in the relative mention-ratio among the stocks. The stocks' relative weight in the mention-weighting portfolios can be written as:

$$Mention{-}weighted\,portfolio\ = \frac{no.mentions\ stock_1}{total\ no.\ mentions} + \frac{no.mentions\ stock_2}{total\ no.\ mentions}\ + \ldots \tag{4.25}$$

By mention-weighting the portfolios with daily rebalancing, this facilitates a theoretical approach to hypothesizing what return an investor would achieve if the investor monitored the sentiment on WallStreetBets daily and rebalanced the portfolio based on the relative mention-weights. The first of the mentioned-weighted strategies we developed is called "Mentions P". This hypothesized strategy is based on using today's mention-weights and implementing the strategy the same day. That means that the investor will enter positions equal to the relative mention-weights on WallStreetBets on a daily basis and get the return on the stocks the same day. The return of this daily rebalanced mention-weighting

portfolio, $r_{t\ \text{Mentions P}}$, on day t is formalized in the following equation, where $r_{\text{stock i}}$ is the daily logarithmic return of stock i:

$$r_{t\ Mentions\ P} = r_{t\ stock\ 1} * \frac{no.mentions\ stock\ 1_t}{total\ no.\ mentions_t} + r_{t\ stock\ 2} * \frac{no.mentions\ stock\ 2_t}{total\ no.\ mentions_t} + \dots \tag{4.26}$$

To check for potential lagged effects of sentiment on return, we also constructed a mention-weighted trading strategy in addition to the strategy outlined above. This lagged return strategy, i.e. strategy "Mentions P_{t+1}" is a strategy where the strategy's return on day t is based on if you implement a trading strategy based on the relative mention-weights on WallStreetBets on day t, and then use stock returns on day t+1 as the foundation for calculating the strategy's return. Practically, this implies that you calculate the relative mention-weights on WallStreetBets at the end of the day and implement these weights at the stock market opening the next day. This is probably the most feasible trading strategy to implement in reality. The daily return of this strategy can be formulated as follows:

$$r_{t\ Mentions\ P\ t+1} = r_{t+1\ stock\ 1} * \frac{no.mentions\ stock\ 1_t}{total\ no.\ mentions_t} + r_{t+1\ stock\ 2} * \frac{no.mentions\ stock\ 2_t}{total\ no.\ mentions_t} + \dots \tag{4.27}$$

We have also constructed two mention-weighted portfolios similar to Mentions P and Mentions P_{t+1}, but where the effects of AMC, BB, GME and NOK are excluded by excluding these stocks from the portfolios. By constructing portfolios excluding these mentioned-based heavyweights on WallStreetBets, the resulting portfolios better capture the general effects mentions on WallStreetBets have had on a more general set of stocks. These strategies are called MentionsX P and MentionsX P_{t+1}, respectively. The returns of these strategies can be formalized as follows:

$$r_{t\ MentionsX\ P} = r_{t\ stock\ 1} * \frac{no.mentions\ stock\ 1_t}{total\ no.\ mentions_t} + r_{t\ stock\ 2} * \frac{no.mentions\ stock\ 2_t}{total\ no.\ mentions_t} + \dots \tag{4.28}$$

$$r_{t\ MentionsX\ P\ t+1} = r_{t+1\ stock\ 1} * \frac{no.mentions\ stock\ 1_t}{total\ no.\ mentions_t} + r_{t+1\ stock\ 2} * \frac{no.mentions\ stock\ 2_t}{total\ no.\ mentions_t} + \dots \tag{4.29}$$

We have also added two control strategies. The first, called Market P, is to invest simply in the value-weighted market portfolio, i.e. the Mkt.RF factor collected from French (2021). The daily return of this strategy is independent of relative mention-weights on WallStreetBets, and can be formulated as follows:

$$r_{t\ Market\ P} = r_{t\ Mkt.RF} \tag{4.30}$$

The second control strategy is to weigh the 18 stocks by relative value-weights instead of mention-weights. This portfolio is also rebalanced daily, where daily market capitalization is calculated by multiplying the daily stock price against the number of shares outstanding at the given day. The return on this strategy, which we have called Market Weight P, can be formalized as follows:

$$r_{t\ Market\ Weight\ P} = r_{t\ stock\ 1} * \frac{market\ cap.\ stock\ 1_t}{tot.\ market\ cap._t} + r_{t\ stock\ 2} * \frac{market\ cap.\ stock\ 2_t}{tot.\ market\ cap._t} + \dots \tag{4.31}$$

The value v of USD 100 invested at January 02, 2020, on day t with strategy j can be formulated as follows:

$$v_{t\ strategy\ j} = 100 * \prod_{i=0}^{t} (1 + r_{t\ strategy\ j}) \tag{4.32}$$

4.7 Event studies

To assess what happened when activity surrounding a stock was high, we wanted to perform an event study on the days with the highest relative individual stock mention ratio on WallStreetBets. To perform the event studies, we started by calculating the relative mention-weight for each stock each day between January 02, 2020, and March 12, 2021. The calculation was done by dividing the number of mentions considering the

specific stock both in submissions and comments by the total number of submissions and comments that day that could be connected to revolve around one or more stocks. Then we find which day this metric is at its highest for each particular stock. This day is then used as the basis for the further analysis.

Then we regressed the stock's return on the market portfolio's return, as well as the Fama-French factors SMB, HML, RMW, CMA to estimate the coefficients we then utilized to calculate the stock's expected return that day. Column 5 in the tables containing the results, i.e. "Expected Return", shows the calculated expected return based on these regression coefficients. The formula used for calculations of expected returns can be formulated as follows:

$$E\left(r_t\right) = \beta_0 + \beta_{Mkt.RF} r_{Mkt.RF\,t} + \beta_{SMB} r_{SMB\,t} + \beta_{HML} r_{HML\,t} + \beta_{RMW} r_{RMW\,t} + \beta_{CMA} r_{CMA\,t} \tag{4.33}$$

Abnormal return is defined as the return we observe over the expected return calculated by a market model, in our case the Fama-French 5 factor model. Abnormal return, AR, for stock i on day t can be formulated as follows, where $r_{i,t}$ is the actual return of stock i on day t and $E(r_{i,t})$ is the expected return for stock i on day t:

$$\mathrm{AR}_{i,t} = r_{i,t} - \ E(r_{i,t}) \tag{4.34}$$

To test whether or not the abnormal returns are statistically significant at the 5% significance level, an hypothesis test is performed where we divide the abnormal return of stock i by the standard error of the initial regression, to calculate a t-value. The calculation of the t-value, where σ is the standard deviation of the initial regression, can be formulated as follows:

$$t = \frac{AR_{i,t}}{\sigma} \tag{4.35}$$

5 Results

5.1 Market response to January 2021 rallies

In January 2021, many stocks with high short ratios experienced remarkably high returns, while many stocks with low short ratios experienced low or negative returns. Many of the stocks experiencing surging stock prices in January 2021 were subject to short and gamma squeezes, which parallelled a significant upturn in trading volume overall as investors all over the world established positions in the rallies to try to get their share of huge potential gains. In this chapter we will look into the difference in returns between the least and most shorted stocks on the Russell 3000 index from January 12 to January 29, 2021. This period was the most significant in terms of extreme returns, activity on WallStreetBets and media coverage, and serves as a starting point to visualize how the market reacted to the rallies in January 2021.

5.1.1 Most and least shorted stocks return January 2021

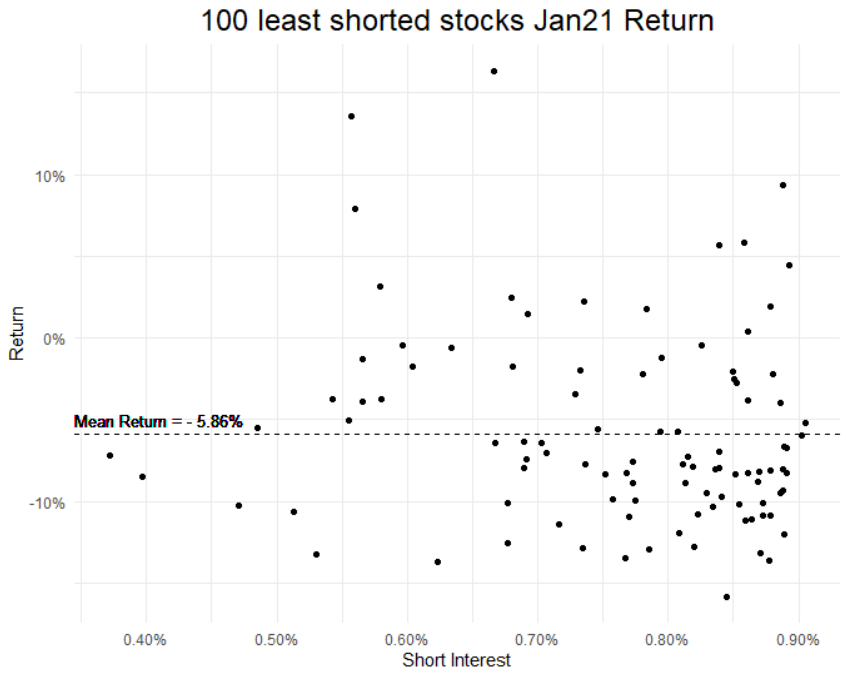

Figure 5.1: 100 least shorted stocks - Jan 2021

Figure 5.1 shows a plot of the return on the 100 least shorted stocks on Russell 3000 from January 12 to January 29, 2021, i.e. the most extreme rally period in the time interval we explore in this thesis when GME's and other stock prices surged before trading restrictions effectively ended the rallies. The mean return of the 100 least shorted stocks on Wall Street was -5.86% in this period, and as the figure shows there were large variations in these stocks' returns.

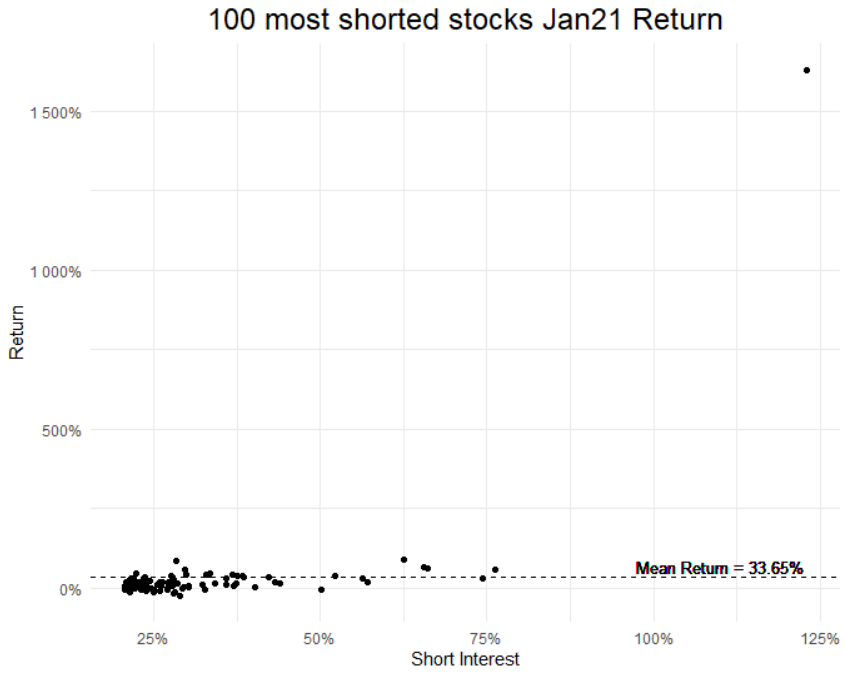

Figure 5.2: 100 most shorted stocks - Jan 2021

Figure 5.2 shows the return on the 100 most shorted stocks on Wall Street from January 12, 2021, to January 29, 2021. In contrast to the 100 least shorted stocks did the 100 most shorted stocks achieve a mean return of 33.65% in the same period. The extreme outlier in the far upper-right corner is GameStop, which achieved a return in this time period of a stunning 1,629%.

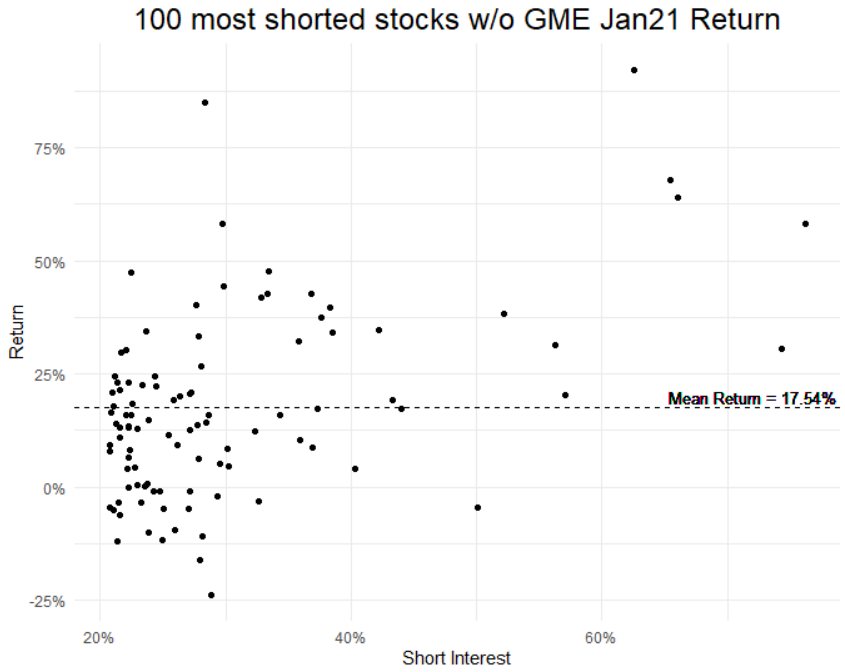

Figure 5.3: 100 most shorted stocks without GME - Jan 2021

As is shown in figure 5.3, even when excluding the effect of GME's extreme return, the 100 most-shorted stocks after GME achieved a mean return of 17.54% between January 12 and January 29, 2021. This is remarkable compared to figure 5.1 with the 100 least-shorted stocks in the same period. Even when excluding the effect of GME's extreme return in this period, the 100 most-shorted stocks outperformed the 100 least-shorted stocks by 23.4% in this period on average.

The cause of the large discrepancy between returns of the least and most shorted stocks can possibly be attributed to the short squeeze rallies, possibly attributed to WallStreetBets, during January 2021 and the snowball effect in the market. Normally the least shorted stocks would be expected to experience a higher return than those who are the most shorted. However, due to the extreme GME rally, when one institutional investor is margin called they need to rebalance by selling out of one of their long positions to cover their short. This then creates a snowball effect where more and more institutional investors need to rebalance, thus creating a price surge in the most shorted stocks, as well as a price drop in the least shorted stocks. Many of the stocks in the graphs above

were not prone to heavy discussions on WallStreetBets, but there are reasons to believe that since stocks with a high short interest were targeted, in order to hedge against risk, most institutional investors reduced their risk by reducing their existing short positions in what could be the next target of the retail investors.

5.1.2 Long short basket returns

Figure 5.4 and 5.5 show the difference in returns on Barclays' and Goldman Sachs' long baskets minus the returns on their short baskets, respectively. The figures show the extreme magnitude of the rallies we explore in this thesis. The time series in figure 5.4 goes back to 2007 and shows that the recent events led to the biggest restructuring since the financial crisis in 2008/09. The data from the Barclays' long/short basket also illustrates the volatility and magnitude of the recent stock rallies and short squeezes. Goldman Sachs data only goes back to 2016, but still illustrates the compelling chain reaction of the rallies.

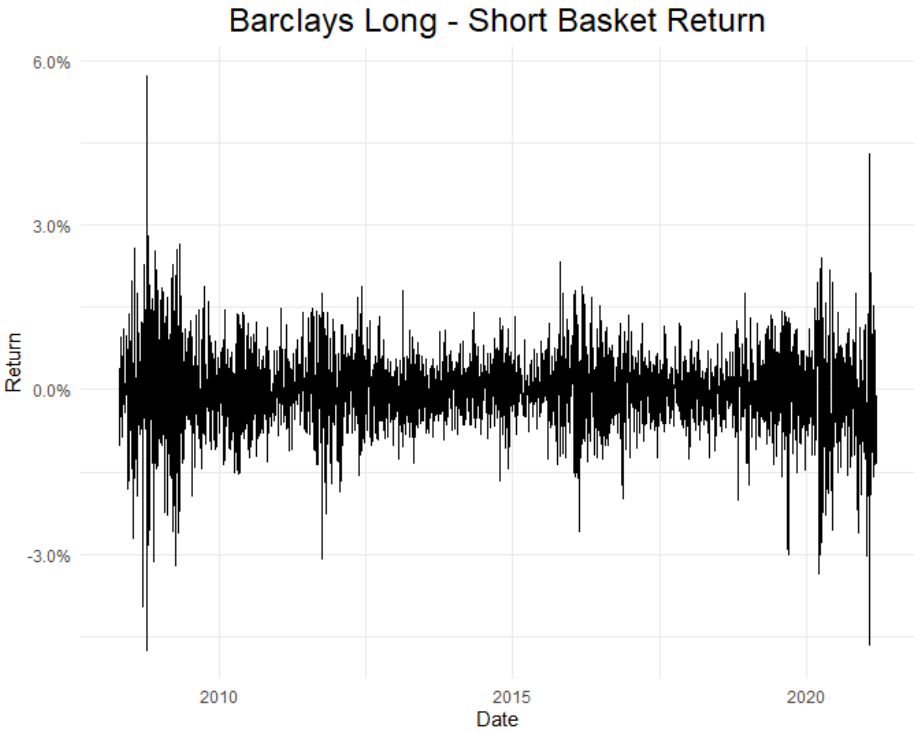

Figure 5.4: Barclays long - short basket return

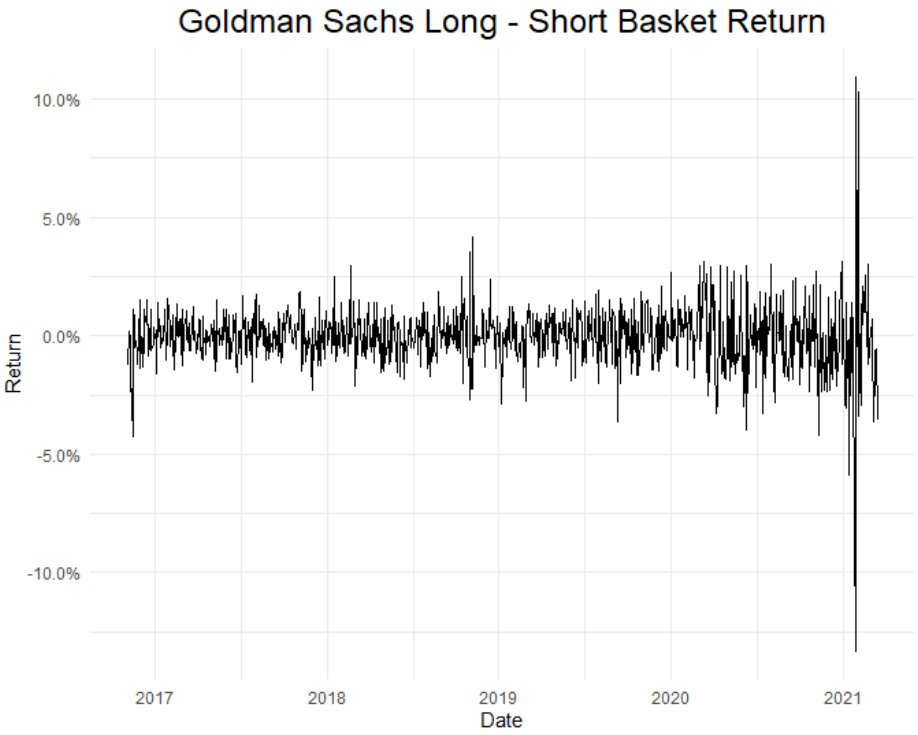

Figure 5.5: Goldman Sachs long - short basket return

The figures in this chapter sets a precedent for how the market reacted to the rallies. It is hard to infer how much of the rallies can be attributed directly to WallStreetBets, but there is no doubt that many financial institutions were either directly or indirectly impacted by what occured in January 2021. In the following chapters we will try to measure WallStreetBets' impact on several stock metrics, and assess whether or not we can link the rallies to the forum.

5.2 Sentiment regression results

5.2.1 Individual stocks

5.2.1.1 Selected findings in the individual stock regressions

To analyze the effect of sentiment on our selected stock and option metrics, we have performed individual regressions on each of the stocks in our sample. The complete

regression outputs can be found in the appendix' chapter A1. Here, we will present and discuss some of the most interesting findings in these regressions.

Table 5.1: PLTR full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.004	9,623.977	0.002	33.988	12.050	−0.000
	t = −1.112	t = 1.343	t = 0.643	t = 0.959	t = 0.801	t = −0.932
Positive sentiment comments	0.609	−1,816,490.000	0.150	−2,171.589	−1,911.710	−0.000
	t = 1.106	t = −1.966**	t = 0.331	t = −0.525	t = −0.969	t = −0.379
Title per day	0.489	1,205,722.000	0.321	10,111.930	4,226.703	0.000
	t = 1.543	t = 2.340**	t = 1.307	t = 3.088***	t = 2.826***	t = 1.959*
Positive sentiment title	0.061	−270,809.100	0.035	−75.045	−269.984	−0.000
	t = 0.519	t = −0.785	t = 0.251	t = −0.045	t = −0.341	t = −0.165
Title text per day	−0.568	−2,156,291.000	−0.445	−16,750.860	−6,487.953	−0.000
	t = −1.706*	t = −3.065***	t = −1.359	t = −3.414***	t = −2.611***	t = −1.516
Positive sentiment title text	0.142	−181,995.900	0.151	308.846	−8.240	0.000
	t = 0.821	t = −0.743	t = 1.246	t = 0.294	t = −0.016	t = 1.574
Mkt.RF	−4.213	−3,196,412.000	0.245	−5,747.076	−1,254.238	0.000
	t = −0.845	t = −0.690	t = 0.105	t = −0.265	t = −0.127	t = 1.386
SMB	10.620					
	t = 1.125					
HML	−9.338					
	t = −1.319					
RMW	2.422					
	t = 0.418					
CMA	−4.133					
	t = −0.436					
MOM	−4.064					
	t = −1.069					
Constant	−61.750	204,454,287.000	71.794	309,756.700	224,560.600	−0.000
	t = −1.037	t = 2.830***	t = 2.421**	t = 1.169	t = 1.578	t = −0.136
Observations	108	108	104	104	104	108
R^2	0.082	0.406	0.394	0.583	0.537	0.060
Adjusted R^2	−0.034	0.364	0.350	0.552	0.503	−0.005
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 5.2: PLTR lagged return

	Lagged return					
	$Return_{t-1}$	Return	$Return_{t+1}$	$Return_{t+2}$	$Return_{t+3}$	$Return_{t+4}$
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.00002	−0.004	0.0004	−0.00004	0.001	0.002
	t = −0.014	t = −1.112	t = 0.144	t = −0.018	t = 0.455	t = 0.658
Positive sentiment comments	0.039	0.609	1.004	1.494	1.108	0.565
	t = 0.390	t = 1.106	t = 1.059	t = 1.010	t = 0.975	t = 0.871
Title per day	0.010	0.489	0.249	0.300	−0.166	−0.018
	t = 0.092	t = 1.543	t = 1.108	t = 0.981	t = −1.072	t = −0.098
Positive sentiment title	−0.037	0.061	−0.251	0.206	0.329	0.046
	t = −0.972	t = 0.519	t = −0.770	t = 0.916	t = 1.065	t = 0.587
Title text per day	−0.024	−0.568	−0.384	−0.448	0.269	−0.076
	t = −0.201	t = −1.706*	t = −0.807	t = −0.826	t = 1.299	t = −0.352
Positive sentiment title text	−0.041	0.142	0.130	−0.138	0.318	0.102
	t = −1.359	t = 0.821	t = 1.104	t = −0.712	t = 1.069	t = 0.516
Mkt.RF	0.009	−4.213	0.235	7.102	−6.491	7.667
	t = 0.010	t = −0.845	t = 0.114	t = 0.879	t = −0.896	t = 0.998
SMB	0.817	10.620	−11.717	−15.865	−2.444	−8.284
	t = 0.544	t = 1.125	t = −0.877	t = −1.071	t = −0.715	t = −0.978
HML	−1.532	−9.338	−0.726	−0.152	−3.494	2.714
	t = −0.678	t = −1.319	t = −0.163	t = −0.032	t = −0.739	t = 0.597
RMW	0.540	2.422	−14.996	−12.620	3.190	1.320
	t = 0.259	t = 0.418	t = −0.931	t = −1.017	t = 0.456	t = 0.312
CMA	0.483	−4.133	15.239	19.673	−2.166	4.957
	t = 0.207	t = −0.436	t = 0.861	t = 0.943	t = −0.274	t = 0.648
MOM	−0.704	−4.064	−5.214	−1.042	−1.905	0.769
	t = −0.423	t = −1.069	t = −1.026	t = −0.291	t = −0.699	t = 0.297
Constant	4.379	−61.750	−64.923	−113.270	−127.977	−56.255
	t = 0.589	t = −1.037	t = −1.214	t = −1.056	t = −1.016	t = −0.852
Observations	107	108	107	106	105	104
R^2	0.039	0.082	0.128	0.161	0.120	0.059
Adjusted R^2	−0.083	−0.034	0.017	0.053	0.006	−0.066
Note:						*p<0.1; **p<0.05; ***p<0.01

Although PLTR was one of the first stock rallies that was connected to WallStreetBets by the media, model (1) in table 5.1 indicates that none of the sentiment variables are statistically significantly correlated with PLTR's return at the 5% level when looking at the time interval overall. The adjusted R-squared in this model is also remarkably low. Neither when introducing lagged return variables in table 5.2, nor when separating the time period into 2020 and 2021 in table A1.50 and A1.51, does sentiment appear to have had a statistically significant impact on PLTR's return. The adjusted R-squareds in table 5.2, i.e. when introducing lagged return variables, nevertheless indicate that future returns are better explained than today's return, but the adjusted R-squareds of all these models are very low.

Given the activity in media and on WallStreetBets for PLTR, this finding is interesting. We do not find any statistically significant relationship between WallStreetBets' sentiment for PLTR and PLTR's return. While none of the sentiment variables are statistically significant in the returns models, when using data for the entire time period in table 5.1 we observe that both Title per day and Title text per day are statistically significant at the 1% level in both model (4) and (5), indicating that sentiment may have had an impact on call and put option volumes. In addition, we observe that multiple sentiment variables are statistically significant in the volume model (2) in table 5.1. Even more striking, in table A1.49 with lagged volume variables we observe that there are statistically significant sentiment variables in all models except for model (4). While the adjusted R-squareds differ across the models and that the R-squared of the volume model today, i.e. model (2), is the highest, the findings in model (3), (5) and (6) indicate that sentiment on WallStreetBets for PLTR may have had an impact on volume the next day, two days on and four days on, respectively. The statistically significant variables in model (1) indicate that yesterday's volume may partially explain sentiment today. We also observe statistically significant sentiment variables in the implied volatility, call option volume and put option volume models both in table A1.50 and A1.51, i.e. when the time period is separated into 2020 and 2021, respectively. The adjusted R-squareds in these models differ, but they are especially high for call and put option volume in 2020. Although we

may well deal with reverse causality problems here, the results do nonetheless suggest that sentiment may have had an impact on several of the stock and option metrics for PLTR. A word of caution is that there are relatively few observations for PLTR in general. The reason for this is that PLTR went public on September 30, 2020, meaning that the number of observations for 2020 is generally lower than for the other stocks in our sample.

Table 5.3: GME full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0001 t = 0.545	307.436 t = 0.942	0.001 t = 2.354**	2.013 t = 1.218	1.589 t = 1.179	0.000 t = 1.229
Positive sentiment comments	0.027 t = 0.821	39,319.100 t = 1.665*	−0.166 t = −1.067	196.904 t = 1.503	129.982 t = 0.961	0.000 t = 1.331
Title per day	−0.019 t = −1.028	8,121.231 t = 0.426	0.037 t = 1.756*	−27.178 t = −0.290	82.296 t = 0.866	−0.000 t = −0.661
Positive sentiment title	0.062 t = 1.858*	87,650.970 t = 3.756***	0.352 t = 3.657***	557.617 t = 4.364***	461.563 t = 3.386***	0.000 t = 1.067
Title text per day	0.043 t = 0.617	−20,545.330 t = −0.515	−0.028 t = −0.326	−32.556 t = −0.173	−17.308 t = −0.045	0.000 t = 0.294
Positive sentiment title text	−0.016 t = −0.443	104,942.400 t = 3.149***	0.242 t = 2.374**	590.938 t = 3.558***	498.559 t = 2.975***	0.000 t = 0.174
Mkt.RF	0.665 t = 1.732*	−423,935.400 t = −1.099	−0.288 t = −0.255	−1,727.850 t = −0.903	−4,493.415 t = −1.341	0.00000 t = 4.802***
SMB	3.031 t = 0.965					
HML	−0.717 t = −0.413					
RMW	0.301 t = 0.129					
CMA	12.558 t = 1.014					
MOM	1.037 t = 0.736					
Constant	−4.046 t = −1.207	−5,258,571.000 t = −2.092**	104.984 t = 7.499***	−37,895.870 t = −2.914***	−33,845.570 t = −2.081**	−0.00000 t = −1.730*
Observations	296	296	296	296	296	296
R^2	0.304	0.396	0.564	0.415	0.574	0.166
Adjusted R^2	0.275	0.382	0.553	0.401	0.564	0.146

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5.4: GME lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0002	0.0001	−0.00004	0.0003	0.0001	−0.0001
	t = 0.639	t = 0.545	t = −0.164	t = 1.246	t = 0.620	t = −0.731
Positive sentiment comments	−0.013	0.027	0.037	0.025	0.024	0.023
	t = −0.604	t = 0.821	t = 1.697*	t = 1.120	t = 1.056	t = 1.077
Title per day	−0.028	−0.019	−0.001	−0.016	0.004	0.036
	t = −0.920	t = −1.028	t = −0.049	t = −0.523	t = 0.090	t = 0.901
Positive sentiment title	0.017	0.062	0.013	−0.012	0.013	−0.030
	t = 0.669	t = 1.858*	t = 0.525	t = −0.442	t = 0.563	t = −1.047
Title text per day	0.045	0.043	0.044	0.012	−0.017	−0.057
	t = 0.829	t = 0.617	t = 0.722	t = 0.249	t = −0.241	t = −0.869
Positive sentiment title text	0.020	−0.016	0.020	0.009	0.017	0.026
	t = 0.681	t = −0.443	t = 0.661	t = 0.275	t = 0.626	t = 0.975
Mkt.RF	−0.014	0.665	−0.569	−0.034	0.117	−0.292
	t = −0.038	t = 1.732*	t = −1.697*	t = −0.096	t = 0.290	t = −0.800
SMB	−1.024	3.031	−0.649	−1.589	−0.069	1.089
	t = −0.497	t = 0.965	t = −0.370	t = −1.116	t = −0.041	t = 1.005
HML	−0.102	−0.717	0.667	0.293	−0.758	−0.601
	t = −0.073	t = −0.413	t = 0.615	t = 0.296	t = −0.584	t = −0.579
RMW	0.232	0.301	3.103	−0.069	1.572	1.006
	t = 0.100	t = 0.129	t = 1.403	t = −0.030	t = 0.659	t = 0.484
CMA	−3.870	12.558	−4.221	−2.178	0.766	−0.034
	t = −0.588	t = 1.014	t = −0.665	t = −0.635	t = 0.311	t = −0.010
MOM	−0.968	1.037	0.371	−0.022	−0.189	0.095
	t = −1.102	t = 0.736	t = 0.419	t = −0.038	t = −0.262	t = 0.142
Constant	−0.602	−4.046	−4.175	−2.301	−3.303	−0.593
	t = −0.240	t = −1.207	t = −1.876*	t = −0.950	t = −1.469	t = −0.255
Observations	295	296	295	294	293	292
R^2	0.103	0.304	0.340	0.314	0.096	0.215
Adjusted R^2	0.064	0.275	0.312	0.285	0.057	0.182
Note:						*p<0.1; **p<0.05; ***p<0.01

As with PLTR, the results for GME are somewhat surprising. Both PLTR and GME have been subject to massive media coverage and public debate, and WallStreetBets has by many been claimed to have made big impacts on these stocks' returns. Contrary to what we initially believed when we began writing this thesis, we find that none of the sentiment variables are statistically significant in model (1) in table 5.3 where we analyze the effect of the sentiment variables on GME's return over the entire time period. The adjusted R-squared of this model is also relatively modest with only 27.5%. Also, in table 5.4 with lagged return variables, none of the sentiment variables are statistically significant in any of the models. Interestingly, we observe that the R-squareds in models (3) and (4) in table 5.4 are higher than in model (2). This indicates that returns one and two days forward are better explained than returns today. While none of the sentiment variables are statistically significant at the 5% level, this is still an interesting finding since it can indicate that the models still capture some of the potential effects of sentiment on returns, although not enough to be statistically significant.

When analyzing the time period overall, in table 5.3 we observe that several of the

sentiment variables are statistically significant at both the 1% and 5% level in model (2), (3), (4) and (5). Sentiment is therefore indicated to have had an impact on GME's trading volume, implied volatility, call option volume and put option volume, respectively. When separating the time period into 2020 and 2021 in table A1.27 and A1.28, we observe statistically significant sentiment variables in model (2), (3), (4) and (5) in 2020, and statistically significant sentiment variables in model (2), (3) and (4) in 2021. In general, this indicates that both in 2020 and 2021 has sentiment had an impact on trading volume, implied volatility and call option volume, while sentiment is only found to have had statistically significant impact on put option volume in 2020.

While we generally consider these option metrics to be less prone to reverse causality problems than stock returns, we advise to exercise more caution when looking at the results for trading volume. We do not believe that volume is as likely as returns to be prone to reverse causality problems, but we consider it to be higher than for the option metrics. The results we therefore observe for GME's trading volume can be due to reverse causality. However, in table A1.26 with lagged volume variables, we observe that several sentiment variables are statistically significant at the 1% and 5% level in both model (1), (2), (3), (4) and (5). This finding in model (1) indicates that yesterday's volume has had an impact on sentiment today, while the findings in model (3), (4) and (5) suggest that sentiment on WallStreetBets for GME has had an impact on trading volume one, two and three days forward, respectively. Also interestingly, the adjusted R-squareds in model (3), (4) and (5) are all higher than the adjusted R-squared in model (2). The lagged volume models therefore better explain the variations in trading volume than the model considering variation in volume today. The magnitude of the coefficients are generally higher in model (2) than in the other models. Combined, these findings suggest that sentiment on WallStreetBets for GME has had an impact on trading volume, and the problem with reverse causality can potentially be smaller for trading volume than for the stock returns because sentiment a given day is found to have had a statistically significant impact on trading volume one, two and three days on. A general note when analyzing GME and WallStreetBets is that many of the forum members did not sell

the stock during the January rally, but are still hoping to this day to see similar price levels in the future. Therefore, as GME stock activity dropped off significantly from the peak, the forums interest in the stock persists and contrary to for example NOK who only captured the forum's interest during the rallies, the models fail to capture the relationship where GME is still talked about to a large degree to this day without experiencing the same highs.

Yet another interesting finding from the individual stock regressions is that AAPL and GOOG generally have higher adjusted R-squareds in the Amihud-models than the other stocks in the sample. AAPL and GOOG are large companies in terms of market capitalization, which may be a contributing explanatory factor for this result. This finding may signal that liquidity is an important factor for these companies, since they are more affected by the market than by sentiment. Another interesting finding for these companies is that although some sentiment variables appear to be statistically significant in the return models on these large cap stocks, we would argue that is probably due to reverse causality. We observe a very low adjusted R-squared for AAPL in model (3) in table A1.2 and a very low adjusted R-squared for GOOG in model (3) in table A1.30. The magnitude of the coefficient Positive sentiment comments in model (3) in table A1.2 for AAPL is very low, and the coefficient on Title text per day in model (3) in table A1.30 for GOOG seems unreasonably high. Therefore, we would argue to exercise utmost caution when interpreting these results, as it seems unreasonable to believe that WallStreetBets have had any noteworthy effect on these stocks' return.

Table 5.5: NOK full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.005	55,849.960	0.013	220.308	46.103	−0.000
	t = −0.579	t = 3.838***	t = 1.580	t = 5.425***	t = 4.238***	t = −0.115
Positive sentiment comments	0.016	1,757.882	−0.041	74.535	−48.478	0.000
	t = 2.796***	t = 0.053	t = −1.309	t = 1.001	t = −2.070**	t = 1.303
Title per day	0.175	−109,226.200	0.120	437.905	−312.026	0.000
	t = 0.604	t = −0.209	t = 0.620	t = 0.505	t = −0.820	t = 0.157
Positive sentiment title	0.019	−15,528.770	0.016	711.891	145.457	0.000
	t = 1.038	t = −0.145	t = 0.446	t = 3.154***	t = 1.940*	t = 1.750*
Title text per day	−0.442	338,847.500	−0.497	−4,046.766	66.754	−0.000
	t = −0.630	t = 0.267	t = −1.024	t = −1.939*	t = 0.077	t = −0.172
Positive sentiment title text	0.014	98,813.420	0.076	256.228	−48.588	−0.000
	t = 0.648	t = 1.091	t = 1.701*	t = 1.411	t = −0.996	t = −0.350
Mkt.RF	0.928	−825,073.600	−0.571	533.901	−525.643	0.000
	t = 7.638***	t = −1.671*	t = −0.722	t = 0.558	t = −2.109**	t = 7.652***
SMB	0.428					
	t = 1.434					
HML	−0.306					
	t = −1.409					
RMW	0.441					
	t = 1.290					
CMA	0.534					
	t = 0.371					
MOM	−0.026					
	t = −0.113					
Constant	−3.221	26,681,947.000	51.350	−22,292.550	4,729.229	−0.000
	t = −2.845***	t = 4.714***	t = 14.208***	t = −1.659*	t = 1.620	t = −1.908*
Observations	296	296	296	296	296	296
R^2	0.673	0.952	0.637	0.967	0.844	0.561
Adjusted R^2	0.659	0.951	0.628	0.967	0.840	0.551

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5.6: NOK lagged volume

	Lagged volume					
	$Volume_{t-1}$	Volume	$Volume_{t+1}$	$Volume_{t+2}$	$Volume_{t+3}$	$Volume_{t+4}$
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	25,525.090	55,849.960	194,041.800	193,933.100	79,317.200	32,738.970
	t = 0.523	t = 3.838***	t = 1.455	t = 4.266***	t = 0.173	t = 0.153
Positive sentiment comments	51,989.850	1,757.882	−16,933.010	−49,257.990	−898.296	−59,450.000
	t = 0.563	t = 0.053	t = −0.355	t = −0.925	t = −0.014	t = −0.619
Title per day	4,358.558	−109,226.200	−3,401,141.000	−95,739.130	972,719.000	390,582.800
	t = 0.002	t = −0.209	t = −0.696	t = −0.067	t = 0.065	t = 0.056
Positive sentiment title	589,010.900	−15,528.770	97,625.980	117,321.200	337,788.700	706,382.800
	t = 1.398	t = −0.145	t = 0.819	t = 0.881	t = 0.916	t = 2.227**
Title text per day	108,511.500	338,847.500	6,500,478.000	−6,011,236.000	−6,097,481.000	−2,481,192.000
	t = 0.025	t = 0.267	t = 0.546	t = −1.834*	t = −0.165	t = −0.153
Positive sentiment title text	302,458.600	98,813.420	−304,258.100	−183,882.100	287,982.500	−69,522.780
	t = 1.057	t = 1.091	t = −0.858	t = −1.322	t = 0.248	t = −0.131
Mkt.RF	−648,582.000	−825,073.600	1,031,258.000	252,132.000	3,375,789.000	2,024,403.000
	t = −0.926	t = −1.671*	t = 0.827	t = 0.346	t = 0.806	t = 0.833
Constant	−18,661,216.000	26,681,947.000	40,908,186.000	37,313,874.000	−29,205.040	7,963,464.000
	t = −0.454	t = 4.714***	t = 2.288**	t = 4.101***	t = −0.0005	t = 0.268
Observations	295	296	295	294	293	292
R^2	0.342	0.952	0.900	0.904	0.352	0.091
Adjusted R^2	0.326	0.951	0.898	0.902	0.336	0.069

Note: *p<0.1; **p<0.05; ***p<0.01

One of the most interesting findings in the individual stock regressions is the strong results for NOK. What makes the NOK example particularly interesting is the generally high adjusted R-squareds in the regressions. Model (2) in table 5.5 has an adjusted R-squared of 95.1%, meaning that it describes 95.1% of the variation in NOK's trading volume. Also model (4) in table 5.5 has a very high adjusted R-squared of 96.7%, meaning

that we also manage to explain most of the variation in call option volume. Even more interestingly, when introducing lagged volume in table 5.6 we observe in models (3) and (4) that the models explain approximately 90% of the volume each of the two succeeding days. The differences in adjusted R-squareds in the models using data for 2020 and those using data for 2021 are particularly interesting. While the adjusted R-squareds in 2020 are very low in general, they are remarkably high in 2021 with several adjusted R-squareds close to 98%. Although one should be cautious using this for inference, the striking difference in adjusted R-squareds from 2020 to 2021 may have some merit in describing the potential impact of sentiment on WallStreetBets on stock and option metrics. NOK surged in 2021, and these models carry important insights as to what may have, at least partially, contributed to drive this surge.

Table 5.7: NOK 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.004 t = −0.263	109,262.200 t = 0.420	0.034 t = 1.372	144.711 t = 0.372	14.844 t = 0.318	−0.000 t = −0.299
Positive sentiment comments	0.012 t = 2.164**	27,060.700 t = 0.505	−0.046 t = −1.354	144.755 t = 1.865*	−14.154 t = −0.939	0.000 t = 0.889
Title per day	0.360 t = 0.661	3,732,993.000 t = 0.626	−1.939 t = −1.276	10,657.550 t = 1.152	−849.667 t = −0.659	0.000 t = 1.749*
Positive sentiment title	0.021 t = 0.686	−178,160.200 t = −1.063	−0.006 t = −0.117	−129.714 t = −0.558	3.320 t = 0.094	0.000 t = 0.243
Title text per day	0.546 t = 0.932	−1,865,386.000 t = −0.284	0.175 t = 0.074	1,813.761 t = 0.174	1,508.679 t = 0.671	0.000 t = 0.508
Positive sentiment title text	−0.007 t = −0.562	106,408.800 t = 1.075	0.050 t = 0.808	133.913 t = 0.775	1.783 t = 0.071	−0.000 t = −0.758
Mkt.RF	0.970 t = 8.091***	−806,917.200 t = −1.551	−0.664 t = −0.816	188.525 t = 0.360	−340.455 t = −2.359**	0.000 t = 7.853***
SMB	0.319 t = 1.642					
HML	−0.262 t = −1.142					
RMW	0.275 t = 0.919					
CMA	−0.327 t = −0.597					
MOM	−0.133 t = −0.773					
Constant	−2.102 t = −1.694*	29,969,353.000 t = 3.567***	54.127 t = 11.839***	14,942.150 t = 1.077	6,184.203 t = 2.693***	−0.000 t = −0.550
Observations	252	252	252	252	252	252
R^2	0.475	0.153	0.046	0.250	0.049	0.590
Adjusted R^2	0.449	0.129	0.018	0.228	0.022	0.578

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5.8: NOK 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.004	52,266.740	0.011	200.725	43.086	−0.000
	t = −0.596	t = 2.370**	t = 1.258	t = 5.527***	t = 1.812*	t = −0.199
Positive sentiment comments	0.061	−24,415.330	0.063	132.527	−330.789	0.000
	t = 0.958	t = −0.077	t = 0.752	t = 0.147	t = −1.122	t = 2.604***
Title per day	0.150	−89,138.550	0.142	536.752	−325.168	0.000
	t = 0.661	t = −0.113	t = 0.570	t = 0.568	t = −0.383	t = 0.177
Positive sentiment title	−0.028	17,165.610	0.146	1,174.159	443.937	0.000
	t = −0.586	t = 0.083	t = 3.651***	t = 2.165**	t = 2.015**	t = 0.975
Title text per day	−0.391	371,620.400	−0.516	−3,859.314	183.803	−0.000
	t = −0.735	t = 0.196	t = −0.854	t = −1.661*	t = 0.093	t = −0.189
Positive sentiment title text	0.023	−167,763.500	0.019	−220.537	−315.481	−0.000
	t = 0.590	t = −0.880	t = 0.389	t = −0.393	t = −1.594	t = −0.193
Mkt.RF	0.238	−1,104,793.000	1.290	1,208.854	−6,689.982	0.000
	t = 0.206	t = −0.209	t = 0.757	t = 0.068	t = −1.046	t = 1.574
SMB	−0.230					
	t = −0.157					
HML	0.305					
	t = 0.204					
RMW	1.089					
	t = 0.590					
CMA	3.263					
	t = 0.605					
MOM	1.548					
	t = 0.861					
Constant	−3.892	54,254,022.000	40.283	23,728.990	31,201.970	−0.00000
	t = −0.706	t = 2.054**	t = 5.696***	t = 0.311	t = 1.311	t = −3.677***
Observations	44	44	44	44	44	44
R^2	0.906	0.989	0.979	0.983	0.908	0.341
Adjusted R^2	0.869	0.987	0.975	0.979	0.891	0.213

Note: *p<0.1; **p<0.05; ***p<0.01

In terms of statistically significant correlation between sentiment and stock and option metrics, when using data from the entire time interval we find statistically significant relationships between up to several of the sentiment variables and NOK's return, call option volume and put option volume. When using data for only 2020 the only statistically significant sentiment variable in any of the models is in model (1) in table 5.7, where Positive sentiment comments is statistically significantly correlated with NOK's return at the 5% level. When using data for 2021 alone, the results are much stronger. Sentiment is indicated to have had a statistically significant impact on both trading volume, implied volatility, call option volume and put option volume, and the NOK example thereby serves as one of the clearest findings from the individual stock regressions.

Table 5.9: NVDA full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.003 t = 3.142***	13,244.420 t = 3.249***	0.002 t = 0.406	169.125 t = 4.050***	61.261 t = 2.670***	0.000 t = 1.177
Positive sentiment comments	0.010 t = 0.926	−14,793.920 t = −0.725	−0.012 t = −0.248	−50.607 t = −0.142	−205.150 t = −1.864*	0.000 t = 1.612
Title per day	−0.235 t = −1.982**	−53,835.330 t = −0.172	−0.181 t = −0.482	958.698 t = 0.277	542.971 t = 0.343	−0.000 t = −0.820
Positive sentiment title	0.001 t = 0.103	−35,332.700 t = −2.622***	−0.045 t = −1.709*	−111.127 t = −0.696	−123.808 t = −1.854*	0.000 t = 0.735
Title text per day	−0.011 t = −0.133	116,054.800 t = 0.410	0.575 t = 1.564	−535.359 t = −0.177	20.726 t = 0.015	−0.000 t = −0.416
Positive sentiment title text	0.002 t = 0.435	−23,094.300 t = −1.840*	−0.051 t = −1.679*	129.657 t = 1.068	17.488 t = 0.309	0.000 t = 0.619
Mkt.RF	1.506 t = 23.786***	−245,120.200 t = −1.141	−1.122 t = −1.502	−540.902 t = −0.492	−1,455.929 t = −2.540**	0.000 t = 12.823***
SMB	0.251 t = 1.501					
HML	−0.690 t = −4.142***					
RMW	0.161 t = 0.656					
CMA	−0.178 t = −0.646					
MOM	0.252 t = 2.761***					
Constant	−1.020 t = −1.173	14,510,007.000 t = 8.064***	56.281 t = 13.210***	89,308.880 t = 3.054***	72,992.430 t = 8.182***	−0.000 t = −1.740*
Observations	296	296	296	296	296	296
R^2	0.783	0.230	0.084	0.326	0.250	0.623
Adjusted R^2	0.774	0.211	0.062	0.310	0.232	0.613
Note:						*p<0.1; **p<0.05; ***p<0.01

Another particularly interesting finding in the individual stock regressions is the results for NVDA. In the return model in table 5.9, Comments per day and Title per day are statistically significant at the 1% and 5% level, respectively. The same model has an adjusted R-squared of 77.4%, which is relatively high compared to most of the other individual stock regression models. Mkt.RF, HML and CMA are also statistically significant at the 1% level in the model, giving reason to believe that these factors explain much of the relatively high adjusted R-squared. Comments per day is also statistically significant in the volume model, call option volume model and put option volume model. When limiting the data to 2020 alone, as is done in table A1.47, Comments per day and Positive sentiment comments are both statistically significant at the 5% level in the return model, and the adjusted R-squared increases to 80.1%. This is remarkably strong compared to its comparables, and also much higher than the adjusted R-squared in the return model when using data from 2021 alone in table A1.48. In the return model for 2021, we also observe that none of the sentiment variables are statistically significant. Also, the adjusted R-squareds in the 2021-based models are generally lower than the

2020-based models. Comments per day is still statistically significant in the call option volume model, but there are no other statistically significant variables in any of the other models using data from 2021.

For NVDA, overall, the results are stronger for 2020 than for 2021, which is no surprise since NVDA was frequently discussed in 2020, but was replaced with e.g. GME, BB and AMC in 2021. Still, when combining the data for 2020 and 2021, as is done in table 5.9, sentiment is suggested to have had a statistically significant impact on return, volume, call option volume and put option volume. However, again, using these results to conclude that sentiment has had an causal effect on NVDA's stock and option metrics is not necessarily correct both because we cannot rule out reverse causality in the models with lagged variables, and we do neither know how much of the total variation is explained by the return on the benchmark portfolio and Fama-French factors.

Table 5.10: RKT lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.005	0.002	0.004	−0.004	0.0002	0.001
	t = −1.171	t = 0.424	t = 0.963	t = −0.888	t = 0.096	t = 0.506
Positive sentiment comments	−0.023	0.072	0.133	−0.002	0.040	−0.015
	t = −0.360	t = 0.767	t = 2.000**	t = −0.029	t = 0.786	t = −0.268
Title per day	0.676	−0.365	−0.126	0.359	−0.044	0.049
	t = 1.347	t = −0.477	t = −0.256	t = 0.735	t = −0.260	t = 0.333
Positive sentiment title	0.029	0.007	0.021	−0.041	−0.006	−0.061
	t = 1.450	t = 0.256	t = 0.809	t = −1.233	t = −0.197	t = −1.681*
Title text per day	−0.892	0.368	−0.708	0.494	0.185	−0.700
	t = −1.918*	t = 0.352	t = −0.791	t = 1.379	t = 0.803	t = −1.560
Positive sentiment title text	−0.003	0.015	0.015	−0.034	−0.031	0.059
	t = −0.153	t = 0.747	t = 0.610	t = −1.024	t = −1.514	t = 1.899*
Mkt.RF	0.237	0.713	0.381	−0.196	0.694	−0.928
	t = 0.420	t = 0.738	t = 0.547	t = −0.235	t = 0.675	t = −1.245
SMB	0.413	0.097	0.061	−3.952	−0.149	0.624
	t = 0.261	t = 0.070	t = 0.031	t = −2.139**	t = −0.131	t = 0.481
HML	−1.472	0.359	−1.883	2.833	−3.206	2.122
	t = −0.963	t = 0.156	t = −1.020	t = 1.341	t = −2.194**	t = 0.928
RMW	0.836	−0.440	−1.353	−3.697	−0.393	1.673
	t = 0.476	t = −0.222	t = −0.700	t = −2.017**	t = −0.143	t = 0.447
CMA	−0.651	1.397	0.830	−0.809	3.295	−3.609
	t = −0.261	t = 0.510	t = 0.315	t = −0.319	t = 1.359	t = −1.481
MOM	−0.487	0.615	−0.883	0.120	−1.448	0.153
	t = −0.657	t = 0.574	t = −0.663	t = 0.095	t = −1.763*	t = 0.112
Constant	0.735	−7.059	−11.771	3.889	−0.639	1.862
	t = 0.150	t = −0.946	t = −1.942*	t = 0.607	t = −0.143	t = 0.436
Observations	145	146	145	144	143	142
R^2	0.295	0.063	0.331	0.181	0.049	0.140
Adjusted R^2	0.231	−0.022	0.270	0.106	−0.039	0.060
Note:						*p<0.1; **p<0.05; ***p<0.01

Another particularly interesting finding from the individual stock regressions is found in table 5.10, which estimates the effect of the sentiment variables as well the benchmark portfolio and Fama-French factors on lagged return variables for RKT. Interestingly,

comparing adjusted R-squared in model (2) and (3), return tomorrow is better explained than return today, and Positive sentiment comments is the only statistically significant variable in model (3). This finding suggests that sentiment for RKT on a given day has had a statistically significant impact on RKT's return the following day, indicating that sentiment might have driven future stock returns. In isolation, this indicates that WallStreetBets actually may exercise market power in terms of being able to move stock returns. Caution should be exercised when using this isolated result as a conclusion. The adjusted R-squared is still relatively low, and this isolated result gives no conclusive evidence on whether WallStreetBets actually exercises market power.

Table 5.11: RKT lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−11,394.930	14,652.660	−9,833.137	−28,209.250	−13,187.830	−16,775.290
	t = −0.476	t = 0.850	t = −1.032	t = −1.208	t = −2.191**	t = −1.513
Positive sentiment comments	11,276.840	215,647.200	−10,250.670	−287,391.500	−320,847.300	−260,370.100
	t = 0.064	t = 1.468	t = −0.093	t = −1.388	t = −1.562	t = −1.874*
Title per day	1,410,331.000	−1,840,373.000	1,418,408.000	2,658,529.000	1,050,440.000	976,678.300
	t = 0.440	t = −0.788	t = 1.243	t = 1.061	t = 1.428	t = 0.913
Positive sentiment title	44,642.970	65,741.710	−23,671.230	−36,446.250	−66.604	−121,606.200
	t = 1.013	t = 1.523	t = −0.606	t = −0.389	t = −0.001	t = −0.991
Title text per day	324,550.800	3,873,774.000	1,533,670.000	2,176,019.000	1,601,274.000	3,465,744.000
	t = 0.096	t = 1.198	t = 1.052	t = 1.033	t = 1.990**	t = 2.207**
Positive sentiment title text	−25,113.250	−17,629.960	−21,733.910	−104,199.400	−43,796.390	22,207.460
	t = −0.489	t = −0.321	t = −0.461	t = −0.915	t = −0.704	t = 0.194
Mkt.RF	284,445.400	−1,692,148.000	181,103.800	−364,842.800	3,278,314.000	2,671,219.000
	t = 0.122	t = −0.626	t = 0.112	t = −0.110	t = 0.761	t = 1.016
Constant	11,240,935.000	−8,660,851.000	12,246,655.000	40,321,448.000	39,692,544.000	36,390,062.000
	t = 0.766	t = −0.831	t = 1.298	t = 1.741*	t = 2.135**	t = 2.753***
Observations	145	146	145	144	143	142
R^2	0.169	0.647	0.807	0.325	0.080	0.153
Adjusted R^2	0.126	0.629	0.797	0.290	0.032	0.109
Note:						*p<0.1; **p<0.05; ***p<0.01

When comparing model (2) and (3) in table 5.11, i.e. in the lagged volume models for RKT, we also observe that volume tomorrow is better explained by sentiment today. The adjusted R-squared in model (2) is 62.9%, while the adjusted R-squared in model (3) is 79.7%. There are no statistically significant variables in neither of the models, but the difference in adjusted R-squared in the two models is still interesting. The difference in explanatory power between 2020 and 2021 is also striking. All models except the Amihud model in table A1.59, where data for only 2021 is used, have higher adjusted R-squareds than their comparables in table A1.58, i.e. when using data for only 2020. No sentiment variables are statistically significant in the 2021-based models, so we do not argue that

sentiment is found to have had an impact on the stock and option metrics for RKT in 2021. We still want to highlight this particular finding, because RKT experienced a significant surge in February 2021 when it was frequently discussed at WallStreetBets. The generally higher adjusted R-squareds in the models for 2021 compared to 2020 seem to capture this feature, although the models fail to indicate any statistically significant impact of the sentiment variables on this.

Table 5.12: TSLA full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0002 t = −0.429	3,779.714 t = 1.844*	0.004 t = 3.291***	122.774 t = 2.758***	123.938 t = 3.167***	0.000 t = 0.939
Positive sentiment comments	0.159 t = 3.297***	−7,700.705 t = −0.041	−0.538 t = −2.889***	1,768.738 t = 0.277	−20,413.150 t = −3.308***	0.000 t = 1.587
Title per day	0.051 t = 1.342	279,964.900 t = 1.537	−0.232 t = −1.540	2,921.466 t = 0.578	−1,813.428 t = −0.510	−0.000 t = −0.653
Positive sentiment title	0.001 t = 0.073	64,355.070 t = 1.798*	−0.090 t = −2.189**	−2,764.940 t = −1.991**	−3,151.955 t = −2.751***	0.000 t = 0.337
Title text per day	−0.040 t = −0.997	−19,015.680 t = −0.075	0.295 t = 2.320**	−3,172.163 t = −0.525	−2,693.718 t = −0.499	0.000 t = 0.243
Positive sentiment title text	−0.008 t = −0.541	8,394.529 t = 0.140	−0.074 t = −1.594	−5,621.732 t = −2.334**	−4,093.105 t = −2.466**	−0.000 t = −0.587
Mkt.RF	1.403 t = 7.699***	−276,312.300 t = −0.743	−0.646 t = −0.707	14,058.140 t = 1.092	−12,766.800 t = −0.971	0.000 t = 4.985***
SMB	0.629 t = 1.441					
HML	−0.346 t = −0.738					
RMW	−1.107 t = −1.882*					
CMA	−1.228 t = −1.490					
MOM	0.232 t = 0.767					
Constant	−8.601 t = −2.740***	11,970,589.000 t = 1.230	113.351 t = 9.152***	1,363,823.000 t = 3.656***	2,392,145.000 t = 6.676***	−0.000 t = −1.238
Observations	296	296	296	296	296	296
R^2	0.407	0.322	0.250	0.156	0.206	0.358
Adjusted R^2	0.382	0.305	0.231	0.136	0.187	0.342
Note:						*p<0.1; **p<0.05; ***p<0.01

Furthermore, the results for TSLA are especially interesting. Throughout 2020 and the beginning of 2021 TSLA's stock price surged, and the market capitalization of the company has seen a dramatic rise. As of May 23, 2021, TSLA's market capitalization is larger than the combined market capitalization of the four next largest automakers by market capitalization (CompaniesMarketCap.com, 2021). TSLA has also been one of the most frequently discussed stocks on WallStreetBets for a long time. Given the large market capitalization of TSLA as well as WallStreetBets' interest in TSLA , the results for TSLA are particularly interesting. In model (1) in table 5.12, Positive sentiment comments is statistically significant at the 1% level, indicating that sentiment may have

had an impact on TSLA's return when using data for the entire time interval. Model (3), (4) and (5) also contain several statistically significant sentiment variables. This is a striking finding, indicating that sentiment may have had an impact on TSLA's implied volatility, call option volume and put option volume, respectively. Although the adjusted R-squareds of these models are relatively low, these results nonetheless indicate that sentiment on WallStreetBets actually may have had affected these option metrics, as well as TSLA's return. However, this may again be a result of reverse causality. In table A1.70 with lagged return variables, we observe that return today in model (2) has a much higher adjusted R-squared than any of the other lead or lag variables. One should pay particular attention to the extreme drop in adjusted R-squared from model (2) to (3), which can be argued to indicate reverse causality. Positive sentiment on WallStreetBets a given day for TSLA may therefore mirror TSLA's returns that day, meaning that we are unable to distinguish what has affected what. This may also be true for the option metrics. We therefore advise to exercise caution when using these results, although there are several statistically significant sentiment variables in the different models for TSLA.

5.2.1.2 General trends

We find that sentiment is indicated to have had a statistically significant effect on several of the stocks' returns. When introducing leading and lagged return variables, we observe for some stocks that the estimated effect of sentiment on return is limited to be statistically significant only in the models with return for the same day. These findings on returns where especially the variables describing positive sentiment are statistically significant, instead of comment frequency, may therefore indicate reverse causality problems, i.e. a stock's return a given day may explain the sentiment on WallStreetBets for the stock the same day, not vice versa. As such, we cannot conclude that sentiment has had a causal effect on several of the stocks' returns even though we find that one or more of the sentiment variables are statistically significant with several of the stocks' returns the same day.

On the other hand, we observe several stocks where sentiment variables are statistically

significant on one day, but also experience a higher adjusted R-squared the succeeding days which might suggest WallStreetBets sentiment could actually drive stock returns. We therefore cannot exclude that this relationship holds either, i.e. sentiment can potentially have had an actual impact on stocks' returns. Not being able to distinguish between these relationships, we advise to exercise caution when interpreting the results in the regression tables. While we find that sentiment may have had an impact on some of the stocks' returns, the effect of sentiment on volume seems in general to be stronger than the effect of sentiment on returns. We find that sentiment variables are often statistically significant in the volume models in the individual stock regressions, and for several of the stocks we also observe that sentiment on WallStreetBets on a given day may have had an effect on the trading volume the following days. In general, therefore, volume seems to be more sensitive to sentiment on WallStreetBets for more of the stocks than returns. However, again, we argue to exercise caution when making inference with these results.

The strongest pattern in the individual stock regressions is that sentiment appears to have had a general effect on the option metrics, i.e. implied volatility, call option volume and put option volume. Most often we find statistically significant sentiment variables in the option-related models, which indicates that sentiment on WallStreetBets actually may have had a larger effect on these metrics than stock returns and trading volume. As we previously have discussed, users on WallStreetBets often advocate to trade in options instead of trading in the stocks directly, and these findings are therefore particularly interesting as they also seem to apply to more stocks than returns and trading volume. However, we have not applied any lagged option-related variables, so these results can also potentially be a result of reverse causality, but we would argue that this is much less likely to be a problem for the option metrics than for return. The pattern is nevertheless stronger than for returns and trading volume, which is indicative of a more general effect of sentiment on these metrics. Also, the adjusted R-squareds in these models are generally very high compared to the other metrics' models. Therefore, we find that the effect of sentiment on any of the metrics we analyze in this thesis is indicated to be strongest on implied volatility, call option volume and put option volume. When dividing the results

into 2020 and 2021, we find that the results differ. While the return, trading volume and option metrics are better explained for some stocks in 2020, these metrics are better explained for other stocks in 2021. This is no surprise, since there are large variations as to when the different stocks have been heavily discussed on WallStreetBets. For example, AMC, BB, NOK and GME, which all were heavily discussed on the forum in 2021 in general have the strongest results in 2021, while others, such as NVDA, in general have the strongest results in 2020. Remark, however, that there are large differences also at the individual stock level which year the different models yield the strongest results in.

5.2.2 Sentiment variables on an aggregated stock sample

Having performed analyses on each of the stocks individually, in this subchapter we will perform analyses on the stock selection overall. First, we will perform an analysis on the time period overall. Then, we will present a regression table with lagged return variables. Next, a regression table with lagged volume is presented. Thereafter, we present a regression table based on data from only 2020 to see whether the sentiment variables display statistically significant correlations with return or the other stock metrics. Lastly, we repeat the process we did for 2020, but replace the data with data from only 2021. This makes us able to compare potential changes and trends from one year to the other, in addition to making comparisons of whether sentiment can be claimed to partially explain more of variation in the stock and option metrics in one year rather than the other.

Table 5.13: Full aggregated regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.00002	−1,691.337	0.0005	−0.142	1.870	0.000
	t = 0.118	t = −1.856*	t = 1.545	t = −0.067	t = 1.466	t = 0.869
Positive sentiment comments	0.031	39,940.160	−0.085	358.151	−325.393	0.000
	t = 7.078***	t = 2.045**	t = −4.123***	t = 3.784***	t = −4.936***	t = 2.829***
Title per day	−0.009	246,801.900	0.090	314.451	124.038	0.000
	t = −0.460	t = 1.839*	t = 2.574**	t = 1.136	t = 1.485	t = 0.316
Positive sentiment title	0.025	248,762.800	0.090	853.156	−264.086	0.000
	t = 5.317***	t = 9.476***	t = 5.088***	t = 3.641***	t = −1.793*	t = 2.391**
Title text per day	0.031	−185,350.600	−0.059	−243.999	−53.555	−0.000
	t = 0.610	t = −0.680	t = −0.617	t = −0.434	t = −0.188	t = −0.147
Positive sentiment title text	0.011	125,916.400	0.045	336.033	−42.651	0.000
	t = 2.495**	t = 6.097***	t = 2.695***	t = 2.259**	t = −0.425	t = 2.100**
Mkt.RF	1.159	−770,716.000	−0.934	−1,044.032	−4,815.201	0.00000
	t = 18.774***	t = −2.767***	t = −2.791***	t = −0.605	t = −2.938***	t = 9.757***
SMB	1.262					
	t = 5.231***					
HML	−0.713					
	t = −3.778***					
RMW	−0.466					
	t = −1.816*					
CMA	1.940					
	t = 2.339**					
MOM	−0.026					
	t = −0.181					
Constant	0.690	−46,484,692.000	120.406	−1,645,041.000	−936,698.100	−0.00000
	t = 0.839	t = −7.939***	t = 33.546***	t = −28.275***	t = −28.537***	t = −2.523**
TickerAMD	−0.103	−21,734,181.000	18.246	−1,444,848.000	−833,383.900	−0.000
	t = −0.347	t = −5.983***	t = 18.439***	t = −24.561***	t = −25.333***	t = −0.448
TickerAPHA	0.548	−66,950,305.000	64.077	−1,695,872.000	−960,372.900	−0.000
	t = 1.116	t = −18.598***	t = 37.896***	t = −29.475***	t = −29.521***	t = −0.837
TickerBB	0.337	−65,149,642.000	38.810	−1,680,342.000	−957,515.100	−0.000
	t = 0.648	t = −16.417***	t = 20.500***	t = −29.096***	t = −29.362***	t = −0.245
TickerGME	1.063	−71,543,680.000	93.182	−1,684,801.000	−938,722.600	−0.00000
	t = 1.308	t = −18.097***	t = 33.300***	t = −28.781***	t = −28.271***	t = −1.479
TickerGOOG	1.974	−61,010,479.000	0.705	−1,650,266.000	−972,199.400	0.000
	t = 4.520***	t = −16.172***	t = 0.503	t = −30.821***	t = −29.876***	t = 2.269**
TickerNIO	0.884	31,320,522.000	78.826	−1,412,580.000	−835,502.400	0.000
	t = 2.018**	t = 5.149***	t = 58.073***	t = −23.567***	t = −24.903***	t = 0.061
TickerNKLA	0.270	−51,670,691.000	103.640	−1,614,554.000	−912,064.300	0.000
	t = 0.342	t = −13.778***	t = 37.072***	t = −28.374***	t = −27.771***	t = 1.128
TickerNOK	−0.003	−37,810,346.000	19.285	−1,649,236.000	−953,604.500	−0.000
	t = −0.007	t = −8.157***	t = 15.140***	t = −28.317***	t = −29.320***	t = −0.499
TickerNVDA	0.304	−67,430,386.000	12.913	−1,609,933.000	−903,602.800	0.000
	t = 0.849	t = −19.217***	t = 12.064***	t = −27.949***	t = −27.710***	t = 0.005
TickerPLTR	−3.701	−11,664,150.000	63.825	−1,385,946.000	−814,281.500	−0.000
	t = −0.958	t = −1.698*	t = 30.885***	t = −21.041***	t = −22.944***	t = −1.742*
TickerPLUG	1.172	−50,681,100.000	59.791	−1,658,713.000	−945,138.400	0.000
	t = 2.532**	t = −13.853***	t = 45.064***	t = −28.813***	t = −28.917***	t = 1.127
TickerRKT	−0.894	−64,955,060.000	45.758	−1,643,886.000	−935,829.900	−0.000
	t = −1.020	t = −14.342***	t = 24.912***	t = −27.548***	t = −28.369***	t = −2.313**
TickerSPCE	0.200	−62,653,226.000	79.928	−1,631,864.000	−928,083.500	−0.000
	t = 0.431	t = −17.776***	t = 40.956***	t = −28.024***	t = −28.366***	t = −0.842
TickerTLRY	0.498	−60,544,828.000	90.710	−1,632,735.000	−944,590.000	−0.00000
	t = 0.735	t = −15.893***	t = 39.522***	t = −28.339***	t = −28.895***	t = −3.498***
TickerTSLA	0.432	−47,804,315.000	39.780	−533,776.500	−90,904.130	0.000
	t = 0.831	t = −12.049***	t = 26.368***	t = −7.379***	t = −1.979**	t = 0.381
TickerTSM	0.380	−66,003,433.000	2.721	−1,681,117.000	−952,629.800	0.000
	t = 0.983	t = −18.463***	t = 2.549**	t = −29.304***	t = −29.304***	t = 0.101
Constant	−4.713	52,942,225.000	34.167	1,622,214.000	1,005,548.000	−0.00000
	t = −7.325***	t = 12.021***	t = 16.040***	t = 30.202***	t = 29.690***	t = −4.188***
Observations	4,884	4,884	4,880	4,880	4,880	4,884
R^2	0.096	0.450	0.707	0.669	0.678	0.119
Adjusted R^2	0.091	0.448	0.706	0.667	0.677	0.115

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5.13 is a regression output with categorical variables, which is based on the entire time period being subject for analysis in this thesis, i.e. January 02, 2020, to March 15, 2021. When aggregating the stocks in a general regression model, we observe in model (1) that Positive sentiment comments and Positive sentiment title both are statistically

significant at the 1% level, and Positive sentiment title text is statistically significant at the 5% level. All the statistically significant sentiment variables indicate a positive effect on returns. In addition, we observe that all the statistically significant variables are direct measures of the positive sentiment in the comments, submission titles or submission texts, not variables explaining the frequency of submissions and comments being made about the stocks in our sample. Positive sentiment is therefore indicated to have had an impact on stocks' return overall, even when controlling for that Mkt.RF, SMB and HML are all significant at the 1% level, while the CMA factor is statistically significant at the 5% level.

The positive coefficient on the SMB factor indicates that small companies are relatively over-weighted in our sample. The negative coefficient on the HML indicates that the firms in our sample generally have low book-to-market ratios, while the positive coefficient on the CMA factor indicates that firms with conservative investment policies are relatively over-weighted in the sample.

The stock-specific categorical variables differ in terms of statistical significance. The adjusted R-squared of 9.1% is also relatively low. When taking into account the illiquidity factor AMIHUD in model (6), we can infer that a larger variation of the returns could potentially be explained by liquidity rather than WallStreetBets sentiment and Fama-French alone. Overall, the results nonetheless suggest that sentiment has had an impact on stocks' return, but the magnitude of the coefficients on the statistically significant sentiment variables indicate that the effect of sentiment on stock returns in general is relatively low. However, we cannot conclude directly that sentiment has had an impact on stock returns, since these effects may be a result of reverse causality, especially considering that positive sentiment might be attributed to positive returns on the day stocks are discussed. We have not included any lagged variables in this model, meaning that we are unable to distinguish whether sentiment has had an impact on returns or vice versa. Therefore, although model (1) indicates that sentiment has had an impact on stock returns overall, we are unable to use this as conclusive evidence for sentiment having had a causal effect on stock returns.

In model (2) in table 5.13, we observe that Positive sentiment title and Positive sentiment title text are statistically significant at the 1% level, while Positive sentiment comments is significant at the 5% level. All categorical variables except TickerPLTR are significant at the 1% level. The adjusted R-squared of this model is 44.8%, indicating a higher explanatory power than model (1). The statistically significant sentiment variables generally suggest a positive impact on trading volume, although the net stock-specific effect when controlling for the general negative coefficient on the categorical variables is not clear-cut. However, the model nevertheless suggests that sentiment has had an impact on trading volume, and sentiment seems to better explain variations in trading volume. When not controlling for lagged volume variables, we are unable to distinguish between whether trading volume drives sentiment or whether sentiment drives trading volume. In general, therefore, we advise to exercise caution when interpreting this result in isolation.

The option-related models, i.e. model (3), (4) and (5), all contain statistically significant sentiment variables. They also generally display higher adjusted R-squareds than the stock-related models, i.e. model (1) and (2). In the implied volatility model, i.e. model (3), Positive sentiment comments, Positive sentiment title and Positive sentiment title text are all statistically significant at the 1% level, while Title per day is significant at the 5% level. The coefficients on Title per day, Positive sentiment title and Positive sentiment title text are positive, while the coefficient on Positive sentiment comments is negative. The total net effect on implied volatility is therefore not clear, but the model nevertheless suggests that sentiment in general has had an impact on implied volatility. In model (4), all statistically significant sentiment variables have positive coefficients, indicating that positive sentiment has had a positive effect on call option volume. This finding is interesting, since WallStreetBets often is flooded by posts encouraging its users to trade in call options. Bear in mind, that the net direction of the effect is not clear-cut given the general negative coefficient on the categorical variables. In model (5), the only statistically significant sentiment variable is Positive sentiment comments, which is significant at the 1% level. The direction of the coefficient is negative, indicating that

increased positive sentiment lowers put option volume, which is the direction one in general would intuitively expect to observe.

As with model (1), the results from these option-related models, i.e. model (3), (4) and (5), can be results of reverse causality problems. We do not know whether sentiment drives the option metrics or whether the variations in the option metrics drive sentiment. However, we would argue that the effect of sentiment on these option-related metrics in general seems stronger than the effect on the stock-related metrics. The adjusted R-squareds are generally much higher in the option-related models, which we also observed in the individual stock regressions, and we would also argue that the problem of reverse causality in general in these models seems less likely than in the stock-related models. Although highly possible due to both direct or indirect effects, we believe that investors' sentiment for a stock is generally less affected by variations in implied volatility and option volumes than returns and trading volume. Note, that variations in implied volatility and option volumes can create spillover effects on stock returns and trading volumes and thereover on sentiment, so we cannot statistically rule out this possibility. Still, we would argue that this seems less likely than for the stock-related metrics, and this relationship therefore seems in general to be stronger than in the stock-related models.

Table 5.14: Lagged return aggregated regression

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0002	0.00002	−0.00004	0.0003	0.0001	−0.0001
	t = 0.938	t = 0.118	t = −0.212	t = 1.566	t = 1.184	t = −1.451
Positive sentiment comments	0.016	0.031	0.024	0.018	0.014	0.010
	t = 3.154***	t = 7.078***	t = 4.618***	t = 2.882***	t = 2.545**	t = 2.218**
Title per day	−0.026	−0.009	−0.005	−0.006	0.007	0.037
	t = −1.590	t = −0.460	t = −0.326	t = −0.401	t = 0.275	t = 1.782*
Positive sentiment title	0.006	0.025	0.010	0.012	0.016	0.006
	t = 1.173	t = 5.317***	t = 1.679*	t = 2.398**	t = 2.634***	t = 1.325
Title text per day	0.040	0.031	0.055	0.004	−0.022	−0.063
	t = 1.251	t = 0.610	t = 1.704*	t = 0.128	t = −0.450	t = −1.660*
Positive sentiment title text	0.007	0.011	0.019	0.002	0.010	0.013
	t = 1.643	t = 2.495**	t = 3.589***	t = 0.518	t = 1.970**	t = 2.298**
Mkt.RF	−0.259	1.159	−0.307	0.268	0.145	−0.088
	t = −4.090***	t = 18.774***	t = −4.371***	t = 3.224***	t = 1.694*	t = −1.224
SMB	0.347	1.262	−0.419	0.188	−0.610	0.298
	t = 1.811*	t = 5.231***	t = −1.986**	t = 0.936	t = −3.020***	t = 1.748*
HML	−0.276	−0.713	0.440	0.325	−0.033	0.007
	t = −1.496	t = −3.778***	t = 2.612***	t = 1.592	t = −0.182	t = 0.039
RMW	0.342	−0.466	−0.272	−0.226	−0.917	0.111
	t = 1.476	t = −1.816*	t = −0.849	t = −0.746	t = −3.396***	t = 0.427
CMA	−1.255	1.940	−1.550	0.050	1.100	−0.189
	t = −2.623***	t = 2.339**	t = −3.047***	t = 0.135	t = 2.692***	t = −0.462
MOM	−0.214	−0.026	−0.216	0.463	−0.118	0.063
	t = −1.694*	t = −0.181	t = −1.447	t = 3.643***	t = −0.938	t = 0.465
Constant	1.030	0.690	0.217	0.312	0.124	−0.855
	t = 1.068	t = 0.839	t = 0.277	t = 0.361	t = 0.132	t = −0.997
TickerAMD	0.101	−0.103	0.052	0.038	−0.129	−0.279
	t = 0.338	t = −0.347	t = 0.174	t = 0.129	t = −0.413	t = −0.917
TickerAPHA	0.281	0.548	0.705	0.488	0.344	−0.005
	t = 0.793	t = 1.116	t = 1.582	t = 1.088	t = 0.699	t = −0.011
TickerBB	0.506	0.337	0.348	0.218	0.128	−0.370
	t = 1.016	t = 0.648	t = 0.762	t = 0.442	t = 0.222	t = −0.684
TickerGME	1.244	1.063	0.489	−0.946	−0.009	0.214
	t = 1.619	t = 1.308	t = 0.627	t = −1.069	t = −0.014	t = 0.296
TickerGOOG	0.765	1.974	1.314	1.073	1.145	0.233
	t = 1.753*	t = 4.520***	t = 3.043***	t = 2.530**	t = 2.018**	t = 0.515
TickerNIO	0.954	0.884	0.982	0.862	0.708	0.402
	t = 2.176**	t = 2.018**	t = 2.332**	t = 2.019**	t = 1.537	t = 0.902
TickerNKLA	0.197	0.270	0.326	−0.086	−0.071	−0.521
	t = 0.245	t = 0.342	t = 0.404	t = −0.107	t = −0.084	t = −0.619
TickerNOK	0.273	−0.003	0.083	0.022	−0.155	−0.736
	t = 0.623	t = −0.007	t = 0.209	t = 0.051	t = −0.282	t = −1.489
TickerNVDA	0.354	0.304	0.381	0.331	0.121	−0.224
	t = 1.057	t = 0.849	t = 1.203	t = 1.044	t = 0.336	t = −0.658
TickerPLTR	0.974	−3.701	−3.632	−3.453	−3.520	−3.671
	t = 1.289	t = −0.958	t = −0.949	t = −0.884	t = −0.880	t = −0.924
TickerPLUG	1.131	1.172	1.335	1.072	0.909	0.540
	t = 2.413**	t = 2.532**	t = 2.919***	t = 2.370**	t = 1.822*	t = 1.132
TickerRKT	0.175	−0.894	−0.605	−0.681	−0.826	−1.075
	t = 0.252	t = −1.020	t = −0.703	t = −0.780	t = −0.910	t = −1.201
TickerSPCE	0.380	0.200	0.375	0.284	0.106	−0.154
	t = 0.785	t = 0.431	t = 0.792	t = 0.602	t = 0.212	t = −0.313
TickerTLRY	0.462	0.498	0.622	0.339	0.154	−0.239
	t = 0.673	t = 0.735	t = 0.945	t = 0.510	t = 0.216	t = −0.347
TickerTSLA	0.396	0.432	0.199	0.276	0.547	0.841
	t = 0.902	t = 0.831	t = 0.483	t = 0.687	t = 1.327	t = 2.046**
TickerTSM	0.377	0.380	0.586	0.308	0.221	−0.155
	t = 1.112	t = 0.983	t = 1.835*	t = 0.971	t = 0.573	t = −0.447
Constant	−2.165	−4.713	−3.789	−2.285	−2.590	−1.581
	t = −3.272***	t = −7.325***	t = −6.745***	t = −3.527***	t = −2.834***	t = −2.325**
Observations	4,866	4,884	4,866	4,848	4,830	4,812
R^2	0.049	0.096	0.117	0.066	0.032	0.063
Adjusted R^2	0.044	0.091	0.112	0.060	0.026	0.057
Note:						*p<0.1; **p<0.05; ***p<0.01

In table 5.14, we introduce lead and lag return variables to examine the effect of sentiment on WallStreetBets on stocks' return. In the previous section, where there were no lead or lag return variables, we found that several sentiment variables were statistically significant, but whether this was due to sentiment actually having had an impact on returns or

whether this was a result of reverse causality problems was unclear. When applying lead and lag variables we facilitate a more comprehensive analysis of the general effect of sentiment on returns.

Positive sentiment comments is statistically significant at the 1% level in model (1) in table 5.14. Also Mkt.RF and CMA are statistically significant at the 1% level in this model. This indicates that sentiment on WallStreetBets on a given day may be influenced by the previous day's returns, but the adjusted R-squared of this model is very low with only 4.4%. Model (2) is the same model as model (2) in table 5.14 in the previous section, and shows that several sentiment variables are statistically significantly correlated with today's return, while the adjusted R-squared of this model is 9.1%. In model (3) we observe that Positive sentiment comments and Positive sentiment title text both are statistically significant at the 1% level, also when controlling for the fact that Mkt.RF, HML and CMA are statistically significant at the 1% and that the SMB factor is statistically significant at the 5% level. The adjusted R-squared in model (3) is also higher than the adjusted R-squared in model (2). Overall, these findings suggest that sentiment on WallStreetBets on a given day may have had an impact on stocks' return the next day, and we are able to explain more of the variation in returns' the next day than the same day. These findings are particularly interesting when we account for the fact that the Fama-French factors have not been lagged, thus we can attribute more of the results we see to the actual sentiment on WallStreetBets. Even though the adjusted R-squared is low, these results give more indicative evidence that sentiment on WallStreetBets may have had an effect on stock returns on an aggregated basis. We also observe statistically significant sentiment variables in model (4), (5) and (6), which indicate that sentiment on WallStreetBets on a given day also can explain variations in returns two, three and four days forward, respectively. It is important, however, to note that the adjusted R-squareds in these models are generally very low.

Overall, these findings suggest that sentiment on WallStreetBets a given day may explain same day stock returns, but the problem of reverse causality is a critical drawback factor for making this conclusion. When introducing lagged variables, we observe that sentiment

on WallStreetBets a given day may have had an impact on stock returns over the next days and that the direction of this effect is positive. Since we have applied lagged return variables, we are able to control for parts of the reverse causality problems, which makes these results more robust than the results for the effect of sentiment on same day's returns. We are not able to completely rule these out, and the generally low adjusted R-squareds also suggest that caution should be exercised when interpreting these results. In general, these results do nonetheless give more indicative and striking evidence on sentiment on WallStreetBets having had an impact on stock returns.

Table 5.15: Lagged volume aggregated regression

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−902.427	−1,691.337	−917.744	−213.581	−294.519	−689.270
	t = −1.818*	t = −1.856*	t = −1.349	t = −0.524	t = −0.659	t = −1.210
Positive sentiment comments	67,555.990	39,940.160	10,988.340	19,960.630	15,487.700	12,417.180
	t = 2.883***	t = 2.045**	t = 0.549	t = 0.907	t = 0.788	t = 0.614
Title per day	137,075.000	246,801.900	141,217.500	123,553.600	150,028.200	160,395.500
	t = 1.731*	t = 1.839*	t = 1.520	t = 1.413	t = 1.409	t = 1.376
Positive sentiment title	264,101.200	248,762.800	239,064.000	235,220.300	225,992.000	194,012.800
	t = 8.073***	t = 9.476***	t = 8.689***	t = 7.826***	t = 8.128***	t = 6.212***
Title text per day	−84,630.280	−185,350.600	−36,157.060	−169,044.900	−255,800.800	−232,071.600
	t = −0.481	t = −0.680	t = −0.155	t = −1.047	t = −1.290	t = −1.117
Positive sentiment title text	117,508.100	125,916.400	109,245.600	109,789.500	118,651.100	98,750.100
	t = 4.485***	t = 6.097***	t = 5.201***	t = 5.013***	t = 4.887***	t = 4.589***
Mkt.RF	−370,409.700	−770,716.000	54,036.050	−343,835.000	15,966.110	266,270.200
	t = −1.603	t = −2.767***	t = 0.205	t = −1.252	t = 0.061	t = 1.107
Constant	−42,530,572.000	−46,484,692.000	−44,459,795.000	−42,925,807.000	−43,437,419.000	−45,281,457.000
	t = −6.408***	t = −7.939***	t = −7.962***	t = −6.983***	t = −7.371***	t = −7.815***
TickerAMD	−21,598,033.000	−21,734,181.000	−20,871,051.000	−20,882,368.000	−21,011,927.000	−20,764,568.000
	t = −6.068***	t = −5.983***	t = −5.832***	t = −5.885***	t = −5.847***	t = −5.735***
TickerAPHA	−66,427,249.000	−66,950,305.000	−65,813,382.000	−66,309,793.000	−66,783,228.000	−67,225,038.000
	t = −19.942***	t = −18.598***	t = −19.170***	t = −20.047***	t = −19.415***	t = −19.239***
TickerBB	−63,680,027.000	−65,149,642.000	−63,636,118.000	−63,163,345.000	−63,412,533.000	−64,104,424.000
	t = −16.821***	t = −16.417***	t = −17.043***	t = −16.415***	t = −15.786***	t = −15.550***
TickerGME	−69,765,337.000	−71,543,680.000	−71,544,136.000	−71,965,491.000	−71,471,517.000	−70,064,135.000
	t = −19.502***	t = −18.097***	t = −19.377***	t = −20.025***	t = −19.759***	t = −19.632***
TickerGOOG	−59,477,702.000	−61,010,479.000	−60,742,589.000	−61,312,446.000	−62,248,115.000	−64,554,621.000
	t = −15.817***	t = −16.172***	t = −17.406***	t = −17.118***	t = −16.668***	t = −16.715***
TickerNIO	31,929,608.000	31,320,522.000	32,417,360.000	32,263,338.000	31,848,017.000	31,349,204.000
	t = 5.336***	t = 5.149***	t = 5.379***	t = 5.341***	t = 5.224***	t = 5.080***
TickerNKLA	−50,762,954.000	−51,670,691.000	−51,256,353.000	−51,220,595.000	−51,483,648.000	−52,186,809.000
	t = −13.854***	t = −13.778***	t = −14.152***	t = −14.109***	t = −13.806***	t = −13.879***
TickerNOK	−36,357,240.000	−37,810,346.000	−35,741,786.000	−35,661,547.000	−36,154,951.000	−36,768,490.000
	t = −6.810***	t = −8.157***	t = −6.730***	t = −6.349***	t = −6.167***	t = −6.148***
TickerNVDA	−66,849,325.000	−67,430,386.000	−66,348,668.000	−66,626,641.000	−67,071,406.000	−67,214,779.000
	t = −20.506***	t = −19.217***	t = −19.781***	t = −20.645***	t = −20.119***	t = −19.876***
TickerPLTR	−13,009,399.000	−11,664,150.000	−10,281,125.000	−8,879,213.000	−8,671,363.000	−8,376,766.000
	t = −2.012**	t = −1.698*	t = −1.489	t = −1.261	t = −1.212	t = −1.173
TickerPLUG	−49,959,954.000	−50,681,100.000	−49,752,444.000	−50,076,185.000	−50,485,602.000	−50,926,220.000
	t = −14.623***	t = −13.853***	t = −14.228***	t = −14.722***	t = −14.401***	t = −14.268***
TickerRKT	−65,176,933.000	−64,955,060.000	−63,451,675.000	−63,685,483.000	−63,933,882.000	−63,496,746.000
	t = −14.936***	t = −14.342***	t = −14.364***	t = −14.369***	t = −13.996***	t = −13.812***
TickerSPCE	−62,249,521.000	−62,653,226.000	−61,571,607.000	−61,743,581.000	−62,042,596.000	−62,094,921.000
	t = −18.629***	t = −17.776***	t = −18.056***	t = −18.471***	t = −18.053***	t = −17.891***
TickerTLRY	−59,749,346.000	−60,544,828.000	−59,618,152.000	−60,066,668.000	−60,614,961.000	−61,118,736.000
	t = −16.875***	t = −15.893***	t = −16.419***	t = −17.085***	t = −16.626***	t = −16.426***
TickerTSLA	−48,084,352.000	−47,804,315.000	−48,817,326.000	−47,895,235.000	−47,238,479.000	−47,161,640.000
	t = −13.110***	t = −12.049***	t = −12.596***	t = −13.619***	t = −13.132***	t = −12.953***
TickerTSM	−65,555,527.000	−66,003,433.000	−64,833,416.000	−65,322,253.000	−65,770,282.000	−66,405,551.000
	t = −19.884***	t = −18.463***	t = −19.144***	t = −20.061***	t = −19.381***	t = −19.246***
Constant	50,063,777.000	52,942,225.000	55,227,700.000	55,301,677.000	56,044,233.000	59,696,544.000
	t = 9.825***	t = 12.021***	t = 13.816***	t = 12.743***	t = 12.631***	t = 13.448***
Observations	4,866	4,884	4,866	4,848	4,830	4,812
R^2	0.357	0.450	0.388	0.332	0.332	0.327
Adjusted R^2	0.354	0.448	0.385	0.329	0.328	0.324

Note: *p<0.1; **p<0.05; ***p<0.01

In table 5.15 we have introduced lead and lag volume variables. We observe that model (2) has several statistically significant sentiment variables, and has also the highest adjusted R-squared of the models with 44.8%. This indicates that sentiment on WallStreetBets a given day has had the highest impact on the same day's trading volume. Model (1) also contains several statistically significant sentiment variables, indicating that the preceding day's volume can partially explain sentiment on WallStreetBets the next day. We would argue that high volume the previous day does not itself seem like an intuitive explanatory factor for sentiment on WallStreetBets the next day, but knowing that high volume is often observed when returns are high, this finding can potentially reflect parts of the previous day's returns' effect on the next day's sentiment on WallStreetBets.

Another interesting finding in table 5.15 is that both Positive sentiment title and Positive sentiment title text are statistically significant at the 1% level in model (3), and that the adjusted R-squared of this model does not stand much back from the adjusted R-squared in model (2). This indicates that sentiment on WallStreetBets one day may have an effect on the next day's trading volume. Both of these statistically significant sentiment variables have positive coefficients, indicating that positive sentiment on WallStreetBets a given day is associated with an increase in the trading volume the next day. In other words, positive sentiment for stocks on WallStreetBets appears to induce people to trade more heavily the next day. Although the adjusted R-squareds in model (4), (5) and (6) fall from model (2) and (3), they also all contain statistically significant sentiment variables, meaning that this positive effect of sentiment on volume seems to sustain over the next few days. This is interesting since it suggests that the sentiment itself on WallStreetBets has had an impact on trading volume, not the frequency of submissions and comments posted on the forum. While the finding is not surprising, it carries important insights for the comprehensive understanding of the effects of sentiment on WallStreetBets.

Table 5.16: 2020 aggregated regression

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0003	16,818.060	0.008	190.934	95.618	0.000
	t = −0.464	t = 3.704***	t = 4.568***	t = 4.849***	t = 4.155***	t = 0.666
Positive sentiment comments	0.023	5,694.574	−0.101	391.644	−279.144	0.000
	t = 5.677***	t = 0.397	t = −4.961***	t = 4.544***	t = −4.511***	t = 2.653***
Title per day	0.096	−48,203.530	−0.262	−4,002.893	−1,175.271	−0.000
	t = 1.784*	t = −0.148	t = −2.583***	t = −1.472	t = −0.561	t = −0.594
Positive sentiment title	0.021	129,747.900	0.013	541.158	−429.143	0.000
	t = 4.607***	t = 5.716***	t = 0.695	t = 2.038**	t = −2.474**	t = 1.751*
Title text per day	−0.109	−268,553.600	0.183	−1,853.698	−2,424.835	0.000
	t = −2.005**	t = −0.662	t = 1.411	t = −0.497	t = −0.759	t = 0.507
Positive sentiment title text	0.010	46,617.610	−0.015	−90.909	−222.264	0.000
	t = 2.421**	t = 2.782***	t = −0.884	t = −0.542	t = −1.921*	t = 1.705*
Mkt.RF	1.188	−300,124.000	−0.817	486.607	−4,239.483	0.00000
	t = 20.967***	t = −1.816*	t = −2.355**	t = 0.306	t = −2.615***	t = 9.729***
SMB	0.679					
	t = 3.923***					
HML	−0.430					
	t = −2.733***					
RMW	−0.431					
	t = −1.775*					
CMA	−0.809					
	t = −3.286***					
MOM	−0.291					
	t = −2.232**					
Constant	−0.783	−47,611,052.000	118.680	−1,661,150.000	−987,315.300	−0.00000
	t = −1.313	t = −13.586***	t = 29.845***	t = −27.495***	t = −24.781***	t = −2.455**
TickerAMD	−0.434	−5,045,977.000	23.223	−1,470,386.000	−887,136.500	−0.000
	t = −1.362	t = −1.465	t = 20.120***	t = −24.276***	t = −24.115***	t = −0.221
TickerAPHA	−0.560	−51,398,204.000	63.888	−1,672,148.000	−992,410.800	−0.000
	t = −1.196	t = −15.073***	t = 37.717***	t = −27.559***	t = −24.957***	t = −0.872
TickerBB	−0.519	−50,666,425.000	35.954	−1,672,198.000	−995,936.900	−0.000
	t = −1.216	t = −14.179***	t = 20.857***	t = −27.594***	t = −25.077***	t = −0.073
TickerGME	−0.233	−60,422,798.000	89.187	−1,725,005.000	−1,011,096.000	−0.00000
	t = −0.433	t = −16.260***	t = 36.375***	t = −27.037***	t = −25.716***	t = −1.450
TickerGOOG	0.792	−49,736,834.000	1.530	−1,637,462.000	−1,009,855.000	0.000
	t = 1.908*	t = −14.165***	t = 0.915	t = −28.617***	t = −25.791***	t = 1.878*
TickerNIO	0.239	44,310,313.000	86.276	−1,457,907.000	−896,816.900	0.000
	t = 0.449	t = 7.096***	t = 53.096***	t = −23.541***	t = −22.901***	t = 0.128
TickerNKLA	−0.701	−38,699,313.000	113.108	−1,630,863.000	−958,855.100	0.000
	t = −0.721	t = −10.635***	t = 34.591***	t = −27.281***	t = −25.250***	t = 1.080
TickerNOK	−0.787	−26,949,332.000	23.293	−1,651,552.000	−986,943.300	−0.000
	t = −1.907*	t = −7.519***	t = 15.253***	t = −27.134***	t = −24.719***	t = −0.194
TickerNVDA	−0.415	−49,246,494.000	19.063	−1,594,612.000	−937,487.600	0.000
	t = −1.092	t = −14.740***	t = 14.061***	t = −26.443***	t = −24.050***	t = 0.035
TickerPLTR	−7.308	−34,597,724.000	59.870	−1,730,088.000	−1,019,681.000	−0.000
	t = −1.009	t = −4.140***	t = 22.127***	t = −23.883***	t = −23.307***	t = −1.451
TickerPLUG	0.427	−35,426,112.000	62.744	−1,638,875.000	−982,275.000	0.000
	t = 0.891	t = −10.120***	t = 37.908***	t = −27.001***	t = −24.683***	t = 1.282
TickerRKT	−1.992	−51,496,371.000	51.719	−1,667,294.000	−988,810.700	−0.000
	t = −2.070**	t = −15.192***	t = 24.424***	t = −26.645***	t = −25.169***	t = −1.786*
TickerSPCE	−0.550	−47,471,617.000	83.344	−1,635,102.000	−970,491.900	−0.000
	t = −1.067	t = −14.408***	t = 38.848***	t = −26.967***	t = −25.494***	t = −0.707
TickerTLRY	−0.728	−47,159,978.000	89.311	−1,616,005.000	−981,556.500	−0.00000
	t = −1.135	t = −13.496***	t = 37.942***	t = −26.650***	t = −24.589***	t = −3.393***
TickerTSLA	0.641	−61,316,290.000	34.636	−699,596.000	−185,428.700	0.000
	t = 1.594	t = −13.972***	t = 22.245***	t = −8.413***	t = −3.807***	t = 0.026
TickerTSM	−0.448	−48,786,644.000	6.964	−1,658,230.000	−985,114.700	0.000
	t = −1.104	t = −14.232***	t = 5.012***	t = −27.266***	t = −24.670***	t = 0.203
Constant	−3.050	48,134,289.000	37.859	1,628,865.000	1,049,212.000	−0.00000
	t = −5.853***	t = 12.749***	t = 17.076***	t = 28.305***	t = 26.240***	t = −3.751***
Observations	4,092	4,092	4,088	4,088	4,088	4,092
R^2	0.114	0.555	0.700	0.731	0.733	0.125
Adjusted R^2	0.107	0.552	0.698	0.729	0.731	0.120
Note:						*p<0.1; **p<0.05; ***p<0.01

Model (1) in table 5.16 shows that four of the sentiment variables are statistically significant at the 1% and 5% level in 2020, while also the return on the benchmark portfolio and all the Fama-French factors except RMW are statistically significant at the 1% and 5% level. The adjusted R-squared of this model is relatively low, however,

with only 10.7%. The net direction of the coefficients on the statistically significant sentiment variables in model (1) is ambiguous, as the coefficients on Positive sentiment comments, Positive sentiment title and Positive sentiment title text are all positive, while the coefficient on Title text per day is negative. In model (2) we observe that Comments per day, Positive sentiment title and Positive sentiment title text are significant at the 1% level, and that the direction of the coefficients are all positive. Although this indicates that both the number of comments and the sentiment in the submission texts and submissions titles have had a positive effect on trading volume, the magnitude of the coefficient on Comments per day seems unrealistically high as it suggests that an increase of one comment on the forum considering a stock in the sample increases the trading volume with over 16 800 shares. The categorical variables may control for parts of this effect, but we would argue that this estimated coefficient seems to be too high. Also implied volatility, call option volume and put option volume are all indicated to have been affected by sentiment on WallStreetBets in 2020. The adjusted R-squareds of these models are generally much higher than the adjusted R-squared in the return model (1), indicating that these findings can potentially be more robust than the associated impact of sentiment on stocks returns in 2020.

Table 5.17: 2021 aggregated regression

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
	0.00001	−1,628.212	0.0001	−1.476	0.570	0.000
Comments per day	t = 0.028	t = −1.607	t = 0.232	t = −0.785	t = 0.681	t = 0.878
	0.090	46,338.050	−0.022	428.908	−477.928	0.000
Positive sentiment comments	t = 3.923***	t = 0.405	t = −0.459	t = 1.114	t = −2.859***	t = 2.812***
	−0.010	225,531.200	0.087	375.000	156.536	−0.000
Title per day	t = −0.516	t = 1.749*	t = 2.551**	t = 1.368	t = 2.609***	t = −1.130
	0.024	192,952.700	0.038	569.362	42.602	0.000
Positive sentiment title	t = 2.312**	t = 3.519***	t = 1.349	t = 2.427**	t = 0.360	t = 0.673
	0.031	−187,064.900	−0.059	−476.967	−145.664	0.000
Title text per day	t = 0.646	t = −0.760	t = −0.765	t = −1.050	t = −0.735	t = 0.523
	0.015	129,174.000	0.074	650.026	158.484	0.000
Positive sentiment title text	t = 1.208	t = 2.256**	t = 2.357**	t = 3.055***	t = 1.506	t = 1.588
	−0.006	−7,331,162.000	−2.117	−19,730.550	−13,006.010	0.000
Mkt.RF	t = −0.010	t = −2.005**	t = −1.612	t = −2.125**	t = −3.210***	t = 5.653***
	0.619					
SMB	t = 0.951					
	−1.083					
HML	t = −1.297					
	−1.243					
RMW	t = −1.368					
	6.722					
CMA	t = 3.172***					
	1.003					
MOM	t = 1.362					
	4.352	32,379,266.000	160.142	−821,523.100	−323,112.200	0.000
Constant	t = 1.013	t = 1.254	t = 20.325***	t = −10.197***	t = −10.370***	t = 0.560
	−0.164	−67,487,109.000	12.823	−848,613.600	−327,895.800	−0.000
TickerAMD	t = −0.147	t = −13.257***	t = 10.708***	t = −14.504***	t = −12.544***	t = −0.262
	2.660	−83,187,746.000	94.983	−1,095,474.000	−455,591.800	0.000
TickerAPHA	t = 1.364	t = −13.662***	t = 18.442***	t = −19.199***	t = −18.034***	t = 2.614***
	0.664	−77,682,377.000	84.306	−1,010,787.000	−428,812.000	0.000
TickerBB	t = 0.284	t = −8.237***	t = 12.293***	t = −16.689***	t = −16.738***	t = 0.092
	5.885	−90,141,150.000	162.585	−1,000,037.000	−302,889.800	0.000
TickerGME	t = 0.747	t = −3.322***	t = 10.421***	t = −14.560***	t = −6.699***	t = 0.035
	2.957	−94,087,350.000	−1.826	−1,087,164.000	−452,579.400	0.000
TickerGOOG	t = 1.987**	t = −13.469***	t = −0.777	t = −17.779***	t = −15.550***	t = 1.199
	0.711	8,039,667.000	61.753	−615,752.300	−251,876.700	0.000
TickerNIO	t = 0.548	t = 0.771	t = 40.426***	t = −9.350***	t = −8.625***	t = 1.034
	1.705	−92,476,662.000	79.440	−1,089,845.000	−445,752.800	0.000
TickerNKLA	t = 1.356	t = −16.839***	t = 36.059***	t = −18.844***	t = −16.909***	t = 0.022
	0.177	−23,064,120.000	27.796	−884,686.200	−434,755.100	0.000
TickerNOK	t = 0.100	t = −1.157	t = 8.835***	t = −10.505***	t = −16.773***	t = 0.398
	1.330	−99,170,173.000	9.783	−994,481.600	−394,098.700	0.000
TickerNVDA	t = 1.057	t = −18.509***	t = 8.131***	t = −16.988***	t = −15.028***	t = 1.795*
	0.290	−23,894,164.000	70.004	−655,940.300	−258,344.700	−0.000
TickerPLTR	t = 0.262	t = −2.074**	t = 26.384***	t = −8.939***	t = −7.749***	t = −0.954
	1.111	−66,958,418.000	72.196	−1,038,962.000	−408,226.700	−0.000
TickerPLUG	t = 0.718	t = −10.040***	t = 30.079***	t = −18.028***	t = −15.637***	t = −0.350
	0.623	−86,249,418.000	48.190	−1,019,360.000	−423,424.700	−0.000
TickerRKT	t = 0.304	t = −8.440***	t = 14.199***	t = −16.331***	t = −15.722***	t = −1.263
	1.535	−89,170,911.000	88.434	−1,039,498.000	−428,591.200	0.000
TickerSPCE	t = 1.081	t = −15.454***	t = 18.799***	t = −17.795***	t = −16.612***	t = 0.477
	2.887	−65,048,688.000	128.027	−977,555.500	−405,498.800	0.000
TickerTLRY	t = 1.088	t = −7.957***	t = 20.038***	t = −16.571***	t = −15.004***	t = 2.069**
	0.506	−74,168,672.000	33.850	−478,077.000	33,813.320	0.000
TickerTSLA	t = 0.408	t = −12.753***	t = 22.363***	t = −6.627***	t = 0.795	t = 1.113
	0.648	−92,589,794.000	7.143	−1,069,025.000	−433,692.000	0.000
TickerTSM	t = 0.521	t = −16.947***	t = 5.156***	t = −18.512***	t = −16.690***	t = 0.435
	−9.330	85,339,313.000	29.395	1,037,598.000	486,979.800	−0.000
Constant	t = −4.570***	t = 7.142***	t = 6.737***	t = 14.670***	t = 14.714***	t = −4.224***
Observations	792	792	792	792	792	792
R^2	0.172	0.519	0.844	0.660	0.717	0.106
Adjusted R^2	0.141	0.504	0.839	0.649	0.708	0.079

Note: *p<0.1; **p<0.05; ***p<0.01

The regression output in table 5.17 is based on data from 2021 alone. In model (1) we observe that Positive sentiment comments and Positive sentiment title are statistically significant at the 1% and 5% level, respectively. Compared to the adjusted R-squared in model (1) in table 5.16 where data from 2020 was used, we observe that the adjusted

R-squared has increased in this model when using data from 2021. We also observe that neither the return on the benchmark portfolio nor any of the Fama-French factors are statistically significant except the CMA factor, which is in sharp contrast to the findings when using data from only 2020. The direction of the coefficients in model (1) in table 5.17 indicates that sentiment has had a positive effect on return in 2021, while the net effect of sentiment on stock returns in 2020 was unclear given the negative coefficient on Title text per day in model (1) in table 5.16. Overall, therefore, sentiment is suggested to explain more of the variation in stock returns in 2021 compared to 2020, and the direction of the effect of sentiment in 2021 is not as ambiguous as it was in 2020 when the net direction of the coefficients on the sentiment variables was unclear.

In model (2) in table 5.17 we observe that Positive sentiment title and Positive sentiment title text are statistically significant at the 1% and 5% level, respectively. Mkt.RF is also statistically significant at the 5% level. While Comments per day was statistically significant in model (2) when using data from 2020 alone, this is no longer significant when using data from only 2021. This is an interesting finding, as it suggests that each comment in 2020 was indicated to potentially have had an impact on trading volume. Note, however, that the magnitude of the coefficient in 2020 seems unreasonably high, but this finding generally points towards that the heightened level of posts on WallStreetBets in 2021 compared to 2020 overall made each comment relatively less important, while the overall sentiment in the submissions became relatively more important factors for explaining variations in trading volume. Also note that the adjusted R-squared falls in model (2) in 2021 compared to 2020, but the magnitude of the drop is small.

Title per day and Positive sentiment title text are the only statistically significant variables at the 5% level in model (3) using data for 2021. The adjusted R-squared of this model is 83.9%. This indicates that sentiment can be argued to explain much of the variation in implied volatility for the stock sample in 2021, and the adjusted R-squared is also relatively much higher in 2021 compared to 2020. The adjusted R-squareds in model (4) and (5) fall from 2020 to 2021, but there are several statistically significant sentiment variables in the models both from 2020 and 2021. The impact of sentiment on call and

put option volume therefore seems to apply to the stock sample also on an aggregated basis, which is a particularly interesting finding as it seems more robust than the findings on stock returns. Option strategies are commonly discussed on WallStreetBets, and users are frequently advocating to trade in options instead of stocks directly so as to maximize potential gains per dollar invested. This common practice on the forum is indicated to have had a general impact on option volumes in both 2020 and 2021, and is one the most striking findings in these regressions.

5.3 Minute-by-minute regressions

AMC, BB, GME and NOK were all subject to massive discussion on WallStreetBets and media coverage during the rallying weeks between January and February 2021. In this subchapter we will analyze potential effects of sentiment on WallStreetBets on these stocks' return and trading volume on a minute-by-minute basis between January and February 2021. Using comments data on WallStreetBets on a minute-by-minute basis and combining this with return and volume data facilitates an unique opportunity to analyze if activity and sentiment had immediate effects on the mentioned stocks' return and trading volume.

For the regressions in this chapter, Comments per minute is the number of comments posted on WallStreetBets concerning the specific stock a given minute.

5.3.1 Individual regressions

Table 5.18: Minute regressions - Individual stocks

	Dependent variable:							
	AMC Return	AMC Volume	BB Return	BB Volume	GME Return	GME Volume	NOK Return	NOK Volume
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comments per minute	−0.001	11,284.960	−0.001	7,297.967	0.001	170.433	−0.0002	24,804.190
	t = −1.280	t = 20.605***	t = −0.538	t = 19.576***	t = 1.975**	t = 7.546***	t = −0.154	t = 14.059***
Positive sentiment comments	0.003	453.672	0.001	98.690	0.003	−1,437.322	−0.0001	1,926.911
	t = 1.809*	t = 0.461	t = 1.236	t = 0.558	t = 0.933	t = −3.572***	t = −0.315	t = 2.805***
Constant	−0.162	465,292.300	−0.047	140,387.100	−0.379	195,834.300	0.001	346,387.500
	t = −1.112	t = 6.349***	t = −1.157	t = 9.779***	t = −1.580	t = 7.453***	t = 0.037	t = 6.050***
Observations	2,466	2,487	3,611	3,616	3,495	3,531	2,393	2,397
R^2	0.001	0.199	0.0005	0.243	0.004	0.029	0.0001	0.233
Adjusted R^2	0.00004	0.199	−0.0001	0.242	0.003	0.028	−0.001	0.232

Note: *p<0.1; **p<0.05; ***p<0.01

In table 5.18 we assess the effect of real-time WallStreetBets sentiment on a minute-by-minute basis from January 19, 2021 to February 3, 2021. Unfortunately due to missing data, we cannot assess the days leading up to the peak on January 25 and 26 which would be extremely interesting to analyze, although we have the data surrounding these dates. When looking at the return for the stocks, we can see a remarkably low adjusted R-squared on all of the models, meaning sentiment has little explanatory power when it comes to return. Only for the GME return do we see a statistically significant relationship with the Comments per minute variable above the 5% level, indicating that comment frequency might have a small effect on returns.

Further, from the Comments per minute variable we observe a statistically significant relationship on all of the stocks' volume, with a strong correlation. These results indicates that forum participants discussing the stock might also buy the stock at the same time. Interestingly, GME compared to the others has a lower t-value on the variable than the other stocks, although still very significant, but also a significantly lower adjusted R-squared of only 3% compared to the other stocks hovering around 20%. This might indicate that GME was driven by other external forces rather than WallStreetBets exclusively, or it could be due to the fact that GME was mentioned in general more than the other stocks, as well as having a higher price inducing a higher barrier of entry, thus many forum participants holding the stock could be talking about the stock without actually buying more. Alternatively, since GME experienced one of the largest short squeeze rallies in history, we cannot rule out that other market participants also took part in the rallies.

When looking at how positive sentiment affected volume, we observe in model (6) and (8) that the variable is statistically significant at the 1% level for both GME and NOK, but that their directions differ. For GME, positive sentiment seems to reduce stock volume, whilst for NOK positive sentiment has a positive relationship with stock volume. Although it is hard to infer why GME has a negative relationship between positive sentiment and volume, this might simply be a result of how some words in the dictionary such as "Diamond hands", which was frequently used during the surges, have a positive

dictionary score, but was mostly used when the stock was experiencing negative returns to signal that even though the stock is dropping, the forum members will not sell the stock.

5.3.2 Aggregated regressions

Table 5.19: Minute regressions - Aggregated sample

	Dependent variable:	
	Return	Volume
	(1)	(2)
Comments per minute	0.001	781.083
	t = 1.899*	t = 18.240***
Positive sentiment comments	0.001	2,856.764
	t = 1.593	t = 8.212***
Constant	0.011	−799,889.900
	t = 0.253	t = −36.453***
TickerGME	−0.089	−1,081,100.000
	t = −1.564	t = −45.499***
TickerNOK	0.019	−294,819.100
	t = 0.410	t = −8.761***
Constant	−0.080	836,562.900
	t = −1.424	t = 24.802***
Observations	11,965	12,031
R^2	0.001	0.205
Adjusted R^2	0.001	0.205
Note:	*p<0.1; **p<0.05; ***p<0.01	

When aggregating the data in a factor regression, we observe many of the same trends observed in table 5.18. For returns, both sentiment variables seem to have a slightly positive relationship with returns, yet they are not statistically significant enough to infer any real relationship between the variables, which ultimately would be supported by the negligible adjusted R-squared of 0.1%. On the other hand, all our sentiment variables as well as the ticker factors are statistically significant at the 1% level for volume, with strong t-values indicating that the correlation is strong. Compared to GME on an individual basis, positive sentiment overall seems to have a positive relationship with

stock volume. With an adjusted R-squared of about 20.5%, we can say that a decent proportion of the volume variation can be explained by WallStreetBets sentiment.

5.3.3 Interpretation

Ultimately, there does not seem to be a clear relationship between WallStreetBets sentiment and returns on a minute-by-minute basis. However, volume seems to be affected to a larger degree with an adjusted R-squared on the aggregated sample of 20.5%, as well as statistically significant sentiment variables. There is no clear confirmation from this data alone on whether WallStreetBets were the catalyst for the rallies or if the forum simply served as an echo chamber intensifying the discussions around the stocks in question. Ultimately, from the data we have analyzed we cannot from a statistical perspective conclude that WallStreetBets has not had an impact on the rallies we observed in January 2021.

Although we have not found a statistical relationship for returns on a minute-by-minute basis, there are some indications of a causal relationship between WallStreetbets sentiment and indirect stock price increases for the selected stocks in the sample. Since many of the stocks subject to the rallies had large short interests, a volume increase in these stocks with a positive direction meant that many investors who were short had to exit their positions by buying shares, thus fueling these rallies even more. With the findings from the individual regressions, indicating that sentiment on WallStreetBets also increases call option volume (see table A1.10, table A1.25, table 5.8 and table A1.28), one could infer that this has also lead to an increase in the price when stocks have surged and market makers have had to gamma hedge, thus more stocks have been bought which in combination with shorters being squeezed out of their positions have lead to the surges we have observed.

5.4 Mention-weighted portfolios

In order to better assess whether stock activity on WallStreetBets could be translated to potential returns we constructed mention-weighted portfolios consisting of all the stocks in our sample. This way we could visualize how stock activity on the forum translated into stock movements, and see whether following WallStreetBets investment advice could yield better or worse returns than the market portfolio.

Figure 5.6 graphs the value of investing USD 100 from January 02, 2020, to March 12, 2021, by using the different trading strategies we developed in chapter 4.6.

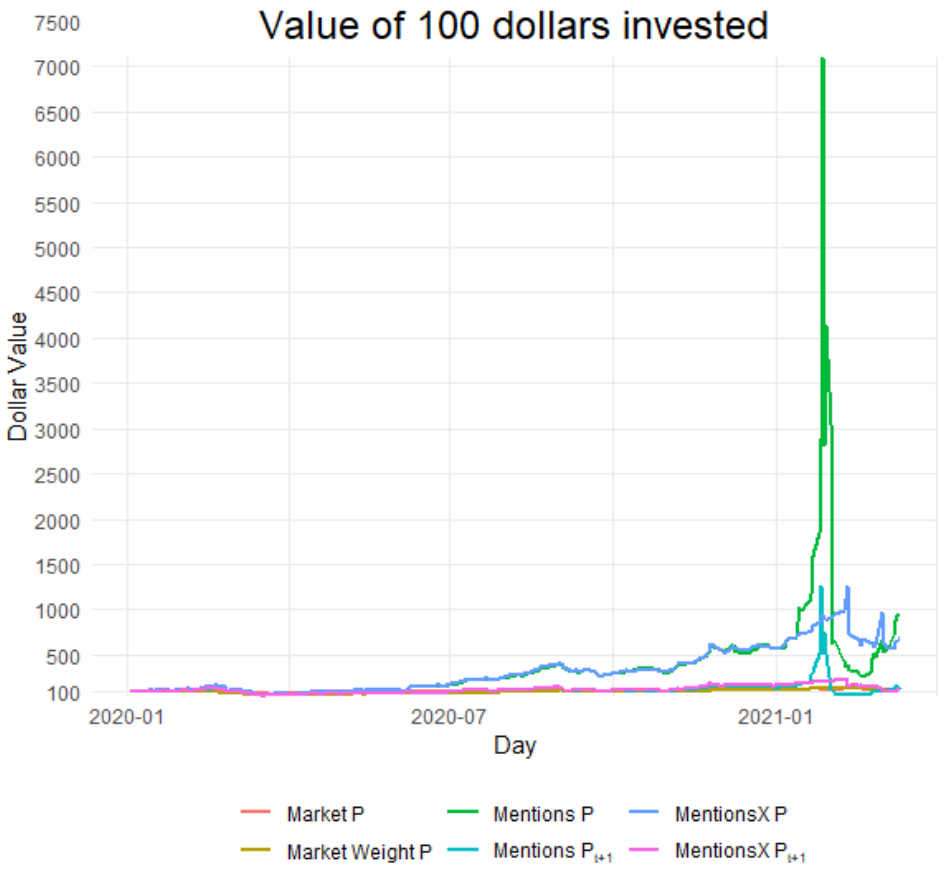

Figure 5.6: Portfolio simulation - January 02, 2020, to March 12, 2021

Figure 5.6 shows that Mentions P gives the highest overall return during the time interval. Although the strategy is somewhat theoretical, it gives an indication of what returns

could be achieved by monitoring the sentiment on WallStreetBets a given day and implementing a mention-weighted portfolio the same day. We would like to underscore that the calculated return on this portfolio should be used as an indication, not as a definite result given that the strategy has its drawbacks. The main drawback is that sentiment and returns can change over the course of the given day, but both returns and sentiment are based on the returns and accumulated sentiment over the whole day. However, we would nevertheless argue that the results still may have some indicative merit. The figure indicates that by investing USD 100 on January 02, 2020 using this strategy, on March 12, 2021 this would have become USD 944.34, implying a return of 844.34%. We also see that MentionsX P yields the second highest return, a remarkable return of 573.46%, which is particularly interesting given that we have taken out the effects AMC, BB, GME and NOK.

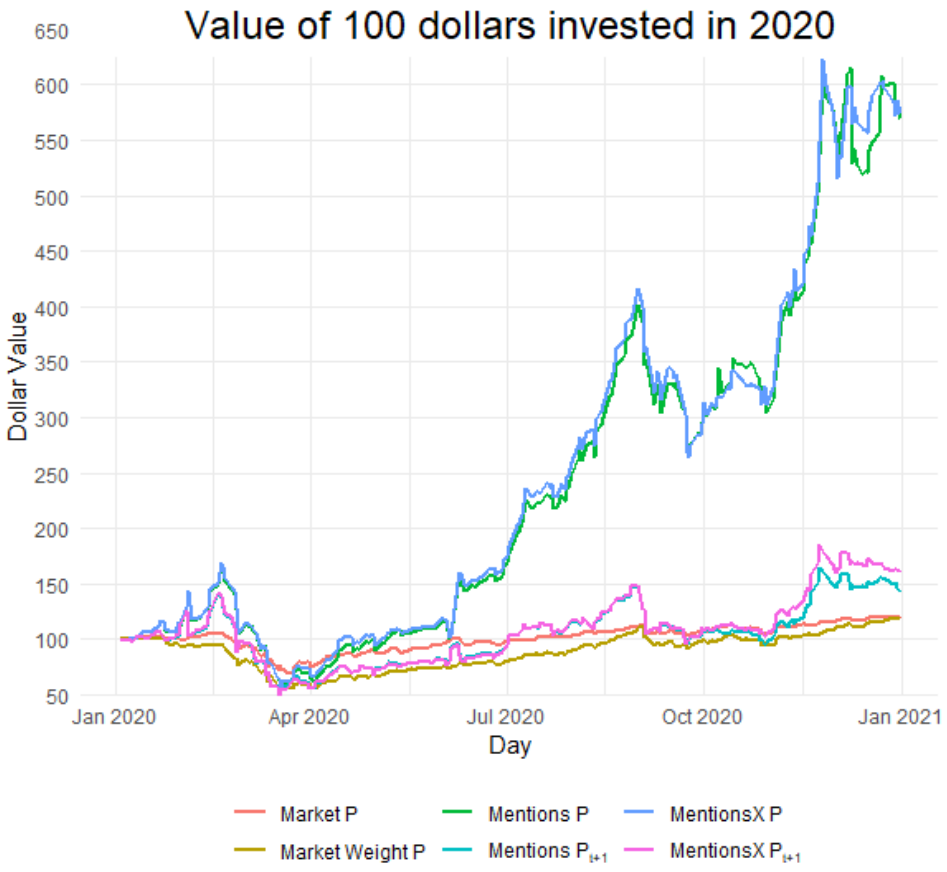

Figure 5.7: Portfolio simulation 2020

Figure 5.7 shows what an investment of USD 100 on January 02, 2020 would become on December 31, 2020. As the figure shows, Mentions P and MentionsX P both yield astonishing returns compared to the other strategies. Both Mentions P_{t+1} and MentionsX P_{t+1} also yield higher returns than the market portfolio and the value-weighted stock portfolio. This indicates that investing based on today's relative mention-weights yields the highest return, but also that investing tomorrow based on today's relative mention-weights would yield remarkably higher returns than by investing in the market portfolio.

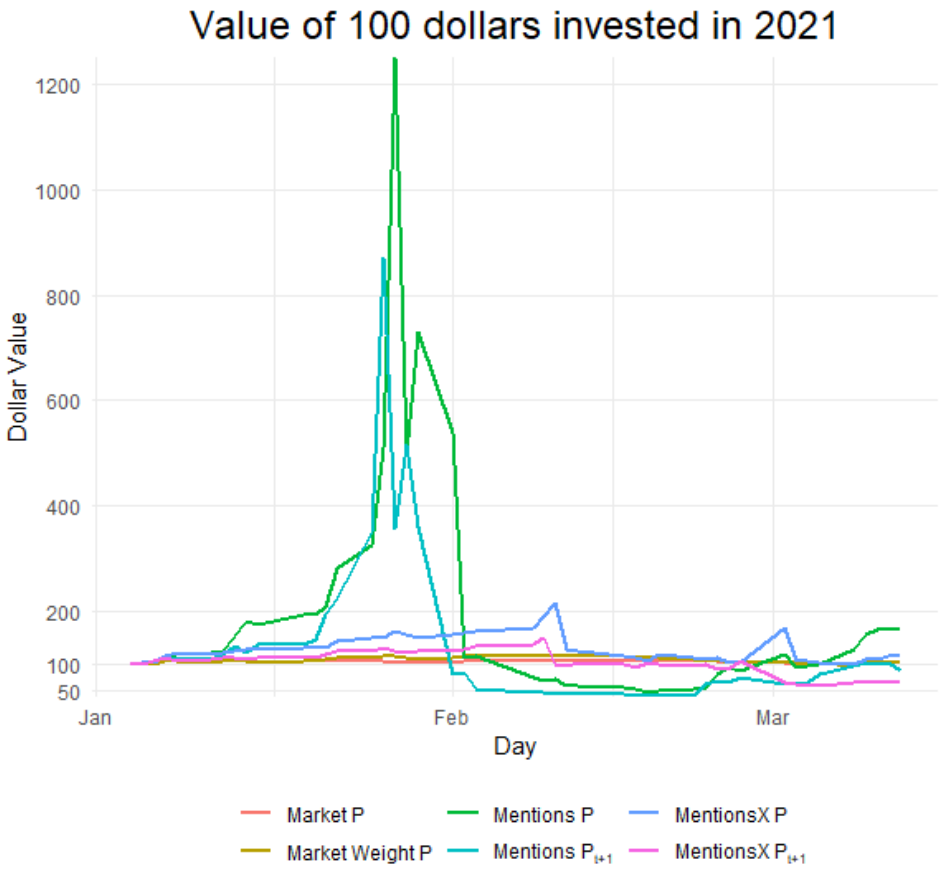

Figure 5.8: Portfolio simulation 2021

Figure 5.8 displays what an investment of USD 100 on January 04, 2021, would become on March 12, 2021. The Mentions P strategy yields the highest return, followed by MentionsX P. Both Mentions P_{t+1} and MentionsX P_{t+1} underperforms relative to the market portfolio and the value-weighted stock portfolio. This indicates that in 2021 alone,

the strategies based on investing based on today's mention-weights yield the highest returns.

The following tables show the value of a USD 100 investment and return of the strategies over the different courses of time. Table 5.20 shows the value of the investment and the return of the different strategies over the whole time period, i.e. from January 03, 2020, to March 12, 2021. Table 5.21 focuses only on 2020 and shows the value of a USD 100 investment and return of the different strategies from January 3, 2020, to December 31, 2020. Table 5.22 is based on data from 2021 alone and shows the value of a USD 100 investment and return of the different strategies from January 04, 2021, to March 12, 2021.

Table 5.20: Portfolio simulation - value of USD 100 invested

Date	Market P	MVW P	Mentions P	Mentions P_{t+1}	MentionsX P	MentionsX P_{t+1}
1/3/2020	100.00	100.00	100.00	100.00	100.00	100.00
3/12/2021	122.56	121.00	944.34	123.94	673.46	112.20
Return	22.56%	21.00%	844.34%	23.94%	573.46%	12.20%

Table 5.21: 2020 portfolio simulation - value of USD 100 invested

Date	Market P	MVW P	Mentions P	Mentions P_{t+1}	MentionsX P	MentionsX P_{t+1}
1/3/2020	100.00	100.00	100.00	100.00	100.00	100.00
12/31/2020	120.79	118.73	570.05	142.7	573.8	163.69
Return	20.79%	18.73%	470.05%	42.70%	473.80%	63.69%

Table 5.22: 2021 portfolio simulation - value of USD 100 invested

Date	Market P	MVW P	Mentions P	Mentions P_{t+1}	MentionsX P	MentionsX P_{t+1}
1/4/2021	100.00	100.00	100.00	100.00	100.00	100.00
3/12/2021	102.91	102.88	166.40	85.20	114.86	67.00
Return	2.91%	2.88%	66.40%	-14.80%	14.86%	-33.00%

A general note to the results in this subchapter is that a large portion of the return in 2020 is explained by Tesla, a WallStreetBets favorite, which has experienced a higher percentage share of mentions than its market value weight. However, the tables highlight

that using mention-weighted trading strategies can yield potentially very high returns. Strategies Mentions P and MentionsX P yield total returns of 844.34% and 573.46% over the entire time interval respectively, but we once again underscore that these should be used as indications for a hypothesized return, not definite returns given the previously discussed drawbacks of these strategies in practice. The more realistic strategies to implement in practice, Mentions P_{t+1} and MentionsX P_{t+1}, yield lower returns over the entire time interval, but still higher returns than the market portfolio and the value-weighted stock portfolio in 2020. A final remark when looking at the results presented in this chapter is to note that most of the stocks in our sample have experienced some kind of a surge during the time period, and subsequently the stocks are not necessarily representative of a mention-weighted portfolio consisting of a broader set of stocks. Overall, mention-weighted strategies may have some merit, and they may potentially yield extremely high returns.

5.5 Event studies

In this subchapter, we will perform event studies on 11 of the 18 stocks in this thesis' list of stocks subject to analysis. The 7 remaining event studies are found in the appendix' chapter A3. These studies enable a more in-depth analysis of whether sentiment on WallStreetBets has had statistically significant impacts on individual stocks' return. The event studies seek to calculate and test for abnormal return of stocks in the period they are relatively most mentioned at WallStreetBets. The rationale behind this is to analyze whether an abnormal return is achieved on the same day, or up to two prior or succeeding days from this day, as the stock is relatively most mentioned. We do not seek to do an event study with a hold-out window, and an event window spanning multiple days, rather, we want to explore what movements occur on the days surrounding activity peaks.

In the following tables, the day at which the relative mention-weight is at its peak is the observation in the middle of the tables. The two upper observations are the two trading days before the main observation, while the two lower observations are the two subsequent trading days to the main observation.

Table 5.23: AMC event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/2/2021	-53.11%	71.73%	18.42%	-19.82%	-33.29%	-2.42	**Yes**
2/3/2021	13.72%	81.67%	16.82%	2.36%	11.36%	0.83	**No**
2/8/2021	-37.26%	61.59%	27.61%	12.04%	-49.29%	-3.58	**Yes**
2/9/2021	-11.66%	63.36%	5.80%	-2.61%	-9.05%	-0.66	**No**
2/10/2021	5.31%	66.42%	3.32%	-2.20%	7.51%	0.55	**No**

Table 5.24: APHA event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/9/2021	22.42%	73.30%	4.40%	1.96%	20.46%	3.53	**Yes**
2/10/2021	10.20%	69.22%	9.61%	-0.40%	10.60%	1.83	**No**
2/11/2021	-44.34%	62.54%	15.28%	0.54%	-44.88%	-7.75	**Yes**
2/12/2021	0.35%	58.08%	15.25%	1.25%	-0.90%	-0.15	**No**
2/16/2021	24.90%	67.47%	1.46%	0.13%	24.77%	4.27	**Yes**

Table 5.25: BB event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
1/13/2021	-2.52%	82.69%	0.42%	-1.77%	-0.75%	-0.125	**No**
1/14/2021	20.25%	87.33%	5.48%	4.21%	16.04%	2.649	**Yes**
1/15/2021	7.71%	82.78%	24.47%	-3.78%	11.49%	1.898	**No**
1/19/2021	22.72%	82.54%	9.67%	3.40%	19.32%	3.191	**Yes**
1/20/2021	3.50%	81.72%	15.57%	-0.38%	3.88%	0.641	**No**

Table 5.26: GME event study

Date	Return%	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
3/8/2021	34.51%	67.90%	79.96%	9.58%	24.93%	1.76	**No**
3/9/2021	23.86%	66.48%	88.46%	2.59%	21.27%	1.51	**No**
3/10/2021	7.07%	61.34%	93.28%	4.16%	2.92%	0.21	**No**
3/11/2021	-1.90%	61.84%	89.36%	-2.04%	0.14%	0.01	**No**
3/12/2021	1.72%	65.45%	91.04%	8.14%	-6.42%	-0.45	**No**

Table 5.27: NOK event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
1/21/2021	0.48%	50.00%	0.70%	1.29%	-0.82%	-0.21	**No**
1/22/2021	-0.48%	0.00%	0.20%	2.81%	-3.29%	-0.83	**No**
1/27/2021	44.44%	37.28%	10.73%	4.66%	39.77%	10.02	**Yes**
1/28/2021	-33.40%	100.00%	2.67%	-7.43%	-25.97%	-6.55	**Yes**
1/29/2021	-2.81%	100.00%	1.40%	-1.27%	-1.54%	-0.39	**No**

Table 5.28: PLTR event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
11/24/2020	12.41%	72.11%	44.40%	-16.68%	29.09%	0.73	**No**
11/25/2020	19.85%	69.87%	65.25%	-0.24%	20.09%	0.51	**No**
11/27/2020	-4.90%	66.78%	78.96%	2.55%	-7.46%	-0.19	**No**
11/30/2020	-2.01%	62.94%	39.75%	-9.85%	7.84%	0.20	**No**
12/1/2020	-5.46%	66.38%	31.74%	-12.17%	6.71%	0.17	**No**

Table 5.29: PLUG event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/14/2020	-1.50%	100.00%	0.12%	0.46%	-1.96%	-0.37	**No**
2/18/2020	8.68%	73.91%	5.38%	1.08%	7.60%	1.44	**No**
2/19/2020	15.90%	76.55%	14.49%	1.42%	14.48%	2.74	**Yes**
2/20/2020	-4.34%	72.50%	0.89%	0.21%	-4.55%	-0.86	**No**
2/21/2020	0.48%	75.00%	0.18%	-0.77%	1.25%	0.24	**No**

Table 5.30: RKT event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/26/2021	9.33%	65.86%	1.26%	-1.16%	10.49%	1.05	**No**
3/2/2021	64.41%	71.80%	41.28%	-1.88%	66.28%	6.61	**Yes**
3/3/2021	-39.56%	67.52%	53.97%	-2.94%	-36.62%	-3.65	**Yes**
3/4/2021	-4.20%	61.18%	11.78%	-2.85%	-1.34%	-0.13	**No**
3/5/2021	-6.77%	65.26%	5.52%	0.97%	-7.74%	-0.77	**No**

Table 5.31: SPCE event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/12/2020	4.47%	74.23%	47.10%	1.23%	3.24%	0.51	**No**
2/13/2020	1.88%	70.47%	41.93%	-0.15%	2.04%	0.32	**No**
2/14/2020	18.64%	67.26%	58.07%	-0.30%	18.94%	2.97	**Yes**
2/18/2020	4.89%	63.65%	56.33%	0.04%	4.85%	0.76	**No**
2/19/2020	20.32%	66.23%	39.62%	0.83%	19.48%	3.06	**Yes**

Table 5.32: TLRY event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/9/2021	34.18%	71.30%	3.71%	4.09%	30.09%	3.12	**Yes**
2/10/2021	41.15%	67.59%	5.42%	-1.17%	42.32%	4.40	**Yes**
2/11/2021	-68.68%	62.62%	12.50%	0.24%	-68.92%	-7.16	**Yes**
2/12/2021	-10.34%	63.09%	6.13%	1.73%	-12.07%	-1.25	**No**
2/16/2021	17.74%	76.92%	1.21%	0.50%	17.24%	1.79	**No**

Table 5.33: TSLA event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
2/3/2020	17.55%	61.57%	75.97%	3.47%	14.07%	3.14	**Yes**
2/4/2020	12.26%	60.92%	87.72%	2.63%	9.63%	2.15	**Yes**
2/5/2020	-19.44%	57.15%	89.07%	-0.59%	-18.85%	-4.21	**Yes**
2/6/2020	1.32%	55.91%	73.01%	1.32%	0.00%	0.00	**No**
2/7/2020	-0.72%	59.50%	72.80%	-0.08%	-0.64%	-0.14	**No**

As can be observed in the tables above, 8 out of the 11 stocks display statistically significant abnormal return on the day they peak in terms of relative mention-weight. 4 out of 11 stocks display statistically significant abnormal return two days prior to their respective relative mention-peak, while also 4 out of 11 stocks display statistically significant abnormal return one day prior to their respective relative mention-peak. Only 2 out of 11 stocks display statistically significant abnormal return the subsequent day to their respective relative mention-peak, and 2 out of 11 stocks display statistically significant abnormal return two days after their respective relative mention-peak. These results indicate that on the respective days the stocks subject for analysis in this subchapter reached their relative mention-peaks, most of them display a statistically significant abnormal return. 4 out of the 8 stocks with statistically significant abnormal returns on the same day as their respective mention-peaks have a negative abnormal return, while the other 4 stocks have a positive abnormal return. These results might suggest that activity is not driven entirely by positive market news, but mentions can also be high when stocks are experiencing negative returns. A characteristic of the forum, originating from their high risk strategies, has long been to celebrate large losses as well as large gains, so these results might serve to emphasize this distinction.

In sum, the event studies indicate that on days where the stocks were relatively much discussed on WallStreetBets, most stocks displayed abnormal returns the same day. Note, however, that we have only analyzed the exact date +/- two days at which the relative mention-weight for each stock reached its peak in the time interval we explore in this thesis. Also, abnormal return may itself be a reason for a stock being heavily discussed on WallStreetBets, not vice versa. As underlined with our regressions, positive sentiment,

rather than pure activity, seems to be the main correlation between returns and forum influence which could be a result of reverse causality. The results for the two days prior to or succeeding the date at which the stocks reached their relative mention-peaks therefore could give a more valuable insight into how sentiment is generated on the forum. All of the 4 stocks displaying statistically significant returns the day prior to the relative mention-peak have positive abnormal returns. While not statistically proven, this indicates in isolation that a stock's prior return on the previous day may be reflected in the relative mention-weight on WallStreetBets the next day.

6 Conclusion

6.1 Conclusion

In this thesis, we have tested whether sentiment on the WallStreetBets forum has exercised some form of market power during the research period. By applying in-depth sentiment analysis techniques combined with financial data, we find ambiguous results on whether or not WallStreetBets have had an impact on stock returns in the time interval we have limited this analysis to. For some individual stocks we find that WallStreetBets sentiment has had a statistically significant impact on returns, and that returns can also be explained for succeeding days. On the contrary, for other stocks we observe only statistically robust relationships between positive sentiment and returns the same day, and for the same stocks we can observe a significant drop off in adjusted R-squared when we lag the returns. These findings may suggest that for many of the stocks in our sample we cannot rule out the effect of reverse causality, but for other stocks in our sample returns can be partially explained by sentiment on the forum. On an aggregated basis only the sentiment variables connected to positive sentiment are statistically significant, which again suggests a problem of reverse causality, alongside a low adjusted R-squared. We therefore cannot rule out that a large proportion of the variation in returns is explained by our control variables, and that sentiment has had a small impact on returns today. However, when supplementing the model with lagged variables we find that sentiment today better explains returns tomorrow even when accounting for non-lagged control variables. These findings might suggest there is some merit to the notion that WallStreetBets actually has the ability to influence stock returns.

While sentiment on WallStreetBets does not have a clear relationship with return, the correlation is stronger when looking at volume. For many individual stocks we see that when activity and positive sentiment is high, volume has a positive relationship with the sentiment variables. Only 4 out of the 18 stocks we analyzed in this thesis did not have any statically significant relationships between sentiment and volume. Overall, the results

differ on an individual basis, where for some stocks, stock volume can be explained to a larger degree than for others. On an aggregated basis can WallStreetBets sentiment be claimed to explain a moderate portion of volume variations, with the degree of positive sentiment being the most correlated explanatory variable with volume. Overall, we can not attribute all the observed volume variations to WallStreetBets, but the relationship is still significant enough to conclude that when a stock is frequently discussed on the site, forum members or observers tend to trade the affected stock.

When testing the impact of sentiment on WallStreetBets on the option-related metrics, we find that sentiment explains relatively much of the variation in stocks' implied volatility, call option volume, and put option volume. In general, these effects seem to be larger than the general effect of sentiment on volume and return. Both when analyzing these metrics on stocks individually and on an aggregated basis, the effect of sentiment on WallStreetBets on these metrics is the most prominent finding of sentiment on any of the factors we have sought to analyze in this thesis.

When assessing whether WallStreetBets can be attributed to have impacted the January 2021 rallies, we find in our minute-by-minute regressions that much of the variation in volume can be attributed to WallStreetBets sentiment. Although we did not find any statistical relationships between return and sentiment when looking at real-time impacts, we could infer that the increased volume as a result of WallStreetBets sentiment could have an indirect effect on the returns for the affected stocks.

In the thesis, we have also shown that the most-shorted stocks on Russell 3000 achieved substantial gains on average in January 2021, while the least-shorted stocks on average achieved a negative return. We also observed how this impacted two of the largest financial institutions' returns, leading to a restructuring between short to long baskets not seen since the financial crisis in 2008. In our view, this could be the direct, or indirect result of WallStreetBets activity.

We have also found that many of the stocks in the thesis achieved abnormal returns at their relative mention-peaks. In addition, we have shown that implementing trading

strategies based on mention-weights on WallStreetBets could have yielded remarkable returns in the period we have explored. This underscores the effect of sentiment on WallStreetBets on stock returns, as it shows that by investing in stocks frequently discussed on the forum, one could potentially have achieved large returns. To what extent these strategies would be possible to implement in reality has been discussed, and they differ in terms of both returns and practical feasibility.

In sum, we would argue that sentiment of WallStreetBets has had a measured effect on several of the stock metrics we have researched, thus we would conclude that the forum does hold some form of market power. For returns, the results are ambiguous and we can not draw any clear empirically grounded conclusions based on our analysis alone. For volume, the relationship is clearer and indicates that when stocks are subject to WallStreetBets activity, more people trade the stocks. Finally, for the option-related metrics we can conclude that WallStreetBets has had a direct effect on implied volatility as well as call and put option volume. Although it is unclear how strongly they have influenced the market just by this analysis alone, it would be hard to claim that WallStreetBets have not had a decisive role in the magnitude of the rallies we have analyzed in this thesis.

Our mission has not been to identify causal relationships between sentiment on WallStreetBets and stock metrics, but to investigate whether sentiment on WallStreetBets can be claimed to have had an isolated impact on these metrics. Sentiment is found to have had a statistically significant impact on several of the stock metrics, but one can not use this study to conclude that sentiment, therefore, has a causal relationship with these metrics. Our results suggest that sentiment explains variations in these metrics, not that sentiment tells the whole story. Therefore, we would advise exercising caution when making inferences with the results in this thesis.

6.2 Limitations and suggestions for future research

In this thesis, we have not sought to explain through what exact mechanisms sentiment on WallStreetBets has translated into statistically significant correlations with returns, volume, and option-related metrics, but rather to investigate whether sentiment on WallStreetBets has had an impact on these metrics overall. This means that we have not sought to explain whether the observed movements in returns, volume, and option-related metrics can be attributed to either retail or institutional investor activity. Although retail investors are believed to make up the majority of the forum's users, institutional investors may well have been monitoring sentiment on WallStreetBets and implemented trading strategies thereof. A suggestion for further research is therefore to drill into stock and option markets order flows and analyze who has been driving the movements in returns, stock volumes, and option volumes in the thesis' selected time interval.

Also, we have limited our analyses to describing the effect of sentiment on WallStreetBets alone, not incorporating sentiment data from other sources. The media's coverage of the rallies in January/February 2021 was massive, which possibly contributed significantly to fuel those rallies. However, we have not analyzed the effect of media coverage nor the sentiment in news articles in this thesis. Neither have we incorporated data from any other social media platform than WallStreetBets. A suggestion for further research is therefore to apply the insights from this thesis and supplement these with models linking e.g. media coverage, sentiment in news articles, and data from other social media to financial data to analyze what effects sentiment in these platforms had on stock returns, volume, implied volatility and option volumes, and whether WallStreetBets serves as a catalyst to media sources or if the forum simply serves as an echo chamber for potential rallies. Although we have not only sought to look at whether the forum influenced the rallies in this thesis, the potential omitted variable bias introduced by excluding these external factors on the January 2021 rallies cannot be ignored.

A limitation on how the WallStreetBets data is collected was based upon how the Reddit API does not allow direct access to its data. Thus, with pushshift.io being the best

alternative we had for data extraction, there were still days not present in our dataset. Most notably, data from January 25 and 26, 2021 was missing from the dataset, vital days in terms of explaining how sentiment was able to fuel the short squeeze rallies experienced during the end of January 2021. The missing days might be linked to the fact that WallStreetBets went private on January 27. As imputing this data would alter our results, these dates are not included in any portion of our analysis. We did an experiment to control for their omitted effects by constructing regressions where sentiment on January 27 was assigned to the two preceding days, which demonstrated that the missing data thankfully did not change our overall results in a significant way. As the majority of the missing data were linked to days in 2021, research revolving around that year in isolation could be flawed and should be interpreted with that in mind.

Our original idea was to capture the overall effect of WallStreetBets where we tried to identify all stocks with CRSP share code 10 and 11 from NYSE, AMEX, and NASDAQ from our WallStreetBets dataset. However, due to computational limits and time constraints, we could only look at a sample of stocks. A suggestion for future research is therefore to research WallStreetBets effect on all stocks that are mentioned on the forum, and see whether the measured effect can be larger for smaller stocks that experience lower volumes. These could be more prone to manipulation or so-called pump-and-dump schemes than the more popular stocks we have taken on in this thesis. This would also allow the creation of more thorough portfolios to assess whether trading strategies originating from the forum could have merit to it.

We opted to exclude weekend activity from our research as the drop in activity would suggest fewer participants on the forum, and thus a lower overall potential market effect. Incorporating this data in future research could potentially result in a clearer picture of WallStreetBets' influence on the market overall.

Finally, we look forward to seeing what literature spawns in the future surrounding this topic, and hope to have contributed at least partially to the emerging theory surrounding the democratization of finance and investment forums' influence on the financial markets.

7 References

Aggarwal, R. K., & Wu, G. (2003). Stock Market Manipulation - Theory & Evidence. In AFA 2004 San Diego Meetings.

AksjeNorge. (2021). Nordmenn og aksjer: Årsstatistikk 2020.

Alexa.com. (2021). https://www.alexa.com/siteinfo/reddit.com. Downloaded February 16, 2021.

Alfonso III, F. (2021, January 29). Class-action lawsuit filed against Robinhood following outrage over GameStop stock restriction. CNN Business. https://edition.cnn.com/2021/01/28/investing/lawsuit-robinhood-gamestop-wallstreetbets/index.html

Aliaj, O., & Fletcher, L. (2021, February 7). Reddit-led market turmoil hits big quant hedge funds. Financial Times. https://www.ft.com/content/1ed2b0de-ea10-4a50-8f33-9f0a1cd38be9

Aliaj, O., Fletcher, L., & Mackenzie, M. (2021, February 6). Melvin Capital, GameStop and the road to disaster. Financial Times. https://www.ft.com/content/3f6b47f9-70c7-4839-8bb4-6a62f1bd39e0

Allen, F., & Gale, D. (1992). Stock-Price Manipulation. The Review of Financial Studies, 5(3), 503-529.

Allen, F., Haas, M., Nowak, E., & Tengulov, A. (2019). Market Efficiency and Limits to Arbitrage: Evidence from the Volkswagen Short Squeeze. Swiss Finance Institute Research Paper, 17-64.

Angel, J. J., & McCabe, D. M. (2009). The business ethics of short selling and naked short selling. Journal of Business Ethics, 85(1), 239-249.

Anthony, K. S. (2020, March 2). Irrational Exuberance: An Interview With WallStreetBets Founder Jaime Rogozinski. SumZero. https://sumzero.com/

headlines/business_services/463-wsb

Armstrong, R. (2021, January 21). Q&A: A basic guide to the GameStop squeeze. Financial Times. https://www.ft.com/content/f2929fcc-cc62-4d20-8397-2040ce3f595e

Asquith, P., Pathak, P. A., & Ritter, J. R. (2005). Short interest, institutional ownership, and stock returns. Journal of Financial Economics, 78(2), 243-276.

Auer, M., Malischnig, D., & Griffiths, M. D. (2020). Gambling Before and During the COVID-19 Pandemic Among European Regular Sports Bettors: An Empirical Study Using Behavioral Tracking Data. International Journal of Mental Health and Addiction, 1-8.

Avis, D. (2021, March 17). Democrats Question Game-Like Trading Apps at House Hearing. Bloomberg. https://www.bloomberg.com/news/articles/2021-03-17/democrats-eye-game-like-trading-apps-at-house-hearing-on-markets

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. The Journal of Finance, 61(4), 1645-1680.

Barber, B. M., Odean, T., & Zhu, N. (2009a). Do retail trades move markets?. The Review of Financial Studies, 22(1), 151-186.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2009b). Just how much do individual investors lose by trading?. The Review of Financial Studies, 22(2), 609-632.

Black, F. (1975). Fact and Fantasy in the Use of Options. Financial Analysts Journal, 31(4), 36-41.

Board of Governors of The Federal Reserve System. (2020, March 15). Federal Reserve issues FOMC statement [Press release]. https://www.federalreserve.gov/newsevents/pressreleases/monetary20200315a.htm

Boulton, T. J., & Braga-Alves, M. V. (2009). Naked short selling and market returns.

The Journal of Portfolio Management, 38(3), 133-142.

Brent, A., Morse, D., & Stice, E. K. (1990). Short interest: Explanations and tests. Journal of Financial and Quantitative Analysis, 25(2), 273-289.

Brettell, K., Pierog, K., & Ranasinghe, D. (2020, March 9). U.S. Treasury yields tumble to record lows on virus, oil fears. Reuters. https://www.reuters.com/article/us-usa-bonds/u-s-treasury-yields-tumble-to-record-lows-on-virus-oil-fears-idUKKBN20W18G

Brown, A. (2021, February 4). Reddit Traders Have Lost Millions Over GameStop. But Many Are Refusing To Quit. Forbes. https://www.forbes.com/sites/abrambrown/2021/02/04/reddit-traders-have-lost-millions-over-gamestop-but-many-are-refusing-to-quit/?sh=738bba792d8f

Burch, T. R., Emery, D. R., & Fuerst, M. E. (2016). Who moves markets in a sudden marketwide crisis? Evidence from 9/11. Journal of Financial and Quantitative Analysis, 51(2), 463-487.

Bursztynsky, J. (2021, January 26). GameStop jumps after hours as Elon Musk tweets out Reddit board that's hyping stock. CNBC. https://www.cnbc.com/2021/01/26/gamestop-jumps-as-elon-musk-tweets-out-reddit-board-thats-hyping-stock.html

Cambridge Dictionary. (n.d.). sentiment. https://dictionary.cambridge.org/dictionary/english/sentiment. Downloaded March 11, 2021.

Canny, W. (2021, February 24). These short-sellers are making a killing in defiance of GameStop's retail army. Financial News.

Charles Schwab. (2021). We've Adjusted Margin Requirements on Certain Securities. https://www.schwab.com/margin-updates. Downloaded May 4, 2021.

Chiah, M., & Zhong, A. (2020). Trading from home: The impact of COVID-19 on trading volume around the world. Finance Research Letters, 37, 101784.

Christian, J. W., Shapiro, R., & Whalen, J. P. (2006). Naked short selling: How exposed are investors. Houston Law Review, 43, 1033-1090.

CNBC.com. (2008, October 28). VW Briefly Rivals Exxon for Biggest Market Cap. CNBC. https://www.cnbc.com/id/27414247

CompaniesMarketCap.com. (2021). Largest automakers by market capitalization. https://companiesmarketcap.com/automakers/largest-automakers-by-market-cap/. Downloaded May 23, 2021.

Corrado, C. J., & Su, T. (1997). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. The European Journal of Finance, 3(1), 73-85.

Craft.co. (2021). Robinhood. https://craft.co/robinhood. Downloaded February 01, 2021.

Cridlin, J. (2021, February 1). After GameStop stock frenzy, Tampa firms join rush to sue Robinhood trading app. Tampa Bay Times. https://www.tampabay.com/news/business/2021/02/01/after-gamestop-stock-frenzy-tampa-firms-join-rush-to-sue-robinhood-trading-app/

Davies, R. (2021, January 28). GameStop shares plunge after ban by Robinhood app. The Guardian. https://www.theguardian.com/business/2021/jan/28/gamestop-shares-robinhood-app-ban-blackberry-amc-nokia-reddit

Dawson, C., & Popina, E. (2021, February 8). Tesla Bets on Bitcoin in Blue-Chip Boost to Cryptocurrency. Bloomberg. https://www.bloomberg.com/news/articles/2021-02-08/tesla-bets-1-5-billion-on-bitcoin-in-new-policy-crypto-surges

Day, C. C. (2004). Is there a tulip in your future?: ruminations on tulip mania and the innovative dutch futures markets. Journal des Economistes et des Etudes Humaines, 14(2), 151-170.

De Long, B., Shleifer, A., Summers, L., & Waldmann, R. (1990). Positive feedback

investment strategies and destabilizing rational speculation. The Journal of Finance, 45(2), 379-395.

Dechow, P. M., Hutton, A. P., Meulbroek, L., & Sloan, R. G. (2001). Short-sellers, fundamental analysis, and stock returns. Journal of Financial Economics, 61(1), 77-106.

DeMarzo, P. M., Kaniel, R., & Kremer, I. (2008). Relative wealth concerns and financial bubbles. The Review of Financial Studies, 21(1), 19-50.

Dorn, D., Huberman, G., & Sengmueller, P. (2008). Correlated trading and returns. The Journal of Finance, 63(2), 885-920.

Duffy, C. (2021, January 28). Robinhood customer files class action suit over GameStop trading restrictions. CNN Business. https://edition.cnn.com/business/live-news/wallstreetbets-reddit-vs-wall-street/h_4f2aeccd71962c141ab3f83b7ad735b2

Finnerty, J. D. (2005). Short selling, death spiral convertibles, and the profitability of stock manipulation. Available at SSRN: https://ssrn.com/abstract=687282

Fitzgerald, M. (2020, May 21). Many Americans used part of their coronavirus stimulus check to trade stocks. CNBC. https://www.cnbc.com/2020/05/21/many-americans-used-part-of-their-coronavirus-stimulus-check-to-trade-stocks.html

Fitzgerald, M. (2021, January 28). Robinhood restricts trading in GameStop, other names involved in frenzy. CNBC. https://www.cnbc.com/2021/01/28/robinhood-interactive-brokers-restrict-trading-in-gamestop-s.html

Foucault, T., Pagano, M. & Röell, A. (2013). Market Liquidity: Theory, Evidence, and Policy. Oxford University Press.

Franck, T., & Li, Y. (2020, March 8). 10-year Treasury yield hits new all-time low of 0.318% amid historic flight to bonds. CNBC. https://www.cnbc.com/2020/03/

09/10-year-treasury-yield-plunges.html

Frederick, R. (2020, May 19). How to Understand Option Greeks. Charles Schwab. https://www.schwab.com/resource-center/insights/content/how-to-understand-options-greeks

French, K. (2021). Current Research Returns. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

French, K. (n.d.). Description of Fama/French 5 Factors (2x3). https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5_factors_2x3.html

Gambrell, D. (2021, February 11). A Brief History of Elon Musk's Recent Market-Moving Tweets. Bloomberg. https://www.bloomberg.com/news/articles/2021-02-11/how-elon-musk-s-tweets-moved-gamestop-gme-bitcoin-dogecoin-and-other-stocks

Gao, X., & Lin, T. C. (2015). Do individual investors treat trading as a fun and exciting gambling activity? Evidence from repeated natural experiments. The Review of Financial Studies, 28(7), 2128-2166.

Garber, P. M. (1990). Famous First Bubbles. Journal of Economic perspectives, 4(2), 35-54.

Ghosh, P. (2021a, January 27). GameStop Loser Citron Says White House Should Stay Out Of The Reddit-Fueled Stock Mania. Forbes. https://www.forbes.com/sites/palashghosh/2021/01/27/gamestop-citron-says-white-house-should-stay-out-of-the-reddit-fueled-stock-frenzy/?sh=30eff8b33fe2

Ghosh, S. (2021b, January 29). Reddit group WallStreetBets hits 6 million users overnight after a wild week of trading antics. Insider. https://www.businessinsider.com/wallstreetbets-fastest-growing-subreddit-hits-58-million-users-2021-1?r=USIR=T

Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. The Journal of Finance, 64(2), 549-578.

Google Finance. (2021). Signal Advance Inc. https://www.google.com/finance/quote/SIGL:OTCMKTS?window=6M. Downloaded May 4, 2021.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. The Journal of Finance, 63(6), 2557-2600.

Han, B., & Kumar, A. (2013). Speculative retail trading and asset prices. Journal of Financial and Quantitative Analysis, 48(2), 377-404.

He, L., & Wang, S. (2021, January 29). China's army of small investors want in on the GameStop mania. CNN Business. https://edition.cnn.com/2021/01/29/investing/china-gamestop-investors-intl-hnk/index.html

Hirschey, M. (1998). How much is a tulip worth? Financial Analysts Journal, 54(4), 11-17.

Hong, H., & Stein, J. C. (2007). Disagreement and the stock market. Journal of Economic Perspectives, 21(2), 109-128.

Hopland, S. (2021, January 28). Nordnet etter GameStop-drama: Vil ikke hindre kjøp av aksjen. E24. https://e24.no/boers-og-finans/i/wevpqG/nordnet-etter-gamestop-drama-vil-ikke-hindre-kjoep-av-aksjen

Imisiker, S., & Tas, B. K. O. (2013). Which firms are more prone to stock market manipulation?. Emerging Markets Review, 16, 119-130.

Jackson, A. (2003). The Aggregate Behavior of Individual Investors. Available at SSRN: http://ssrn.com/abstract=536942.

Jain, D., & Singh, M. (2021, February 8). US STOCKS-Wall Street at record highs on stimulus bets. Reuters. https://www.reuters.com/article/usa-stocks-idUSL4N2KE3NO

Jones, Z. E., & Carissimo, J. (2021, January 28). Robinhood restricts trading in

GameStop, drawing criticism from lawmakers. CBS News. https://www.cbsnews.com/news/robinhood-block-trading-of-gamestop-stock/

Kawa, L. (2020a, February 25). Retail Option Traders Re-Up Bets After Stock Darlings Plummet. Bloomberg. https://www.bloomberg.com/news/articles/2020-02-25/retail-option-traders-re-up-bets-after-stock-darlings-plummet

Kawa, L. (2020b, February 26). Reddit's Profane, Greedy Traders Are Shaking Up the Stock Market. Bloomberg. https://www.bloomberg.com/news/articles/2020-02-26/reddit-s-profane-greedy-traders-are-shaking-up-the-stock-market

Keshner, A. (2021, January 31). GameStop's wild ride will leave some investors with a huge tax bill. MarketWatch. https://www.marketwatch.com/story/taxes-are-usually-an-afterthought-gamestops-wild-ride-will-leave-some-investors-with-a-huge-tax-bill-11611866789

Kindleberger, C. (1978). Manias, Panics, and Crashes. New York: Basic.

Kowsmann, P. (2021, May 4). Negative Rates Push Cautious Germans Into the Stock Market. The Wall Street Journal. https://www.wsj.com/articles/cautious-german-savers-brave-the-stock-market-11620120602

Kumar, A. (2009). Who gambles in the stock market? The Journal of Finance, 64(4), 1889-1933.

Kumar, A., & Lee, C. M. (2006). Retail investor sentiment and return comovements. The Journal of Finance, 61(5), 2451-2486.

Kumar, N. (2021, January 21). GameStop Skeptics Citron, Melvin Succumb to Epic Short Squeeze. Bloomberg. https://www.bloomberg.com/news/articles/2021-01-27/melvin-capital-closes-out-its-gamestop-position-cnbc-reports

Lamont, O. A. (2012). Go down fighting: Short sellers vs. firms. The Review of Asset Pricing Studies, 2(1), 1-30.

Li, G. (2012). Gamblers as personal finance activists. FEDS Working Paper No.

2012-18.

Li, Y. (2021, February 2). As GameStop plunges, Volkswagen's 2008 short squeeze gives an idea of how painful it will get. CNBC. https://www.cnbc.com/2021/02/02/as-gamestop-plunges-volkswagens-2008-short-squeeze-gives-an-idea-of-how-painful-it-will-get.html

Li, Y., & Pound, J. (2021, January 28). GameStop's stock closes down more than 40% after brokers place restrictions on trades. CNBC. https://www.cnbc.com/2021/01/28/gamestop-reverses-losses-and-surges-another-30percent-in-the-premarket-to-450-as-mania-continues.html

Lipschultz, B. (2020, February 19). Option Market Darlings Virgin Galactic, Plug Power Are Surging Again. Bloomberg. https://www.bloomberg.com/news/articles/2020-02-19/options-market-darlings-virgin-galactic-plug-power-are-surging

Liu, B. & Xu, W. (2016). Short Squeezes. Available at SSRN: https://ssrn.com/abstract=2019361

Lyons, K. (2021, January 22). GameStop stock halts trading after Reddit drama. The Verge. https://www.theverge.com/2021/1/22/22244900/game-stop-stock-halted-trading-volatility

Musk, E. [@elonmusk]. (2021a, January 7). Use Signal [Tweet]. Twitter. https://twitter.com/elonmusk/status/1347165127036977153

Musk, E. [@elonmusk]. (2021b, January 26). Gamestonk!! [Tweet]. Twitter. https://twitter.com/elonmusk/status/1354174279894642703

Musk, E. [@elonmusk]. (2021c, January 26). I kinda love Etsy [Tweet]. Twitter. https://twitter.com/elonmusk/status/1354027651468550144

Musk, E. [@elonmusk]. (2021d, January 31). On Clubhouse tonight at 10pm LA time [Tweet]. Twitter. https://twitter.com/elonmusk/status/1355983231988862978

Nagel, S. (2004). Short sales, institutional investors and the cross-section of stock returns. Journal of Financial Economics, 78(2), 277-309.

Niu, E. (2021, January 26). GameStop's Gargantuan Gamma Squeeze. The Motley Fool. https://www.fool.com/investing/2021/01/26/gamestops-gargantuan-gamma-squeeze/

Ortmann, R., Pelster, M., & Wengerek, S. T. (2020). COVID-19 and investor behavior. Finance Research Letters, 37, 101717.

Panigirtzoglou, N., Inkinen, M., Poddar, N., & Agarwal, E. (2021, February 16). Flows Liquidity: The retail impulse remains strong. J.P. Morgan Global Markets Strategy.

Pedersen, L. H. (2021). Game on: Social networks and markets. Available at SSRN: https://ssrn.com/abstract=3794616

Phillips, M. (2021, March 17). Congress hears testimony again on GameStop, focusing on the financial plumbing behind the frenzy. The New York Times. https://www.nytimes.com/2021/03/17/business/gamestop-hearing.html

Ponciano, J. (2021, February 10). Meme Stock Saga Officially Over? GameStop Short Interest Plunged 70% Amid $20 Billion Loss. Forbes. https://www.forbes.com/sites/jonathanponciano/2021/02/10/meme-stock-saga-officially-over-gamestop-short-interest-plunged-70-amid-20-billion-loss/?sh=138c22eb2130

Ponczek, S., Gemmell, K., & Wells, C. (2021, February 2). Meme Stocks Lose $167 Billion as Reddit Crowd Preaches Defiance. Bloomberg. https://www.bloomberg.com/news/articles/2021-02-02/moonshot-stocks-lose-167-billion-as-crowd-preaches-defiance

Porsche. (2005, September 28). Porsche Acquires over 10 Per Cent of VM Common Stock [Press release]. https://www.porsche.com/usa/aboutporsche/pressreleases/pag/?id=2005-09-28pool=international-de

Porsche. (2008a, October 26). Porsche heads for domination agreement: Interest in Volkswagen increased to 42.6 percent [Press release]. https://www.porsche-se.com/en/news/press-releases/details/news/detail/News/porsche-heads-for-domination-agreement-1

Porsche. (2008b, October 30). Partial Settlement of Hedging Transactions Planned [Press release]. https://www.porsche-se.com/en/news/press-releases/details/news/detail/News/short-sellers-responsible-for-extreme-price-movements-in-volkswagen-ordinary-shares-1

Powell, J., & Stafford, P. (2020, February 20). Reddit: bull attack. Financial Times. https://www.ft.com/content/5a5aa2a6-0888-49b9-92f5-8fbf967bf283

Ramey, J. (2021, January 28). Here's How the GameStop Short Squeeze Is like the VW Squeeze of 2008. Yahoo! Finance. https://finance.yahoo.com/news/heres-gamestop-short-squeeze-vw-232300412.html

Rao, S. (2021, February 1). GameStop short-squeeze losses at $12.5 billion YTD - Ortex data. Reuters. https://www.reuters.com/article/us-retail-trading-gamestop-losses-idUSKBN2A12XA

Reddit.com. (2020). https://www.redditinc.com. Downloaded January 12, 2021.

Rega, S. (2020, October 7). How Robinhood and Covid opened the floodgates for 13 million amateur stock traders. CNBC. https://www.cnbc.com/2020/10/07/how-robinhood-and-covid-introduced-millions-to-the-stock-market.html

Reuters Staff. (2021, January 27). Nokia not aware of any reason for share surge. Reuters. https://www.reuters.com/article/us-nokia-shares-idUSKBN29W2H1

Robinhood. (2021, January 29). What happened this week [Press release]. https://blog.robinhood.com/news/2021/1/29/what-happened-this-week. Downloaded February 2, 2021.

Robinhood. (n.d.a). About us. https://robinhood.com/us/en/about-us/.

Downloaded February 10, 2021.

Robinhood. (n.d.b). How we make money. https://robinhood.com/us/en/about-us/how-we-make-money/. Downloaded February 1, 2021.

Robinhood. (n.d.c.). Our Products. https://robinhood.com/us/en/about/our-products. Downloaded February 23, 2021.

Rooney, K. (2020, June 17). Fintech app Robinhood is driving a retail trading renaissance during the stock market's wild ride. CNBC. https://www.cnbc.com/2020/06/17/robinhood-drives-retail-trading-renaissance-during-markets-wild-ride.html

Sarlin, J. (2021, January 30). Inside the Reddit army that's crushing Wall Street. CNN Business. https://edition.cnn.com/2021/01/29/investing/wallstreetbets-reddit-culture/index.html

Scott, G. (2021, February 7). Gamma Hedging. Investopedia. https://www.investopedia.com/terms/g/gamma-hedging.asp

SEC. (2015, April 8). Key Points About Regulation SHO. https://www.sec.gov/investor/pubs/regsho.htm

SEC. (2021, January 29). Statement of Acting Chair Lee and Commissioners Peirce, Roisman, and Crenshaw Regarding Recent Market Volatility [Public statement]. https://www.sec.gov/news/public-statement/joint-statement-market-volatility-2021-01-29?utm_medium=emailutm_source=govdelivery

Sheetz, M. (2020, February 22). Crazy speculative trading is driving some stocks inexplicably higher, a 'late cycle' market sign. CNBC. https://www.cnbc.com/2020/02/22/crazy-speculative-trading-is-driving-some-stocks-inexplicably-higher-a-late-cycle-market-sign.html

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. Journal of Economic Perspectives, 4(2), 19-33.

Sornette, D. & Cauwels, P. (2014). Financial bubbles: mechanisms and diagnostics. arXiv preprint arXiv:1404.2140.

Stankiewicz, K. (2021, February 17). Interactive Brokers chair says financial system came 'dangerously close' to failure during GameStop mania. CNBC. https://www.cnbc.com/2021/02/17/interactive-brokers-chairman-thomas-peterffy-on-gamestop-frenzy.html

Subreddit Stats. (2021). https://subredditstats.com/r/wallstreetbets. Downloaded April 1, 2021.

Summa, J. (2021, February 23). Option Greeks: 4 Factors for Measuring Risk. Investopedia. https://www.investopedia.com/trading/getting-to-know-the-greeks/

[SuperShortSqueezer]. (2021, May 11). $BNGO earnings tomorrow! + My Expectations for Earnings [Online forum post]. Reddit. https://www.reddit.com/r/wallstreetbets/comments/n9zd4d/bngo_earnings_tomorrow_my_expectations_for/

TD Ameritrade. (2021, April 30). Securities with increased margin requirements and trading strategy limits. https://www.tdameritrade.com/td-ameritrade-trading-restrictions-stocks.html

U.S. House Committee on Financial Services. (2021, February 18). Virtual Hearing – Game Stopped? Who Wins and Loses When Short Sellers, Social Media, and Retail Investors Collide. https://financialservices.house.gov/calendar/eventsingle.aspx?EventID=407107

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A., Jetten, J., . . . Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. Nature human behaviour, 4(5), 460-471.

Wang, L., & Hajric, V. (2020, February 20). Mom and pop investors are on an epic stock buying spree, fueled by free trades. Los Angeles Times. https://www.latimes.com/business/story/2020-02-21/small-investors-pile-into-stocks

Warmbrodt, Z. (2021, February 1). Robinhood CEO expected to testify before Maxine Waters' panel on GameStop. Politico. https://www.politico.com/news/2021/02/01/robinhood-ceo-testify-maxine-waters-464491

Whiteman, L. (2021, January 25). Why Shares of Palantir Technologies Are Up Again Today. The Motley Fool. https://www.fool.com/investing/2021/01/25/why-shares-of-palantir-technologies-are-up-again-t/

Winck, B. (2021, January 27). GameStop short-sellers Melvin Capital and Citron surrender bearish bets after 700% rally drives huge losses. Markets Insider. https://markets.businessinsider.com/news/stocks/gamestop-stock-short-sellers-melvin-capital-citron-surrender-bets-gme-2021-1-1030010382

Xu, W., & Zheng, Y. (2016). The Short Squeeze: The "Invisible" Cost of Short Sales. Available at SSRN: https://ssrn.com/abstract=2783374

Zweig, J. (2020, June 12). Playing the Market Has a Whole New Meaning. The Wall Street Journal. https://www.wsj.com/articles/playing-the-market-has-a -whole-new-meaning-11591974010

Appendix

A1 Sentiment regressions on individual stocks

A1.1 AAPL - Apple, Inc.

A1.1.1 Full regression

Table A1.1: AAPL full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0001 t = 0.415	16,717.870 t = 2.301**	0.001 t = 1.588	509.547 t = 3.040***	169.380 t = 2.169**	0.000 t = 0.130
Positive sentiment comments	0.022 t = 1.816*	−105,273.800 t = −0.303	−0.178 t = −2.085**	2,091.438 t = 0.348	−18,696.700 t = −3.897***	0.000 t = 1.415
Title per day	0.063 t = 1.249	1,179,907.000 t = 1.251	−0.153 t = −1.532	13,956.790 t = 0.821	−7,926.693 t = −0.886	0.000 t = 0.499
Positive sentiment title	0.002 t = 0.616	194,044.700 t = 2.110**	−0.003 t = −0.147	−1,208.661 t = −0.726	−1,864.472 t = −1.668*	−0.000 t = −1.817*
Title text per day	−0.071 t = −1.977**	621,519.300 t = 0.499	0.368 t = 3.173***	−21,265.930 t = −0.976	756.031 t = 0.063	−0.000 t = −0.591
Positive sentiment title text	−0.003 t = −1.121	27,574.240 t = 0.322	−0.022 t = −0.829	−2,126.413 t = −1.220	−2,042.698 t = −1.595	−0.000 t = −1.313
Mkt.RF	1.270 t = 24.122***	−1,959,715.000 t = −1.964**	−1.022 t = −1.497	−10,099.700 t = −0.418	−37,628.410 t = −1.948*	0.000 t = 18.379***
SMB	−0.109 t = −0.976					
HML	−0.659 t = −5.133***					
RMW	0.658 t = 3.803***					
CMA	1.322 t = 4.555***					
MOM	0.129 t = 1.348					
Constant	−1.271 t = −1.727*	44,640,430.000 t = 2.134**	47.899 t = 8.544***	1,430,829.000 t = 3.719***	2,293,731.000 t = 8.056***	−0.000 t = −0.431
Observations	296	296	296	296	296	296
R^2	0.814	0.308	0.142	0.178	0.127	0.651
Adjusted R^2	0.806	0.291	0.121	0.158	0.106	0.642
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.1.2 Lagged return

Table A1.2: AAPL lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
	0.001	0.0001	0.0002	0.0002	0.0001	0.0004
Comments per day	t = 1.618	t = 0.415	t = 0.732	t = 0.432	t = 0.302	t = 0.977
	−0.001	0.022	0.046	0.008	−0.022	0.041
Positive sentiment comments	t = −0.025	t = 1.816*	t = 2.023**	t = 0.311	t = −0.874	t = 1.415
	0.002	0.063	0.069	−0.003	0.046	−0.006
Title per day	t = 0.041	t = 1.249	t = 1.601	t = −0.065	t = 0.899	t = −0.094
	−0.002	0.002	0.005	−0.004	0.007	−0.002
Positive sentiment title	t = −0.347	t = 0.616	t = 1.054	t = −0.650	t = 1.340	t = −0.379
	−0.099	−0.071	−0.072	0.004	−0.040	−0.023
Title text per day	t = −1.879*	t = −1.977**	t = −1.047	t = 0.070	t = −0.857	t = −0.456
	0.003	−0.003	−0.003	−0.003	0.015	0.005
Positive sentiment title text	t = 0.407	t = −1.121	t = −0.391	t = −0.475	t = 2.138**	t = 0.695
	−0.438	1.270	−0.484	0.235	0.096	−0.155
Mkt.RF	t = −3.207***	t = 24.122***	t = −3.026***	t = 1.220	t = 0.631	t = −1.073
	0.425	−0.109	0.015	0.293	−0.390	0.547
SMB	t = 1.189	t = −0.976	t = 0.048	t = 0.944	t = −1.055	t = 1.637
	−0.162	−0.659	0.245	0.139	−0.446	−0.037
HML	t = −0.573	t = −5.133***	t = 0.731	t = 0.303	t = −1.205	t = −0.093
	0.284	0.658	0.101	−0.047	−0.803	0.676
RMW	t = 0.767	t = 3.803***	t = 0.253	t = −0.137	t = −2.184**	t = 1.924*
	−0.566	1.322	−0.522	0.276	0.973	−0.805
CMA	t = −1.175	t = 4.555***	t = −1.120	t = 0.526	t = 1.737*	t = −1.298
	−0.042	0.129	0.036	0.204	−0.539	0.179
MOM	t = −0.256	t = 1.348	t = 0.191	t = 0.539	t = −2.018**	t = 0.592
	0.091	−1.271	−2.995	−0.221	−0.244	−2.944
Constant	t = 0.061	t = −1.727*	t = −1.977**	t = −0.127	t = −0.151	t = −1.662*
Observations	295	296	295	294	293	292
R^2	0.119	0.814	0.136	0.056	0.094	0.064
Adjusted R^2	0.082	0.806	0.099	0.016	0.056	0.024

Note: *p<0.1; **p<0.05; ***p<0.01

A1.1.3 Lagged volume

Table A1.3: AAPL lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
	2,256.591	16,717.870	6,496.114	6,733.283	−1,542.807	−445.273
Comments per day	t = 0.339	t = 2.301**	t = 0.880	t = 0.857	t = −0.201	t = −0.056
	9,272.617	−105,273.800	24,595.640	9,693.490	143,378.200	257,155.200
Positive sentiment comments	t = 0.026	t = −0.303	t = 0.066	t = 0.029	t = 0.391	t = 0.663
	1,318,574.000	1,179,907.000	513,130.100	139,638.100	−20,975.830	−127,032.800
Title per day	t = 1.631	t = 1.251	t = 0.675	t = 0.170	t = −0.027	t = −0.151
	275,565.800	194,044.700	290,189.400	312,757.400	233,426.400	251,379.700
Positive sentiment title	t = 3.005***	t = 2.110**	t = 3.151***	t = 3.324***	t = 2.614***	t = 2.517**
	1,837,799.000	621,519.300	1,487,432.000	1,152,372.000	1,890,992.000	1,411,828.000
Title text per day	t = 1.838*	t = 0.499	t = 1.466	t = 1.078	t = 1.621	t = 1.177
	−4,618.636	27,574.240	58,135.200	36,107.590	46,701.260	24,230.070
Positive sentiment title text	t = −0.051	t = 0.322	t = 0.657	t = 0.371	t = 0.451	t = 0.211
	−1,389,623.000	−1,959,715.000	131,240.900	−156,698.500	−63,270.020	248,620.500
Mkt.RF	t = −1.392	t = −1.964**	t = 0.125	t = −0.147	t = −0.059	t = 0.251
	38,283,887.000	44,640,430.000	34,575,077.000	39,886,123.000	38,750,681.000	34,992,905.000
Constant	t = 1.805*	t = 2.134**	t = 1.520	t = 1.910*	t = 1.776*	t = 1.506
Observations	295	296	295	294	293	292
R^2	0.279	0.308	0.200	0.138	0.095	0.069
Adjusted R^2	0.261	0.291	0.180	0.117	0.073	0.046

Note: *p<0.1; **p<0.05; ***p<0.01

A1.1.4 2020 regression

Table A1.4: AAPL 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.00005 t = −0.172	18,358.270 t = 1.773*	0.001 t = 0.978	589.201 t = 2.755***	191.924 t = 1.855*	−0.000 t = −0.721
Positive sentiment comments	0.023 t = 1.871*	−428,904.100 t = −1.197	−0.214 t = −2.075**	9,542.717 t = 1.302	−16,261.720 t = −2.965***	0.000 t = 1.497
Title per day	0.111 t = 2.456**	979,697.300 t = 0.710	−0.170 t = −0.959	11,245.300 t = 0.391	−10,866.960 t = −0.730	0.000 t = 1.209
Positive sentiment title	0.001 t = 0.472	188,968.800 t = 2.052**	−0.002 t = −0.092	−215.012 t = −0.110	−1,111.656 t = −0.879	−0.000 t = −1.765*
Title text per day	−0.090 t = −2.052**	901,119.800 t = 0.671	0.385 t = 2.882***	−28,817.160 t = −1.199	−408.645 t = −0.031	−0.000 t = −0.418
Positive sentiment title text	−0.003 t = −1.064	−11,893.240 t = −0.140	−0.022 t = −0.773	−1,480.445 t = −0.807	−1,553.864 t = −1.174	−0.000 t = −1.224
Mkt.RF	1.242 t = 22.837***	−1,483,153.000 t = −1.492	−1.009 t = −1.419	−10,799.890 t = −0.441	−37,699.250 t = −1.967**	0.000 t = 17.883***
SMB	−0.151 t = −1.262					
HML	−0.547 t = −4.050***					
RMW	0.799 t = 4.210***					
CMA	1.021 t = 3.869***					
MOM	0.143 t = 1.503					
Constant	−1.276 t = −1.652*	60,571,745.000 t = 2.735***	50.293 t = 7.643***	955,551.900 t = 2.030**	2,132,869.000 t = 6.594***	−0.000 t = −0.542
Observations	252	252	252	252	252	252
R^2	0.837	0.341	0.140	0.202	0.102	0.669
Adjusted R^2	0.829	0.322	0.115	0.179	0.077	0.659
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.1.5 2021 regression

Table A1.5: AAPL 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0001 t = −0.178	4,835.607 t = 0.309	0.002 t = 1.077	148.867 t = 0.600	73.839 t = 0.680	0.000 t = 1.088
Positive sentiment comments	0.0003 t = 0.007	−496,052.300 t = −0.615	0.094 t = 0.786	3,991.748 t = 0.415	−3,089.509 t = −0.540	0.000 t = 0.704
Title per day	−0.006 t = −0.059	1,655,474.000 t = 1.439	0.016 t = 0.104	11,486.390 t = 0.789	6,202.621 t = 0.911	−0.000 t = −0.429
Positive sentiment title	0.0003 t = 0.026	79,950.380 t = 0.344	0.026 t = 1.006	−1,966.616 t = −0.769	−1,290.585 t = −0.878	−0.000 t = −0.483
Title text per day	0.005 t = 0.079	−1,731,433.000 t = −0.901	0.211 t = 1.044	5,013.995 t = 0.218	−4,637.103 t = −0.405	0.000 t = 0.401
Positive sentiment title text	−0.002 t = −0.188	−239,930.100 t = −0.779	−0.009 t = −0.260	640.473 t = 0.243	−3.521 t = −0.003	−0.000 t = −0.787
Mkt.RF	1.183 t = 2.686***	−8,588,422.000 t = −1.607	−0.928 t = −1.281	−59,779.060 t = −0.962	−44,717.290 t = −1.234	0.000 t = 2.918***
SMB	−0.562 t = −1.142					
HML	−0.940 t = −2.041**					
RMW	0.157 t = 0.282					
CMA	2.264 t = 2.421**					
MOM	0.194 t = 0.402					
Constant	0.467 t = 0.164	155,351,084.000 t = 2.650***	25.798 t = 3.056***	726,275.000 t = 1.133	697,821.600 t = 1.807*	−0.000 t = −0.501
Observations	44	44	44	44	44	44
R^2	0.734	0.259	0.410	0.352	0.312	0.512
Adjusted R^2	0.631	0.115	0.295	0.226	0.178	0.417

Note: *p<0.1; **p<0.05; ***p<0.01

A1.2 AMC - AMC Entertainment Holdings, Inc.

A1.2.1 Full regression

Table A1.6: AMC full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.001 t = −0.225	1,363.954 t = 0.082	0.003 t = 0.352	9.029 t = 0.777	3.356 t = 0.809	−0.000 t = −0.002
Positive sentiment comments	0.039 t = 1.302	288,125.700 t = 2.815***	−0.396 t = −1.943*	661.892 t = 3.090***	152.557 t = 1.989**	0.000 t = 1.036
Title per day	0.041 t = 0.211	357,248.800 t = 0.631	0.099 t = 0.402	382.253 t = 0.892	349.328 t = 2.302**	−0.000 t = −0.014
Positive sentiment title	0.029 t = 0.727	647,510.400 t = 3.543***	0.171 t = 1.035	1,955.864 t = 4.433***	681.913 t = 4.480***	0.000 t = 0.581
Title text per day	−0.031 t = −0.093	−319,215.100 t = −0.326	−0.102 t = −0.238	−738.664 t = −0.830	−620.934 t = −1.894*	0.000 t = 0.068
Positive sentiment title text	0.023 t = 0.843	158,026.400 t = 1.287	0.042 t = 0.278	479.263 t = 1.678*	224.911 t = 2.341**	0.000 t = 1.809*
Mkt.RF	1.174 t = 2.457**	−1,269,714.000 t = −0.512	−2.286 t = −0.561	−1,479.827 t = −0.439	−1,239.921 t = −1.106	0.00000 t = 6.190***
SMB	3.580 t = 0.870					
HML	−2.118 t = −0.941					
RMW	2.232 t = 0.815					
CMA	7.803 t = 0.428					
MOM	−0.703 t = −0.336					
Constant	−5.542 t = −1.976**	−39,284,858.000 t = −2.735***	166.746 t = 9.986***	−116,492.100 t = −3.825***	−31,665.100 t = −2.955***	−0.00000 t = −2.620***
Observations	296	296	296	296	296	296
R^2	0.255	0.756	0.273	0.449	0.687	0.277
Adjusted R^2	0.223	0.750	0.255	0.436	0.679	0.259

Note: *p<0.1; **p<0.05; ***p<0.01

A1.2.2 Lagged return

Table A1.7: AMC lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.001	−0.001	−0.0003	0.002	0.002	−0.001
	t = 0.366	t = −0.225	t = −0.086	t = 0.558	t = 0.578	t = −0.729
Positive sentiment comments	0.054	0.039	0.016	0.007	0.027	−0.053
	t = 1.451	t = 1.302	t = 0.700	t = 0.276	t = 0.969	t = −1.587
Title per day	−0.047	0.041	−0.013	−0.045	−0.021	0.093
	t = −0.470	t = 0.211	t = −0.124	t = −0.407	t = −0.154	t = 0.896
Positive sentiment title	0.068	0.029	0.020	−0.026	0.021	0.0005
	t = 1.503	t = 0.727	t = 0.512	t = −0.544	t = 0.598	t = 0.015
Title text per day	0.069	−0.031	0.094	0.034	−0.010	−0.149
	t = 0.406	t = −0.093	t = 0.547	t = 0.198	t = −0.042	t = −0.841
Positive sentiment title text	0.024	0.023	0.032	−0.007	−0.040	−0.024
	t = 0.753	t = 0.843	t = 1.137	t = −0.231	t = −1.597	t = −0.955
Mkt.RF	0.023	1.174	−0.546	−0.212	0.337	0.181
	t = 0.068	t = 2.457**	t = −1.172	t = −0.396	t = 0.637	t = 0.512
SMB	0.601	3.580	−0.942	−0.357	0.691	1.193
	t = 0.248	t = 0.870	t = −0.418	t = −0.156	t = 0.271	t = 0.605
HML	−1.609	−2.118	1.652	1.842	−0.801	−0.150
	t = −0.993	t = −0.941	t = 1.234	t = 1.129	t = −0.451	t = −0.105
RMW	1.242	2.232	0.652	−0.964	1.044	0.024
	t = 0.572	t = 0.815	t = 0.348	t = −0.334	t = 0.320	t = 0.011
CMA	−2.554	7.803	−9.799	1.906	3.882	−1.883
	t = −0.299	t = 0.428	t = −0.956	t = 0.215	t = 0.503	t = −0.282
MOM	−0.707	−0.703	−0.477	1.384	0.628	0.285
	t = −0.578	t = −0.336	t = −0.448	t = 1.290	t = 0.509	t = 0.268
Constant	−8.114	−5.542	−4.324	0.676	−1.500	3.515
	t = −1.727*	t = −1.976**	t = −1.620	t = 0.230	t = −0.560	t = 1.226
Observations	295	296	295	294	293	292
R^2	0.093	0.255	0.410	0.201	0.144	0.301
Adjusted R^2	0.055	0.223	0.385	0.166	0.108	0.271

Note: *p<0.1; **p<0.05; ***p<0.01

A1.2.3 Lagged volume

Table A1.8: AMC lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	7,088.092	1,363.954	18,509.200	21,915.230	7,168.928	−4,996.468
	t = 1.206	t = 0.082	t = 1.477	t = 2.138**	t = 0.539	t = −0.430
Positive sentiment comments	461,580.000	288,125.700	156,331.000	225,476.100	220,590.500	123,096.500
	t = 2.261**	t = 2.815***	t = 1.927*	t = 2.341**	t = 2.022**	t = 1.045
Title per day	19,696.970	357,248.800	−137,171.300	−52,155.240	402,970.200	644,761.900
	t = 0.086	t = 0.631	t = −0.431	t = −0.195	t = 1.015	t = 1.301
Positive sentiment title	861,607.100	647,510.400	531,103.500	613,201.000	694,728.900	705,925.300
	t = 2.749***	t = 3.543***	t = 3.528***	t = 3.238***	t = 3.283***	t = 2.853***
Title text per day	67,187.070	−319,215.100	193,448.400	−442,534.200	−1,029,963.000	−1,121,149.000
	t = 0.157	t = −0.326	t = 0.427	t = −0.965	t = −1.656*	t = −1.346
Positive sentiment title text	173,030.700	158,026.400	7,551.169	115,581.800	169,915.000	249,861.000
	t = 0.745	t = 1.287	t = 0.077	t = 0.816	t = 1.345	t = 1.785*
Mkt.RF	467,480.000	−1,269,714.000	1,307,779.000	−1,053,314.000	859,769.300	1,046,310.000
	t = 0.390	t = −0.512	t = 0.782	t = −0.695	t = 0.458	t = 0.501
Constant	−57,971,160.000	−39,284,858.000	−20,403,959.000	−33,107,764.000	−37,674,909.000	−34,661,763.000
	t = −1.870*	t = −2.735***	t = −2.074**	t = −2.601***	t = −2.640***	t = −2.176**
Observations	295	296	295	294	293	292
R^2	0.392	0.756	0.835	0.761	0.607	0.492
Adjusted R^2	0.377	0.750	0.831	0.755	0.597	0.479

Note: *p<0.1; **p<0.05; ***p<0.01

A1.2.4 2020 regression

Table A1.9: AMC 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.017 t = 1.016	50,476.350 t = 1.920*	0.290 t = 3.046***	42.552 t = 1.266	89.722 t = 2.830***	0.000 t = 1.069
Positive sentiment comments	0.008 t = 0.369	38,499.310 t = 1.571	−0.580 t = −3.044***	128.501 t = 2.135**	0.391 t = 0.011	0.000 t = 0.822
Title per day	−0.061 t = −0.069	5,142,062.000 t = 2.622***	−7.221 t = −1.323	8,800.494 t = 2.772***	3,300.869 t = 1.682*	0.000 t = 0.023
Positive sentiment title	0.018 t = 0.646	24,272.290 t = 0.416	0.016 t = 0.086	42.886 t = 0.393	24.633 t = 0.420	0.000 t = 0.156
Title text per day	−0.034 t = −0.034	−1,280,038.000 t = −0.748	6.126 t = 1.348	−731.163 t = −0.319	−1,142.209 t = −0.551	−0.000 t = −0.339
Positive sentiment title text	0.014 t = 0.659	1,414.483 t = 0.039	−0.111 t = −0.676	−33.274 t = −0.507	−43.395 t = −1.055	0.000 t = 1.539
Mkt.RF	1.322 t = 5.571***	360,529.500 t = 1.406	−1.965 t = −0.480	826.534 t = 1.391	−2.166 t = −0.007	0.00000 t = 6.297***
SMB	0.802 t = 0.966					
HML	−0.348 t = −0.400					
RMW	1.043 t = 0.902					
CMA	−4.793 t = −3.134***					
MOM	−1.675 t = −2.203**					
Constant	−3.947 t = −1.861*	1,867,117.000 t = 0.544	176.075 t = 10.363***	4,754.293 t = 0.746	8,766.049 t = 2.345**	−0.00000 t = −2.394**
Observations	252	252	252	252	252	252
R^2	0.395	0.394	0.156	0.278	0.400	0.289
Adjusted R^2	0.365	0.377	0.132	0.258	0.382	0.268
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.2.5 2021 regression

Table A1.10: AMC 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
	−0.0004	−782.822	0.001	2.420	1.054	−0.000
Comments per day	t = −0.118	t = −0.052	t = 0.186	t = 0.487	t = 0.747	t = −0.326
	0.126	1,267,211.000	0.903	−2,178.285	−1,578.814	0.000
Positive sentiment comments	t = 0.261	t = 0.976	t = 1.513	t = −0.785	t = −1.441	t = 2.570**
	0.040	327,113.800	0.092	303.284	323.648	−0.000
Title per day	t = 0.243	t = 0.690	t = 0.369	t = 2.067**	t = 7.425***	t = −0.283
	−0.003	2,396,216.000	0.266	6,448.231	1,833.645	0.000
Positive sentiment title	t = −0.013	t = 2.830***	t = 0.756	t = 3.373***	t = 3.128***	t = 1.074
	−0.051	−266,788.100	−0.084	−558.917	−563.578	0.000
Title text per day	t = −0.163	t = −0.333	t = −0.200	t = −1.699*	t = −5.822***	t = 0.258
	0.047	424,125.600	0.014	646.793	566.610	0.000
Positive sentiment title text	t = 0.198	t = 0.559	t = 0.044	t = 0.339	t = 0.731	t = 0.949
	−2.216	−38,521,223.000	−7.245	−59,622.080	−29,287.310	0.000
Mkt.RF	t = −0.347	t = −0.802	t = −0.288	t = −1.267	t = −2.720***	t = 0.856
	−2.758					
SMB	t = −0.312					
	−0.392					
HML	t = −0.046					
	3.810					
RMW	t = 0.309					
	39.386					
CMA	t = 0.653					
	9.823					
MOM	t = 0.549					
	−11.931	−186,140,820.000	100.923	−36,327.610	64,951.200	−0.00000
Constant	t = −0.239	t = −1.251	t = 1.535	t = −0.118	t = 0.555	t = −1.974**
Observations	44	44	44	44	44	44
R^2	0.715	0.802	0.756	0.465	0.778	0.382
Adjusted R^2	0.605	0.763	0.708	0.361	0.735	0.262

Note: *p<0.1; **p<0.05; ***p<0.01

A1.3 AMD - Advanced Micro Devices, Inc.

A1.3.1 Full regression

Table A1.11: AMD full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.001 t = 0.841	36,562.760 t = 5.249***	0.004 t = 1.992**	156.398 t = 4.380***	80.575 t = 4.850***	0.000 t = 0.417
Positive sentiment comments	0.022 t = 1.061	−901,661.400 t = −4.698***	−0.385 t = −3.349***	1,573.523 t = 1.935*	−962.612 t = −2.341**	0.000 t = 1.660*
Title per day	0.076 t = 0.689	323,432.400 t = 0.446	−0.447 t = −2.395**	9,002.442 t = 2.278**	−21.901 t = −0.013	0.000 t = 0.878
Positive sentiment title	0.005 t = 1.080	−35,507.990 t = −0.847	−0.034 t = −1.722*	233.672 t = 1.040	97.029 t = 1.173	0.000 t = 0.580
Title text per day	−0.149 t = −1.228	−1,434,452.000 t = −2.218**	0.095 t = 0.522	−5,338.762 t = −1.463	−369.196 t = −0.270	−0.000 t = −1.285
Positive sentiment title text	−0.004 t = −0.859	−88,039.970 t = −2.007**	−0.030 t = −1.086	−120.258 t = −0.540	−87.647 t = −0.957	−0.000 t = −0.187
Mkt.RF	1.306 t = 15.009***	−1,716,047.000 t = −2.157**	−1.263 t = −1.668*	−2,630.964 t = −1.222	−6,251.241 t = −3.573***	0.000 t = 17.021***
SMB	−0.039 t = −0.177					
HML	−0.471 t = −1.886*					
RMW	−0.183 t = −0.575					
CMA	0.040 t = 0.106					
MOM	0.202 t = 1.366					
Constant	−1.545 t = −1.130	115,852,177.000 t = 9.472***	85.521 t = 11.520***	96,688.320 t = 1.889*	159,517.900 t = 5.787***	−0.000 t = −1.756*
Observations	296	296	296	296	296	296
R^2	0.568	0.361	0.204	0.393	0.374	0.526
Adjusted R^2	0.550	0.345	0.185	0.379	0.359	0.515
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.3.2 Lagged return

Table A1.12: AMD lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
	0.0002	0.001	0.002	0.001	0.002	0.002
Comments per day	t = 0.230	t = 0.841	t = 1.842*	t = 1.029	t = 1.811*	t = 1.464
	−0.076	0.022	0.058	−0.006	0.024	0.016
Positive sentiment comments	t = −2.180**	t = 1.061	t = 1.730*	t = −0.194	t = 0.634	t = 0.360
	−0.014	0.076	−0.109	0.051	0.012	−0.069
Title per day	t = −0.198	t = 0.689	t = −1.198	t = 0.641	t = 0.164	t = −0.736
	0.006	0.005	0.009	−0.005	0.0004	0.002
Positive sentiment title	t = 0.861	t = 1.080	t = 1.151	t = −0.652	t = 0.057	t = 0.239
	−0.236	−0.149	−0.098	−0.069	−0.145	−0.013
Title text per day	t = −2.390**	t = −1.228	t = −1.217	t = −0.752	t = −1.795*	t = −0.135
	0.019	−0.004	−0.003	0.018	0.004	0.0003
Positive sentiment title text	t = 1.932*	t = −0.859	t = −0.325	t = 1.985**	t = 0.341	t = 0.035
	−0.604	1.306	−0.459	0.372	0.022	−0.191
Mkt.RF	t = −3.999***	t = 15.009***	t = −2.371**	t = 1.805*	t = 0.116	t = −1.195
	0.244	−0.039	0.115	0.221	−0.643	0.525
SMB	t = 0.707	t = −0.177	t = 0.284	t = 0.635	t = −1.572	t = 1.116
	0.230	−0.471	−0.147	0.437	0.076	0.099
HML	t = 0.601	t = −1.886*	t = −0.335	t = 0.984	t = 0.159	t = 0.161
	0.053	−0.183	0.091	0.548	−1.248	−0.344
RMW	t = 0.098	t = −0.575	t = 0.191	t = 1.113	t = −2.620***	t = −0.747
	−1.093	0.040	−0.456	0.114	1.002	0.132
CMA	t = −2.276**	t = 0.106	t = −0.784	t = 0.197	t = 1.586	t = 0.195
	−0.031	0.202	−0.192	0.431	−0.395	0.216
MOM	t = −0.095	t = 1.366	t = −0.660	t = 1.081	t = −1.169	t = 0.460
	4.129	−1.545	−4.240	−0.848	−2.084	−1.554
Constant	t = 1.926*	t = −1.130	t = −1.863*	t = −0.381	t = −0.790	t = −0.560
Observations	295	296	295	294	293	292
R^2	0.155	0.568	0.093	0.117	0.072	0.049
Adjusted R^2	0.120	0.550	0.055	0.079	0.032	0.008
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.3.3 Lagged volume

Table A1.13: AMD lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
	26,175.460	36,562.760	21,339.470	16,325.830	20,844.790	22,055.940
Comments per day	t = 3.969***	t = 5.249***	t = 2.785***	t = 1.865*	t = 2.470**	t = 2.536**
	−889,711.300	−901,661.400	−772,434.300	−729,831.000	−725,646.600	−709,819.500
Positive sentiment comments	t = −4.280***	t = −4.698***	t = −3.744***	t = −3.494***	t = −3.676***	t = −3.398***
	−683,798.000	323,432.400	−456,922.300	−257,579.900	−1,266,656.000	−2,221,258.000
Title per day	t = −1.501	t = 0.446	t = −0.550	t = −0.396	t = −2.537**	t = −3.473***
	−46,154.700	−35,507.990	−62,630.630	−84,431.990	−23,171.240	−59,942.070
Positive sentiment title	t = −0.829	t = −0.847	t = −1.251	t = −1.592	t = −0.426	t = −1.232
	−625,830.700	−1,434,452.000	−111,135.500	−49,603.620	−82,884.440	712,396.800
Title text per day	t = −1.078	t = −2.218**	t = −0.204	t = −0.064	t = −0.133	t = 1.229
	−105,082.600	−88,039.970	−33,881.880	−62,741.690	−59,583.330	−42,568.790
Positive sentiment title text	t = −2.048**	t = −2.007**	t = −0.678	t = −1.084	t = −1.107	t = −0.762
	−843,760.900	−1,716,047.000	−423,193.500	−300,912.100	−271,017.000	120,515.400
Mkt.RF	t = −1.091	t = −2.157**	t = −0.477	t = −0.343	t = −0.350	t = 0.134
	121,456,608.000	115,852,177.000	109,272,070.000	111,489,475.000	108,180,784.000	107,851,798.000
Constant	t = 9.082***	t = 9.472***	t = 8.439***	t = 8.166***	t = 8.887***	t = 8.111***
Observations	295	296	295	294	293	292
R^2	0.203	0.361	0.149	0.128	0.119	0.155
Adjusted R^2	0.184	0.345	0.129	0.107	0.097	0.134
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.3.4 2020 regression

Table A1.14: AMD 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.001 t = 0.513	41,706.280 t = 5.866***	0.004 t = 1.604	175.676 t = 4.510***	102.899 t = 6.548***	−0.000 t = −0.335
Positive sentiment comments	0.024 t = 0.993	−1,053,222.000 t = −5.112***	−0.542 t = −4.439***	1,686.700 t = 1.905*	−1,192.846 t = −2.642***	0.000 t = 1.306
Title per day	0.116 t = 0.725	209,432.100 t = 0.253	−0.530 t = −2.215**	7,896.826 t = 1.634	−1,928.806 t = −1.087	0.000 t = 1.340
Positive sentiment title	0.004 t = 0.779	−17,387.560 t = −0.367	−0.027 t = −1.228	174.511 t = 0.663	60.535 t = 0.654	0.000 t = 0.210
Title text per day	−0.140 t = −1.013	−1,909,489.000 t = −2.978***	−0.017 t = −0.086	−5,955.442 t = −1.458	−656.509 t = −0.449	−0.000 t = −1.093
Positive sentiment title text	−0.005 t = −0.933	−53,643.810 t = −1.145	−0.019 t = −0.638	−88.546 t = −0.364	−44.897 t = −0.486	−0.000 t = −0.375
Mkt.RF	1.304 t = 13.706***	−1,743,475.000 t = −2.235**	−1.363 t = −1.835*	−1,559.872 t = −0.728	−5,619.300 t = −3.237***	0.000 t = 15.782***
SMB	−0.177 t = −0.692					
HML	−0.408 t = −1.392					
RMW	−0.072 t = −0.189					
CMA	−0.422 t = −0.837					
MOM	0.161 t = 0.964					
Constant	−1.586 t = −0.967	124,098,796.000 t = 9.219***	96.111 t = 12.124***	85,253.340 t = 1.500	168,539.400 t = 5.476***	−0.000 t = −1.151
Observations	252	252	252	252	252	252
R^2	0.564	0.404	0.251	0.392	0.438	0.536
Adjusted R^2	0.542	0.386	0.230	0.374	0.422	0.523

Note: *p<0.1; **p<0.05; ***p<0.01

A1.3.5 2021 regression

Table A1.15: AMD 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
	0.0004	5,097.233	0.004	103.349	19.093	0.000
Comments per day	t = 0.275	t = 0.444	t = 0.919	t = 1.109	t = 0.556	t = 0.389
	−0.001	122,306.900	0.159	1,810.239	606.067	0.000
Positive sentiment comments	t = −0.021	t = 0.454	t = 1.396	t = 0.760	t = 0.514	t = 0.028
	0.046	107,276.600	−0.009	6,334.767	480.798	0.000
Title per day	t = 0.367	t = 0.099	t = −0.021	t = 0.802	t = 0.172	t = 0.488
	0.003	−63,631.050	−0.051	262.978	−45.624	0.000
Positive sentiment title	t = 0.189	t = −0.816	t = −1.388	t = 0.529	t = −0.153	t = 1.950*
	−0.214	385,684.200	0.154	1,192.107	2,271.692	−0.000
Title text per day	t = −1.236	t = 0.261	t = 0.277	t = 0.109	t = 0.386	t = −0.899
	−0.004	−71,429.870	0.006	−47.940	57.276	0.000
Positive sentiment title text	t = −0.276	t = −0.793	t = 0.152	t = −0.070	t = 0.131	t = 0.307
	1.000	−2,751,975.000	−0.484	−24,000.600	−20,786.840	0.000
Mkt.RF	t = 2.485**	t = −1.373	t = −0.658	t = −1.411	t = −1.865*	t = 6.204***
	−0.004					
SMB	t = −0.008					
	−0.438					
HML	t = −0.730					
	−0.668					
RMW	t = −0.891					
	0.741					
CMA	t = 0.834					
	0.531					
MOM	t = 0.881					
	0.514	44,531,900.000	40.337	101,380.800	84,239.390	−0.000
Constant	t = 0.187	t = 2.604***	t = 6.536***	t = 0.680	t = 1.035	t = −1.080
Observations	44	44	44	44	44	44
R^2	0.722	0.209	0.194	0.480	0.273	0.579
Adjusted R^2	0.614	0.055	0.037	0.379	0.131	0.497

Note: *p<0.1; **p<0.05; ***p<0.01

A1.4 APHA - Aphria, Inc.

A1.4.1 Full regression

Table A1.16: APHA full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.007	12,758.120	0.013	13.424	7.387	−0.000
	t = −0.627	t = 0.655	t = 0.319	t = 0.343	t = 1.783*	t = −0.428
Positive sentiment comments	0.016	−12.325	−0.047	18.410	−6.144	0.000
	t = 1.592	t = −0.001	t = −0.979	t = 0.610	t = −1.007	t = 1.380
Title per day	0.713	603,247.000	1.382	1,537.108	401.364	0.000
	t = 0.659	t = 0.316	t = 0.342	t = 0.387	t = 1.309	t = 0.165
Positive sentiment title	0.030	99,029.730	0.129	290.310	25.178	0.000
	t = 0.908	t = 1.592	t = 0.929	t = 2.112**	t = 1.200	t = 2.406**
Title text per day	−0.273	−367,775.200	−1.004	2,808.757	−255.137	0.000
	t = −0.294	t = −0.221	t = −0.295	t = 0.817	t = −0.404	t = 0.659
Positive sentiment title text	0.017	76,211.410	0.079	128.136	17.916	−0.000
	t = 0.642	t = 1.776*	t = 0.800	t = 1.418	t = 1.050	t = −1.130
Mkt.RF	1.048	7,507.914	−0.740	175.533	−33.017	0.00000
	t = 5.149***	t = 0.057	t = −0.856	t = 0.545	t = −0.566	t = 5.553***
SMB	0.767					
	t = 1.969**					
HML	−0.098					
	t = −0.184					
RMW	−1.630					
	t = −2.508**					
CMA	−0.078					
	t = −0.096					
MOM	−0.309					
	t = −0.796					
Constant	−3.567	−1,798,390.000	88.569	−10,404.390	735.515	−0.00000
	t = −1.700*	t = −0.464	t = 9.798***	t = −1.204	t = 0.482	t = −2.158**
Observations	296	296	296	296	296	296
R^2	0.335	0.712	0.341	0.679	0.724	0.405
Adjusted R^2	0.307	0.705	0.325	0.671	0.717	0.390

Note: *p<0.1; **p<0.05; ***p<0.01

A1.4.2 Lagged return

Table A1.17: APHA lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0001	−0.007	0.002	0.002	0.003	0.002
	t = 0.098	t = −0.627	t = 0.226	t = 0.568	t = 0.486	t = 1.235
Positive sentiment comments	−0.0002	0.016	0.022	0.008	0.018	0.0002
	t = −0.020	t = 1.592	t = 1.553	t = 0.627	t = 1.099	t = 0.013
Title per day	−0.051	0.713	−0.576	0.037	0.090	0.127
	t = −0.752	t = 0.659	t = −0.408	t = 0.106	t = 0.161	t = 0.751
Positive sentiment title	0.020	0.030	0.068	0.019	0.026	0.014
	t = 1.526	t = 0.908	t = 1.206	t = 0.683	t = 0.641	t = 0.601
Title text per day	0.104	−0.273	0.741	0.244	−0.344	−0.448
	t = 1.126	t = −0.294	t = 0.640	t = 0.511	t = −0.547	t = −0.973
Positive sentiment title text	−0.007	0.017	0.009	0.004	0.008	0.021
	t = −0.753	t = 0.642	t = 0.286	t = 0.227	t = 0.291	t = 1.053
Mkt.RF	−0.538	1.048	−0.302	0.530	0.172	−0.036
	t = −3.534***	t = 5.149***	t = −1.220	t = 2.126**	t = 0.713	t = −0.183
SMB	0.263	0.767	0.066	1.222	−0.048	0.514
	t = 0.794	t = 1.969**	t = 0.118	t = 2.305**	t = −0.094	t = 0.928
HML	0.160	−0.098	0.742	−1.264	0.152	0.473
	t = 0.418	t = −0.184	t = 1.481	t = −1.822*	t = 0.311	t = 0.603
RMW	0.177	−1.630	0.029	0.889	−0.615	−0.586
	t = 0.319	t = −2.508**	t = 0.035	t = 1.328	t = −0.913	t = −0.681
CMA	−1.023	−0.078	−0.923	−0.398	0.358	1.573
	t = −2.048**	t = −0.096	t = −1.183	t = −0.394	t = 0.407	t = 1.684*
MOM	−0.058	−0.309	0.243	−0.204	0.268	0.259
	t = −0.175	t = −0.796	t = 0.612	t = −0.610	t = 0.855	t = 0.405
Constant	−0.633	−3.567	−5.648	−2.271	−3.184	−1.826
	t = −0.479	t = −1.700*	t = −1.511	t = −1.058	t = −1.182	t = −0.954
Observations	295	296	295	294	293	292
R^2	0.096	0.335	0.128	0.177	0.115	0.111
Adjusted R^2	0.057	0.307	0.091	0.141	0.077	0.073

Note: *p<0.1; **p<0.05; ***p<0.01

A1.4.3 Lagged volume

Table A1.18: APHA lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	9,895.708	12,758.120	11,354.880	947.508	51.076	2,589.196
	t = 0.416	t = 0.655	t = 1.435	t = 0.113	t = 0.004	t = 0.283
Positive sentiment comments	3,355.157	−12.325	−7,890.697	−14,435.390	−21,634.610	−15,685.620
	t = 0.219	t = −0.001	t = −0.469	t = −0.794	t = −0.951	t = −0.640
Title per day	778,321.900	603,247.000	521,328.600	1,460,714.000	838,487.000	45,007.530
	t = 0.306	t = 0.316	t = 0.742	t = 1.506	t = 0.815	t = 0.069
Positive sentiment title	89,735.070	99,029.730	68,557.290	18,646.370	−3,147.259	439.355
	t = 1.078	t = 1.592	t = 1.103	t = 0.235	t = −0.035	t = 0.005
Title text per day	−741,528.000	−367,775.200	53,324.700	−699,663.600	−282,687.200	635,153.600
	t = −0.367	t = −0.221	t = 0.031	t = −0.412	t = −0.229	t = 0.399
Positive sentiment title text	75,014.890	76,211.410	36,275.640	67,043.040	88,592.440	95,227.490
	t = 1.398	t = 1.776*	t = 0.857	t = 1.437	t = 1.672*	t = 1.825*
Mkt.RF	264,044.000	7,507.914	85,287.780	80,956.210	−106,473.100	−198,211.600
	t = 1.434	t = 0.057	t = 0.662	t = 0.530	t = −0.626	t = −1.012
Constant	−988,523.800	−1,798,390.000	2,809,476.000	4,872,252.000	5,927,707.000	4,899,014.000
	t = −0.203	t = −0.464	t = 0.659	t = 0.933	t = 1.068	t = 0.885
Observations	295	296	295	294	293	292
R^2	0.539	0.712	0.586	0.369	0.149	0.112
Adjusted R^2	0.528	0.705	0.576	0.353	0.128	0.090

Note: *p<0.1; **p<0.05; ***p<0.01

A1.4.4 2020 regression

Table A1.19: APHA 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.002	39,931.910	0.064	99.493	14.955	−0.000
	t = −0.205	t = 2.858***	t = 2.088**	t = 3.045***	t = 3.102***	t = −0.134
Positive sentiment comments	0.015	425.518	−0.020	10.048	0.730	0.000
	t = 1.595	t = 0.049	t = −0.445	t = 0.502	t = 0.172	t = 1.491
Title per day	−0.312	−1,059,222.000	−2.406	−2,974.212	−734.207	−0.000
	t = −0.337	t = −0.996	t = −1.332	t = −1.248	t = −2.208**	t = −0.134
Positive sentiment title	0.045	59,355.760	0.080	174.188	27.177	0.000
	t = 1.488	t = 1.634	t = 1.146	t = 2.037**	t = 2.308**	t = 1.994**
Title text per day	0.853	54,687.020	1.268	3,007.485	360.322	0.000
	t = 0.847	t = 0.040	t = 0.597	t = 1.016	t = 0.835	t = 0.388
Positive sentiment title text	−0.021	13,368.610	−0.067	14.149	−7.372	−0.000
	t = −0.925	t = 0.716	t = −1.091	t = 0.329	t = −0.655	t = −1.220
Mkt.RF	0.992	−13,145.050	−0.783	73.796	−45.647	0.00000
	t = 5.008***	t = −0.172	t = −0.835	t = 0.395	t = −1.226	t = 5.504***
SMB	0.386					
	t = 1.025					
HML	−0.406					
	t = −1.032					
RMW	−0.365					
	t = −0.587					
CMA	−0.324					
	t = −0.430					
MOM	−0.498					
	t = −1.471					
Constant	−2.710	2,220,275.000	93.550	−1,083.130	890.121	−0.00000
	t = −1.630	t = 1.177	t = 17.024***	t = −0.218	t = 1.156	t = −2.256**
Observations	252	252	252	252	252	252
R^2	0.255	0.395	0.080	0.526	0.206	0.422
Adjusted R^2	0.218	0.378	0.054	0.512	0.183	0.405

Note: *p<0.1; **p<0.05; ***p<0.01

A1.4.5 2021 regression

Table A1.20: APHA 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.003 t = −0.340	15,817.770 t = 0.879	0.029 t = 0.980	17.943 t = 0.603	5.761 t = 2.404**	−0.000 t = −0.420
Positive sentiment comments	0.054 t = 0.470	−36,420.330 t = −0.243	−0.562 t = −1.029	103.363 t = 0.350	−108.397 t = −0.691	0.000 t = 0.658
Title per day	0.476 t = 0.399	273,073.300 t = 0.140	0.165 t = 0.047	940.859 t = 0.236	441.003 t = 3.198***	0.000 t = 0.203
Positive sentiment title	0.014 t = 0.169	67,853.950 t = 0.503	0.033 t = 0.123	230.433 t = 0.820	4.523 t = 0.080	−0.000 t = −0.781
Title text per day	−1.335 t = −0.702	−1,429,170.000 t = −0.418	−5.536 t = −1.045	1,002.053 t = 0.200	−134.278 t = −0.245	0.000 t = 0.234
Positive sentiment title text	0.064 t = 0.940	118,150.800 t = 0.996	0.191 t = 0.863	59.991 t = 0.265	22.761 t = 0.474	0.000 t = 0.185
Mkt.RF	0.985 t = 0.405	−67,159.810 t = −0.042	−0.120 t = −0.044	492.342 t = 0.111	116.974 t = 0.148	0.000 t = 1.412
SMB	1.498 t = 0.463					
HML	1.943 t = 0.350					
RMW	−4.975 t = −1.596					
CMA	−1.950 t = −0.505					
MOM	−0.047 t = −0.011					
Constant	−5.644 t = −0.547	9,899,808.000 t = 0.626	149.560 t = 3.380***	15,019.540 t = 0.475	13,061.160 t = 1.147	−0.000 t = −0.067
Observations	44	44	44	44	44	44
R^2	0.646	0.758	0.612	0.682	0.787	0.210
Adjusted R^2	0.509	0.711	0.536	0.620	0.745	0.056
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.5 BB - BlackBerry Limited

A1.5.1 Full regression

Table A1.21: BB full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.002 t = −0.406	−51.242 t = −0.004	0.002 t = 1.036	10.696 t = 0.284	3.730 t = 0.999	−0.000 t = −1.137
Positive sentiment comments	0.017 t = 1.240	14,438.460 t = 0.533	−0.205 t = −2.992***	114.440 t = 1.430	−76.270 t = −1.638	0.000 t = 0.041
Title per day	0.157 t = 0.480	534,523.300 t = 0.631	0.531 t = 5.977***	1,291.750 t = 0.509	312.374 t = 1.298	0.000 t = 1.379
Positive sentiment title	0.021 t = 0.918	450,259.500 t = 2.749***	0.592 t = 7.093***	1,362.903 t = 4.570***	422.746 t = 3.797***	−0.000 t = −0.003
Title text per day	−0.272 t = −0.496	−794,517.400 t = −0.541	−0.862 t = −6.057***	−2,309.961 t = −0.508	−496.713 t = −1.270	−0.000 t = −1.170
Positive sentiment title text	0.027 t = 1.929*	76,105.330 t = 1.735*	0.087 t = 1.536	327.998 t = 2.488**	34.513 t = 0.970	0.000 t = 1.392
Mkt.RF	1.247 t = 4.946***	89,945.010 t = 0.240	−0.723 t = −0.691	−297.533 t = −0.329	−550.902 t = −1.517	0.00000 t = 7.543***
SMB	1.119 t = 1.464					
HML	−1.165 t = −1.598					
RMW	−0.129 t = −0.175					
CMA	2.440 t = 0.871					
MOM	−0.349 t = −0.632					
Constant	−3.926 t = −2.191**	−21,399,633.000 t = −2.165**	46.103 t = 6.831***	−80,031.640 t = −4.311***	−15,021.050 t = −3.671***	−0.00000 t = −0.775
Observations	296	296	296	296	296	296
R^2	0.528	0.657	0.720	0.787	0.757	0.323
Adjusted R^2	0.508	0.649	0.713	0.782	0.751	0.306

Note: *p<0.1; **p<0.05; ***p<0.01

A1.5.2 Lagged return

Table A1.22: BB lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.003	−0.002	0.0002	0.001	0.001	−0.001
	t = 0.678	t = −0.406	t = 0.071	t = 1.174	t = 1.165	t = −0.336
Positive sentiment comments	−0.004	0.017	0.026	−0.004	0.011	0.011
	t = −0.297	t = 1.240	t = 1.603	t = −0.282	t = 0.716	t = 0.690
Title per day	−0.119	0.157	−0.084	0.006	−0.003	0.070
	t = −0.378	t = 0.480	t = −0.516	t = 0.100	t = −0.040	t = 0.639
Positive sentiment title	0.005	0.021	0.029	−0.005	−0.004	0.015
	t = 0.181	t = 0.918	t = 0.962	t = −0.169	t = −0.165	t = 0.589
Title text per day	0.130	−0.272	0.213	−0.043	−0.025	−0.113
	t = 0.214	t = −0.496	t = 0.819	t = −0.289	t = −0.170	t = −0.712
Positive sentiment title text	−0.023	0.027	0.006	−0.024	0.017	−0.003
	t = −1.370	t = 1.929*	t = 0.423	t = −1.581	t = 0.863	t = −0.124
Mkt.RF	−0.092	1.247	−0.431	0.168	0.341	0.156
	t = −0.270	t = 4.946***	t = −1.532	t = 0.675	t = 1.406	t = 0.584
SMB	0.572	1.119	0.645	−0.245	−0.130	0.193
	t = 0.646	t = 1.464	t = 1.230	t = −0.320	t = −0.147	t = 0.228
HML	−0.308	−1.165	0.384	1.026	−1.583	0.166
	t = −0.478	t = −1.598	t = 0.544	t = 1.146	t = −2.450**	t = 0.241
RMW	−0.853	−0.129	1.199	−0.874	0.338	−0.881
	t = −1.198	t = −0.175	t = 1.100	t = −0.792	t = 0.321	t = −0.897
CMA	1.371	2.440	−2.766	1.400	4.028	−1.300
	t = 0.493	t = 0.871	t = −1.995**	t = 0.729	t = 1.318	t = −0.463
MOM	0.041	−0.349	0.195	0.605	−0.527	−0.057
	t = 0.112	t = −0.632	t = 0.407	t = 0.917	t = −1.054	t = −0.112
Constant	0.996	−3.926	−3.796	1.591	−1.903	−1.676
	t = 0.580	t = −2.191**	t = −1.553	t = 0.663	t = −1.231	t = −0.803
Observations	295	296	295	294	293	292
R^2	0.371	0.528	0.334	0.116	0.135	0.095
Adjusted R^2	0.344	0.508	0.306	0.078	0.098	0.057

Note: *p<0.1; **p<0.05; ***p<0.01

A1.5.3 Lagged volume

Table A1.23: BB lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	6,489.838	−51.242	9,537.154	9,905.820	3,374.564	1,915.569
	t = 0.228	t = −0.004	t = 0.738	t = 0.500	t = 0.192	t = 0.141
Positive sentiment comments	85,037.410	14,438.460	8,425.309	17,634.250	42,628.910	28,171.080
	t = 2.138**	t = 0.533	t = 0.376	t = 0.573	t = 1.042	t = 0.645
Title per day	251,691.300	534,523.300	−48,550.890	324,638.400	442,635.100	173,976.400
	t = 0.127	t = 0.631	t = −0.058	t = 0.267	t = 0.402	t = 0.210
Positive sentiment title	384,657.700	450,259.500	393,643.800	323,683.200	317,437.300	592,183.900
	t = 2.570**	t = 2.749***	t = 2.317**	t = 1.674*	t = 2.492**	t = 3.372***
Title text per day	−553,708.600	−794,517.400	125,171.300	−865,197.300	−926,175.500	−289,084.700
	t = −0.148	t = −0.541	t = 0.087	t = −0.461	t = −0.544	t = −0.257
Positive sentiment title text	6,948.354	76,105.330	60,443.260	87,803.180	139,832.600	102,285.900
	t = 0.111	t = 1.735*	t = 1.592	t = 1.365	t = 1.151	t = 1.133
Mkt.RF	−315,688.200	89,945.010	69,386.700	151,003.400	627,130.900	397,333.800
	t = −0.586	t = 0.240	t = 0.242	t = 0.332	t = 0.891	t = 0.693
Constant	−18,302,459.000	−21,399,633.000	−17,713,512.000	−15,707,888.000	−18,309,020.000	−30,224,023.000
	t = −2.025**	t = −2.165**	t = −1.963**	t = −1.359	t = −2.341**	t = −2.268**
Observations	295	296	295	294	293	292
R^2	0.381	0.657	0.626	0.500	0.326	0.263
Adjusted R^2	0.366	0.649	0.617	0.487	0.309	0.245

Note: *p<0.1; **p<0.05; ***p<0.01

A1.5.4 2020 regression

Table A1.24: BB 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.026	16,123.100	0.105	102.625	37.787	−0.000
	t = −2.473**	t = 0.360	t = 1.618	t = 2.475**	t = 2.454**	t = −1.209
Positive sentiment comments	0.010	−18,800.270	−0.188	53.248	−7.033	0.000
	t = 0.894	t = −0.903	t = −3.107***	t = 1.417	t = −0.559	t = 0.048
Title per day	1.443	8,235,735.000	−1.999	7,859.602	1,474.350	0.00000
	t = 3.553***	t = 1.734*	t = −1.373	t = 1.787*	t = 2.098**	t = 1.503
Positive sentiment title	−0.026	−272,040.600	0.178	−201.320	−21.292	0.000
	t = −1.296	t = −1.496	t = 1.727*	t = −1.119	t = −0.630	t = 0.104
Title text per day	0.199	384,365.500	0.247	2,037.746	−213.363	−0.000
	t = 0.416	t = 0.111	t = 0.083	t = 0.542	t = −0.237	t = −0.387
Positive sentiment title text	0.014	26,472.030	0.064	84.532	−5.637	0.000
	t = 0.888	t = 0.681	t = 1.079	t = 1.048	t = −0.348	t = 0.805
Mkt.RF	1.195	−67,203.930	−0.924	−232.705	−325.946	0.00000
	t = 5.993***	t = −0.324	t = −0.960	t = −0.633	t = −1.879*	t = 7.596***
SMB	0.145					
	t = 0.308					
HML	−0.630					
	t = −1.009					
RMW	0.025					
	t = 0.060					
CMA	−1.132					
	t = −1.664*					
MOM	−0.635					
	t = −1.233					
Constant	0.346	16,780,851.000	61.574	6,870.266	2,382.238	−0.00000
	t = 0.244	t = 1.771*	t = 8.992***	t = 0.709	t = 1.224	t = −0.370
Observations	252	252	252	252	252	252
R^2	0.527	0.834	0.310	0.831	0.808	0.356
Adjusted R^2	0.503	0.829	0.290	0.827	0.803	0.337

Note: *p<0.1; **p<0.05; ***p<0.01

A1.5.5 2021 regression

Table A1.25: BB 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.001	−663.501	0.001	6.603	3.950	−0.000
	t = −0.367	t = −0.050	t = 0.276	t = 0.184	t = 0.752	t = −0.304
Positive sentiment comments	0.002	145,654.700	−1.014	1,385.802	−1,981.968	0.000
	t = 0.009	t = 0.426	t = −2.917***	t = 0.795	t = −2.520**	t = 1.485
Title per day	0.095	541,612.200	0.569	1,160.213	347.261	0.000
	t = 0.466	t = 0.651	t = 4.005***	t = 0.496	t = 1.021	t = 0.175
Positive sentiment title	0.161	287,642.700	0.274	1,167.563	487.094	−0.000
	t = 1.993**	t = 2.002**	t = 1.109	t = 2.057**	t = 1.902*	t = −0.351
Title text per day	−0.164	−792,567.800	−0.904	−1,980.207	−557.853	−0.000
	t = −0.517	t = −0.563	t = −3.974***	t = −0.473	t = −0.960	t = −0.192
Positive sentiment title text	0.016	153,230.600	0.172	852.713	107.953	0.000
	t = 0.331	t = 1.647*	t = 1.345	t = 2.931***	t = 0.989	t = 0.796
Mkt.RF	0.225	1,909,425.000	5.088	−948.450	−2,301.095	0.000
	t = 0.084	t = 0.303	t = 1.344	t = −0.053	t = −0.359	t = 2.066**
SMB	3.045					
	t = 1.188					
HML	−4.510					
	t = −1.350					
RMW	−1.308					
	t = −0.260					
CMA	11.707					
	t = 1.612					
MOM	−0.439					
	t = −0.133					
Constant	−13.920	−23,480,874.000	136.516	−161,227.100	116,968.300	−0.00000
	t = −0.890	t = −0.683	t = 3.865***	t = −1.069	t = 2.362**	t = −1.461
Observations	44	44	44	44	44	44
R^2	0.824	0.906	0.887	0.805	0.791	0.255
Adjusted R^2	0.756	0.888	0.865	0.767	0.751	0.111

Note: *p<0.1; **p<0.05; ***p<0.01

A1.6 GME - GameStop Corp.

A1.6.1 Lagged volume

Table A1.26: GME lagged volume

	Lagged volume					
	$Volume_{t-1}$	Volume	$Volume_{t+1}$	$Volume_{t+2}$	$Volume_{t+3}$	$Volume_{t+4}$
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	251.983	307.436	194.127	294.892	491.151	109.839
	t = 0.959	t = 0.942	t = 0.875	t = 2.257**	t = 3.434***	t = 0.748
Positive sentiment comments	53,611.540	39,319.100	38,884.890	20,684.680	14,999.150	19,053.600
	t = 2.113**	t = 1.665*	t = 1.368	t = 0.795	t = 0.624	t = 0.756
Title per day	3,252.855	8,121.231	4,475.316	−7,116.890	−9,764.814	25,874.460
	t = 0.233	t = 0.426	t = 0.306	t = −0.675	t = −0.614	t = 1.273
Positive sentiment title	80,397.480	87,650.970	86,409.260	82,274.810	36,638.520	32,114.930
	t = 2.578***	t = 3.756***	t = 3.185***	t = 3.132***	t = 0.984	t = 0.784
Title text per day	−14,133.530	−20,545.330	18,966.710	36,125.880	−6,066.443	−46,324.980
	t = −0.649	t = −0.515	t = 0.336	t = 1.472	t = −0.180	t = −0.934
Positive sentiment title text	113,869.000	104,942.400	89,420.590	76,314.420	82,488.140	92,798.010
	t = 3.137***	t = 3.149***	t = 2.673***	t = 2.880***	t = 2.518**	t = 2.575**
Mkt.RF	58,524.880	−423,935.400	−512,258.500	−183,832.000	−281,970.800	−61,589.700
	t = 0.175	t = −1.099	t = −1.016	t = −0.508	t = −0.706	t = −0.125
Constant	−5,294,596.000	−5,258,571.000	−4,198,999.000	−2,411,093.000	379,598.800	758,848.500
	t = −1.735*	t = −2.092**	t = −1.303	t = −0.834	t = 0.127	t = 0.274
Observations	295	296	295	294	293	292
R^2	0.228	0.396	0.427	0.473	0.426	0.268
Adjusted R^2	0.210	0.382	0.413	0.460	0.412	0.250

Note: *p<0.1; **p<0.05; ***p<0.01

A1.6.2 2020 regression

Table A1.27: GME 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.002	2,571.117	−0.004	20.562	14.088	−0.000
	t = −1.453	t = 0.733	t = −1.684*	t = 1.093	t = 0.971	t = −0.066
Positive sentiment comments	0.021	1,222.543	−0.206	23.924	−27.083	0.000
	t = 1.292	t = 0.134	t = −1.515	t = 0.559	t = −0.925	t = 1.281
Title per day	0.424	150,082.500	−0.040	1,437.559	1,270.339	−0.000
	t = 2.486**	t = 0.440	t = −0.086	t = 0.673	t = 0.804	t = −0.758
Positive sentiment title	0.043	40,006.400	0.188	223.178	138.319	0.000
	t = 2.450**	t = 2.840***	t = 2.114**	t = 3.430***	t = 2.748***	t = 0.936
Title text per day	−0.513	−237,888.400	1.911	−2,301.957	−2,299.988	0.000
	t = −1.896*	t = −0.643	t = 2.690***	t = −0.861	t = −1.155	t = 0.970
Positive sentiment title text	0.014	22,332.830	0.031	57.221	34.817	0.000
	t = 0.697	t = 1.434	t = 0.321	t = 0.992	t = 0.833	t = 0.038
Mkt.RF	1.018	−44,419.240	0.162	122.748	−378.803	0.00000
	t = 3.967***	t = −0.442	t = 0.186	t = 0.265	t = −0.986	t = 4.715***
SMB	0.987					
	t = 1.799*					
HML	−0.110					
	t = −0.157					
RMW	−1.100					
	t = −1.041					
CMA	0.651					
	t = 0.576					
MOM	−0.259					
	t = −0.483					
Constant	−4.303	1,431,626.000	119.052	746.207	2,068.108	−0.00000
	t = −2.428**	t = 1.193	t = 9.635***	t = 0.168	t = 0.513	t = −1.621
Observations	252	252	252	252	252	252
R^2	0.253	0.273	0.140	0.588	0.492	0.172
Adjusted R^2	0.215	0.252	0.115	0.576	0.478	0.149

Note: *p<0.1; **p<0.05; ***p<0.01

A1.6.3 2021 regression

Table A1.28: GME 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
	0.00004	176.279	0.0004	1.162	0.372	0.000
Comments per day	t = 0.185	t = 0.510	t = 1.982**	t = 0.752	t = 0.291	t = 0.585
	1.146	3,207,892.000	2.628	13,257.550	7,957.519	0.000
Positive sentiment comments	t = 0.919	t = 2.560**	t = 1.087	t = 2.464**	t = 1.211	t = 3.070***
	−0.002	5,869.403	0.033	−29.646	99.132	−0.000
Title per day	t = −0.070	t = 0.275	t = 1.756*	t = −0.301	t = 0.768	t = −2.355**
	−0.114	−1,818,415.000	−0.423	−9,203.418	−6,527.222	−0.000
Positive sentiment title	t = −0.192	t = −1.751*	t = −0.198	t = −2.343**	t = −1.462	t = −0.002
	0.013	−20,801.310	−0.016	−37.376	−31.798	0.000
Title text per day	t = 0.259	t = −0.636	t = −0.259	t = −0.264	t = −0.101	t = 1.601
	0.102	48,469.460	−1.205	−380.378	−784.365	−0.000
Positive sentiment title text	t = 0.218	t = 0.090	t = −0.866	t = −0.178	t = −0.253	t = −1.661*
	−2.134	−3,217,191.000	−0.474	−11,901.130	−74,148.580	0.000
Mkt.RF	t = −0.305	t = −0.502	t = −0.036	t = −0.506	t = −1.557	t = 0.634
	−6.591					
SMB	t = −1.124					
	7.304					
HML	t = 0.829					
	−2.590					
RMW	t = −0.276					
	42.835					
CMA	t = 1.386					
	12.477					
MOM	t = 1.071					
	−72.417	−37,029,657.000	156.495	70,559.200	247,726.400	−0.00000
Constant	t = −0.987	t = −0.490	t = 1.200	t = 0.213	t = 0.716	t = −0.798
Observations	44	44	44	44	44	44
R^2	0.738	0.324	0.580	0.305	0.533	0.178
Adjusted R^2	0.637	0.192	0.498	0.170	0.442	0.018

Note: *p<0.1; **p<0.05; ***p<0.01

A1.7 GOOG - Alphabet, Inc.

A1.7.1 Full regression

Table A1.29: GOOG full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002 t = 0.350	4,032.531 t = 1.237	0.009 t = 0.716	41.770 t = 1.035	31.409 t = 0.986	−0.000 t = −0.134
Positive sentiment comments	0.002 t = 0.897	−247.460 t = −0.157	0.011 t = 0.623	2.236 t = 0.185	−4.955 t = −0.561	0.000 t = 0.116
Title per day	−0.062 t = −0.217	89,857.620 t = 0.691	0.258 t = 0.212	3,825.190 t = 1.879*	2,253.060 t = 1.566	−0.000 t = −0.614
Positive sentiment title	0.013 t = 1.255	3,926.276 t = 0.967	0.031 t = 0.581	−3.053 t = −0.069	−0.702 t = −0.019	0.000 t = 1.760*
Title text per day	0.015 t = 0.077	−46,545.890 t = −0.449	0.302 t = 0.377	−223.627 t = −0.157	−153.564 t = −0.157	0.000 t = 0.909
Positive sentiment title text	0.003 t = 0.755	2,183.641 t = 1.051	−0.0001 t = −0.007	33.722 t = 1.564	20.563 t = 1.253	−0.000 t = −0.318
Mkt.RF	1.004 t = 23.675***	−57,319.480 t = −1.283	−0.937 t = −1.620	16.915 t = 0.123	−213.141 t = −2.360**	0.000 t = 15.494***
SMB	−0.126 t = −1.316					
HML	−0.362 t = −3.239***					
RMW	0.497 t = 3.116***					
CMA	−0.718 t = −3.474***					
MOM	−0.146 t = −2.316**					
Constant	−0.469 t = −1.691*	1,660,559.000 t = 11.110***	31.356 t = 20.985***	11,125.530 t = 7.301***	8,446.505 t = 7.133***	0.000 t = 0.404
Observations	296	296	296	296	296	296
R^2	0.757	0.088	0.061	0.132	0.132	0.642
Adjusted R^2	0.747	0.066	0.038	0.111	0.111	0.633
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.7.2 Lagged return

Table A1.30: GOOG lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.006	0.002	0.0005	−0.002	0.005	0.005
	t = 0.996	t = 0.350	t = 0.112	t = −0.537	t = 1.919*	t = 1.550
Positive sentiment comments	0.003	0.002	0.003	0.003	0.001	−0.004
	t = 0.450	t = 0.897	t = 0.497	t = 0.440	t = 0.106	t = −0.782
Title per day	0.148	−0.062	−0.004	−0.392	0.196	−0.338
	t = 0.349	t = −0.217	t = −0.013	t = −1.093	t = 0.625	t = −0.791
Positive sentiment title	−0.001	0.013	−0.001	0.015	−0.019	0.016
	t = −0.117	t = 1.255	t = −0.070	t = 1.714*	t = −1.336	t = 1.614
Title text per day	−0.027	0.015	0.532	−0.028	0.336	−0.059
	t = −0.078	t = 0.077	t = 2.392**	t = −0.136	t = 1.881*	t = −0.271
Positive sentiment title text	−0.001	0.003	−0.005	−0.002	−0.007	0.002
	t = −0.247	t = 0.755	t = −0.736	t = −0.352	t = −1.372	t = 0.302
Mkt.RF	−0.365	1.004	−0.413	0.236	0.083	−0.097
	t = −3.063***	t = 23.675***	t = −2.799***	t = 1.525	t = 0.605	t = −0.784
SMB	0.285	−0.126	−0.152	0.442	−0.433	0.328
	t = 1.017	t = −1.316	t = −0.609	t = 1.661*	t = −1.393	t = 1.202
HML	−0.185	−0.362	0.363	0.062	−0.145	0.150
	t = −0.792	t = −3.239***	t = 1.437	t = 0.156	t = −0.447	t = 0.433
RMW	0.408	0.497	−0.186	0.179	−1.031	0.253
	t = 1.247	t = 3.116***	t = −0.557	t = 0.543	t = −3.484***	t = 0.896
CMA	−0.407	−0.718	−0.226	0.081	1.035	−0.301
	t = −1.003	t = −3.474***	t = −0.521	t = 0.212	t = 2.192**	t = −0.551
MOM	−0.089	−0.146	0.050	0.208	−0.306	0.297
	t = −0.674	t = −2.316**	t = 0.417	t = 0.556	t = −1.321	t = 1.096
Constant	−0.215	−0.469	0.0004	0.008	0.167	0.096
	t = −0.463	t = −1.691*	t = 0.001	t = 0.015	t = 0.366	t = 0.205
Observations	295	296	295	294	293	292
R^2	0.119	0.757	0.126	0.093	0.100	0.050
Adjusted R^2	0.081	0.747	0.089	0.055	0.061	0.010

Note: *p<0.1; **p<0.05; ***p<0.01

A1.7.3 Lagged volume

Table A1.31: GOOG lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	2,091.944	4,032.531	963.638	−575.682	−892.699	−1,659.220
	t = 0.872	t = 1.237	t = 0.657	t = −0.785	t = −1.044	t = −2.068**
Positive sentiment comments	−499.188	−247.460	−435.499	1,234.654	542.287	1,022.987
	t = −0.296	t = −0.157	t = −0.251	t = 0.742	t = 0.338	t = 0.603
Title per day	209,008.300	89,857.620	29,195.760	35,109.290	76,890.430	67,633.200
	t = 1.207	t = 0.691	t = 0.291	t = 0.305	t = 0.639	t = 0.464
Positive sentiment title	−2,058.907	3,926.276	2,456.570	−293.068	750.043	764.781
	t = −0.476	t = 0.967	t = 0.681	t = −0.080	t = 0.181	t = 0.231
Title text per day	19,952.420	−46,545.890	−43,928.170	−41,007.940	−71,596.150	−5,643.442
	t = 0.178	t = −0.449	t = −0.610	t = −0.688	t = −1.147	t = −0.097
Positive sentiment title text	561.545	2,183.641	−869.965	3,161.244	3,138.593	1,789.417
	t = 0.274	t = 1.051	t = −0.391	t = 1.700*	t = 1.696*	t = 0.919
Mkt.RF	−29,455.460	−57,319.480	30,396.550	−28,667.560	1,395.058	4,827.425
	t = −0.640	t = −1.283	t = 0.687	t = −0.740	t = 0.041	t = 0.141
Constant	1,780,321.000	1,660,559.000	1,905,202.000	1,628,579.000	1,677,913.000	1,718,122.000
	t = 12.025***	t = 11.110***	t = 12.119***	t = 12.402***	t = 12.309***	t = 12.114***
Observations	295	296	295	294	293	292
R^2	0.037	0.088	0.015	0.018	0.014	0.012
Adjusted R^2	0.013	0.066	−0.009	−0.006	−0.010	−0.012

Note: *p<0.1; **p<0.05; ***p<0.01

A1.7.4 2020 regression

Table A1.32: GOOG 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.001 t = 0.128	3,290.834 t = 1.089	0.006 t = 0.408	32.450 t = 0.900	25.418 t = 0.864	−0.000 t = −0.353
Positive sentiment comments	−0.0001 t = −0.054	547.160 t = 0.316	0.022 t = 1.036	0.537 t = 0.042	−9.139 t = −0.953	−0.000 t = −0.504
Title per day	0.047 t = 0.166	161,179.900 t = 1.058	0.548 t = 0.406	4,792.695 t = 1.923*	2,132.338 t = 1.385	−0.000 t = −0.601
Positive sentiment title	0.006 t = 1.034	964.099 t = 0.233	0.035 t = 0.570	−40.706 t = −1.080	−34.279 t = −1.344	0.000 t = 1.459
Title text per day	0.115 t = 0.659	−45,592.310 t = −0.448	0.394 t = 0.467	118.331 t = 0.084	−190.693 t = −0.205	0.000 t = 1.083
Positive sentiment title text	−0.0003 t = −0.096	2,422.570 t = 1.095	0.005 t = 0.204	32.076 t = 1.403	18.522 t = 1.126	−0.000 t = −0.508
Mkt.RF	0.996 t = 24.132***	−58,511.660 t = −1.296	−0.946 t = −1.596	−43.840 t = −0.313	−241.676 t = −2.637***	0.000 t = 15.182***
SMB	−0.189 t = −1.945*					
HML	−0.379 t = −3.196***					
RMW	0.627 t = 3.857***					
CMA	−1.036 t = −4.118***					
MOM	−0.191 t = −2.936***					
Constant	−0.284 t = −1.105	1,650,893.000 t = 10.307***	30.827 t = 18.062***	11,239.860 t = 7.152***	8,780.029 t = 7.238***	0.000 t = 0.952
Observations	252	252	252	252	252	252
R^2	0.804	0.074	0.067	0.133	0.107	0.659
Adjusted R^2	0.794	0.047	0.040	0.108	0.081	0.649

Note: *p<0.1; **p<0.05; ***p<0.01

A1.7.5 2021 regression

Table A1.33: GOOG 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.009 t = 0.385	14,009.020 t = 0.899	0.188 t = 3.130***	210.272 t = 1.142	103.081 t = 0.777	0.000 t = 1.412
Positive sentiment comments	0.020 t = 1.822*	−3,944.010 t = −0.917	−0.034 t = −1.188	24.368 t = 0.478	−5.595 t = −0.171	0.000 t = 1.712*
Title per day	0.960 t = 1.670*	−131,174.500 t = −0.231	0.034 t = 0.025	1,282.716 t = 0.168	3,205.031 t = 0.974	0.000 t = 0.562
Positive sentiment title	0.051 t = 0.617	12,633.340 t = 0.611	−0.183 t = −1.326	73.326 t = 0.313	147.022 t = 0.419	−0.000 t = −0.156
Title text per day	−2.795 t = −1.593	−247,281.900 t = −0.515	−1.940 t = −0.680	−6,407.783 t = −1.100	−508.604 t = −0.072	−0.000 t = −1.227
Positive sentiment title text	0.032 t = 0.991	701.108 t = 0.088	−0.009 t = −0.172	23.489 t = 0.221	−28.544 t = −0.219	0.000 t = 0.383
Mkt.RF	1.450 t = 5.416***	−45,884.730 t = −0.426	−0.793 t = −1.394	979.538 t = 0.743	213.134 t = 0.266	0.000 t = 6.399***
SMB	−0.474 t = −1.310					
HML	−0.834 t = −1.897*					
RMW	0.064 t = 0.133					
CMA	0.450 t = 0.729					
MOM	−0.573 t = −1.240					
Constant	−2.334 t = −1.335	1,678,483.000 t = 3.252***	31.381 t = 9.932***	9,989.198 t = 1.511	11,270.000 t = 1.643	−0.000 t = −1.134
Observations	44	44	44	44	44	44
R^2	0.771	0.539	0.469	0.464	0.717	0.613
Adjusted R^2	0.683	0.449	0.366	0.360	0.662	0.538

Note: *p<0.1; **p<0.05; ***p<0.01

A1.8 NIO - NIO, Inc.

A1.8.1 Full regression

Table A1.34: NIO full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002 t = 0.990	78,173.140 t = 3.350***	0.009 t = 1.786*	154.948 t = 2.196**	36.046 t = 0.933	0.000 t = 0.075
Positive sentiment comments	0.044 t = 2.179**	60,644.260 t = 0.419	−0.061 t = −1.497	772.373 t = 1.620	−206.089 t = −0.915	0.000 t = 0.913
Title per day	−0.061 t = −0.430	−1,398,177.000 t = −0.764	−0.546 t = −1.287	2,842.713 t = 0.572	5,039.727 t = 1.573	0.000 t = 0.005
Positive sentiment title	0.035 t = 2.358**	421,554.300 t = 2.616***	−0.021 t = −0.358	1,969.295 t = 3.938***	523.587 t = 3.118***	0.000 t = 0.715
Title text per day	−0.067 t = −0.817	900,337.600 t = 0.546	0.308 t = 1.459	828.073 t = 0.136	475.665 t = 0.173	−0.000 t = −0.720
Positive sentiment title text	0.004 t = 0.295	129,718.100 t = 1.008	−0.079 t = −1.770*	600.809 t = 1.618	15.467 t = 0.107	0.000 t = 0.371
Mkt.RF	0.889 t = 5.068***	72,544.570 t = 0.056	−0.906 t = −1.515	1,631.352 t = 0.378	−650.152 t = −0.386	0.000 t = 4.907***
SMB	0.461 t = 1.017					
HML	0.106 t = 0.183					
RMW	−0.436 t = −0.627					
CMA	−1.414 t = −1.651*					
MOM	0.221 t = 0.474					
Constant	−4.673 t = −2.844***	45,684,629.000 t = 3.504***	125.073 t = 28.679***	15,184.430 t = 0.354	64,606.140 t = 3.239***	−0.000 t = −1.110
Observations	296	296	296	296	296	296
R^2	0.185	0.495	0.056	0.525	0.695	0.147
Adjusted R^2	0.151	0.483	0.033	0.513	0.688	0.127
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.8.2 Lagged return

Table A1.35: NIO lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.001	0.002	0.001	0.001	0.002	0.002
	t = −0.434	t = 0.990	t = 0.364	t = 0.499	t = 1.622	t = 1.343
Positive sentiment comments	−0.004	0.044	0.033	0.017	0.023	0.043
	t = −0.153	t = 2.179**	t = 1.795*	t = 0.896	t = 0.988	t = 2.380**
Title per day	0.142	−0.061	−0.004	−0.037	−0.179	−0.090
	t = 1.453	t = −0.430	t = −0.034	t = −0.314	t = −1.568	t = −1.034
Positive sentiment title	−0.002	0.035	0.005	0.018	−0.001	0.030
	t = −0.101	t = 2.358**	t = 0.318	t = 1.159	t = −0.058	t = 1.862*
Title text per day	−0.209	−0.067	0.032	0.027	0.111	0.086
	t = −2.507**	t = −0.817	t = 0.300	t = 0.221	t = 1.127	t = 0.848
Positive sentiment title text	0.029	0.004	0.025	0.022	0.012	0.001
	t = 2.041**	t = 0.295	t = 1.797*	t = 1.491	t = 0.853	t = 0.067
Mkt.RF	0.030	0.889	−0.046	0.123	0.153	−0.109
	t = 0.148	t = 5.068***	t = −0.261	t = 0.583	t = 0.796	t = −0.461
SMB	0.021	0.461	−0.460	−0.351	−1.120	−0.064
	t = 0.049	t = 1.017	t = −0.772	t = −0.846	t = −2.520**	t = −0.146
HML	0.307	0.106	0.701	0.388	0.250	−0.225
	t = 0.567	t = 0.183	t = 1.346	t = 0.778	t = 0.448	t = −0.365
RMW	0.135	−0.436	0.696	−0.294	−2.971	0.067
	t = 0.170	t = −0.627	t = 0.939	t = −0.415	t = −4.056***	t = 0.091
CMA	−1.291	−1.414	−0.505	−0.372	0.778	0.379
	t = −1.469	t = −1.651*	t = −0.507	t = −0.428	t = 1.028	t = 0.453
MOM	0.142	0.221	0.061	0.173	−0.464	0.079
	t = 0.376	t = 0.474	t = 0.173	t = 0.404	t = −1.194	t = 0.156
Constant	−0.697	−4.673	−3.617	−3.126	−1.760	−4.422
	t = −0.391	t = −2.844***	t = −2.349**	t = −2.155**	t = −0.942	t = −3.080***
Observations	295	296	295	294	293	292
R^2	0.041	0.185	0.074	0.043	0.091	0.063
Adjusted R^2	−0.0001	0.151	0.035	0.002	0.052	0.023

Note: *p<0.1; **p<0.05; ***p<0.01

A1.8.3 Lagged volume

Table A1.36: NIO lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	89,911.600	78,173.140	78,339.550	91,366.890	48,480.910	60,674.690
	t = 3.074***	t = 3.350***	t = 2.192**	t = 2.502**	t = 1.645	t = 2.606***
Positive sentiment comments	128,965.000	60,644.260	−216,356.100	−38,848.180	17,323.490	222,623.600
	t = 0.715	t = 0.419	t = −1.094	t = −0.216	t = 0.114	t = 1.267
Title per day	−2,873,136.000	−1,398,177.000	−2,208,892.000	−4,725,159.000	−1,631,736.000	−3,622,927.000
	t = −1.005	t = −0.764	t = −0.705	t = −1.414	t = −0.495	t = −1.449
Positive sentiment title	556,679.400	421,554.300	388,864.000	382,153.900	641,739.900	249,912.000
	t = 2.868***	t = 2.616***	t = 1.920*	t = 1.590	t = 2.712***	t = 0.922
Title text per day	800,598.600	900,337.600	1,462,105.000	3,494,597.000	2,247,877.000	4,480,116.000
	t = 0.560	t = 0.546	t = 1.035	t = 3.173***	t = 0.953	t = 2.917***
Positive sentiment title text	213,992.700	129,718.100	149,150.000	222,559.300	176,496.800	202,875.700
	t = 1.305	t = 1.008	t = 1.041	t = 1.404	t = 1.198	t = 1.469
Mkt.RF	164,059.400	72,544.570	1,501,964.000	−223,701.700	−1,091,377.000	1,092,712.000
	t = 0.108	t = 0.056	t = 1.050	t = −0.153	t = −0.644	t = 0.769
Constant	31,938,458.000	45,684,629.000	67,180,219.000	53,703,251.000	40,255,326.000	47,443,237.000
	t = 2.194**	t = 3.504***	t = 4.661***	t = 3.489***	t = 2.729***	t = 2.570**
Observations	295	296	295	294	293	292
R^2	0.387	0.495	0.392	0.308	0.263	0.235
Adjusted R^2	0.372	0.483	0.377	0.291	0.245	0.217

Note: *p<0.1; **p<0.05; ***p<0.01

A1.8.4 2020 regression

Table A1.37: NIO 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002 t = 0.625	83,697.470 t = 1.977**	0.011 t = 1.215	112.939 t = 1.055	−1.505 t = −0.026	0.000 t = 0.903
Positive sentiment comments	0.042 t = 2.024**	110,294.500 t = 0.809	−0.088 t = −2.170**	854.327 t = 2.295**	37.768 t = 0.263	0.000 t = 0.893
Title per day	−0.093 t = −0.385	−1,936,760.000 t = −0.681	−0.474 t = −0.684	2,869.058 t = 0.416	6,533.180 t = 1.460	−0.000 t = −0.831
Positive sentiment title	0.035 t = 1.917*	480,808.800 t = 2.348**	0.015 t = 0.234	1,940.819 t = 3.274***	413.282 t = 2.482**	0.000 t = 0.209
Title text per day	−0.054 t = −0.589	1,196,856.000 t = 0.642	−0.065 t = −0.275	5,679.628 t = 0.902	2,361.716 t = 0.779	−0.000 t = −0.431
Positive sentiment title text	0.002 t = 0.150	212,616.700 t = 1.548	−0.104 t = −2.347**	977.326 t = 3.051***	203.570 t = 1.775*	0.000 t = 0.477
Mkt.RF	0.752 t = 4.360***	−61,248.950 t = −0.049	−0.900 t = −1.657*	1,274.537 t = 0.394	−1,184.055 t = −1.099	0.000 t = 4.859***
SMB	0.298 t = 0.545					
HML	0.410 t = 0.640					
RMW	0.573 t = 0.733					
CMA	−2.090 t = −1.883*					
MOM	0.277 t = 0.542					
Constant	−4.218 t = −2.393**	32,387,276.000 t = 2.425**	129.811 t = 29.357***	−50,061.650 t = −1.431	29,118.680 t = 2.112**	−0.000 t = −1.049
Observations	252	252	252	252	252	252
R^2	0.164	0.512	0.076	0.609	0.800	0.153
Adjusted R^2	0.122	0.498	0.049	0.598	0.794	0.129

Note: *p<0.1; **p<0.05; ***p<0.01

A1.8.5 2021 regression

Table A1.38: NIO 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002 t = 0.810	59,119.570 t = 2.558**	0.003 t = 0.683	218.575 t = 2.613***	59.186 t = 1.611	−0.000 t = −0.379
Positive sentiment comments	0.034 t = 0.443	−874,907.700 t = −1.480	0.089 t = 0.450	1,341.879 t = 0.610	−1,956.317 t = −1.010	0.000 t = 2.122**
Title per day	0.181 t = 0.656	2,397,483.000 t = 0.877	−0.129 t = −0.303	5,473.699 t = 0.614	3,986.016 t = 0.885	0.000 t = 1.273
Positive sentiment title	0.036 t = 1.453	206,650.700 t = 0.797	−0.013 t = −0.249	1,161.322 t = 1.131	558.268 t = 1.345	0.000 t = 1.916*
Title text per day	−0.402 t = −1.076	−2,407,564.000 t = −0.652	0.454 t = 0.699	−11,717.930 t = −0.963	−6,138.901 t = −1.024	−0.000 t = −1.020
Positive sentiment title text	−0.019 t = −0.557	−403,900.200 t = −1.107	0.018 t = 0.237	−1,388.709 t = −1.249	−669.243 t = −1.152	−0.000 t = −0.383
Mkt.RF	1.564 t = 1.504	8,968,535.000 t = 0.927	−1.512 t = −1.040	16,517.200 t = 0.542	11,064.540 t = 0.631	0.000 t = 4.547***
SMB	−0.578 t = −0.432					
HML	0.271 t = 0.191					
RMW	−3.014 t = −1.961**					
CMA	1.466 t = 0.920					
MOM	1.315 t = 0.962					
Constant	−4.380 t = −0.893	149,498,028.000 t = 3.025***	89.944 t = 7.302***	353,245.900 t = 2.068**	311,985.500 t = 2.359**	−0.000 t = −3.277***
Observations	44	44	44	44	44	44
R^2	0.758	0.478	0.159	0.393	0.330	0.578
Adjusted R^2	0.664	0.377	−0.004	0.275	0.199	0.496

Note: *p<0.1; **p<0.05; ***p<0.01

A1.9 NKLA - Nikola Corporation

A1.9.1 Full regression

Table A1.39: NKLA full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.005 t = 0.996	16,031.080 t = 2.644***	0.033 t = 2.702***	58.656 t = 3.222***	18.312 t = 1.141	0.000 t = 1.652*
Positive sentiment comments	0.082 t = 1.361	−103,182.400 t = −1.475	−0.351 t = −2.830***	−839.505 t = −2.247**	−907.672 t = −2.861***	−0.000 t = −0.350
Title per day	0.001 t = 0.003	869,243.700 t = 1.652*	0.709 t = 0.801	−386.935 t = −0.203	3,002.240 t = 1.977**	−0.000 t = −1.256
Positive sentiment title	0.038 t = 1.188	25,222.900 t = 0.505	0.128 t = 1.302	503.218 t = 2.226**	28.089 t = 0.259	0.000 t = 0.882
Title text per day	−0.599 t = −1.069	−204,623.100 t = −0.221	−1.631 t = −1.455	23.802 t = 0.009	1,529.445 t = 0.460	0.000 t = 0.601
Positive sentiment title text	0.011 t = 0.499	−1,034.106 t = −0.042	−0.016 t = −0.226	231.632 t = 1.305	23.756 t = 0.303	0.000 t = 0.241
Mkt.RF	0.769 t = 0.802	−930,054.400 t = −0.700	−1.806 t = −0.500	−2,254.171 t = −0.395	−292.605 t = −0.118	0.000 t = 3.560***
SMB	3.113 t = 1.786*					
HML	−3.503 t = −1.827*					
RMW	−1.425 t = −0.505					
CMA	4.571 t = 1.260					
MOM	−1.822 t = −0.944					
Constant	−7.911 t = −1.924*	17,156,046.000 t = 2.924***	143.203 t = 17.146***	73,798.870 t = 2.889***	78,633.150 t = 3.563***	−0.000 t = −0.544
Observations	190	190	190	190	190	190
R^2	0.212	0.536	0.322	0.204	0.546	0.137
Adjusted R^2	0.159	0.518	0.296	0.173	0.529	0.104

Note: *p<0.1; **p<0.05; ***p<0.01

A1.9.2 Lagged return

Table A1.40: NKLA lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.001	0.005	0.011	0.001	−0.002	0.002
	t = 0.175	t = 0.996	t = 1.268	t = 0.210	t = −0.370	t = 0.649
Positive sentiment comments	0.082	0.082	0.032	−0.084	−0.002	−0.031
	t = 1.369	t = 1.361	t = 0.645	t = −1.429	t = −0.056	t = −0.688
Title per day	−0.141	0.001	−0.491	−0.208	−0.184	0.320
	t = −0.482	t = 0.003	t = −1.237	t = −0.886	t = −0.581	t = 1.021
Positive sentiment title	0.028	0.038	0.040	0.019	0.007	0.036
	t = 1.024	t = 1.188	t = 1.315	t = 0.583	t = 0.236	t = 1.370
Title text per day	0.111	−0.599	−0.398	0.049	0.191	−0.866
	t = 0.384	t = −1.069	t = −0.864	t = 0.111	t = 0.447	t = −1.975**
Positive sentiment title text	−0.015	0.011	0.023	−0.013	−0.035	0.025
	t = −0.685	t = 0.499	t = 0.951	t = −0.392	t = −1.487	t = 0.736
Mkt.RF	0.726	0.769	−0.209	−1.164	−1.835	−0.362
	t = 0.801	t = 0.802	t = −0.201	t = −1.628	t = −0.926	t = −0.584
SMB	−0.836	3.113	−0.152	2.110	−0.343	0.589
	t = −0.605	t = 1.786*	t = −0.115	t = 1.701*	t = −0.190	t = 0.533
HML	1.137	−3.503	1.656	0.493	−1.708	−0.996
	t = 0.749	t = −1.827*	t = 0.810	t = 0.265	t = −0.877	t = −0.603
RMW	−0.170	−1.425	0.029	0.729	0.260	−0.165
	t = −0.091	t = −0.505	t = 0.014	t = 0.364	t = 0.120	t = −0.103
CMA	−0.062	4.571	−4.174	−0.808	2.568	−1.346
	t = −0.023	t = 1.260	t = −1.371	t = −0.326	t = 0.998	t = −0.531
MOM	−0.060	−1.822	0.961	1.611	−0.082	−1.310
	t = −0.048	t = −0.944	t = 0.584	t = 1.175	t = −0.067	t = −0.872
Constant	−5.385	−7.911	−6.163	4.239	2.618	−1.371
	t = −1.288	t = −1.924*	t = −1.550	t = 1.167	t = 0.844	t = −0.409
Observations	189	190	189	188	187	186
R^2	0.050	0.212	0.181	0.079	0.110	0.070
Adjusted R^2	−0.015	0.159	0.125	0.016	0.048	0.006

Note: *p<0.1; **p<0.05; ***p<0.01

A1.9.3 Lagged volume

Table A1.41: NKLA lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	5,490.595	16,031.080	14,417.300	607.228	−2,376.663	−7,831.802
	t = 0.729	t = 2.644***	t = 2.099**	t = 0.101	t = −0.406	t = −1.024
Positive sentiment comments	−163,420.400	−103,182.400	−162,828.000	−193,955.700	−28,460.550	−9,004.064
	t = −1.652*	t = −1.475	t = −1.797*	t = −1.755*	t = −0.354	t = −0.113
Title per day	794,716.400	869,243.700	−71,062.620	174,110.600	426,560.700	1,342,268.000
	t = 1.242	t = 1.652*	t = −0.116	t = 0.298	t = 0.833	t = 1.898*
Positive sentiment title	−12,021.790	25,222.900	−54,711.490	−24,235.220	−73,054.180	−112,195.100
	t = −0.240	t = 0.505	t = −0.804	t = −0.457	t = −1.365	t = −1.982**
Title text per day	−129,799.100	−204,623.100	839,036.500	1,598,601.000	1,739,811.000	376,530.400
	t = −0.145	t = −0.221	t = 0.994	t = 2.595***	t = 2.223**	t = 0.282
Positive sentiment title text	−51,511.740	−1,034.106	−21,857.190	20,963.440	14,974.590	21,982.510
	t = −1.428	t = −0.042	t = −0.599	t = 0.436	t = 0.320	t = 0.514
Mkt.RF	−27,457.630	−930,054.400	1,823,549.000	−1,002,051.000	1,088,518.000	3,507,658.000
	t = −0.024	t = −0.700	t = 1.416	t = −0.527	t = 0.734	t = 2.661***
Constant	29,421,962.000	17,156,046.000	26,774,349.000	26,441,382.000	19,614,323.000	22,581,315.000
	t = 3.703***	t = 2.924***	t = 3.889***	t = 3.727***	t = 3.535***	t = 3.760***
Observations	189	190	189	188	187	186
R^2	0.214	0.536	0.340	0.252	0.279	0.223
Adjusted R^2	0.184	0.518	0.315	0.223	0.251	0.193

Note: *p<0.1; **p<0.05; ***p<0.01

A1.9.4 2020 regression

Table A1.42: NKLA 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.007 t = 1.210	15,899.080 t = 2.466**	0.030 t = 2.205**	54.536 t = 2.764***	16.358 t = 0.956	0.000 t = 2.270**
Positive sentiment comments	0.074 t = 0.604	−254,153.200 t = −1.551	−0.065 t = −0.232	−1,008.298 t = −1.075	−1,582.722 t = −1.885*	0.000 t = 0.709
Title per day	−0.143 t = −0.522	871,712.700 t = 1.424	0.705 t = 0.649	−390.787 t = −0.176	2,979.724 t = 1.674*	−0.000 t = −1.227
Positive sentiment title	0.032 t = 0.858	21,900.630 t = 0.373	0.149 t = 1.220	570.234 t = 2.083**	89.877 t = 0.672	0.000 t = 1.374
Title text per day	−0.529 t = −0.898	−320,375.100 t = −0.336	−1.733 t = −1.429	−502.230 t = −0.181	1,270.449 t = 0.365	0.000 t = 0.197
Positive sentiment title text	0.001 t = 0.047	−17,109.500 t = −0.575	−0.020 t = −0.229	225.529 t = 1.005	31.772 t = 0.325	−0.000 t = −1.100
Mkt.RF	1.044 t = 0.936	−1,005,133.000 t = −0.587	−1.778 t = −0.387	−1,583.086 t = −0.225	−253.297 t = −0.088	0.000 t = 3.154***
SMB	3.411 t = 1.411					
HML	−3.830 t = −1.328					
RMW	−0.161 t = −0.044					
CMA	1.488 t = 0.363					
MOM	−2.151 t = −0.764					
Constant	−6.954 t = −1.031	26,608,988.000 t = 2.727***	134.008 t = 9.801***	90,839.970 t = 1.833*	114,502.700 t = 2.514**	−0.000 t = −1.486
Observations	146	146	146	146	146	146
R^2	0.209	0.523	0.273	0.159	0.525	0.174
Adjusted R^2	0.138	0.499	0.237	0.116	0.501	0.133

Note: *p<0.1; **p<0.05; ***p<0.01

A1.9.5 2021 regression

Table A1.43: NKLA 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.011 t = 0.594	53,119.330 t = 1.134	−0.026 t = −0.598	155.992 t = 1.207	42.643 t = 1.134	0.000 t = 0.670
Positive sentiment comments	0.032 t = 0.433	61,750.450 t = 1.044	−0.104 t = −0.979	117.333 t = 0.793	−2.787 t = −0.021	−0.000 t = −0.804
Title per day	0.809 t = 0.319	1,341,003.000 t = 1.568	1.251 t = 1.948*	3,811.910 t = 0.888	2,074.637 t = 1.559	0.000 t = 0.027
Positive sentiment title	0.019 t = 0.296	−4,697.502 t = −0.069	0.147 t = 1.662*	71.841 t = 0.368	17.252 t = 0.168	−0.000 t = −0.624
Title text per day	−0.390 t = −0.236	−399,053.700 t = −0.536	0.799 t = 0.628	1,353.939 t = 0.416	408.007 t = 0.168	−0.000 t = −0.077
Positive sentiment title text	0.023 t = 0.511	65,035.470 t = 1.690*	0.021 t = 0.428	242.555 t = 1.733*	142.668 t = 1.847*	0.000 t = 1.337
Mkt.RF	0.746 t = 0.418	91,254.420 t = 0.085	−2.814 t = −1.394	−1,720.151 t = −0.550	−827.677 t = −0.358	0.000 t = 1.879*
SMB	−0.020 t = −0.008					
HML	−0.460 t = −0.171					
RMW	−3.731 t = −1.086					
CMA	7.930 t = 0.942					
MOM	0.445 t = 0.185					
Constant	−5.976 t = −0.945	71,202.750 t = 0.010	111.094 t = 12.131***	−2,844.262 t = −0.134	6,358.382 t = 0.637	0.000 t = 0.248
Observations	44	44	44	44	44	44
R^2	0.633	0.697	0.343	0.739	0.576	0.236
Adjusted R^2	0.490	0.638	0.216	0.688	0.494	0.087

Note: *p<0.1; **p<0.05; ***p<0.01

A1.10 NOK - Nokia Corporation

A1.10.1 Lagged return

Table A1.44: NOK lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002	−0.005	−0.002	0.005	0.002	−0.001
	t = 0.689	t = −0.579	t = −0.164	t = 0.221	t = 0.233	t = −0.070
Positive sentiment comments	0.019	0.016	0.008	0.020	0.003	−0.005
	t = 1.915*	t = 2.796***	t = 0.932	t = 2.423**	t = 0.383	t = −0.477
Title per day	−0.072	0.175	−0.101	−0.062	0.009	0.032
	t = −0.718	t = 0.604	t = −0.248	t = −0.088	t = 0.036	t = 0.102
Positive sentiment title	−0.0001	0.019	0.016	−0.008	−0.003	−0.006
	t = −0.003	t = 1.038	t = 1.078	t = −0.367	t = −0.193	t = −0.415
Title text per day	0.142	−0.442	0.456	0.064	−0.093	−0.090
	t = 0.522	t = −0.630	t = 0.454	t = 0.036	t = −0.138	t = −0.123
Positive sentiment title text	0.0001	0.014	0.013	−0.021	0.005	0.011
	t = 0.007	t = 0.648	t = 0.439	t = −0.391	t = 0.182	t = 0.534
Mkt.RF	−0.107	0.928	−0.307	0.357	0.109	−0.030
	t = −0.651	t = 7.638***	t = −1.570	t = 1.575	t = 0.538	t = −0.184
SMB	0.910	0.428	0.090	0.390	−0.117	0.259
	t = 1.968**	t = 1.434	t = 0.188	t = 0.726	t = −0.197	t = 0.424
HML	−0.793	−0.306	−0.052	0.291	−0.370	0.565
	t = −1.954*	t = −1.409	t = −0.135	t = 0.655	t = −0.920	t = 0.896
RMW	0.524	0.441	−0.143	−0.487	−0.117	0.155
	t = 1.102	t = 1.290	t = −0.327	t = −0.725	t = −0.164	t = 0.258
CMA	0.478	0.534	−1.776	1.679	0.809	−2.145
	t = 0.417	t = 0.371	t = −0.844	t = 0.471	t = 0.319	t = −0.872
MOM	−0.185	−0.026	−0.401	0.544	−0.151	0.239
	t = −0.921	t = −0.113	t = −1.388	t = 1.188	t = −0.421	t = 0.442
Constant	−1.578	−3.221	−2.401	−0.207	−0.634	−0.019
	t = −0.787	t = −2.845***	t = −1.282	t = −0.069	t = −0.393	t = −0.013
Observations	295	296	295	294	293	292
R^2	0.241	0.673	0.443	0.174	0.045	0.053
Adjusted R^2	0.208	0.659	0.419	0.139	0.004	0.012

Note: *p<0.1; **p<0.05; ***p<0.01

A1.11 NVDA - NVIDIA Corporation

A1.11.1 Lagged return

Table A1.45: NVDA lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0002	0.003	−0.001	0.002	0.003	−0.0001
	t = 0.074	t = 3.142***	t = −0.487	t = 1.104	t = 1.676*	t = −0.077
Positive sentiment comments	0.032	0.010	0.024	0.025	−0.005	0.031
	t = 1.768*	t = 0.926	t = 1.593	t = 1.626	t = −0.370	t = 1.981**
Title per day	0.042	−0.235	0.145	0.096	0.084	−0.094
	t = 0.271	t = −1.982**	t = 0.963	t = 0.517	t = 0.413	t = −0.500
Positive sentiment title	−0.008	0.001	0.006	0.0005	0.009	0.006
	t = −0.853	t = 0.103	t = 0.687	t = 0.049	t = 0.831	t = 0.532
Title text per day	−0.205	−0.011	0.024	−0.087	−0.291	0.244
	t = −1.289	t = −0.133	t = 0.147	t = −0.542	t = −1.777*	t = 1.672*
Positive sentiment title text	0.010	0.002	0.005	0.008	0.009	−0.002
	t = 1.008	t = 0.435	t = 0.389	t = 0.901	t = 1.015	t = −0.197
Mkt.RF	−0.736	1.506	−0.446	0.247	0.104	−0.122
	t = −4.371***	t = 23.786***	t = −2.027**	t = 1.216	t = 0.550	t = −0.662
SMB	0.254	0.251	−0.195	0.787	−0.949	0.362
	t = 0.604	t = 1.501	t = −0.485	t = 2.009**	t = −1.892*	t = 0.832
HML	−0.028	−0.690	0.092	0.404	0.065	0.002
	t = −0.072	t = −4.142***	t = 0.196	t = 0.712	t = 0.155	t = 0.004
RMW	0.432	0.161	−0.122	0.493	−1.662	0.212
	t = 0.786	t = 0.656	t = −0.265	t = 0.972	t = −3.378***	t = 0.449
CMA	−0.989	−0.178	−0.794	−0.233	1.411	0.111
	t = −2.015**	t = −0.646	t = −1.207	t = −0.418	t = 1.898*	t = 0.137
MOM	−0.047	0.252	−0.331	0.513	−0.470	0.250
	t = −0.169	t = 2.761***	t = −1.168	t = 1.050	t = −1.615	t = 0.607
Constant	−1.872	−1.020	−2.228	−2.256	−0.741	−2.370
	t = −1.250	t = −1.173	t = −1.535	t = −1.881*	t = −0.572	t = −1.882*
Observations	295	296	295	294	293	292
R^2	0.179	0.783	0.096	0.127	0.120	0.044
Adjusted R^2	0.144	0.774	0.058	0.090	0.083	0.003
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.11.2 Lagged volume

Table A1.46: NVDA lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	10,688.480	13,244.420	6,827.326	6,913.476	5,214.737	2,417.447
	t = 3.928***	t = 3.249***	t = 2.221**	t = 2.012**	t = 1.655*	t = 0.835
Positive sentiment comments	5,297.134	−14,793.920	7,516.777	−10,219.000	−33,660.850	−18,212.760
	t = 0.254	t = −0.725	t = 0.319	t = −0.428	t = −1.249	t = −0.701
Title per day	89,989.820	−53,835.330	−530,454.300	−678,601.800	−396,000.900	−69,190.280
	t = 0.343	t = −0.172	t = −1.912*	t = −1.809*	t = −1.379	t = −0.222
Positive sentiment title	−39,455.820	−35,332.700	−5,129.183	−20,254.500	−16,203.890	−37,091.710
	t = −2.582***	t = −2.622***	t = −0.330	t = −1.438	t = −0.900	t = −2.035**
Title text per day	195,303.100	116,054.800	553,334.800	526,476.800	303,097.300	133,828.200
	t = 0.698	t = 0.410	t = 1.946*	t = 1.696*	t = 1.036	t = 0.533
Positive sentiment title text	−9,169.648	−23,094.300	−28,080.860	−13,880.930	−16,045.970	−5,640.475
	t = −0.712	t = −1.840*	t = −2.124**	t = −1.105	t = −1.254	t = −0.417
Mkt.RF	−315,063.000	−245,120.200	17,904.490	−64,311.790	−16,115.630	71,533.070
	t = −1.577	t = −1.141	t = 0.096	t = −0.332	t = −0.099	t = 0.402
Constant	12,529,116.000	14,510,007.000	12,290,257.000	13,578,659.000	15,265,173.000	15,013,742.000
	t = 7.199***	t = 8.064***	t = 6.739***	t = 7.300***	t = 7.043***	t = 7.374***
Observations	295	296	295	294	293	292
R^2	0.210	0.230	0.079	0.065	0.041	0.032
Adjusted R^2	0.191	0.211	0.057	0.042	0.017	0.009
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.11.3 2020 regression

Table A1.47: NVDA 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.003 t = 2.529**	12,299.730 t = 2.832***	0.003 t = 0.643	161.331 t = 3.720***	57.764 t = 2.361**	0.000 t = 0.777
Positive sentiment comments	0.017 t = 2.055**	−32,129.540 t = −1.261	−0.053 t = −0.856	275.135 t = 1.812*	−94.479 t = −1.043	0.000 t = 1.941*
Title per day	−0.211 t = −1.496	−61,127.350 t = −0.169	−0.347 t = −0.771	923.854 t = 0.244	491.700 t = 0.267	−0.000 t = −0.703
Positive sentiment title	0.002 t = 0.488	−28,203.350 t = −1.714*	−0.052 t = −1.621	−10.368 t = −0.090	−91.107 t = −1.276	0.000 t = 1.266
Title text per day	−0.003 t = −0.036	191,546.700 t = 0.611	0.605 t = 1.492	433.434 t = 0.136	633.126 t = 0.408	−0.000 t = −0.248
Positive sentiment title text	0.0001 t = 0.029	−34,243.810 t = −2.422**	−0.063 t = −1.779*	1.108 t = 0.009	−43.854 t = −0.713	0.000 t = 0.701
Mkt.RF	1.497 t = 23.491***	−197,850.800 t = −0.928	−1.093 t = −1.450	−205.672 t = −0.246	−1,263.623 t = −2.264**	0.000 t = 12.044***
SMB	0.261 t = 1.588					
HML	−0.705 t = −4.393***					
RMW	0.354 t = 1.366					
CMA	−0.386 t = −1.124					
MOM	0.225 t = 2.533**					
Constant	−1.491 t = −2.525**	16,399,404.000 t = 7.077***	61.095 t = 10.946***	63,007.070 t = 4.966***	64,806.170 t = 8.104***	−0.000 t = −2.279**
Observations	252	252	252	252	252	252
R^2	0.811	0.233	0.097	0.445	0.296	0.641
Adjusted R^2	0.801	0.211	0.072	0.429	0.275	0.631

Note: *p<0.1; **p<0.05; ***p<0.01

A1.11.4 2021 regression

Table A1.48: NVDA 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.006	19,612.870	−0.010	361.525	136.862	0.000
	t = 0.769	t = 1.576	t = −1.403	t = 2.152**	t = 1.893*	t = 1.493
Positive sentiment comments	−0.015	2,481.728	−0.022	−241.959	−284.591	0.000
	t = −0.242	t = 0.064	t = −0.580	t = −0.133	t = −0.793	t = 0.396
Title per day	−0.155	19,866.070	1.104	−3,060.002	−1,572.289	−0.000
	t = −0.376	t = 0.018	t = 2.416**	t = −0.168	t = −0.507	t = −0.430
Positive sentiment title	−0.005	−41,255.690	0.017	−521.229	−210.179	−0.000
	t = −0.245	t = −1.583	t = 0.700	t = −0.972	t = −1.244	t = −0.864
Title text per day	−0.290	−484,606.300	−0.056	−2,924.682	−3,650.597	−0.000
	t = −0.671	t = −1.124	t = −0.095	t = −0.301	t = −0.890	t = −0.170
Positive sentiment title text	0.014	17,247.000	0.002	694.824	308.425	0.000
	t = 0.676	t = 0.935	t = 0.085	t = 1.869*	t = 2.412**	t = 0.063
Mkt.RF	1.736	−426,027.000	−1.411	2,806.507	−841.901	0.000
	t = 2.312**	t = −0.611	t = −3.208***	t = 0.256	t = −0.197	t = 5.303***
SMB	−1.011					
	t = −1.352					
HML	0.272					
	t = 0.256					
RMW	−0.857					
	t = −0.753					
CMA	1.259					
	t = 0.989					
MOM	1.188					
	t = 1.304					
Constant	0.507	9,048,813.000	45.212	118,482.500	78,875.390	−0.000
	t = 0.124	t = 4.208***	t = 17.398***	t = 0.959	t = 3.334***	t = −0.216
Observations	44	44	44	44	44	44
R^2	0.717	0.395	0.314	0.256	0.328	0.492
Adjusted R^2	0.608	0.278	0.181	0.111	0.197	0.393

Note: *p<0.1; **p<0.05; ***p<0.01

A1.12 PLTR - Palantir Technologies, Inc.

A1.12.1 Lagged volume

Table A1.49: PLTR lagged volume

	Lagged volume					
	$Volume_{t-1}$	Volume	$Volume_{t+1}$	$Volume_{t+2}$	$Volume_{t+3}$	$Volume_{t+4}$
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	3,831.076	9,623.977	9,912.435	11,346.610	12,176.670	8,344.649
	t = 0.692	t = 1.343	t = 0.967	t = 1.505	t = 2.681***	t = 1.366
Positive sentiment comments	−1,555,083.000	−1,816,490.000	−2,220,361.000	−2,470,158.000	−2,804,149.000	−2,235,762.000
	t = −2.455**	t = −1.966**	t = −2.426**	t = −1.908*	t = −2.323**	t = −2.162**
Title per day	1,429,903.000	1,205,722.000	−169,210.900	−778,158.200	−167,905.000	68,684.520
	t = 3.401***	t = 2.340**	t = −0.234	t = −1.236	t = −0.385	t = 0.116
Positive sentiment title	−443,149.800	−270,809.100	3,418.359	−143,231.500	−161,223.000	263,134.500
	t = −1.104	t = −0.785	t = 0.009	t = −0.379	t = −0.425	t = 0.691
Title text per day	−2,288,793.000	−2,156,291.000	17,192.910	763,789.600	−371,843.000	−571,891.100
	t = −5.042***	t = −3.065***	t = 0.019	t = 0.883	t = −0.656	t = −0.730
Positive sentiment title text	34,646.270	−181,995.900	109,171.300	152,257.000	−326,720.900	−353,081.300
	t = 0.180	t = −0.743	t = 0.361	t = 0.547	t = −0.998	t = −0.888
Mkt.RF	1,760,669.000	−3,196,412.000	−3,854,206.000	−6,802,111.000	3,747,064.000	−1,085,075.000
	t = 0.445	t = −0.690	t = −0.660	t = −1.015	t = 0.547	t = −0.139
Constant	187,157,980.000	204,454,287.000	202,995,665.000	236,590,565.000	291,081,318.000	224,368,089.000
	t = 2.441**	t = 2.830***	t = 2.840***	t = 2.447**	t = 2.743***	t = 2.884***
Observations	107	108	107	106	105	104
R^2	0.319	0.406	0.214	0.158	0.183	0.137
Adjusted R^2	0.271	0.364	0.159	0.098	0.124	0.074

Note: *p<0.1; **p<0.05; ***p<0.01

A1.12.2 2020 regression

Table A1.50: PLTR 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.007	14,520.580	0.011	113.898	55.474	−0.000
	t = −1.034	t = 2.517**	t = 3.773***	t = 7.827***	t = 8.668***	t = −0.961
Positive sentiment comments	0.969	−1,260,367.000	0.026	886.329	−299.176	−0.000
	t = 0.819	t = −1.457	t = 0.048	t = 0.514	t = −0.666	t = −0.318
Title per day	0.866	−56,311.220	−0.265	−1,102.269	−1,838.162	0.000
	t = 1.236	t = −0.105	t = −1.151	t = −0.774	t = −2.714***	t = 1.675*
Positive sentiment title	−0.016	138,987.000	0.139	621.945	153.573	−0.000
	t = −0.077	t = 0.529	t = 0.850	t = 1.049	t = 0.740	t = −0.052
Title text per day	−0.969	−363,450.900	−0.013	−2,995.422	933.230	−0.000
	t = −1.227	t = −0.634	t = −0.040	t = −1.682*	t = 1.117	t = −1.917*
Positive sentiment title text	0.230	−137,895.600	0.140	94.575	7.147	0.000
	t = 0.831	t = −0.467	t = 1.126	t = 0.210	t = 0.044	t = 1.338
Mkt.RF	−6.389	1,712,168.000	0.618	−2,872.667	−686.707	0.000
	t = −0.720	t = 0.400	t = 0.304	t = −0.322	t = −0.186	t = 0.912
SMB	16.020					
	t = 0.865					
HML	−16.024					
	t = −0.751					
RMW	4.893					
	t = 0.384					
CMA	−3.342					
	t = −0.338					
MOM	−6.375					
	t = −0.653					
Constant	−86.738	122,175,924.000	64.599	−51,771.120	24,531.730	−0.000
	t = −0.907	t = 1.487	t = 1.911*	t = −0.405	t = 0.624	t = −0.181
Observations	64	64	60	60	60	64
R^2	0.100	0.447	0.667	0.882	0.925	0.091
Adjusted R^2	−0.112	0.378	0.622	0.866	0.915	−0.023

Note: *p<0.1; **p<0.05; ***p<0.01

A1.12.3 2021 regression

Table A1.51: PLTR 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0003	11,991.700	−0.005	28.957	3.026	−0.000
	t = −0.410	t = 1.027	t = −1.343	t = 0.608	t = 0.120	t = −0.516
Positive sentiment comments	0.139	−3,709,235.000	0.123	−8,661.742	−4,921.627	0.000
	t = 1.357	t = −1.224	t = 0.376	t = −0.760	t = −1.042	t = 1.168
Title per day	0.088	1,063,754.000	0.478	10,239.710	5,873.819	0.000
	t = 0.703	t = 1.260	t = 2.016**	t = 3.434***	t = 3.030***	t = 1.056
Positive sentiment title	−0.033	−1,270,437.000	−0.244	−3,640.124	−2,062.496	−0.000
	t = −0.911	t = −1.475	t = −2.395**	t = −1.105	t = −1.656*	t = −0.064
Title text per day	0.148	−371,032.200	−0.091	−6,400.280	−8,072.823	−0.000
	t = 0.409	t = −0.150	t = −0.115	t = −0.808	t = −1.372	t = −0.459
Positive sentiment title text	−0.014	−408,183.500	−0.020	−679.672	−649.416	0.000
	t = −0.255	t = −1.072	t = −0.124	t = −0.389	t = −0.766	t = 0.602
Mkt.RF	0.698	−5,867,770.000	−1.251	8,671.880	5,611.044	0.000
	t = 0.594	t = −0.675	t = −0.382	t = 0.230	t = 0.295	t = 0.892
SMB	1.846					
	t = 1.072					
HML	−3.425					
	t = −1.189					
RMW	−1.808					
	t = −1.177					
CMA	6.268					
	t = 3.034***					
MOM	−1.056					
	t = −0.410					
Constant	−8.432	433,090,001.000	117.825	1,158,343.000	661,265.000	−0.000
	t = −1.035	t = 2.845***	t = 4.680***	t = 2.336**	t = 2.300**	t = −1.336
Observations	44	44	44	44	44	44
R^2	0.712	0.555	0.311	0.659	0.558	0.113
Adjusted R^2	0.601	0.468	0.177	0.593	0.472	−0.059

Note: *p<0.1; **p<0.05; ***p<0.01

A1.13 PLUG - Plug Power, Inc.

A1.13.1 Full regression

Table A1.52: PLUG full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.001 t = 0.382	38,221.150 t = 1.781*	0.005 t = 0.621	113.152 t = 2.552**	80.661 t = 2.198**	−0.000 t = −0.726
Positive sentiment comments	0.010 t = 0.692	69,705.860 t = 2.602***	−0.031 t = −0.506	234.718 t = 2.546**	74.639 t = 2.095**	0.000 t = 2.026**
Title per day	0.769 t = 2.692***	2,423,023.000 t = 1.434	1.102 t = 0.823	4,710.334 t = 1.557	1,512.392 t = 0.632	0.000 t = 0.557
Positive sentiment title	0.006 t = 0.303	81,229.430 t = 1.265	0.049 t = 1.111	434.042 t = 2.448**	128.065 t = 1.474	−0.000 t = −1.200
Title text per day	−0.381 t = −1.213	−1,122,287.000 t = −0.651	2.102 t = 1.812*	−1,740.317 t = −0.414	−1,946.847 t = −0.575	0.000 t = 0.002
Positive sentiment title text	−0.0003 t = −0.022	85,921.140 t = 2.098**	0.029 t = 0.838	214.241 t = 1.516	153.271 t = 2.399**	0.000 t = 0.819
Mkt.RF	1.572 t = 9.862***	−254,566.400 t = −0.940	−1.131 t = −1.732*	556.374 t = 0.674	−290.788 t = −0.921	0.00000 t = 7.112***
SMB	1.801 t = 5.081***					
HML	−1.132 t = −2.663***					
RMW	−1.372 t = −2.019**					
CMA	0.416 t = 0.553					
MOM	−0.267 t = −0.794					
Constant	−1.152 t = −0.721	7,293,836.000 t = 1.985**	90.309 t = 17.909***	−11,042.100 t = −1.007	−7,042.533 t = −1.307	−0.00000 t = −1.564
Observations	296	296	296	296	296	296
R^2	0.437	0.535	0.236	0.488	0.544	0.454
Adjusted R^2	0.413	0.523	0.217	0.476	0.533	0.440

Note: *p<0.1; **p<0.05; ***p<0.01

A1.13.2 Lagged return

Table A1.53: PLUG lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
	0.00002	0.001	0.004	0.005	0.001	0.005
Comments per day	t = 0.006	t = 0.382	t = 0.818	t = 1.143	t = 0.393	t = 0.891
	−0.008	0.010	0.006	0.020	−0.012	0.031
Positive sentiment comments	t = −0.389	t = 0.692	t = 0.313	t = 1.100	t = −0.670	t = 1.757*
	0.211	0.769	0.397	0.0004	0.347	0.094
Title per day	t = 0.556	t = 2.692***	t = 0.943	t = 0.001	t = 0.529	t = 0.278
	−0.025	0.006	0.015	0.007	0.044	−0.009
Positive sentiment title	t = −1.192	t = 0.303	t = 0.635	t = 0.274	t = 1.719*	t = −0.494
	−0.211	−0.381	−0.378	−0.387	−0.305	−0.038
Title text per day	t = −0.628	t = −1.213	t = −1.075	t = −1.065	t = −0.548	t = −0.103
	0.028	−0.0003	0.001	0.007	0.004	0.011
Positive sentiment title text	t = 1.994**	t = −0.022	t = 0.094	t = 0.449	t = 0.252	t = 0.701
	−0.575	1.572	−0.245	0.329	0.126	−0.226
Mkt.RF	t = −2.393**	t = 9.862***	t = −0.887	t = 1.335	t = 0.470	t = −0.916
	0.857	1.801	−0.535	0.772	−1.289	0.813
SMB	t = 1.588	t = 5.081***	t = −0.831	t = 1.184	t = −2.088**	t = 1.383
	−0.291	−1.132	0.024	−0.034	0.343	−0.173
HML	t = −0.470	t = −2.663***	t = 0.034	t = −0.049	t = 0.604	t = −0.245
	0.377	−1.372	−0.159	−0.731	−0.909	0.610
RMW	t = 0.509	t = −2.019**	t = −0.195	t = −0.800	t = −0.958	t = 0.667
	−1.392	0.416	−0.260	−0.099	0.659	0.746
CMA	t = −1.573	t = 0.553	t = −0.267	t = −0.086	t = 0.585	t = 0.740
	−0.103	−0.267	−0.354	0.229	−0.153	0.170
MOM	t = −0.230	t = −0.794	t = −0.591	t = 0.414	t = −0.350	t = 0.305
	1.047	−1.152	−1.006	−1.527	−1.481	−1.787
Constant	t = 0.613	t = −0.721	t = −0.551	t = −0.798	t = −0.764	t = −1.053
Observations	295	296	295	294	293	292
R^2	0.070	0.437	0.070	0.069	0.081	0.065
Adjusted R^2	0.030	0.413	0.031	0.029	0.042	0.024
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.13.3 Lagged volume

Table A1.54: PLUG lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
	32,149.820	38,221.150	35,465.090	25,544.240	16,410.110	18,481.920
Comments per day	t = 1.884*	t = 1.781*	t = 1.333	t = 1.383	t = 1.122	t = 0.987
	61,030.490	69,705.860	70,926.730	78,093.900	7,824.763	28,567.740
Positive sentiment comments	t = 1.697*	t = 2.602***	t = 2.145**	t = 2.038**	t = 0.224	t = 0.670
	2,357,025.000	2,423,023.000	69,547.630	511,106.400	618,869.200	−174,790.700
Title per day	t = 1.485	t = 1.434	t = 0.035	t = 0.246	t = 0.370	t = −0.110
	51,410.850	81,229.430	116,589.500	29,071.970	62,949.970	36,193.100
Positive sentiment title	t = 0.864	t = 1.265	t = 1.662*	t = 0.344	t = 0.901	t = 0.382
	−1,290,292.000	−1,122,287.000	261,435.700	−168,864.700	1,054,648.000	1,704,896.000
Title text per day	t = −1.064	t = −0.651	t = 0.114	t = −0.069	t = 0.418	t = 0.703
	79,822.920	85,921.140	116,508.400	89,892.010	115,263.900	70,490.530
Positive sentiment title text	t = 2.275**	t = 2.098**	t = 2.377**	t = 1.763*	t = 2.196**	t = 1.099
	2,760.481	−254,566.400	−275,566.500	−385,510.700	−256,338.700	216,279.000
Mkt.RF	t = 0.011	t = −0.940	t = −0.812	t = −1.003	t = −0.673	t = 0.617
	10,785,091.000	7,293,836.000	4,858,610.000	11,993,583.000	12,505,166.000	15,411,834.000
Constant	t = 2.468**	t = 1.985**	t = 1.045	t = 2.119**	t = 2.403**	t = 1.905*
Observations	295	296	295	294	293	292
R^2	0.397	0.535	0.306	0.168	0.200	0.150
Adjusted R^2	0.382	0.523	0.289	0.148	0.181	0.129
Note:						*p<0.1; **p<0.05; ***p<0.01

A1.13.4 2020 regression

Table A1.55: PLUG 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.014	82,694.230	0.010	300.282	115.032	0.000
	t = 1.265	t = 2.182**	t = 0.410	t = 2.197**	t = 1.876*	t = 1.239
Positive sentiment comments	0.010	51,312.960	−0.047	132.307	40.245	0.000
	t = 0.690	t = 1.881*	t = −0.730	t = 1.487	t = 1.283	t = 1.860*
Title per day	0.748	2,717,257.000	2.748	5,014.731	−176.191	0.000
	t = 1.200	t = 1.631	t = 1.503	t = 1.030	t = −0.075	t = 1.769*
Positive sentiment title	0.005	151,709.700	0.039	610.700	219.125	−0.000
	t = 0.179	t = 1.767*	t = 0.832	t = 2.317**	t = 2.084**	t = −1.944*
Title text per day	−0.980	−5,122,408.000	0.378	−13,727.060	−5,019.101	−0.000
	t = −1.584	t = −1.915*	t = 0.240	t = −1.444	t = −1.228	t = −1.837*
Positive sentiment title text	0.005	117,237.400	0.020	228.275	156.243	0.000
	t = 0.345	t = 1.833*	t = 0.448	t = 0.957	t = 1.643	t = 1.011
Mkt.RF	1.459	−128,347.000	−1.031	633.270	−109.164	0.00000
	t = 8.500***	t = −0.562	t = −1.484	t = 0.874	t = −0.477	t = 7.231***
SMB	1.668					
	t = 4.478***					
HML	−0.902					
	t = −1.999**					
RMW	−0.855					
	t = −1.218					
CMA	0.504					
	t = 0.569					
MOM	−0.165					
	t = −0.472					
Constant	−1.267	2,660,015.000	91.513	−19,322.240	−11,436.910	−0.00000
	t = −0.665	t = 0.562	t = 17.948***	t = −1.234	t = −1.885*	t = −1.144
Observations	252	252	252	252	252	252
R^2	0.398	0.392	0.159	0.391	0.296	0.474
Adjusted R^2	0.368	0.375	0.135	0.373	0.276	0.459

Note: *p<0.1; **p<0.05; ***p<0.01

A1.13.5 2021 regression

Table A1.56: PLUG 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002 t = 0.677	34,641.230 t = 1.169	0.003 t = 0.297	78.169 t = 1.858*	55.327 t = 1.211	0.000 t = 0.686
Positive sentiment comments	0.084 t = 0.735	−83,849.040 t = −0.441	0.064 t = 0.457	478.540 t = 0.790	102.045 t = 0.350	0.000 t = 0.689
Title per day	0.582 t = 1.370	2,491,569.000 t = 1.035	0.506 t = 0.301	5,296.389 t = 1.880*	2,475.686 t = 0.743	−0.000 t = −0.321
Positive sentiment title	−0.002 t = −0.034	−175,047.600 t = −1.432	−0.026 t = −0.306	−365.170 t = −1.311	−235.103 t = −1.235	0.000 t = 1.946*
Title text per day	−0.214 t = −0.422	722,446.600 t = 0.388	2.607 t = 1.993**	9.380 t = 0.003	301.656 t = 0.118	0.000 t = 0.396
Positive sentiment title text	0.008 t = 0.170	−14,229.580 t = −0.137	0.009 t = 0.107	85.319 t = 0.314	75.286 t = 0.439	0.000 t = 0.520
Mkt.RF	2.840 t = 1.696*	−3,317,679.000 t = −1.159	−2.317 t = −1.007	−1,642.574 t = −0.303	−3,894.290 t = −1.079	0.000 t = 5.064***
SMB	1.048 t = 0.388					
HML	0.418 t = 0.140					
RMW	−1.779 t = −0.485					
CMA	1.200 t = 0.347					
MOM	1.496 t = 0.445					
Constant	−8.152 t = −0.750	41,361,585.000 t = 2.454**	96.037 t = 5.367***	57,134.930 t = 0.938	37,702.850 t = 1.294	−0.000 t = −1.682*
Observations	44	44	44	44	44	44
R^2	0.684	0.748	0.521	0.728	0.692	0.579
Adjusted R^2	0.562	0.699	0.428	0.675	0.632	0.497

Note: *p<0.1; **p<0.05; ***p<0.01

A1.14 RKT - Rocket Companies, Inc.

A1.14.1 Full regression

Table A1.57: RKT full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.002 t = 0.424	14,652.660 t = 0.850	0.002 t = 0.348	14.510 t = 0.244	5.236 t = 0.269	−0.000 t = −0.667
Positive sentiment comments	0.072 t = 0.767	215,647.200 t = 1.468	0.134 t = 1.061	466.820 t = 0.855	−5.058 t = −0.030	0.000 t = 2.379**
Title per day	−0.365 t = −0.477	−1,840,373.000 t = −0.788	−0.995 t = −1.305	−1,089.767 t = −0.137	132.143 t = 0.052	0.000 t = 0.643
Positive sentiment title	0.007 t = 0.256	65,741.710 t = 1.523	0.015 t = 0.305	167.591 t = 1.000	14.745 t = 0.283	0.000 t = 1.363
Title text per day	0.368 t = 0.352	3,873,774.000 t = 1.198	4.266 t = 3.386***	5,849.129 t = 0.539	1,072.180 t = 0.315	0.000 t = 0.011
Positive sentiment title text	0.015 t = 0.747	−17,629.960 t = −0.321	0.063 t = 1.435	76.402 t = 0.350	−48.678 t = −0.707	0.000 t = 2.541**
Mkt.RF	0.713 t = 0.738	−1,692,148.000 t = −0.626	−1.003 t = −0.667	−7,247.574 t = −0.786	−3,154.307 t = −1.081	0.000 t = 5.856***
SMB	0.097 t = 0.070					
HML	0.359 t = 0.156					
RMW	−0.440 t = −0.222					
CMA	1.397 t = 0.510					
MOM	0.615 t = 0.574					
Constant	−7.059 t = −0.946	−8,660,851.000 t = −0.831	62.006 t = 6.727***	20,526.450 t = 0.503	19,649.910 t = 1.466	−0.00000 t = −3.549***
Observations	146	146	146	146	146	146
R^2	0.063	0.647	0.460	0.484	0.736	0.298
Adjusted R^2	−0.022	0.629	0.433	0.457	0.722	0.262

Note: *p<0.1; **p<0.05; ***p<0.01

A1.14.2 2020 regression

Table A1.58: RKT 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.007 t = −0.642	18,936.310 t = 1.582	0.004 t = 0.412	120.941 t = 2.235**	14.758 t = 1.606	0.000 t = 0.520
Positive sentiment comments	0.076 t = 0.787	165,277.300 t = 1.693*	0.197 t = 1.263	−35.456 t = −0.075	−92.509 t = −0.899	0.000 t = 2.672***
Title per day	−0.338 t = −0.636	−152,669.500 t = −0.274	1.650 t = 1.270	−6,888.285 t = −1.220	−463.559 t = −0.640	−0.000 t = −0.883
Positive sentiment title	0.023 t = 0.703	65,465.570 t = 1.509	−0.027 t = −0.403	402.008 t = 1.366	32.166 t = 0.816	0.000 t = 1.576
Title text per day	1.226 t = 1.153	574,598.100 t = 0.475	0.840 t = 0.490	−5,550.229 t = −1.535	−417.008 t = −0.340	0.000 t = 0.771
Positive sentiment title text	0.009 t = 0.503	−8,330.243 t = −0.328	0.024 t = 0.425	320.209 t = 1.541	16.432 t = 0.461	0.000 t = 1.876*
Mkt.RF	2.405 t = 3.202***	7,834.867 t = 0.007	−0.323 t = −0.182	−7,112.806 t = −1.292	−1,867.225 t = −2.301**	0.000 t = 6.058***
SMB	0.908 t = 0.727					
HML	3.634 t = 0.919					
RMW	−1.133 t = −0.485					
CMA	−2.973 t = −0.989					
MOM	2.105 t = 0.980					
Constant	−7.980 t = −0.968	−7,925,358.000 t = −0.952	61.330 t = 5.907***	18,507.360 t = 0.440	17,256.220 t = 2.091**	−0.00000 t = −4.051***
Observations	102	102	102	102	102	102
R^2	0.190	0.667	0.529	0.256	0.263	0.412
Adjusted R^2	0.080	0.643	0.494	0.200	0.208	0.368

Note: *p<0.1; **p<0.05; ***p<0.01

A1.14.3 2021 regression

Table A1.59: RKT 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.005 t = 0.573	18,880.400 t = 0.699	0.004 t = 0.454	24.995 t = 0.282	10.640 t = 0.384	−0.000 t = −0.587
Positive sentiment comments	0.260 t = 0.774	361,750.400 t = 0.525	0.062 t = 0.165	1,143.700 t = 0.485	93.393 t = 0.129	0.000 t = 0.498
Title per day	−1.075 t = −0.599	−3,337,492.000 t = −0.599	−1.406 t = −0.763	−7,086.664 t = −0.388	−1,765.849 t = −0.318	0.000 t = 0.387
Positive sentiment title	0.010 t = 0.200	−11,323.210 t = −0.095	−0.010 t = −0.120	59.113 t = 0.160	−25.082 t = −0.166	−0.000 t = −0.020
Title text per day	2.265 t = 0.563	8,211,291.000 t = 0.682	4.930 t = 1.222	26,443.180 t = 0.672	6,570.225 t = 0.538	−0.000 t = −0.128
Positive sentiment title text	0.009 t = 0.116	−39,924.690 t = −0.167	0.181 t = 1.485	25.970 t = 0.035	−137.256 t = −0.419	0.000 t = 0.624
Mkt.RF	−3.987 t = −0.713	−9,416,282.000 t = −0.592	−2.250 t = −0.395	−27,567.830 t = −0.527	−11,533.480 t = −0.662	0.000 t = 0.964
SMB	2.219 t = 0.638					
HML	−3.409 t = −0.579					
RMW	3.947 t = 0.631					
CMA	0.493 t = 0.077					
MOM	−0.080 t = −0.013					
Constant	−17.424 t = −0.736	−6,631,612.000 t = −0.160	62.773 t = 2.318**	1,861.471 t = 0.013	31,531.440 t = 0.631	−0.00000 t = −0.619
Observations	44	44	44	44	44	44
R^2	0.377	0.709	0.605	0.637	0.794	0.087
Adjusted R^2	0.136	0.653	0.529	0.566	0.754	−0.091

Note: *p<0.1; **p<0.05; ***p<0.01

A1.15 SPCE - Virgin Galactic Holdings, Inc.

A1.15.1 Full regression

Table A1.60: SPCE full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.002 t = −0.635	19,049.960 t = 3.292***	0.015 t = 2.038**	95.766 t = 4.136***	58.990 t = 3.268***	−0.000 t = −0.518
Positive sentiment comments	0.069 t = 2.392**	−110,386.800 t = −2.561**	−0.580 t = −4.166***	136.231 t = 0.615	−527.898 t = −5.418***	0.000 t = 0.532
Title per day	0.215 t = 0.866	1,002,851.000 t = 2.665***	0.975 t = 2.093**	5,065.499 t = 3.548***	1,114.347 t = 1.179	0.000 t = 0.164
Positive sentiment title	0.028 t = 1.906*	8,576.951 t = 0.430	−0.121 t = −2.268**	227.969 t = 1.647*	−87.780 t = −1.959*	0.000 t = 2.451**
Title text per day	0.146 t = 0.430	−1,251,354.000 t = −2.240**	−0.406 t = −0.431	−8,078.772 t = −3.091***	−2,864.580 t = −1.799*	0.000 t = 0.730
Positive sentiment title text	0.017 t = 1.383	10,282.270 t = 0.560	−0.131 t = −2.213**	273.324 t = 2.103**	−12.887 t = −0.262	0.000 t = 1.970**
Mkt.RF	1.446 t = 6.482***	−65,547.480 t = −0.293	−0.475 t = −0.366	−1,115.239 t = −0.896	−1,717.257 t = −2.598***	0.000 t = 6.072***
SMB	1.517 t = 2.355**					
HML	−0.529 t = −0.916					
RMW	−0.960 t = −1.197					
CMA	0.621 t = 0.400					
MOM	0.128 t = 0.359					
Constant	−8.328 t = −3.528***	18,253,196.000 t = 5.476***	167.084 t = 15.275***	28,390.760 t = 1.718*	65,728.890 t = 7.657***	−0.00000 t = −2.399**
Observations	296	296	296	296	296	296
R^2	0.349	0.643	0.299	0.440	0.616	0.339
Adjusted R^2	0.321	0.634	0.282	0.427	0.606	0.323

Note: *p<0.1; **p<0.05; ***p<0.01

A1.15.2 Lagged return

Table A1.61: SPCE lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.002	−0.002	0.002	0.002	0.001	0.001
	t = −0.560	t = −0.635	t = 0.488	t = 0.977	t = 0.497	t = 0.648
Positive sentiment comments	0.054	0.069	0.083	0.042	0.120	0.091
	t = 1.685*	t = 2.392**	t = 2.241**	t = 1.309	t = 3.173***	t = 2.714***
Title per day	−0.100	0.215	0.322	0.064	0.032	0.105
	t = −0.503	t = 0.866	t = 0.970	t = 0.476	t = 0.229	t = 0.731
Positive sentiment title	−0.010	0.028	−0.002	0.011	0.005	0.017
	t = −0.654	t = 1.906*	t = −0.146	t = 0.718	t = 0.398	t = 1.087
Title text per day	0.308	0.146	−0.288	−0.042	0.046	−0.102
	t = 0.698	t = 0.430	t = −0.981	t = −0.165	t = 0.166	t = −0.311
Positive sentiment title text	−0.004	0.017	0.009	0.002	0.001	0.002
	t = −0.252	t = 1.383	t = 0.599	t = 0.129	t = 0.045	t = 0.135
Mkt.RF	−0.390	1.446	−0.061	0.473	0.291	−0.392
	t = −1.568	t = 6.482***	t = −0.184	t = 1.506	t = 0.980	t = −1.209
SMB	0.765	1.517	−0.945	0.699	−1.355	−0.110
	t = 1.131	t = 2.355**	t = −1.113	t = 1.075	t = −1.873*	t = −0.192
HML	−0.152	−0.529	1.035	0.912	0.072	0.139
	t = −0.225	t = −0.916	t = 1.303	t = 1.331	t = 0.114	t = 0.167
RMW	0.247	−0.960	−1.008	−0.427	−2.827	−1.265
	t = 0.250	t = −1.197	t = −1.156	t = −0.446	t = −2.638***	t = −1.407
CMA	−2.386	0.621	−2.153	0.817	1.005	−0.877
	t = −1.833*	t = 0.400	t = −1.230	t = 0.735	t = 0.754	t = −0.564
MOM	−0.299	0.128	0.134	1.123	−0.575	−0.035
	t = −0.796	t = 0.359	t = 0.266	t = 2.576***	t = −1.356	t = −0.057
Constant	−2.573	−8.328	−6.644	−4.056	−8.911	−7.584
	t = −1.016	t = −3.528***	t = −2.217**	t = −1.549	t = −3.132***	t = −2.887***
Observations	295	296	295	294	293	292
R^2	0.054	0.349	0.141	0.107	0.108	0.076
Adjusted R^2	0.013	0.321	0.104	0.069	0.070	0.036

Note: *p<0.1; **p<0.05; ***p<0.01

A1.15.3 Lagged volume

Table A1.62: SPCE lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	11,688.110	19,049.960	13,369.120	9,625.541	3,743.413	10,153.210
	t = 2.181**	t = 3.292***	t = 2.051**	t = 1.763*	t = 0.662	t = 0.918
Positive sentiment comments	−98,027.670	−110,386.800	−115,456.700	−114,659.900	−97,207.810	−116,153.800
	t = −2.418**	t = −2.561**	t = −2.350**	t = −2.350**	t = −1.588	t = −1.593
Title per day	350,801.800	1,002,851.000	291,402.800	−35,640.370	−341,402.300	145,002.700
	t = 1.546	t = 2.665***	t = 0.530	t = −0.147	t = −0.707	t = 0.263
Positive sentiment title	15,762.700	8,576.951	−8,519.285	−12,723.390	−24,577.760	−23,099.140
	t = 0.664	t = 0.430	t = −0.367	t = −0.563	t = −0.811	t = −0.784
Title text per day	−95,165.680	−1,251,354.000	−370,582.200	−39,010.410	613,813.100	−729,652.100
	t = −0.174	t = −2.240**	t = −0.545	t = −0.064	t = 0.523	t = −0.525
Positive sentiment title text	−12,461.450	10,282.270	3,009.560	−19,740.330	−35,136.620	−61,506.290
	t = −0.490	t = 0.560	t = 0.144	t = −0.775	t = −1.469	t = −2.577***
Mkt.RF	80,911.040	−65,547.480	67,396.990	−286,522.300	−427,072.900	−542,048.400
	t = 0.357	t = −0.293	t = 0.266	t = −0.686	t = −1.127	t = −1.315
Constant	19,595,777.000	18,253,196.000	21,787,876.000	24,865,768.000	26,391,498.000	29,642,877.000
	t = 6.049***	t = 5.476***	t = 6.072***	t = 6.702***	t = 6.114***	t = 5.470***
Observations	295	296	295	294	293	292
R^2	0.392	0.643	0.322	0.169	0.104	0.095
Adjusted R^2	0.377	0.634	0.306	0.148	0.082	0.072

Note: *p<0.1; **p<0.05; ***p<0.01

A1.15.4 2020 regression

Table A1.63: SPCE 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.003	26,134.440	0.028	104.377	80.227	−0.000
	t = −0.744	t = 4.489***	t = 3.349***	t = 3.873***	t = 4.501***	t = −0.474
Positive sentiment comments	0.071	−82,782.830	−0.706	114.381	−566.174	0.000
	t = 2.226**	t = −1.992**	t = −4.797***	t = 0.407	t = −5.332***	t = 0.396
Title per day	0.352	−79,101.770	−2.265	3,599.285	−1,857.988	0.000
	t = 1.099	t = −0.172	t = −3.727***	t = 1.715*	t = −1.014	t = 0.338
Positive sentiment title	0.017	18,871.860	−0.050	205.760	−48.739	0.000
	t = 1.105	t = 0.900	t = −0.908	t = 1.377	t = −0.935	t = 2.155**
Title text per day	0.073	−972,922.300	1.316	−7,767.955	−2,291.276	0.000
	t = 0.185	t = −1.553	t = 1.422	t = −2.550**	t = −1.149	t = 0.503
Positive sentiment title text	0.016	13,729.610	−0.116	258.560	4.853	0.000
	t = 1.235	t = 0.726	t = −1.854*	t = 2.013**	t = 0.095	t = 2.018**
Mkt.RF	1.413	−91,786.360	−0.435	−860.968	−1,759.385	0.000
	t = 6.231***	t = −0.453	t = −0.355	t = −0.679	t = −2.597***	t = 5.848***
SMB	0.859					
	t = 1.379					
HML	−0.058					
	t = −0.096					
RMW	−0.898					
	t = −1.063					
CMA	−1.391					
	t = −1.336					
MOM	0.030					
	t = 0.080					
Constant	−7.585	15,404,886.000	169.883	32,152.390	64,406.610	−0.00000
	t = −2.920***	t = 4.958***	t = 13.886***	t = 1.630	t = 7.072***	t = −1.893*
Observations	252	252	252	252	252	252
R^2	0.330	0.656	0.345	0.431	0.646	0.340
Adjusted R^2	0.296	0.646	0.326	0.414	0.636	0.321

Note: *p<0.1; **p<0.05; ***p<0.01

A1.15.5 2021 regression

Table A1.64: SPCE 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.010 t = 0.739	41,324.940 t = 1.766*	0.035 t = 0.654	406.838 t = 2.112**	75.195 t = 6.036***	0.000 t = 0.119
Positive sentiment comments	0.035 t = 0.257	−159,618.800 t = −0.983	−0.067 t = −0.301	48.360 t = 0.083	−317.040 t = −0.890	0.000 t = 0.509
Title per day	0.130 t = 0.229	640,938.000 t = 0.872	2.985 t = 2.086**	−3,291.218 t = −0.715	982.918 t = 2.375**	−0.000 t = −1.030
Positive sentiment title	0.074 t = 1.319	73,696.620 t = 1.158	−0.154 t = −0.992	358.741 t = 1.291	24.757 t = 0.309	0.000 t = 0.857
Title text per day	0.138 t = 0.092	−185,166.300 t = −0.094	−5.667 t = −1.579	12,511.850 t = 1.125	−1,957.682 t = −1.409	0.000 t = 1.291
Positive sentiment title text	−0.008 t = −0.127	−31,430.570 t = −0.466	−0.161 t = −1.342	−523.444 t = −1.217	−57.006 t = −0.520	−0.000 t = −0.140
Mkt.RF	1.367 t = 0.541	456,571.000 t = 0.293	−2.125 t = −0.686	−3,380.244 t = −0.399	−583.364 t = −0.217	0.000 t = 3.391***
SMB	2.125 t = 0.658					
HML	−2.709 t = −0.796					
RMW	−1.648 t = −0.448					
CMA	5.537 t = 0.929					
MOM	−0.940 t = −0.271					
Constant	−8.495 t = −0.940	18,941,001.000 t = 1.804*	135.275 t = 10.877***	50,520.730 t = 1.271	48,738.740 t = 2.219**	−0.000 t = −0.830
Observations	44	44	44	44	44	44
R^2	0.610	0.777	0.550	0.805	0.680	0.415
Adjusted R^2	0.459	0.734	0.463	0.767	0.618	0.301

Note: *p<0.1; **p<0.05; ***p<0.01

A1.16 TLRY - Tilray, Inc.

A1.16.1 Full regression

Table A1.65: TLRY full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.018	8,500.225	−0.014	1.976	33.913	−0.000
	t = −0.548	t = 0.162	t = −0.251	t = 0.018	t = 1.813*	t = −0.465
Positive sentiment comments	0.045	69,567.020	−0.004	350.183	32.777	0.000
	t = 2.447**	t = 2.445**	t = −0.059	t = 2.890***	t = 1.041	t = 1.280
Title per day	0.463	1,319,441.000	1.778	3,934.222	1,885.413	0.000
	t = 0.355	t = 0.699	t = 0.628	t = 1.266	t = 3.975***	t = 0.474
Positive sentiment title	0.054	167,441.900	0.272	802.942	141.163	0.000
	t = 0.882	t = 1.904*	t = 1.827*	t = 3.438***	t = 2.383**	t = 2.242**
Title text per day	1.522	4,239,268.000	5.895	13,968.080	2,156.908	0.000
	t = 1.436	t = 2.624***	t = 1.963**	t = 3.162***	t = 1.379	t = 0.519
Positive sentiment title text	−0.004	37,778.820	−0.094	211.097	31.692	0.000
	t = −0.130	t = 1.009	t = −1.069	t = 1.455	t = 0.772	t = 0.235
Mkt.RF	1.133	402,045.300	0.066	2,307.956	−204.835	0.00000
	t = 2.392**	t = 1.052	t = 0.037	t = 1.761*	t = −0.577	t = 2.702***
SMB	1.828					
	t = 1.644					
HML	0.605					
	t = 0.670					
RMW	−3.148					
	t = −3.098***					
CMA	−1.658					
	t = −1.080					
MOM	−0.160					
	t = −0.341					
Constant	−6.290	−5,643,130.000	111.258	−16,728.610	2,494.565	−0.00000
	t = −1.840*	t = −1.160	t = 11.321***	t = −1.152	t = 0.694	t = −3.170***
Observations	296	296	296	296	296	296
R^2	0.315	0.725	0.264	0.566	0.821	0.272
Adjusted R^2	0.286	0.718	0.246	0.556	0.817	0.254

Note: *p<0.1; **p<0.05; ***p<0.01

A1.16.2 Lagged return

Table A1.66: TLRY lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.007	−0.018	0.010	0.001	0.003	0.005
	t = −0.587	t = −0.548	t = 0.475	t = 0.204	t = 0.641	t = 1.086
Positive sentiment comments	0.055	0.045	0.032	0.020	−0.015	0.032
	t = 2.427**	t = 2.447**	t = 1.757*	t = 0.917	t = −0.592	t = 1.577
Title per day	0.185	0.463	−0.708	0.283	0.281	0.154
	t = 0.317	t = 0.355	t = −0.651	t = 0.515	t = 1.938*	t = 0.663
Positive sentiment title	−0.014	0.054	0.043	0.014	0.010	−0.0001
	t = −0.333	t = 0.882	t = 0.770	t = 0.278	t = 0.212	t = −0.002
Title text per day	0.717	1.522	1.358	0.585	−0.922	−1.203
	t = 0.778	t = 1.436	t = 1.288	t = 0.694	t = −1.371	t = −1.134
Positive sentiment title text	0.024	−0.004	0.042	−0.021	0.043	0.050
	t = 0.916	t = −0.130	t = 1.656*	t = −0.860	t = 1.633	t = 1.816*
Mkt.RF	0.346	1.133	−0.258	0.638	0.535	−0.224
	t = 0.715	t = 2.392**	t = −0.440	t = 1.361	t = 1.119	t = −0.540
SMB	0.535	1.828	−0.879	2.647	−0.440	−0.317
	t = 0.434	t = 1.644	t = −0.917	t = 2.269**	t = −0.341	t = −0.244
HML	−0.686	0.605	1.286	−1.558	2.210	0.018
	t = −0.356	t = 0.670	t = 1.335	t = −1.047	t = 1.650*	t = 0.014
RMW	0.676	−3.148	−3.320	2.563	−1.499	−0.170
	t = 0.485	t = −3.098***	t = −2.081**	t = 1.796*	t = −1.141	t = −0.117
CMA	−2.654	−1.658	−1.243	0.799	−0.317	2.250
	t = −1.350	t = −1.080	t = −0.661	t = 0.383	t = −0.173	t = 1.220
MOM	−0.214	−0.160	−0.715	0.590	1.637	0.003
	t = −0.140	t = −0.341	t = −0.603	t = 0.759	t = 1.726*	t = 0.004
Constant	−4.419	−6.290	−7.522	−2.017	−2.001	−4.690
	t = −1.634	t = −1.840*	t = −2.359**	t = −0.606	t = −0.628	t = −1.640
Observations	295	296	295	294	293	292
R^2	0.067	0.315	0.262	0.167	0.124	0.062
Adjusted R^2	0.028	0.286	0.230	0.131	0.087	0.021

Note: *p<0.1; **p<0.05; ***p<0.01

A1.16.3 Lagged volume

Table A1.67: TLRY lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	1,591.281	8,500.225	6,041.674	1,493.103	−16,316.720	−6,250.366
	t = 0.019	t = 0.162	t = 0.538	t = 0.071	t = −0.878	t = −0.457
Positive sentiment comments	78,811.620	69,567.020	35,589.870	−1,942.424	9,149.061	3,222.851
	t = 2.455**	t = 2.445**	t = 1.840*	t = −0.086	t = 0.281	t = 0.127
Title per day	1,337,321.000	1,319,441.000	1,730,382.000	1,632,220.000	1,529,581.000	303,735.300
	t = 0.412	t = 0.699	t = 1.647*	t = 1.313	t = 3.704***	t = 0.716
Positive sentiment title	220,698.600	167,441.900	124,414.200	259,797.400	249,076.200	240,129.400
	t = 1.416	t = 1.904*	t = 1.365	t = 2.290**	t = 2.112**	t = 2.071**
Title text per day	2,764,294.000	4,239,268.000	1,815,379.000	17,027.860	1,651,200.000	3,847,095.000
	t = 1.011	t = 2.624***	t = 1.138	t = 0.010	t = 0.919	t = 1.552
Positive sentiment title text	25,172.290	37,778.820	66,493.160	19,697.970	38,353.860	14,905.770
	t = 0.529	t = 1.009	t = 1.334	t = 0.329	t = 0.676	t = 0.270
Mkt.RF	967,936.600	402,045.300	201,515.100	459,303.100	−111,993.900	−243,040.700
	t = 2.371**	t = 1.052	t = 0.629	t = 1.448	t = −0.326	t = −0.597
Constant	−6,900,978.000	−5,643,130.000	−1,438,114.000	−2,397,281.000	−3,133,670.000	−1,325,061.000
	t = −0.886	t = −1.160	t = −0.276	t = −0.329	t = −0.423	t = −0.197
Observations	295	296	295	294	293	292
R^2	0.472	0.725	0.591	0.384	0.226	0.150
Adjusted R^2	0.459	0.718	0.581	0.369	0.207	0.129

Note: *p<0.1; **p<0.05; ***p<0.01

A1.16.4 2020 regression

Table A1.68: TLRY 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.019 t = −0.319	150,899.800 t = 3.132***	0.208 t = 2.394**	731.542 t = 3.461***	147.232 t = 3.823***	−0.000 t = −1.119
Positive sentiment comments	0.039 t = 2.125**	63,650.000 t = 2.369**	0.010 t = 0.131	334.342 t = 2.787***	34.419 t = 1.112	0.000 t = 1.217
Title per day	1.179 t = 0.528	2,196,881.000 t = 0.915	−3.552 t = −0.478	5,209.457 t = 0.745	1,669.711 t = 1.087	0.00000 t = 0.917
Positive sentiment title	0.021 t = 0.409	−3,205.844 t = −0.056	0.151 t = 0.935	271.483 t = 1.023	29.348 t = 0.489	0.000 t = 1.354
Title text per day	2.181 t = 1.348	1,365,841.000 t = 0.802	5.243 t = 1.357	−3,062.075 t = −0.386	−569.026 t = −0.422	0.00000 t = 0.441
Positive sentiment title text	−0.025 t = −0.746	22,203.990 t = 0.653	−0.086 t = −0.810	243.990 t = 1.605	35.457 t = 0.875	−0.000 t = −0.234
Mkt.RF	1.181 t = 2.627***	243,084.400 t = 0.583	−0.187 t = −0.102	1,334.871 t = 0.779	−333.844 t = −0.949	0.00000 t = 3.016***
SMB	1.485 t = 1.122					
HML	0.470 t = 0.367					
RMW	−2.637 t = −2.366**					
CMA	−2.395 t = −1.396					
MOM	−0.245 t = −0.335					
Constant	−3.517 t = −1.235	1,934,748.000 t = 0.593	111.695 t = 11.440***	−686.808 t = −0.040	6,092.086 t = 1.732*	−0.00000 t = −2.301**
Observations	252	252	252	252	252	252
R^2	0.233	0.331	0.073	0.326	0.223	0.310
Adjusted R^2	0.195	0.311	0.046	0.307	0.201	0.290

Note: *p<0.1; **p<0.05; ***p<0.01

A1.16.5 2021 regression

Table A1.69: TLRY 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.016	10,215.630	0.004	−3.291	36.101	−0.000
	t = −0.448	t = 0.251	t = 0.091	t = −0.064	t = 2.543**	t = −1.221
Positive sentiment comments	0.293	64,768.980	−1.258	536.192	−424.012	0.000
	t = 1.091	t = 0.150	t = −1.742*	t = 0.480	t = −0.802	t = 1.836*
Title per day	0.469	1,191,437.000	1.230	3,623.521	1,699.823	0.000
	t = 0.312	t = 0.691	t = 0.416	t = 2.031**	t = 6.359***	t = 0.725
Positive sentiment title	0.038	218,493.700	0.301	722.933	176.167	0.000
	t = 0.321	t = 1.547	t = 0.928	t = 1.892*	t = 1.155	t = 0.450
Title text per day	0.827	3,230,894.000	0.514	13,226.260	1,183.933	0.000
	t = 0.599	t = 1.804*	t = 0.159	t = 2.313**	t = 0.421	t = 0.910
Positive sentiment title text	0.020	183,689.600	0.098	456.871	149.339	0.000
	t = 0.257	t = 2.570**	t = 0.574	t = 1.528	t = 0.971	t = 0.265
Mkt.RF	−0.955	−1,228,198.000	1.716	−1,384.711	−90.821	0.000
	t = −0.274	t = −0.455	t = 0.432	t = −0.157	t = −0.026	t = 0.519
SMB	4.167					
	t = 0.890					
HML	−0.749					
	t = −0.115					
RMW	−3.890					
	t = −1.169					
CMA	−3.279					
	t = −0.586					
MOM	−1.766					
	t = −0.297					
Constant	−22.944	−12,297,521.000	205.807	−18,678.460	31,459.170	−0.00000
	t = −1.395	t = −0.450	t = 5.499***	t = −0.246	t = 1.005	t = −2.064**
Observations	44	44	44	44	44	44
R^2	0.522	0.859	0.513	0.797	0.909	0.354
Adjusted R^2	0.337	0.831	0.418	0.758	0.891	0.228

Note: *p<0.1; **p<0.05; ***p<0.01

A1.17 TSLA - Tesla, Inc.

A1.17.1 Lagged return

Table A1.70: TSLA lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0004	−0.0002	0.0005	0.0004	0.0004	0.001
	t = −0.998	t = −0.429	t = 1.326	t = 1.169	t = 1.549	t = 2.894***
Positive sentiment comments	0.020	0.159	0.115	0.033	0.080	0.091
	t = 0.315	t = 3.297***	t = 1.893*	t = 0.548	t = 1.409	t = 1.486
Title per day	0.033	0.051	−0.026	0.009	0.002	−0.014
	t = 0.703	t = 1.342	t = −0.828	t = 0.291	t = 0.083	t = −0.556
Positive sentiment title	−0.003	0.001	0.014	0.016	−0.007	0.028
	t = −0.221	t = 0.073	t = 0.900	t = 1.066	t = −0.417	t = 1.956*
Title text per day	−0.047	−0.040	0.025	−0.013	−0.015	−0.002
	t = −0.847	t = −0.997	t = 0.503	t = −0.380	t = −0.388	t = −0.070
Positive sentiment title text	−0.003	−0.008	0.010	−0.003	0.022	−0.028
	t = −0.144	t = −0.541	t = 0.505	t = −0.164	t = 1.166	t = −1.405
Mkt.RF	−0.212	1.403	−0.351	0.478	0.354	−0.253
	t = −0.800	t = 7.699***	t = −1.410	t = 2.065**	t = 1.485	t = −0.972
SMB	0.264	0.629	−0.410	0.871	−0.946	−0.240
	t = 0.447	t = 1.441	t = −0.648	t = 1.371	t = −1.486	t = −0.511
HML	−0.195	−0.346	0.543	−0.296	0.004	0.547
	t = −0.252	t = −0.738	t = 0.978	t = −0.419	t = 0.007	t = 0.779
RMW	0.290	−1.107	0.229	0.402	−2.223	0.414
	t = 0.357	t = −1.882*	t = 0.356	t = 0.479	t = −3.236***	t = 0.658
CMA	−1.147	−1.228	−1.557	0.544	1.981	0.059
	t = −1.341	t = −1.490	t = −1.744*	t = 0.590	t = 2.283**	t = 0.070
MOM	−0.188	0.232	0.089	0.347	−0.360	0.284
	t = −0.345	t = 0.767	t = 0.273	t = 0.685	t = −1.068	t = 0.578
Constant	0.765	−8.601	−8.778	−3.265	−5.884	−5.845
	t = 0.187	t = −2.740***	t = −2.177**	t = −0.846	t = −1.566	t = −1.393
Observations	295	296	295	294	293	292
R^2	0.041	0.407	0.077	0.087	0.091	0.075
Adjusted R^2	−0.0001	0.382	0.038	0.048	0.052	0.035

Note: *p<0.1; **p<0.05; ***p<0.01

A1.17.2 Lagged volume

Table A1.71: TSLA lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	1,184.120	3,779.714	3,475.632	2,140.386	658.934	871.339
	t = 0.942	t = 1.844*	t = 2.585***	t = 1.751*	t = 0.546	t = 0.714
Positive sentiment comments	123,610.900	−7,700.705	−37,391.250	148,009.000	225,547.400	275,322.200
	t = 0.613	t = −0.041	t = −0.175	t = 0.709	t = 1.168	t = 1.220
Title per day	192,908.200	279,964.900	−149,128.900	−141,167.800	−184,196.000	−265,115.200
	t = 1.348	t = 1.537	t = −0.964	t = −0.953	t = −1.364	t = −2.216**
Positive sentiment title	47,328.450	64,355.070	90,212.660	76,262.110	98,770.970	108,827.900
	t = 1.224	t = 1.798*	t = 2.058**	t = 1.839*	t = 2.374**	t = 2.407**
Title text per day	254,680.500	−19,015.680	452,381.100	414,986.300	497,122.400	489,265.900
	t = 1.307	t = −0.075	t = 2.242**	t = 1.693*	t = 2.461**	t = 2.391**
Positive sentiment title text	14,805.530	8,394.529	−25,268.820	−16,126.860	92,941.980	113,569.900
	t = 0.224	t = 0.140	t = −0.349	t = −0.207	t = 1.085	t = 1.592
Mkt.RF	−173,330.100	−276,312.300	82,787.800	−70,041.800	−237,884.000	343,043.300
	t = −0.463	t = −0.743	t = 0.200	t = −0.149	t = −0.491	t = 0.762
Constant	7,678,611.000	11,970,589.000	15,955,786.000	8,756,256.000	−1,925,191.000	−5,610,540.000
	t = 0.677	t = 1.230	t = 1.243	t = 0.744	t = −0.175	t = −0.415
Observations	295	296	295	294	293	292
R^2	0.223	0.322	0.203	0.118	0.100	0.086
Adjusted R^2	0.204	0.305	0.183	0.097	0.078	0.063

Note: *p<0.1; **p<0.05; ***p<0.01

A1.17.3 2020 regression

Table A1.72: TSLA 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0004	4,313.673	0.006	154.530	176.979	0.000
	t = −0.705	t = 1.168	t = 2.832***	t = 2.392**	t = 3.378***	t = 1.020
Positive sentiment comments	0.172	188,190.700	−0.531	8,226.255	−18,537.130	0.000
	t = 3.130***	t = 1.064	t = −2.538**	t = 1.264	t = −2.790***	t = 1.589
Title per day	0.063	204,937.300	−0.309	4,054.134	−3,151.093	−0.000
	t = 1.351	t = 0.764	t = −1.230	t = 0.528	t = −0.727	t = −0.759
Positive sentiment title	0.004	36,452.750	−0.078	−1,532.648	−2,457.994	0.000
	t = 0.473	t = 1.008	t = −1.779*	t = −1.051	t = −2.062**	t = 0.427
Title text per day	−0.038	54,255.240	0.240	−7,184.631	−5,531.649	0.000
	t = −0.843	t = 0.153	t = 1.403	t = −0.947	t = −0.846	t = 0.049
Positive sentiment title text	−0.003	6,077.803	−0.065	−5,668.375	−4,303.594	−0.000
	t = −0.166	t = 0.114	t = −1.296	t = −2.327**	t = −2.551**	t = −0.539
Mkt.RF	1.366	−371,374.700	−0.613	11,448.160	−12,687.050	0.000
	t = 7.234***	t = −1.049	t = −0.641	t = 0.875	t = −0.949	t = 5.004***
SMB	0.407					
	t = 0.816					
HML	−0.112					
	t = −0.215					
RMW	−1.159					
	t = −1.715*					
CMA	−2.376					
	t = −2.794***					
MOM	0.183					
	t = 0.546					
Constant	−9.642	861,680.800	111.464	972,004.700	2,266,401.000	−0.000
	t = −2.750***	t = 0.085	t = 8.082***	t = 2.495**	t = 5.868***	t = −1.315
Observations	252	252	252	252	252	252
R^2	0.408	0.334	0.254	0.174	0.212	0.372
Adjusted R^2	0.378	0.315	0.232	0.150	0.190	0.354

Note: *p<0.1; **p<0.05; ***p<0.01

A1.17.4 2021 regression

Table A1.73: TSLA 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.0001	1,194.824	0.001	45.963	29.319	0.000
	t = −0.093	t = 0.560	t = 1.392	t = 1.124	t = 1.120	t = 0.250
Positive sentiment comments	0.120	−1,566,017.000	−0.237	−15,909.380	−16,268.730	0.000
	t = 0.704	t = −2.005**	t = −0.605	t = −1.804*	t = −1.964**	t = 0.658
Title per day	0.040	564,093.000	0.064	9,512.453	7,587.047	0.000
	t = 0.685	t = 1.430	t = 0.424	t = 2.219**	t = 2.657***	t = 0.430
Positive sentiment title	−0.071	−46,104.780	0.027	1,292.583	1,338.403	−0.000
	t = −1.599	t = −0.289	t = 0.297	t = 0.479	t = 0.609	t = −0.916
Title text per day	−0.056	−518,079.900	0.135	−9,745.618	−9,443.654	−0.000
	t = −0.711	t = −1.278	t = 0.768	t = −2.221**	t = −2.965***	t = −0.199
Positive sentiment title text	−0.075	126,869.200	−0.025	3,632.756	3,746.200	−0.000
	t = −1.680*	t = 0.374	t = −0.144	t = 0.907	t = 1.042	t = −0.718
Mkt.RF	2.427	1,763,143.000	−0.647	62,389.220	4,672.694	0.000
	t = 2.655***	t = 0.562	t = −0.538	t = 1.719*	t = 0.140	t = 3.998***
SMB	−1.228					
	t = −1.283					
HML	−0.417					
	t = −0.406					
RMW	−2.350					
	t = −1.787*					
CMA	3.749					
	t = 1.690*					
MOM	0.571					
	t = 0.589					
Constant	3.942	118,595,566.000	78.156	1,093,495.000	1,027,160.000	−0.000
	t = 0.310	t = 3.406***	t = 3.098***	t = 2.365**	t = 2.697***	t = −0.147
Observations	44	44	44	44	44	44
R^2	0.693	0.522	0.431	0.626	0.556	0.595
Adjusted R^2	0.574	0.429	0.320	0.553	0.469	0.516

Note: *p<0.1; **p<0.05; ***p<0.01

A1.18 TSM - Taiwan Semiconductor Manufacturing Company Limited

A1.18.1 Full regression

Table A1.74: TSM full regression

	Full regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.013 t = 1.719*	77,879.050 t = 4.158***	0.011 t = 0.990	672.220 t = 5.977***	180.024 t = 3.681***	0.000 t = 0.185
Positive sentiment comments	0.010 t = 2.149**	−12,917.620 t = −1.470	−0.023 t = −1.116	152.157 t = 3.068***	15.918 t = 0.686	0.000 t = 2.131**
Title per day	−0.281 t = −0.833	−1,750,584.000 t = −1.430	−0.088 t = −0.102	−10,990.520 t = −1.767*	−1,245.802 t = −0.385	−0.000 t = −0.961
Positive sentiment title	−0.001 t = −0.070	24,768.020 t = 0.485	−0.037 t = −1.037	79.025 t = 0.299	10.690 t = 0.082	−0.000 t = −0.160
Title text per day	−0.456 t = −1.109	−600,205.000 t = −0.533	−2.168 t = −2.148**	−14,571.310 t = −2.307**	−5,651.885 t = −1.485	0.000 t = 0.878
Positive sentiment title text	0.014 t = 1.212	19,613.780 t = 0.609	0.037 t = 1.226	184.881 t = 1.165	52.035 t = 0.460	0.000 t = 0.526
Mkt.RF	0.974 t = 13.969***	−265,431.400 t = −1.929*	−0.658 t = −1.421	311.857 t = 0.554	−405.931 t = −1.637	0.000 t = 14.451***
SMB	0.283 t = 1.401					
HML	−0.218 t = −0.961					
RMW	0.299 t = 1.263					
CMA	−0.015 t = −0.042					
MOM	0.139 t = 0.991					
Constant	−1.558 t = −1.604	7,487,003.000 t = 2.406**	39.939 t = 12.953***	−3,636.824 t = −0.227	4,823.700 t = 0.575	−0.000 t = −1.122
Observations	296	296	296	296	296	296
R^2	0.506	0.279	0.063	0.403	0.299	0.555
Adjusted R^2	0.486	0.261	0.040	0.389	0.282	0.544

Note: *p<0.1; **p<0.05; ***p<0.01

A1.18.2 Lagged return

Table A1.75: TSM lagged return

	Lagged return					
	Return_{t-1}	Return	Return_{t+1}	Return_{t+2}	Return_{t+3}	Return_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	−0.001	0.013	0.008	0.002	0.015	0.002
	t = −0.146	t = 1.719*	t = 1.340	t = 0.277	t = 2.521**	t = 0.233
Positive sentiment comments	0.004	0.010	−0.001	0.001	0.006	−0.001
	t = 0.603	t = 2.149**	t = −0.187	t = 0.204	t = 0.798	t = −0.074
Title per day	0.108	−0.281	−0.723	0.396	−0.452	−0.149
	t = 0.265	t = −0.833	t = −1.856*	t = 1.043	t = −1.274	t = −0.368
Positive sentiment title	−0.007	−0.001	0.034	−0.004	−0.001	0.011
	t = −0.411	t = −0.070	t = 2.003**	t = −0.277	t = −0.078	t = 0.716
Title text per day	−0.401	−0.456	0.371	−0.645	0.220	0.035
	t = −0.922	t = −1.109	t = 0.647	t = −1.265	t = 0.525	t = 0.062
Positive sentiment title text	0.004	0.014	−0.006	0.012	−0.018	0.004
	t = 0.300	t = 1.212	t = −0.448	t = 0.871	t = −1.252	t = 0.284
Mkt.RF	−0.570	0.974	−0.279	0.203	0.116	−0.083
	t = −4.279***	t = 13.969***	t = −1.848*	t = 1.470	t = 0.908	t = −0.574
SMB	0.349	0.283	−0.345	0.355	−0.591	0.349
	t = 1.213	t = 1.401	t = −1.096	t = 1.148	t = −1.424	t = 1.097
HML	0.055	−0.218	−0.032	0.496	−0.018	−0.187
	t = 0.183	t = −0.961	t = −0.108	t = 1.182	t = −0.061	t = −0.500
RMW	0.175	0.299	−0.136	−0.044	−1.189	0.048
	t = 0.503	t = 1.263	t = −0.402	t = −0.122	t = −3.236***	t = 0.132
CMA	−0.545	−0.015	−0.143	0.107	0.856	0.376
	t = −1.356	t = −0.042	t = −0.254	t = 0.225	t = 1.513	t = 0.604
MOM	0.021	0.139	−0.313	0.478	−0.363	0.116
	t = 0.122	t = 0.991	t = −1.679*	t = 1.380	t = −1.758*	t = 0.385
Constant	0.100	−1.558	−1.366	−0.413	0.400	−0.691
	t = 0.094	t = −1.604	t = −1.115	t = −0.352	t = 0.300	t = −0.638
Observations	295	296	295	294	293	292
R^2	0.148	0.506	0.090	0.089	0.093	0.026
Adjusted R^2	0.111	0.486	0.051	0.050	0.054	−0.016

Note: *p<0.1; **p<0.05; ***p<0.01

A1.18.3 Lagged volume

Table A1.76: TSM lagged volume

	Lagged volume					
	Volume_{t-1}	Volume	Volume_{t+1}	Volume_{t+2}	Volume_{t+3}	Volume_{t+4}
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	47,979.620	77,879.050	41,432.610	27,690.280	16,136.920	14,960.230
	t = 3.095***	t = 4.158***	t = 3.437***	t = 2.215**	t = 1.262	t = 0.923
Positive sentiment comments	−78.803	−12,917.620	−10,394.970	−14,392.610	−16,771.950	−11,897.640
	t = −0.009	t = −1.470	t = −1.034	t = −1.363	t = −1.647*	t = −1.039
Title per day	−1,427,884.000	−1,750,584.000	−1,556,431.000	−1,263,623.000	−123,322.200	139,823.300
	t = −1.572	t = −1.430	t = −1.432	t = −1.140	t = −0.125	t = 0.124
Positive sentiment title	36,265.320	24,768.020	53,186.900	16,744.240	5,984.180	−24,116.120
	t = 0.934	t = 0.485	t = 1.194	t = 0.414	t = 0.142	t = −0.623
Title text per day	−254,115.400	−600,205.000	1,044,765.000	860,383.400	−249,801.100	−586,937.200
	t = −0.302	t = −0.533	t = 0.785	t = 0.589	t = −0.242	t = −0.561
Positive sentiment title text	30,063.980	19,613.780	−4,051.414	6,972.531	17,086.580	26,064.740
	t = 1.127	t = 0.609	t = −0.140	t = 0.203	t = 0.584	t = 0.874
Mkt.RF	−145,134.800	−265,431.400	12,542.830	−88,831.140	−70,764.910	26,411.430
	t = −1.128	t = −1.929*	t = 0.091	t = −0.712	t = −0.650	t = 0.233
Constant	5,909,677.000	7,487,003.000	7,508,321.000	9,478,534.000	9,861,363.000	10,646,010.000
	t = 2.499**	t = 2.406**	t = 2.484**	t = 3.347***	t = 3.315***	t = 3.544***
Observations	295	296	295	294	293	292
R^2	0.125	0.279	0.123	0.055	0.027	0.022
Adjusted R^2	0.104	0.261	0.102	0.032	0.003	−0.002

Note: *p<0.1; **p<0.05; ***p<0.01

A1.18.4 2020 regression

Table A1.77: TSM 2020

	2020 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.017	78,834.550	−0.010	539.814	105.065	0.000
	t = 1.594	t = 2.611***	t = −0.522	t = 4.007***	t = 3.135***	t = 0.063
Positive sentiment comments	0.009	−15,172.250	−0.027	138.990	7.725	0.000
	t = 1.954*	t = −1.820*	t = −1.222	t = 3.055***	t = 0.569	t = 2.099**
Title per day	−0.467	−2,181,133.000	0.444	−9,609.789	−36.570	−0.000
	t = −1.200	t = −1.762*	t = 0.440	t = −1.549	t = −0.013	t = −1.100
Positive sentiment title	0.002	58,856.360	−0.061	131.523	52.137	0.000
	t = 0.125	t = 1.039	t = −1.199	t = 0.446	t = 0.447	t = 0.114
Title text per day	−0.401	−477,248.200	−1.657	−10,399.040	−4,550.619	0.000
	t = −0.889	t = −0.510	t = −1.499	t = −2.164**	t = −1.861*	t = 1.001
Positive sentiment title text	0.011	18,178.960	0.039	174.816	76.200	0.000
	t = 0.983	t = 0.759	t = 1.029	t = 1.327	t = 1.432	t = 0.293
Mkt.RF	0.931	−232,108.800	−0.657	384.684	−221.982	0.000
	t = 11.528***	t = −1.566	t = −1.334	t = 0.806	t = −1.305	t = 13.968***
SMB	0.244					
	t = 1.132					
HML	−0.208					
	t = −0.813					
RMW	0.697					
	t = 2.936***					
CMA	0.203					
	t = 0.502					
MOM	0.174					
	t = 1.062					
Constant	−1.504	5,628,440.000	40.991	−6,966.679	674.214	−0.000
	t = −1.165	t = 1.600	t = 10.631***	t = −0.414	t = 0.107	t = −0.966
Observations	252	252	252	252	252	252
R^2	0.535	0.233	0.066	0.260	0.218	0.560
Adjusted R^2	0.511	0.211	0.040	0.238	0.196	0.548

Note: *p<0.1; **p<0.05; ***p<0.01

A1.18.5 2021 regression

Table A1.78: TSM 2021

	2021 regression					
	Return	Volume	Implied Volatility	Call Volume	Put Volume	AMIHUD
	(1)	(2)	(3)	(4)	(5)	(6)
Comments per day	0.0003	53,364.680	0.001	596.687	161.425	0.000
Positive sentiment comments	0.029	17,080.650	0.028	415.236	40.569	0.000
Title per day	0.599	2,252,990.000	0.705	6,936.177	6,210.327	0.000
Positive sentiment title	−0.011	−89,254.380	−0.036	−433.817	−231.729	−0.000
Title text per day	−2.657	1,033,784.000	−0.494	−15,131.120	16,832.460	−0.000
Positive sentiment title text	0.058	−34,673.440	−0.008	−64.053	−440.973	0.000
Mkt.RF	1.193	−293,469.800	−0.530	2,194.810	−1,615.156	0.000
SMB	0.754					
HML	−0.129					
RMW	−0.581					
CMA	−0.560					
MOM	−0.025					
Constant	−4.524	16,417,791.000	41.610	37,061.010	51,088.130	−0.000
Observations	44	44	44	44	44	44
R^2	0.594	0.419	0.169	0.548	0.593	0.563
Adjusted R^2	0.436	0.306	0.008	0.460	0.514	0.478
Note:						*p<0.1; **p<0.05; ***p<0.01

A2 Minute-by-minute regression

Table A2.1: Minute-by-minute regressions - Individual stocks shorter interval

	Dependent variable:							
	AMC Return	AMC Volume	BB Return	BB Volume	GME Return	GME Volume	NOK Return	NOK Volume
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comments per minute	−0.001	12,977.030	−0.001	6,883.858	0.001	60.929	−0.0002	22,915.930
	t = −1.089	t = 19.049***	t = −0.516	t = 17.439***	t = 2.108**	t = 6.245***	t = −0.134	t = 13.277***
Positive sentiment comments	0.005	−647.227	0.001	294.688	0.004	2,070.337	−0.0001	1,620.056
	t = 1.178	t = −0.315	t = 0.933	t = 1.258	t = 0.487	t = 7.080***	t = −0.293	t = 1.779*
Constant	−0.346	394,497.000	−0.069	121,935.500	−0.536	−80,501.730	0.002	508,016.500
	t = −0.876	t = 2.485**	t = −1.159	t = 6.664***	t = −1.087	t = −4.151***	t = 0.042	t = 6.508***
Observations	2,053	2,074	2,099	2,103	1,997	2,029	1,912	1,916
R^2	0.001	0.213	0.0004	0.269	0.005	0.053	0.0001	0.205
Adjusted R^2	−0.0001	0.212	−0.001	0.268	0.004	0.052	−0.001	0.204
Note:								*p<0.1; **p<0.05; ***p<0.01

In order to better assess what happened during the stocks' peak, we cut the time frame to between January 27 and February 3, 2021. Compared to the tables in 5.3 there are not all too many differences. For the individual regressions, positive sentiment on GME now has a positive relationship with volume with a stronger t-value; 7 against -3,5 for the longer

time frame. The adjusted R-squared is also increased from 3% to 5%, meaning that a larger variation of the stocks volume can be attributed to WallStreetBets sentiment. Both BB and AMC volume models also saw increases in their adjusted R-squareds with an increase of between 2 and 3%. These results might indicate that some people jumped on the bandwagon after reading about the forum in the media or through other sources, or simply that the influence of the forum increased as forum subscribers increased from around 1.9 million on January 19 to around 4.5 million on the January 27.

Table A2.2: Minute-by-minute regressions - Aggregated sample shorter interval

	Dependent variable:	
	Return	Volume
	(1)	(2)
Comments per minute	0.001	841.763
	t = 2.070**	t = 17.323***
Positive sentiment comments	0.001	3,176.969
	t = 0.928	t = 6.234***
Constant	0.010	−829,663.000
	t = 0.182	t = −32.795***
TickerGME	−0.189	−1,258,740.000
	t = −2.205**	t = −43.226***
TickerNOK	0.038	−172,039.700
	t = 0.672	t = −4.329***
Constant	−0.099	857,047.900
	t = −1.296	t = 18.210***
Observations	8,061	8,122
R^2	0.002	0.214
Adjusted R^2	0.001	0.213
Note:	*p<0.1; **p<0.05; ***p<0.01	

On the aggregated regressions we see an increase in correlation between Comments per minute and return, where the variable is now significant on a 5% level. The GME factor on return is also now statistically significant at a 5% level, and even though it has a negative coefficient, given that factor variables can be considered as dummy variables and that GME experienced a larger majority of comments than the other stocks, this should overall signal a positive relationship between WallStreetBets sentiment and GME return.

A3 Event study - excluded stocks

These stocks were placed in the appendix due to the reasons explained below.

Table A3.1: AAPL event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
7/29/2020	1.90%	66.96%	13.63%	1.24%	0.66%	0.51	**No**
7/30/2020	1.20%	72.51%	34.36%	0.49%	0.71%	0.54	**No**
7/31/2020	9.96%	72.60%	56.26%	2.95%	7.01%	5.37	**Yes**
8/3/2020	2.49%	71.97%	38.26%	1.94%	0.56%	0.43	**No**
8/4/2020	0.67%	60.67%	24.23%	0.48%	0.19%	0.15	**No**

AAPL had a stock split the day with the most mentions, thus the abnormal return can be explained by other factors.

Table A3.2: AMD event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
1/3/2020	-1.62%	67.81%	40.13%	-1.01%	-0.62%	-0.25	**No**
1/6/2020	-1.03%	71.24%	46.95%	0.79%	-1.82%	-0.72	**No**
1/7/2020	-0.89%	56.77%	53.02%	-0.23%	-0.66%	-0.26	**No**
1/8/2020	-1.47%	62.54%	26.18%	1.01%	-2.48%	-0.98	**No**
1/9/2020	1.76%	66.57%	23.29%	1.15%	0.61%	0.24	**No**

AMD did not have any significant stock movements during the event period.

Table A3.3: GOOG event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
4/24/2020	0.23%	54.55%	2.34%	1.09%	-0.85%	-0.69	**No**
4/27/2020	-0.27%	63.64%	0.86%	0.66%	-0.93%	-0.75	**No**
4/28/2020	-3.36%	46.88%	6.55%	-1.01%	-2.35%	-1.90	**No**
4/29/2020	8.38%	59.40%	4.67%	3.62%	4.76%	3.84	**Yes**
4/30/2020	0.53%	62.50%	0.85%	-0.27%	0.81%	0.65	**No**

GOOG did not have too many mentions overall during the time frame we explored. However, a note can be taken on the significance the day after the most mentions.

Table A3.4: NIO event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
11/11/2020	3.62%	65.52%	32.57%	1.66%	1.96%	0.34	**No**
11/12/2020	11.44%	62.37%	53.82%	0.24%	11.19%	1.93	**No**
11/13/2020	-8.06%	57.17%	67.58%	1.28%	-9.34%	-1.61	**No**
11/16/2020	2.26%	57.86%	54.67%	0.75%	1.52%	0.26	**No**
11/17/2020	2.19%	59.68%	26.89%	1.46%	0.74%	0.13	**No**

NIO did not have any significant stock movements during the event period.

Table A3.5: NKLA event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
6/4/2020	-28.61%	60.59%	45.60%	-1.27%	-27.34%	-2.70	**Yes**
6/5/2020	6.37%	62.32%	15.18%	2.87%	3.50%	0.35	**No**
6/8/2020	71.15%	60.27%	56.16%	3.76%	67.38%	6.65	**Yes**
6/9/2020	8.45%	58.51%	52.41%	-4.87%	13.32%	1.31	**No**
6/10/2020	-20.41%	58.80%	21.57%	-4.66%	-15.75%	-1.55	**No**

We opted to not include NKLA as the company had an IPO on June 4, 2020, thus the abnormal returns can be explained by post-IPO volatility.

Table A3.6: NVDA event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
5/19/2020	0.63%	67.48%	16.84%	-0.27%	0.90%	0.50	**No**
5/20/2020	1.85%	75.30%	13.00%	1.76%	0.10%	0.05	**No**
5/21/2020	-2.20%	59.08%	32.54%	-1.36%	-0.83%	-0.46	**No**
5/22/2020	2.82%	52.81%	24.85%	1.24%	1.58%	0.88	**No**
5/26/2020	-3.48%	71.67%	8.42%	-2.31%	-1.17%	-0.65	**No**

NVDA did not have any significant stock movements during the event period.

Table A3.7: TSM event study

Date	Return	Sentiment	Mentions	Expected Return	Abnormal Return	AR-test	Significance?
1/14/2020	-0.17%	89.36%	1.54%	0.06%	-0.23%	-0.11	**No**
1/15/2020	-3.85%	79.71%	3.59%	0.49%	-4.35%	-2.00	**Yes**
1/16/2020	0.01%	64.67%	10.96%	1.06%	-1.04%	-0.48	**No**
1/17/2020	-0.89%	63.41%	5.74%	0.16%	-1.05%	-0.48	**No**
1/21/2020	-1.18%	69.23%	1.28%	-0.36%	-0.82%	-0.38	**No**

TSM did not have any stock movements during the event period that we deemed of interest.