
The Gamification of Financial Markets

Effects from Robinhood, Reddit, and the Gamestop Short Squeeze

MIDN 1/C Francis R. Carlos
SA475E: Research Seminar in Quantitative Economics

Abstract

This paper assesses the impact of social media traction, particularly through Google searches and user signup on the internet community known as Reddit, onto the closing price of various “meme stocks”. Meme stocks are defined as stocks affected by the initial Gamestop short squeeze, which occurred in late January of 2021. Using a General Additive Model, the data and regressions demonstrated that user activity in specific subcommunities, known as Subreddits, actually had a statistically significant detrimental effect on meme stock closing prices. In particular, this community is the Subreddit called *r/wallstreetbets*. *r/wallstreetbets* is a Subreddit notorious for its irrational and impulsive user base, which commonly invests disproportionate amounts of their savings into stocks which have already begun increasing in price, or out-of-the-money call options dangerously close to expiration. As nearly all other Subreddit activity demonstrated a statistically significant positive correlation to a stock, it can be argued that online retail investors that took the time to educate and broaden their investing scope proved beneficial to a meme stock’s performance, while retail investors who were interested in *r/wallstreetbets* did not. Also assessed were the Subreddit performances on log relative return and volume, which all demonstrated similar consistent patterns. A look into the anomaly of *r/wallstreetbets* demonstrates a clear example of retail investors acting irrationally, against expected market efficiency.

Introduction

Beginning roughly January 13, 2021, the Gamestop stock had spiked in both price and volume. Seen in several other securities such as AMC, BlackBerry, and Nokia, these stocks went through what is known as a short squeeze. Before this, hedge funds had been profiting on the steady decline in these stocks through a method known as short selling. They would borrow stocks from others, sell them, and purchase them again at a lower price, thus making profit off of the differential. The entire time, financial leaders had been broadcasting the decline of Gamestop on social media, encouraging others to short the stock as well. Unbeknownst to them, individual investors, known as retail investors, had a different goal in mind.

A retail investor is defined as “a non-professional investor who buys and sells securities”. Their portfolio is much smaller than institutional investors, and typically they invest without a formal education. These investors open brokerage accounts through established services, and invest their own capital for the sake of their own personal financial gain. Since the late 2010s, the surge in online brokerages has made retail investing increasingly accessible, especially to younger audiences. Especially crucial was the advent of smartphones, which allowed mobile platforms for investing to develop. One platform in particular seemed to dominate the industry: Robinhood. Founded in 2013 by Vlad Tenev and Baiju Bhatt, Robinhood stood out by being one of the first brokerages to offer commission free trading, which had initially imposed a daunting barrier to entry for those interested.

As more people began to invest exclusively online, naturally, internet communities sprang up to discuss potential strategies and noteworthy stocks. One such community that became particularly active and dedicated was on the website Reddit. Reddit is an online forum, where users can post text articles or photos. In response, other users comment on and deliberate

the posts. Specific subjects and topics are talked about in sections known as Subreddits.

Subreddits dedicate themselves to talking about a particular video game, sport, or in this case, financial market. One such Subreddit is known as *r/wallstreetbets*. Describing itself “Like 4chan found a Bloomberg Terminal”, wallstreetbets users exclusively place trades that are unnecessarily risky or volatile, such as investing their entire life savings into an out-of-the-money call option very close to expiration. Ironically enough, when users experience significant losses, especially their entire investment or trading portfolio, they remain eager and proud to share this with the community, to be lauded by others.

Given their reckless and impulsive trading habits, wallstreetbets users did not appreciate the ongoing establishment toting Gamestop as a stock to be shorted. The inherent nature of short selling depends upon the underlying asset to lose its value, with no potential to rise in the near future. Therefore, by rallying together to collectively purchase the stock, as well as out-of-the-money call options, theoretically their sheer numbers should be enough to drive Gamestop in the other direction, thus resulting in significant gains for retail investors, and significant losses for established hedge funds.

The purpose of this research is to quantify the potential effect, through stock closing price, found in select securities that experienced this phenomenon, and regress it on user traffic based primarily on Reddit, in order to quantify the potential effect observed. Multiple regressions will be made, one for each stock, and one including both the Google search data and one without. A regression will also be performed with the stock relative return, in attempt to bypass the potential random walk prevalent in stock data. The dataset will include stock data, acquired from the NASDAQ website, and will be using simple linear regressions to model the closing price and log relative return. Ultimately, the regressions will attempt to prove if there was a quantifiable

effect from the trends reflected on various Reddit subreddits onto the stock performances, and if this can provide any context to market implications on retail investors, social media, and market efficiency.

Literature Review

Many academic papers view the Gamestop short squeeze as a prime example of counter-hegemony: a collective resistance to confront the current financial status quo. If retail investors are David, then hedge funds are Goliath. Dr. Usman W. Chohan notes that Citron Research and Melvin Capital were “aggressively short-positioned to crush Gamestop’s stock”. Before this, there was no true incentive for retail investors to push towards this short squeeze. There is no evidence or mention of any collective of individuals who had been continuously losing money due to Gamestop short selling. For that reason, the motivation behind the convergence on Gamestop must have been to simply fight the establishment. Andrew Left, the financial analyst mentioned earlier who had been encouraging retail investors to also short Gamestop, was in fact writing as an author for Citron Research at the time. Ultimately, because of the short squeeze, he inevitably stepped down from his position.

Due to the primary usage of social media as a tool to propagate the Gamestop sentiment, the question arises as to whether or not this form of crowdsourcing efforts is legal. By definition, stock-price manipulation is simply illegal in the United States. At the same time, a quick visit to *r/wallstreetbets* shows millions of users commenting nothing but repeated rocket ship and diamond emojis, chanting “We like the stock.” There is no coordination, or collective planning to continue to buy more shares of any of these stocks. That being said, many users have made posts on *r/wallstreetbets* claiming the Subreddit convinced them to bolster their portfolio with these

stocks. Could it then be considered a form of legal manipulation, as protected by the 1st Amendment?

While collective user input on social media garners questions about legality, democratizing investment research is not new. Unlike institutional investors, with both formal training and analytic tools to assess the performance of financial markets, retail investors are left to their own devices. This leaves them dependent upon what current event publications they can find on the markets they desire to invest in. Just as there are mainstream sources, such as large newspapers or economic magazines, there are smaller, crowd-sourced websites that offer insight. One such website that has previously existed and was popular among retail investors is Seeking Alpha. Seeking Alpha sets itself apart from mainstream media by allowing unaffiliated authors to publish their content on the website, and be monetized by the amount of traction their articles gain, as opposed to a formal system of employment. These “technological innovations have helped democratize access to investment research with important implications for firm liquidity”, according to Michael Farrell, Clifton Green, Russell Jame, and Stanimir Markov.

Farrell, Green, Jame, and Markov in 2018 performed their research into determining the extent Seeking Alpha articles correlated to investment activity. As it turns out, Seeking Alpha coverage is negatively related to institutional ownership and positively related to the number of unique shareholders in a particular stock. Also, days when articles were published did in fact produce a greater trading volume. These findings prove that articles specifically catering to retail investors can prove to have a significant impact on their subsequent trading habits. The democratization of investment research has helped improve market liquidity, which is consistent with a reduction in information asymmetry between retail investors and institutional investors. That being said, the only difference from Seeking Alpha and Reddit is the level of user

engagement. With Seeking Alpha, only supposed leaders in these groups have the opportunity to share their comments and analysis. Articles to read and comment on available on Seeking Alpha differ little from *r/wallstreetbets* posts.

With the onset of the COVID-19 pandemic, people around the world were confined to only their homes, and activities that could be done in isolation. Evidently, one such activity that gained increasing popularity was retail investing, especially on the Robinhood platform. From 2018 to 2020, Robinhood offered an application programming interface (API) for software engineers to extract data on their website from its users, anonymously. This data was consolidated onto a website known as Robintrack. Ivo Welch performed an analysis on this data, in order to get an aggregated sense of how Robinhood users invested and what kinds of returns they were seeing. Welch was able to conclude that Robinhood investors purchased their stocks systematically, often times when stock prices increased or decreased greatly. Purchases of shares were made only in reaction to price, as opposed to being speculative in nature.

While few other examples of short squeezes exist in history, The Volkswagen short squeeze is another close example to the one on Gamestop. Apparent in this previous instance of a short squeeze is, once again, the importance of the question of market regulation. More importantly than Reddit, however, is the regulations of brokerages themselves, as opposed to retail investor discussion. At the height of the short squeeze on January 28, Robinhood, as well as other brokerages such as TD Ameritrade and Charles Schwab, halted the trading of both shares and options on stocks such as Gamestop, AMC, BlackBerry, citing unprecedented volatility on the market.



Figure 1. Tweet from Senator Alexandria Ocasio-Cortez regarding the decision of Robinhood to halt trading of selected stocks.

Source: Ocasio-Cortez, Alexandria. Twitter Post. January 28, 2021.

This decision was met with massive backlash from both their consumer base and U.S. policymakers. The potential effects on stock returns from this ban would also be worth researching, to determine how much of a loss it had put retail investors at. Going back to the Volkswagen short squeeze, Franklin Allen, Marlene Haas, Eric Nowak, and Angel Tengulov focused their model on the implications of Porsche's acquisition of Volkswagen, and evidence of the short squeeze based on potential manipulation there.

When Porsche had announced their intent to acquire a stake of one fifth in Volkswagen, this decision was not publicly announced. Part of this acquisition involved “the building up of derivative positions consisting of a synthetic combination of cash-settled call and put options” in order to fund their takeover. As rumours of this acquisition began to surface, Porsche had consistently denied their investment strategy. This was at the turn of the 2008 financial crisis,

and as their short position built up, many investors plunged into Volkswagen, bringing it to the most valuable company in the world by market capitalization. The authors' research focused primarily on the effects on the stock price based on these manipulative effects. Because their evidence supports the idea that this subterfuge on Porsche's side led to a short squeeze, further research into the effects of Robinhood's trading halt would also potentially prove necessary.

Ultimately, previous literature on relevant topics focus on the advent of retail investing as increasingly accessible (especially online, with the advent of COVID-19) and potential legality concerns with short squeezes. Furthermore, short squeezes also raise the question about market efficiency. In theory, when people act rationally, short squeezes should not work, as members of the financial markets would not be interested in buying or holding the stock. That being said, if the short squeeze were to in fact happen, those who induced acted irrationally, against how a typical member of the market would.

Data

A myriad of data was available from several sources. The opening and closing price of stocks (by ticker) GME, AMC, BB BBW, NOK, and PLTR were all acquired directly from the NASDAQ website, exported in spreadsheet format. Another variable offered will be the Google search data for each of these stocks. This is provided by the Google Trends service. It gathers the search trends for the particular Google search term, in this case each of the stocks, and puts it on a scale of 1 to 100. 100 is defined as its most searched period, while a score of 1 indicates little to no searches for the term. Finally, there is data for the subscriber counts for each of the subreddits. This was provided by the website frontpagemetrics.com upon request. It is simply the number of registered Reddit users subscribed to each subreddit by the end of a given day.

The Reddit subscriber count data was used in two ways. First of all, it was taken at face value, as total subscriber count. Also, it was used by counting the new subscribers for that day alone: taking the new subscriber count for that day and subtracting from it the day prior. The stock data was used both from simply its closing value, to its log relative return, to be defined later in the model section.

The only potential complication with the data would be the difference in scale for the subscriber count data versus the Google search data. While the subscriber count data from Reddit is simply in users, the Google search data is in relation to the specific search term, and could potentially vary when using it for separate terms.

Theoretical Model

Initially, one can simply observe the results of the short squeeze by taking a look at the closing price of the stocks leading up to the event. Figure 2 depicts the closing price of Gamestop shares, while Figure 3 depicts the closing price of other shares alleged to have also been affected by the Gamestop short squeeze. There was no formalized method in deciding which stocks were chosen to be assessed; the other five stocks were selected based on general Reddit sentiment. It can be seen clearly that on January 27, the price had soared in each stock to varying degrees. That being said, the effect of the trading halt the day after is also seen, as limitations to retail investors led to a decline of equal magnitude.

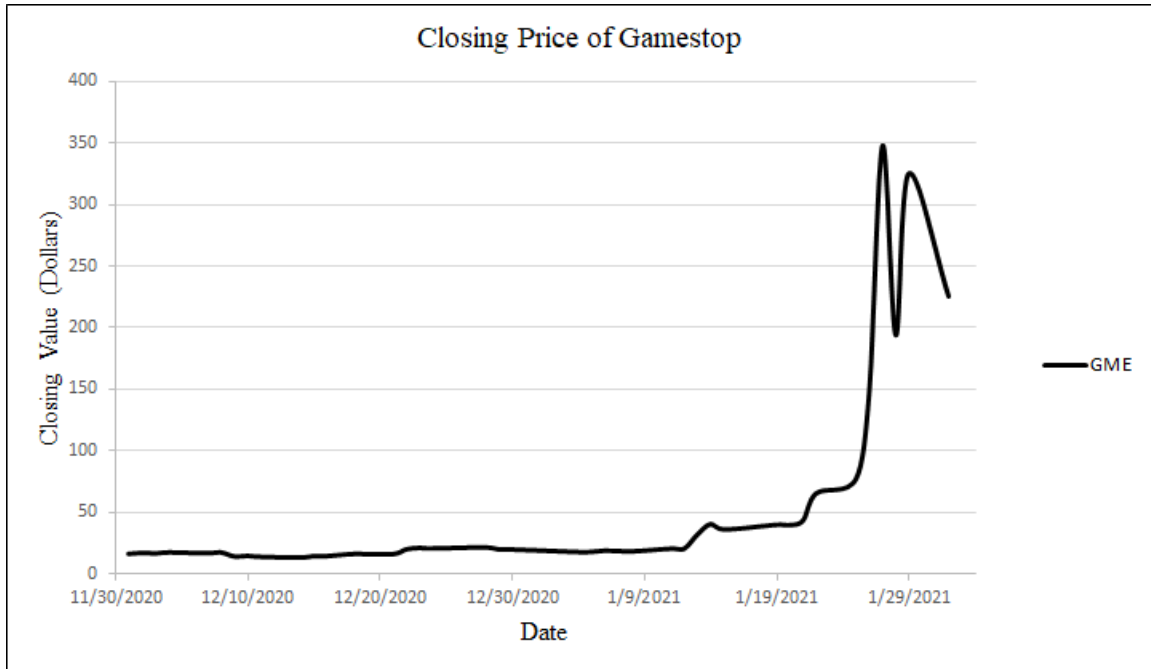


Figure 2. Closing price of Gamestop over time.
Source: NASDAQ.

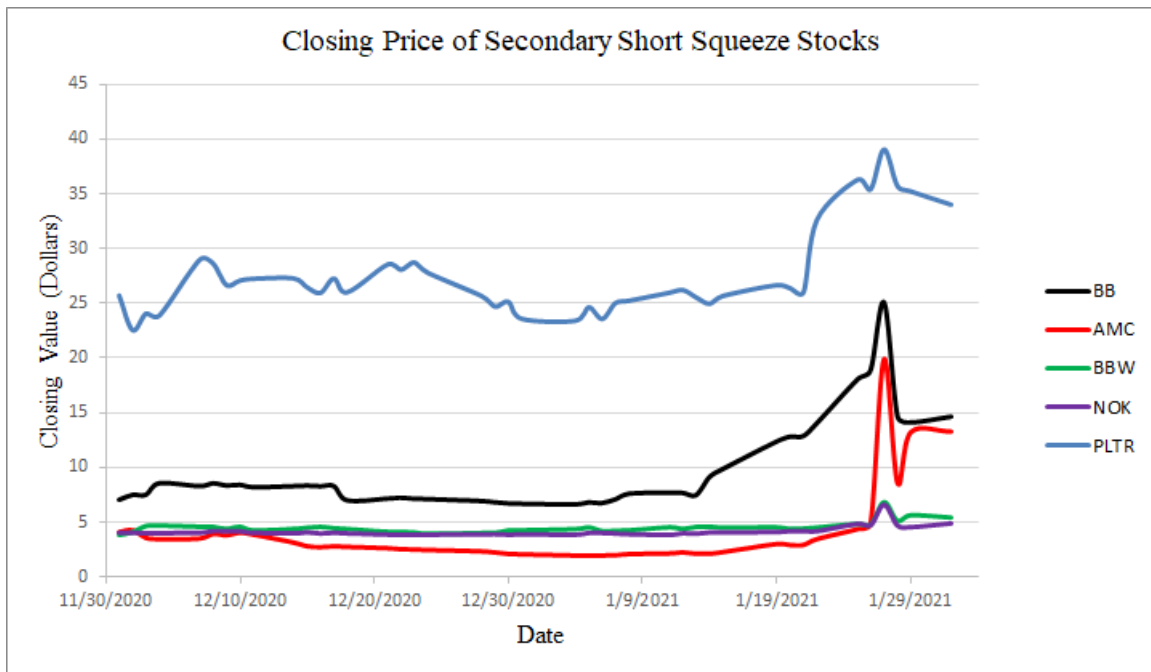


Figure 3. Closing price of other stocks affected by the short squeeze.
Source: NASDAQ.

The model is designed to attempt to quantify the relationship between changes in a stock's closing price, and changes in online engagement of the topic itself. The primary variable regressed will be the closing price of the stock to be assessed. Initially, the daily log relative return for stocks was considered. The log relative return of a stock is defined as

$$\text{Loading} \dots$$

Where $\text{Loading} \dots$ is the closing price of a stock, and $\text{Loading} \dots$ is its opening price. Intuitively, as short sellers make profit off of the stock value decreasing, a negative value indicates that hedge funds are making more gains, and retail investors are at a loss. Conversely, retail investors profit off of the increase of the stock, so a positive value denotes a gain.

This will be first regressed on the closing price of Gamestop itself. The desired effect is intended to be captured exclusively through the change in social media traction.

One of the more easily assessed datasets was the trend in Google searches for the relevant topic. Google trends are provided by the search engine itself, and are based on a scale from 0 to 100. A trend score of 100 indicates its peak popularity in relation to itself in the given time frame, and a 0 indicates little to no popularity at all. In theory, as more Google searches sprang up about the specific topic, more retail investors should have bought into the stock, thus increasing the price.

Aside from Google Trends, the user base for specific Subreddits is also considered. This is based off of the number of users registered to Reddit who have made the decision to subscribe to these specific Subreddits. These Subreddits include *r/dogecoin*, *r/options*, *r/robinhood*, *r/stocks*, and *r/wallstreetbets*. Ideally, only the active traction of a Subreddit is preferred, as opposed to the total number of users signed up. Unfortunately, this data is not readily available or disclosed by Reddit, nor any Subreddit staff. The user data itself for registered users was

provided by frontpagemetrics.com upon request, and even that was difficult to come by. An alternative approach is to assess the new users who have signed up that day. This will be defined as

Loading...

Where *subs* is the subscriber count for the particular Subreddit *i*, and *t* denotes the day. Models for both total user count and new user count will be considered in order to determine which provides a more accurate representation of the data. In a similar concept to the Google search trends, greater traction in the user base should mean more retail investors got involved in the market, thus driving up the price.

An issue with regressing the stock's closing value is the random walk theory. This argues that the changes in a stock price is independent of its past, and other stocks, and estimating stocks with technical analysis is unreliable. Because of this, two alternative regressions onto Gamestop will be presented. Linear models onto the log relative returns of Gamestop, as well as Gamestop's daily volume will be considered. The log relative return will consolidate the randomness in the stock's performance, instead assessing its performance in relation to the investor. This assumes that they would have to reinvest the earnings into the stocks themselves for the day. To circumvent the potential issue of random walk entirely, the volume of Gamestop will also be the dependent variable in the model. This will make no attempt in predicting the performance of the stock at all, instead how much of it is traded in a given day. This will attempt to prove whether or not social media traction had an effect on the amount traded, which is independent of its performance, thus circumventing the consideration of the random walk theory.

Empirical Model

The overall model to determine the significance becomes

Loading...

Where i denotes the specific stock being assessed, and n denotes whether or not the total subscriber count or new subscriber count is being used. Additionally, the scale of Google searches differs from the total subscriber count. Google search trends were not based in total numbers like the subscriber count was. This differential in order of magnitude may cause a problem in modeling the data. Therefore, additional models that do not include Google search trends, and only include Reddit user data, will also be assessed. The model to regress the log relative return will be reflected as

Loading...

Whereas the model to regress the Volume will be

Loading...

Models for both the log relative return and the volume will also both include regressions without Google, to determine whether or not the difference in order of magnitude affected the significance and coefficients of the Subreddit data.

Following are all of the summary statistics, as well as regressions.

Table 1. Financial Market Summary Statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
GME Closing Price	50.06	79.41497	12.72	347.51
AMC Closing Price	3.980	3.537355	1.980	19.900
BB Closing Price	9.554	4.021337	6.580	25.100
BBW Closing Price	4.513	0.4956616	3.850	6.780
NOK Closing Price	4.171	0.4567403	3.870	6.550
PLTR Closing Price	27.52	3.946891	22.51	39.00

42 total observations.

Table 2. Google Search Summary Statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
GME	8.764	22.89961	0.100	100.000
AMC	9.524	21.35443	2.000	100.000
BB	7.869	19.71599	0.100	100.000
BBW	4.833	21.42499	0.100	100.000
NOK	6.505	18.26226	0.100	100.000
PLTR	56.90	20.46748	33.00	100.00

42 total observations.

Table 3. Reddit Summary Statistics (Total Subscriber Count)

Variable	Mean	Std. Dev.	Minimum	Maximum
r/robinhood	404067	30830.73	380935	563739
r/stocks	407161	171422.6	367937	641236
r/options	1051427	56084.51	938950	1860611
r/wallstreetbets	2082892	1132403	1638439	7686846

43 total observations.

Table 4. Reddit Summary Statistics (New Subscriber Count)

Variable	Mean	Std. Dev.	Minimum	Maximum
r/robinhood	4370.6	14824.75	437	94270
r/stocks	22003	59195.73	1810	298145
r/options	6529	15690.27	633	67250
r/wallstreetbets	144163	441468.4	2340	2067348

43 total observations.

Table 5. Gamestop (GME) Regressions on Closing Value

Closing Value	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	2.999e+00** (8.914e-01)		1.310e+01*** (6.357e-01)	
r/robinhood (Total Users)	-1.113e-03 (8.350e-04)	-3.119e-03*** (6.590e-04)		
r/stocks (Total Users)	1.198e-03* (5.112e-04)	4.547e-04 (5.199e-04)		
r/options (Total Users)	-1.578e-03 (1.367e-03)	2.258e-03* (8.504e-04)		
r/wallstreetbets (Total Users)	-6.858e-05* (3.100e-05)	-4.166e-05 (3.378e-05)		
r/robinhood (New Users)			-3.608e-03*** (9.955e-04)	0.0098735*** (0.0026466)
r/stocks (New Users)			7.980e-03*** (5.389e-04)	0.0012846 (0.0015164)
r/options (New Users)			-1.712e-02*** (1.898e-03)	0.0175106*** (0.0031091)
r/wallstreetbets (New Users)			-8.584e-04*** (6.468e-05)	-0.0008962*** (0.0002281)
Constant	-3.583e-03** (1.065e-03)	-8.361e-09 (5.277e-09)	1.096e+01*** (2.655e+00)	-6.4775304 (8.8790806)
R-squared	.81	0.759	.984	.796

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 6. (AMC) Regressions on Closing Value.

Closing Value	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	1.549e-01*** (3.500e-02)		3.217e-01*** (4.641e-02)	
r/robinhood (Total Users)	2.925e-05 (3.666e-05)	-3.663e-05 (4.088e-05)		
r/stocks (Total Users)	2.790e-05 (2.910e-05)	-2.602e-05 (3.225e-05)		
r/options (Total Users)	-9.262e-05 (6.177e-05)	1.026e-04 (5.275e-05)		
r/wallstreetbets (Total Users)	-4.523e-07 (1.811e-06)	2.096e-06 (2.095e-06)		
r/robinhood (New Users)			-2.482e-06 (1.326e-04)	6.010e-04** (1.507e-04)
r/stocks (New Users)			3.575e-04*** (8.117e-05)	-4.094e-05 (8.635e-05)
r/options (New Users)			-1.373e-04 (1.816e-04)	8.229e-04*** (1.770e-04)
r/wallstreetbets (New Users)			-5.062e-05*** (8.971e-06)	-3.335e-05* (1.299e-05)
Constant	-1.375e-04*** (3.108e-05)	1.834e-10 (3.273e-10)	1.255e+00*** (3.413e-01)	1.689e+00** (5.056e-01)
R-squared	.686	0.532	.853	.666

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 7. BlackBerry (BB) Regressions on Closing Value.

Closing Value	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	1.645e-01*** (1.681e-02)		2.084e-01*** (4.146e-02)	
r/robinhood (Total Users)	-7.297e-05* (2.969e-05)	-2.126e-04*** (4.868e-05)		
r/stocks (Total Users)	1.048e-04*** (5.112e-04)	1.214e-04** (3.840e-05)		
r/options (Total Users)	-1.457e-04*** (3.558e-05)	-3.092e-05 (6.281e-05)		
r/wallstreetbets (Total Users)	-6.334e-06*** (1.371e-06)	-9.404e-06*** (2.495e-06)		
r/robinhood (New Users)			-6.297e-05 (2.672e-04)	1.093e-03*** (1.752e-04)
r/stocks (New Users)			2.247e-04* (9.742e-05)	-6.860e-05 (1.004e-04)
r/options (New Users)			-6.537e-05 (3.523e-04)	1.512e-03*** (2.058e-04)
r/wallstreetbets (New Users)			-2.668e-05 (1.435e-05)	-6.822e-05*** (1.510e-05)
Constant	-7.683e-05*** (7.848e-06)	-1.282e-09** (3.898e-10)	7.518e+00*** (5.219e-01)	6.249e+00*** (5.877e-01)
R-squared	.8853	0.487	.789	.651

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 8. Build-A-Bear (BBW) Workshop Regressions on Closing Value.

Closing Value	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	1.014e-02*** (2.790e-03)		2.426e-02*** (5.659e-03)	
r/robinhood (Total Users)	6.632e-06 (5.562e-06)	2.435e-06 (6.254e-06)		
r/stocks (Total Users)	-9.472e-07 (4.370e-06)	-3.926e-06 (4.934e-06)		
r/options (Total Users)	7.885e-06 (7.738e-06)	1.970e-05* (8.071e-06)		
r/wallstreetbets (Total Users)	-2.070e-07 (2.790e-07)	-1.748e-07 (3.205e-07)		
r/robinhood (New Users)			-4.359e-05 (3.833e-054)	9.877e-05*** (2.321e-05)
r/stocks (New Users)			8.313e-05*** (2.243e-05)	-7.482e-07 (1.330e-05)
r/options (New Users)			-1.960e-05 (4.424e-05)	1.437e-04*** (2.726e-05)
r/wallstreetbets (New Users)			-8.993e-06*** (1.724e-06)	-6.846e-06** (2.000e-06)
Constant	-4.586e-06*** (1.262e-06)	7.556e-11 (5.008e-11)	4.182e+00*** (6.475e-02)	4.146e+00*** (7.786e-02)
R-squared	.578	0.442	.725	.596

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 9. Nokia (NOK) Regressions on Closing Value.

Closing Value	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	1.574e-02* (6.458e-03)		8.585e-04 2.188e-02)	
r/robinhood (Total Users)	8.224e-06 (6.514e-06)	1.866e-06 (6.346e-06)		
r/stocks (Total Users)	4.724e-06 (5.288e-06)	-1.142e-06 (5.006e-06)		
r/options (Total Users)	-7.677e-06 (1.145e-05)	1.298e-05 (8.189e-06)		
r/wallstreetbets (Total Users)	-5.261e-07 (3.175e-07)	-3.191e-07 (3.252e-07)		
r/robinhood (New Users)			1.369e-04*** (3.811e-05)	1.381e-04*** (1.923e-05)
r/stocks (New Users)			-2.106e-05 (3.544e-05)	-2.238e-05* (1.102e-05)
r/options (New Users)			1.716e-04** (5.903e-05)	1.738e-04*** (2.260e-05)
r/wallstreetbets (New Users)			-6.358e-06* (2.968e-06)	-6.262e-06*** (1.658e-06)
Constant	-1.413e-05* (5.795e-06)	4.810e-11 (5.082e-11)	3.826e+00*** (7.254e-02)	3.828e+00*** (6.453e-02)
R-squared	.402	0.324	.664	.674

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 10. Palantir (PLTR) Regressions on Closing Value.

Closing Value	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	8.290e-02*** (1.857e-02)		3.730e-02 . (1.917e-02)	
r/robinhood (Total Users)	-9.283e-05* (4.097e-05)	-3.863e-05 (4.790e-05)		
r/stocks (Total Users)	4.173e-05 (3.300e-05)	-1.037e-05 (3.779e-05)		
r/options (Total Users)	6.510e-05 (1.367e-03)	1.452e-04* (6.181e-05)		
r/wallstreetbets (Total Users)	-4.832e-06* (2.076e-06)	-2.434e-06 (2.455e-06)		
r/robinhood (New Users)			6.765e-04*** (1.773e-04)	7.971e-04*** (1.723e-04)
r/stocks (New Users)			-6.424e-06 (1.005e-04)	-6.927e-05 (9.869e-05)
r/options (New Users)			1.026e-03*** (2.073e-04)	1.162e-03*** (2.024e-04)
r/wallstreetbets (New Users)			-4.720e-05** (1.432e-05)	-4.678e-05** (1.484e-05)
Constant	4.854e-07*** (1.087e-07)	2.666e-10 3.836e-10)	2.269e+01*** (1.182e+00)	2.472e+01*** (5.779e-01)
R-squared	.656	0.484	.674	.649

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 11. Gamestop Regressions on Log Relative Return

Log Rel. Ret.	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	-3.852e-03 . (1.933e-03)		9.626e-03 * (4.444e-03)	
r/robinhood (Total Users)	-4.613e-06* (1.811e-06)	-2.036e-06 (1.316e-06)		
r/stocks (Total Users)	1.262e-06 (1.109e-06)	2.216e-06* (1.038e-06)		
r/options (Total Users)	2.038e-06 (2.965e-06)	-2.890e-06. (6.181e-05)		
r/wallstreetbets (Total Users)	-1.410e-07* (6.724e-08)	-1.756e-07 (6.746e-08)		
r/robinhood (New Users)			-1.407e-05. (1.773e-04)	-4.158e-06 (5.500e-06)
r/stocks (New Users)			1.198e-05** (3.767e-06)	7.060e-06 * (3.151e-06)
r/options (New Users)			-2.605e-05 . (1.327e-05)	-5.997e-07 (6.461e-06)
r/wallstreetbets (New Users)			-8.882e-07 . (4.522e-07)	-9.160e-07 . (4.740e-07)
Constant	4.603e-06 . (2.310e-06)	-2.184e-11 * (1.054e-11)	-2.317e-02 (1.856e-02)	-3.599e-02 . (1.845e-02)
R-squared	.318	0.265	.386	.325

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Table 12. Gamestop Regressions on Volume (By Trading Day)

GME Volume	Total w/ Google	Total w/o Google	New w/ Google	New w/o Google
Google Search History	2306980.8 * (885513.7)		7101429.1 (2697613.8)	
r/robinhood (Total Users)	-2167.5* (829.5)	-3.711e+03*** (6.232e+02)		
r/stocks (Total Users)	3198.2*** (507.8)	2.627e+03*** (4.917e+02)		
r/options (Total Users)	-4996.1*** (1358.2)	-2.045e+03* (8.043e+02)		
r/wallstreetbets (Total Users)	-209.3*** (30.8)	-1.886e+02*** (3.194e+01)		
r/robinhood (New Users)			1049.3 (4224.6)	8359.6* (3429.3)
r/stocks (New Users)			4598.5 . (2286.8)	967.9 (1964.9)
r/options (New Users)			-5722.3 (8053.4)	13053.2 ** (4028.6)
r/wallstreetbets (New Users)			-766.5** (274.5)	-787.0* (295.5)
Constant	-2756.5 * (1058.0)	-2.772e-02*** (4.991e-03)	16787549.7 (11268855.5)	7331451.3 (11505141.8)
R-squared	.55	0.481	.288	.174

42 total observations.

Robust Standard Errors in Parentheses

p<0 '***' p<0.001 '**' p<0.01 '*' p<0.05 '.' p<0.1 ' ' p<1

Results

Given the flurry of regressions, the single most important regression is the Gamestop regression with the highest R-squared value. This is because Gamestop is the primary stock assessed, and the intent is to look for the most representative model for it. Technically, it should be the new user data coupled with the Google data, but ultimately the Google data should have a recommendation to be thrown out. Its variable is so many orders of magnitude below the values for Gamestop, that changes in its value always had a far more profound effect than Reddit user data. It would have been more worthwhile to have the data for the true number of Google searches performed. As such, the regression including only new users by day will be assessed, and this regression's effect will be examined for all other stocks. The other regressions are still included, in order to see how closely the coefficients from the predictors matched the performance on Gamestop's.

Surprisingly, an unintended result emerged from this regression. As it turns out, new subscribers to *r/wallstreetbets* turned out to be detrimental to the closing performance of Gamestop. While this may not be quantitatively proven, the nature of *r/wallstreetbets* inherently involves entirely irrational and uninformed trading behavior. Because the Subreddit is intended for this style of trading, users who only subscribed to *r/wallstreetbets* could have demonstrated poorer trading performance than other more informed users, who took the time to subscribe to a greater variety of Subreddits. If a user would take the time to ensure they were well-informed, and visit these more formalized and traditional communities, then presumably this would reflect in their trading habits.

AMC proved to be the closest in terms of R-squared value to GME. This makes sense, as AMC was the closest in terms of traction on Reddit. The lowest R-squared value, and thus the

stock least affected by the meme stocks from those listed, is NOK with a .324 for the total subscriber count with Google regression.

At least for the regression accounting for new subscriber count alone, all R-squared values denoted that these variables accounted for at least half of the variance in the stock closing value. Also, all of the stocks consistently give *r/wallstreetbets* a negative correlation to their value with statistical, further proving that it is likely that their impetuous theme draws attention to poorly fated investors. The only other Subreddit with a negative correlation was *r/stocks*, which also demonstrated statistical significance only rarely. Both *r/options* and *r/robinhood* had a consistent, statistically significant positive correlation for each of the stocks, potentially implying a more discerning audience.

A potential issue with assessing exclusively the stock closing price is the random walk theory. The random walk theory suggests that a change in a stock's price is independent of other stocks, all sharing the same distribution. Essentially, with this unquantifiable direction, it would be impractical to estimate the stocks itself. That is why the log relative return regression for Gamestop was included. These alternative regressions did not share the same statistical significance, R-squared values, or even coefficient signs as the closing values. It alternatively suggested that there were no statistically significant correlations between social media activity and the returns of these stocks. Most estimated coefficients appeared to be negative, and differed amongst each other when the Google search data was removed. That being said, log relative return of stocks places the assumption that an investor is reinvesting the earnings made on the rate of return to the stock itself. Finally, with such a great standard error in relation to the order of magnitude to the coefficient, there is a significant room for deviation within these terms. When assessing the regressions of GME Volume, less of the deviance was explained, but the

coefficients became far more statistically significant once again, in a similar manner as the regressions on Closing Price. However, there was far more variability in the signs of the coefficients between regressions. The total subscriber count, without Google, demonstrated the most statistically significant coefficients. They explained that *r/robinhood*, *r/options*, and *r/wallstreetbets* all demonstrated a negative effect on the volume. In other words, as more users were already subscribed to these Subreddits, there was a detrimental effect on the volume of GME. This could be interpreted in a similar manner to the closing price regression conclusions, as *r/robinhood* and *r/options* are also considered within the Reddit community to introduce inexperienced and potentially irrational investors. The new subscriber count without Google, assessed in the Closing Value regressions, did not have as much statistical significance, but did share the same pattern of coefficient signs for each predictor, potentially confirming the previous hypothesis.

Conclusion

The short squeeze on Gamestop was an unprecedented financial anomaly for a number of reasons. Not often has the short interest for a stock been so high to reverse the direction of a stock so strongly. Never before has a collective of individuals used their aggregate buying power so remotely to allow such an event to occur. Seldom have brokerages in the past flat out cherry picked which stocks they would like to cease the trading of to deliberately slow their losses. Nonetheless, it still happened, and to the date this paper was finished many are still in the fight, with their unwavering diamond hands.

The advent of mobile investment platforms effectively broke down the last few barriers to entry for any individual to aspire to become a full-fledged retail investor. When commission free trading from Robinhood was introduced, it revolutionized the market, even convincing other

brokerages to adapt their model just to remain competitive. Coupling this with the COVID-19 pandemic, confining the world to their homes led to boredom, and desire for an outlet. Many found this outlet in gains and losses, whether it be solace or sorrow.

Using Reddit as a primary source of variables, the closing price of select meme stocks were chosen in order to fit a generalized additive model to the stocks. Using data from Google search activity was also considered, but given the nature that it was provided, did not prove beneficial to the model. The hypothesis was that there was a statistically significant correlation on stock closing price coming from Reddit user traffic. Furthermore, this correlation was presumed to be positive, as intuitively, more retail investors getting involved would lead to greater market activity and long positions.

Something that was not considered was the potential for a certain community to actually be detrimental to stock performance. Upon finishing and assessing the model, it turned out that *r/wallstreetbets* had in fact a statistically significant negative impact on the stock. At first, this may not seem like the truth; a community rising from under 2 to nearly 7 million users in the span of just a few months should cause enough of a conversation to convince many other retail investors to go long.

The truth can potentially lie in the nature of *r/wallstreetbets* itself, as a community. Members of the community pride themselves on doing the exact opposite of what is expected in a financial market, and relish in both winning and losing it all. From dumping their entire college loan into shares of a single asset, to incredible amounts of out-of-money calls just days before expiration, *r/wallstreetbets* quite simply does not follow the traditional establishment of how one should invest. Therefore, it could be rationalized that an influx of users into *r/wallstreetbets*

alone had a negative correlation to stock returns, due to the intended behavior of those who join the community.

In almost every other case, increased user activity in different Subreddits had a positive correlation. These Subreddits were *r/stocks*, *r/options*, and *r/Robinhood*, far more traditional communities, who do not pride themselves in reckless behavior. When a user should join these communities, as opposed to *r/wallstreetbets*, they are assumed to do a bit more research to some degree, and act more rationally within the market.

All this considered, the log relative return regressions potentially tell a different story, or not one at all. Given their statistical insignificance, as well as the poorness of fit of the models themselves, they argue there is potentially little effect from Subreddit traction to relative return performance. Perhaps a different unit of time should be estimated, or a different approach should be used in attempting to quantify these effects. Using social media data alone was in specific consideration to the random walk theory: the trends of the stocks were not predicted with their past movement, but entirely unrelated predictors to the stocks at all (with the exception of their Google data). Granted, these models are still fundamentally an estimation of stock performance, which still inevitably fall under this fallacy.

In also attempting to model the changes in volume, the issue of the random walk theory was passed over. This is because there was no attempt to regress stock performance, but rather stock volume, indicative of increased traction and activity in GME. If the initial hypothesis was to be correct, then similar increases in different Subreddits should have similar increases in the stock volume as well. The patterns demonstrated in stock volume shared similar coefficient signs as the log relative return and stock closing price, demonstrating that specific Subreddit activity had a consistent effect on closing price, log relative return, and volume.

Future Work

Further questions can be asked on the topic of social media traction having an effect on stock market performance. For example, there are many other platforms that exist, beyond Reddit. These include Twitter, Discord, and other websites that allow for the specific discussion of financial markets. It would be worthwhile to determine if some websites had a more profound effect than others, and if any interaction terms would be significant. Just as important as the short squeeze and spike in price was the subsequent fall shortly after, in response to Robinhood and other brokerages halting the trading of meme stocks on their platforms. Perhaps attempting to find ways to model the losses would also add insight. After all, several class action lawsuits and inquiries are ongoing into Robinhood and others for this decision, and being able to quantitatively prove that retail investors were hurt would aid in this endeavor. Finally, if social media usage was modeled in a statistical learning perspective, it could be determined whether or not social media activity could actually predict the potential direction a stock might go.

On another note, the regressions included only assessed the period of time involving the short squeeze. If the Reddit data available went as far back to include potentially a year prior, the same regressions could be performed on these stocks in a separate, unrelated year in order to potentially determine if there were similar effects on the stocks before the advent of COVID-19 and retail investing. Furthermore, the data can be followed into the future, further into 2021, to determine if the effect of social media was sustained as the traction of the short squeeze began to dwindle. Potentially a General Additive Model would be preferred in this case, as to determine if the effect slowed as time passed.

Investing has now been designed to become as accessible as possible to users. Some might even argue far too accessible. Where millions could be won or lost from anywhere on the

comfort of a mobile device, the very landscape of financial markets are changing. Whether this change is an evolution or a mutation lies yet to be seen.

References

- Allen, Franklin, Marlene Haas, Eric Nowak, and Angel Tengulov. "Market efficiency and limits to arbitrage: Evidence from the biggest short squeeze in history." (2017).
<https://www.ifin.usi.ch/assets/docs/seminars/f-allen-324362.pdf>
- Carney, Michael. "Robinhood Gets \$3M to Take from Wall St. and Give to Main St. with Its Mobile-first, Zero-commission Brokerage." Pando. December 18, 2013. Accessed April 15, 2021.
<https://pando.com/2013/12/18/robinhood-gets-3m-to-take-from-wall-street-and-give-to-main-street-with-its-mobile-first-zero-commission-brokerage/>.
- Chohan, Usman W. "Counter-Hegemonic Finance: The Gamestop Short Squeeze." *Available at SSRN* (2021). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3775127
- Farrell, Michael, Clifton Green, Russell Jame, and Stanimir Markov. "The democratization of investment research: Implications for retail investor profitability and firm liquidity." (2018). http://128.171.57.22/bitstream/10125/59256/HARC_2019_paper_36.pdf
- Hayes, Adam, and Gordon Scott. "Retail Investor." Investopedia. February 17, 2021. Accessed April 15, 2021. <https://www.investopedia.com/terms/r/retailinvestor.asp>.
- Lyócsa, Štefan, Eduard Baumöhl, and Tomáš Vřrost. "YOLO trading: Riding with the herd during the GameStop episode." (2021). <https://www.econstor.eu/handle/10419/230679>
- McElhaney, Alicia. "Amid the GameStop Frenzy, Andrew Left Calls It Quits On Short Reports After 20 Years." *Institutional Investor*, January 29, 2021. Accessed April 15, 2021.
<https://www.institutionalinvestor.com/article/b1qbfm5tftymfw/Amid-the-GameStop-Frenzy-Andrew-Left-Calls-It-Quits-On-Short-Reports-After-20-Years>
- Morrow, Allison. "Confused about This GameStop Saga? Here Are the 5 Things You Need to Know." CNN Business. January 30, 2021. Accessed April 15, 2021.
<https://www.cnn.com/2021/01/30/business/gamestop-reddit-rebellion-explained/index.html>.
- Ocasio-Cortez, Alexandria. Twitter Post. January 28, 2021, 11:36 AM.
<https://twitter.com/AOC/status/1354830697459032066?>
- "Wallstreetbets." Reddit. Accessed April 15, 2021. <https://www.reddit.com/r/wallstreetbets/>.

Welch, Ivo. "The Wisdom of the Robinhood Crowd." *NBER Working Paper* 27866 (2020).

https://www.nber.org/system/files/working_papers/w27866/w27866.pdf?source=content_type%3Areact%7Cfirst_level_url%3Aarticle%7Csection%3Amain_content%7Cbutton%3Abody_link