



Università
Ca' Foscari
Venezia

Single Cycle Master Degree in Finance

Double Degree in Financial Analytics with Stevens
Institute of Technology University

Final Thesis

GME: A PUMP AND DUMP SCHEME?

Supervisor

Ch. Prof. Monica Billio

Abroad activity Supervisor

Dr. Ionut Florescu

Graduand

Francesco Forner

Matriculation Number 882030

Academic Year

2020 / 2021

Dedicated to my family.

Contents

1	Abstract	5
2	Literature Review	6
3	Introduction	7
4	REDDIT r/Wallstreetbets	8
5	Data Pooling	9
5.1	Reddit PUSHIFT script	9
5.1.1	Lack of data	9
5.2	Stock stevens lab: tick level data	10
5.2.1	Lack of data	10
6	SENTIMENT ANALYSIS	11
6.1	Preliminary idioms analysis	11
6.2	Sentiment methodology	12
6.3	Primary results	12
6.3.1	What is the relation between the sentiment of the subreddit and GME price trend?	12
6.3.2	Which neural language processing dictionary is the best representation of r/Wallstreetbets Reddit group?	14
6.3.3	What is the dominance in word used?	14
6.4	Results	15
7	FAMA AND FRENCH	18
7.1	The three (plus one) factors model.	18
7.1.1	Is GME event related to economic reason?	19
7.2	Results	19
8	THE TRIGGER	21
8.1	MARKET MANIPULATION	21
8.2	"Pump and Dump" Fraud	21
8.2.1	Is the market manipulation the trigger of GME event?	21
9	DISCORD: The use of a social gaming platform to perform fraud	22
9.1	Discord	22
9.2	Pump and dump Groups	22
9.3	The GME "CASE"	24
10	CONCLUSION	27

Appendices	28
A “SentimentAnalysis” RStudio package (Nicolas and Proellochs 2021)	28
B "WordCloud"	28
B.1 Common words before 01-28	28
B.2 Common words after 01-28	29
C Damped Harmonic Hoscillator motion	29
D Appendix : Discord scam group netiquette	29
D.1 The Invite Chain	29
D.2 PUMP and DUMP details	31
D.3 Social media guide	32
E Bid - Volume	33

1 Abstract

At the end of January 2021, an important page in financial history was written for the records as a moment in which private investors attacked hedge funds. The operation "headquarter" was Wallstreetbets, a forum of people looking for high return investments and it played a central role in the GME rally. From an economical perspective, the main reason of the exponential increase of the stock price is Short-squeeze; to make it happen, what is need is pressure from the market. This research focused on finding a relationship between activity on the Reddit platform and the change in GME stock price. Many different analysis were used, starting with a Sentiment analysis measuring the similarity beten different NLP dictionaries and verbiage within the Reddit platform. Lastly, a Fama and French model was run to find whether or not, there was an economical reason behind the return amounts for this event. Relevant data was found from all the tests conducted and the result led to believe that this event was a pump and dump scheme.

2 Literature Review

The dictionaries proposed by the ‘Package “SentimentAnalysis”’ presented a multiple way of find the sentiment analysis. This help to understand different way on how the people communicate. By the use Nicolas2021 “Stock Returns and Investor Sentiment: Textual Analysis and Social Media”, a better understating of how to use textual research on social media messages could be implemented. The investigation present the relationship between investor sentiment and stock returns.

Furthermore, for a behavioral analysis point of view, the reading provided by Mamun2015 helped distinguish rational investors from the irrational. All the methods and the implementation of correlation and the Cross correlation concepts are studied in both “An Efficient Method for Computing Local Cross-Correlations of Multi-Dimensional Signals.

Mean reversion and harmonic oscillator ideas came from the reading of Junior2011, where a relation between concepts from the physic are applied to the stock market.

The economical evaluation conducted by Dhaoui2017, showed how the sentiment an other factor can implement a Fama and French model.

Regarding the market manipulation analysis, LaMorgia2020 research help to understand how scam groups are settle down and operated.

All the declaration about the reasons why rWallstreetbets have been banned have been analysed from the reading of website article “Discord Bans rWallstreetbets Server over Hate Speech as the Group Drives GameStop Shares through the Roof.”¹

All the information in the Appendix D derive from online groups²

From the combination of all this technique and analysis, the focus will try to demonstrate a manipulation over Game Stop Corp. during its epic stock run at the end of January.

1. Sonnemaker, Tyler. “Discord Bans r/Wallstreetbets Server over Hate Speech as the Group Drives GameStop Shares through the Roof.” Business Insider. Business Insider, January 27, 2021 <https://www.businessinsider.com/discord-bans-wallstreetbets-server-hate-speech-reddit-gamestop-gme-2021-1>

2. The decision it is to not cite any reference

3 Introduction

By looking at the last five years of the price history for Game Stop Corporation, it is easy to notice that something unique happened after a very long time with no activity.

While GameStop shares did not exceed \$17 at the beginning of the year, in the last week of January the stock rose to \$347 with an increase by 860%. Everyone is wondering what caused this sudden and impressive surge. Although this phenomenon would only have been possible thanks to the intervention of the big names in speculation, in this case, the rise occurred for the action of thousands of small investors.

Two are the actors: hedge funds and the Redditors. The former are speculative funds managed by insurance companies or private entities to obtain the greatest profit with their operations, while the latter are individual investors. This all began when some institutional investors, including the hedge fund Melvin Capital, saw the opportunity to profit from short-selling on GameStop shares. This investing technique is a trading strategy that aims to anticipate the drop in the price of a share or other asset listed on the markets. The investor opens a position by borrowing the shares of the stock, which are believed to decrease in value in a future date. Then, the same person will sell these borrowed shares to other investors willing to pay the market price; before the borrowed shares have to be repaid (on the future date set), the investor bets that the price will continue to fall and that the shares can be purchased at a lower cost.

Some users on Reddit – who gathered on the "r/WallStreetBets" forum - decided to buy GameStop shares in bulk (through direct purchase or options) initially because they thought the title was undervalued, but later on the goal became to send a strong and clear message to short sellers. They pushed the GameStop share price to very high levels by acting altogether, causing huge losses to short sellers.

At the end of January, due to a rally over the stock, a short squeeze impacted GME's Short-seller traders that tried to close their short positions as quickly as possible to reduce their losses. A short squeeze indeed occurs when a market price suddenly rises, exceeding the expectations of analysts and investors. However, most of the time, it impacts investors who adopt the short-selling strategy with borrowed shares as they may spend more money to pay off and deliver the borrowed stock than expected. The result of this rare event caused an additional boost to the stock price of Game Stop.

Did private investors and Redditors decide together to purchase GME stock at the same moment? Were there a few of them that already knew about what was going to happen in the stock market and chose to manipulate naïve investors? This paper will attempt to show a relationship between social media scam activity and GM stock price, therefore, demonstrate this event could have been premeditated.

4 REDDIT r/Wallstreetbets

Reddit is a social news aggregator, a mixture of a discussion platform and link distributor divided into subreddit forums. Users are called "Redditors," they are content creators, consumers and curators. Reddit uses a point rank system consisting of upvotes and downvotes; the community determines which content and discussions are important (these will appear at the top of the page). The platform is mainly self-managed; the moderators act as managers of the forum. Reddit identifies itself as "The Front Page of the Internet".³

On the front page you can find the most important trending posts, news, and topics of the day just like the front page of a newspaper. Content is controlled by the "Redditors" who choose what appears on the "front page." The platform collects social news in an impressive number of subreddits controlled by the users themselves. Typically, a forum deals with a central topic and breaks it down into sub-topics. In this regard, there are directories in subforums. Within the subreddits users make new posts about a particular topic. Humorous and creative contributions are often encountered here, but there is also a presence of controversial discussions.

In this paper, it is taken into consideration r/WallStreetBets. This subreddit is populated by people looking at the market without fundamental analysis. They are in trying to find the riskiest bets that can have a huge payoff. In this community, they share their ideas and thoughts about trades, their stock goals and investments that are high-risk high-reward. The "GameStop Stonk" event is a clear example of how this subreddit can show interest. r/Wallstreetbets in that period started going all-in on GameStop and a few other meme stocks, causing a significant shock in the market.

3. "reddit: the front page of the internet". reddit.com. May 15, 2021. <https://www.reddit.com/> .

5 Data Pooling

This analysis focuses on finding a relationship between r/Wallstreetbets and Game Stop stock prices with social media interactions. This analysis will be focused around these two data sets:

- Reddit Post of r/Wallstreetbets subreddit.
- Prices for Game Stop Corp.

5.1 Reddit PUSHIFT script

With its official API⁴, Reddit allows all the users to access the historical database records on the platform. The procedure is straightforward, after requesting the token security key it is possible to send a download request for a specific subreddit. However, this procedure lacks efficiency. Reddit limits downloadable data by post filter. Reddit posts are divided in these 4 subcategories:

- "Hot" = all the posts that recently got a consistent number of upvotes/comments.
- "New" = the newest posts at the top of the page; posts are sorted by the time of submission.
- "Rising" = represent all the publication rising in term of activity (comments/upvotes) at the moment you are looking at.
- "Top" = posts that have received the most upvotes over a set period.

Here after, further details re downloaded for creating a complete dataset:

- Comments numbers per post.
- Images or not = if the post has an image or not.
- Pinned or not = was the post in evidence.
- Link: to prove the existence.

5.1.1 Lack of data

The Reddit dataset review showed some relevant insights. Even if a large pool of users worldwide are part of the r/Wallstreetbets subreddit, there are some time frames in which nobody posts anything and others that have multiple publications.

Looking more closely at the data, it is easily noticeable that there are no data for the timeslot from 01/26/2021 (Time 4:29:43 pm) to 01/27/01 (Time 2:56:36 pm). Thus, even though the group went private on Wednesday for less than one hour, the Pushift database did not have any posts available for download in that period.

4. "API documentation". reddit.com. May 15, 2021. <https://www.reddit.com/dev/api/>.

5.2 Stock stevens lab: tick level data

As mentioned above, the r/Wallstreetbets group aimed at some particular types of "meme" stocks with fundamental characteristics. For this particular analysis, for this particular analysis the focus was on the most popular one: Game Stop Corp. The tick first-level data, with frequency by second, for GME, was downloaded from the Thomson Reuters servers:

- Time of trade
- Tick name
- Bid Price
- Bid Volume
- Ask Price
- Volume
- Volume of trade

5.2.1 Lack of data

The data-set corresponding to the tick level data presents missing values related to the market hours. From 1:00am – 9:00am and weekends there are no prices and volumes available for Exchange-traded markets because they are closed during this time.

6 SENTIMENT ANALYSIS

The sentiment analysis in this research focuses on the relationship between the sentiment trend of the subreddit and the price of GME if there is a dominance in terms of word frequency, and what language type could be mainly associated to the Reddit community.

6.1 Preliminary idioms analysis

Social media has changed how a society communicate in the modern age. Sarcasm, emojis, gifs, abbreviations and acronyms are the common way to interact over online platforms.

In Reddit, this new "social media idioms" is emblematic. While evaluating r/Wallstreetbets, it is easy to notice how the members talk to each other. The majority of the posts are a combination of words and emoticons that sometimes could look meaningless but they are not, and the reason behind is simple: they want to create a unique community. A simple visitor will never fully understand all the posts and jokes of the community unless she is a member. Thereby using a sentiment analysis all the specific words or emoticons look meaningless even if they playing an important role in the discussion. In this particular case, some of the recurring ones are:

- Stonks: It is the incorrect spell of "stock" by purpose. If a loss is made, it's often used as a joke word. However, when used in other ways, such as with the side eyes emoji (stonks), it denotes a stock that is worth to monitoring.
- Apes together strong 🐒: a way to show support for other investors working for the same target.
- Rockets 🚀: indicates speculative stocks with high growth potential.
- Paper hands 📄👉: Paper Hands occurs when a trader sells his or her stocks and shares well before reaching their maximum profit potential. On r/wallstreetbets, the words "Paper Hands" and "Weak Hands" are synonymous.
- Diamond hands 💎👉: Having "Diamond Hands" refers to a trader who is willing to hold on to their stocks or securities for an extended period of time. They are so confident in the potential success of their company that they will hang on to it until a profit is realized.
- Whale 🐳: Someone who owns a large enough amount of a stock or cryptocurrency that if they sell, it would cause the asset's price to fall due to a flood of supply.
- YOLO: the acronym of "You Only live one". This word is used when someone risks their entire portfolio on one single stock or options trade.

6.2 Sentiment methodology

Reddit's posts cleaning process is focused on selecting the right environment for the sentiment analysis. Even though they play an important role over a communication perspective the first step was removing all the emojis from the post. Unfortunately, the sentiment score attribution process presents a lack of recognizing the right meaning of the little images.

Once the data cleaning has been finished, it was essential to give a different weight to some words where under a Neuro-linguistic processing dictionary are meaningless but play a role in the Reddit r/Wallstreetbets context. As mentioned before, the social media language is complex, and unluckily, there is no NLP dictionary designed to understand it completely. The "SentimentAnalysis" R Studio package is the tool used to implement the analysis of the sentiment.⁵ This library is not just able to divide a glossary by positive, negative and polarity scores but it presents the possibility to compare multiple different dictionaries (See Appendix A).

6.3 Primary results

6.3.1 What is the relation between the sentiment of the subreddit and GME price trend?

What is the relation between the sentiment of the subreddit and GME price trend? To identify if there is a relation between r/Wallstreetbets sentiment and the price of GME, a Pearson correlation analysis was conducted.⁶ The correlation matrix clearly showed a relationship between each dictionary sentiment and what happened under a financial perspective at GME during the time frame 01/26 to 02/01.

5. Author Nicolas and Maintainer Nicolas Proelochs, "Package 'SentimentAnalysis'," 2021,

6. Samithambe Senthilnathan, "Usefulness of Correlation Analysis," *SSRN Electronic Journal*, no. July (2019).

	Trade Volume	Bid Price	Ask Price	Stock Price	Log Trade Vol.	Stock Ret.
N°WSBuser	-0.18	0.78	0.47	0.67	-0.01	-0.00
SentimentGI	-0.15	0.70	0.36	0.56	0.05	0.00
NegativityGI	-0.16	0.76	0.41	0.61	0.05	0.01
PositivityGI	-0.16	0.77	0.41	0.62	0.05	0.00
SentimentHE	-0.15	0.72	0.36	0.56	0.05	0.02
NegativityHE	-0.15	0.71	0.36	0.56	0.05	0.01
PositivityHE	-0.15	0.72	0.36	0.56	0.05	0.01
SentimentLM	-0.13	0.63	0.31	0.49	0.04	0.02
NegativityLM	-0.15	0.72	0.36	0.56	0.05	0.01
PositivityLM	-0.15	0.72	0.37	0.57	0.05	0.01
SentQDAP	-0.14	0.67	0.35	0.53	0.05	0.02
NegQDAP	-0.16	0.76	0.40	0.61	0.04	0.00
PosQDAP	-0.16	0.76	0.41	0.61	0.05	0.00
Trade Volume	1.00	-0.17	-0.06	-0.11	0.64	0.01
Bid Price	-0.17	1.00	0.36	0.66	0.09	0.00
Ask Price	-0.06	0.36	1.00	0.94	-0.12	-0.35
Stock Price	-0.11	0.66	0.94	1.00	-0.06	-0.28
Log-vol Trades	0.64	0.09	-0.12	-0.06	1.00	0.09
Stock Return	0.01	0.00	-0.35	-0.28	0.09	1.00

The use of the cross-correlation analysis made possible to re-balance the result and significantly increase the magnitude of the relation of the variables by lagging the stock price/volume. Considering two variables x and y, that differ only by a shift on the time axis, the cross-correlation can be calculated to show how much y must be anticipated to make it identical to x. The optimal delay of five hours.⁷

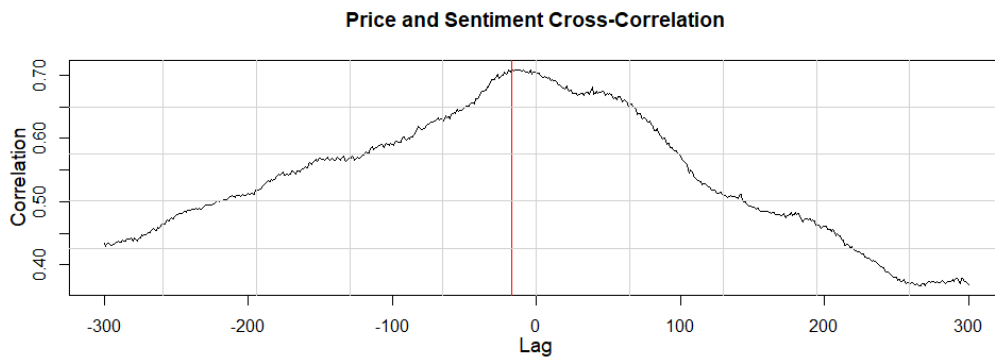


Figure 1: The optimal lag to maximize the correlation.

7. "An efficient method for computing local cross-correlations of multi-dimensional signals," *Colorado School of Mines Consortium Project on Seismic Inverse Methods for Complex Structures*, 2006,

6.3.2 Which neural language processing dictionary is the best representation of r/Wallstreetbets Reddit group?

To answer this question, a consideration towards the dictionaries with the highest correlation to prices could have been significant but not relevant for this paper. In order to make a more consistent analysis, looking at the scatter plot of the price with the sentiment analysis, the choice is to take consideration a linear model and ran a linear regression. Thanks to the Best Selection as validation approach, is possible to recognize the best version of the linear model and detect the variables that best explain the target: the mid-price. By considering the parsimony concepts⁸ for the statistic, the best choice for the predictor's number is made by the minimization of the BIC instead of the maximization of the R^2 . If the latter suggested that nine out of twelve sentiment predictors should have been chosen, the former told us to use three for explaining the mid-price. All of them belong to the General Inquire dictionary.

6.3.3 What is the dominance in word used?

After the data cleaning, the focus moved on the sentiment but also analyzed the most used word in the group. Wordcloud analysis measures the words that were most frequently used. The bigger the word appears, the more that word is used in the posts analyzed. GME is the most frequent word used and is the focal point of the event of interest. It notice that most the words are regarding the stock targeted by the users. AMC, NOK and BB are other stocks that also were aimed during the r/Wallstreetbets action and therefore frequently used in the posts. During the GME event, people keep buy or hold the stocks with the aim of "short squeeze" the hedge found. Most of the private investors were using the more user friendly stock trading platform Robinhood leading it to be mentioned pretty much.

8. Joachim Vandekerckhove, Dora Matzke, and Eric-Jan Wagenmakers, "Model comparison and the principle of parsimony," *Oxford Handbook of*, 2014,

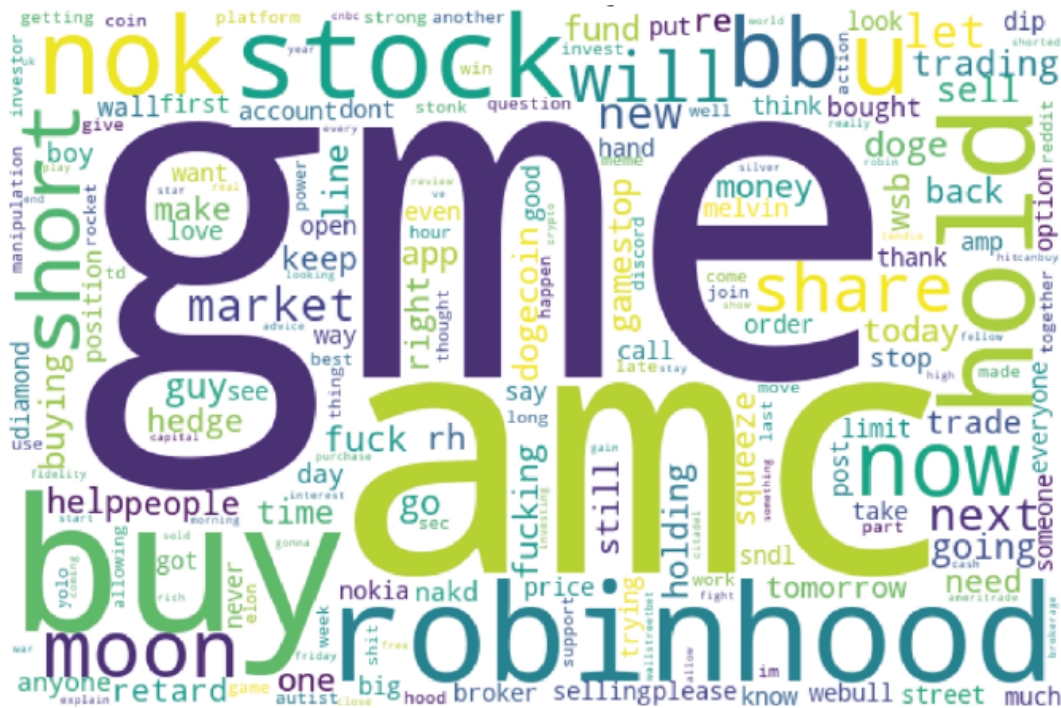


Figure 2: Most common words among post on r/Wallstreetbets 01-26/01-29

6.4 Results

The analysis showed some interesting results. It is possible to understand that most of the people that are part of r/Wallstreetbets do not use economic language. Therefore, even though the group idiom could be associated with a new economical way of communicating, the sentiment analysis showed how the General Inquiry dictionary best represented the idiom used by the users. This result confirms that most of the participants are not experts and can be associated to irrational investors.⁹ These type of individuals make "gut" choices, following their own instinct or "the flock". Most of the time, they align with the behavior of the masses; the "herd-effect" indeed is emblematic. In behavioral finance, this social bias appears when subjects are interacting in a system. In the case of r/Wallstreetbets, they want to be part of a community and contribute in the "free-meal" against the big hedge fund. The consequence was that all the "Retars" or "Apes" (as they call themselves) started following generic ideas without any ratioMcGurk2020. From a correlation perspective, two main different observations were made. First, the analysis showed an output with a high level of noise error: all the sentiment with zero as results. A "zero" in a sentiment analysis could be related to a null combination of the word's score of the post. An attempt was made to omit those values to improve the correlation between the sentiment and prices but it led to lose the entire relation strength (from a 50% to a 5%). One interpretation is the absence in terms of sentiment importance of a hierarchy in the posts. All the publications together

9. Al Mamun, Abu Syeed, and Farida Yasmeen, "Journal of Economic Financial Studies Are investors rational , irrational or normal ?," 03 (2015).

explain a connection of price and sentiment score. Second, by the cross-correlation analysis is possible to understand that by lagging the variable, the outcome is represented by the highest level of relation; in other words a change in one of the variables had an impact over the other variables after five hours. This can tell us that one variable influence the others when its' delayed.

From a word frequency perspective, it is easy to notice that one of the most frequent words is GME, which refers to the ticker for Game Stop Corp. The analysis focused in the evaluation of words and whether they can tell more information about theirs used trend during this time. The group changed the way of posting after some events. Many attempts have been done considering different time frame, but what was found to be most suitable for the analysis was splitting the dataset in only two parts. The most used words before 01/28 (See Appendix B.1) among the Positive Sentiment posts were about euphoria. All the "apes" try to "YOLO" Game Stop and Black Berry ("bb" in the figure). The period considered, corresponded to the peak of the prices. For the words after 01/28 (See Appendix B.2), is remarkable to see that, even if it the most common words for positive posts were taken into consideration, one of the most frequent was "HELP." Probably many people are referring to the losses caused by the gigantic draw-down of GME stock price.



Figure 3: Split of the period on a word frequency basis.

An analysis of the most common negative words in both periods look similar. However, the second period was remarkable because of the presence of two different platforms: Discord and Robinhood. The latter refers to the financial service that enables private investors to access the market, and the former is a server-based social media platform that helps people to exchange ideas. The moderators of Reddit r/Wallstreetbets also have a complementary server in Discords with the same focus of the main group: discuss investing bets. An analysis over Discord will be conducted later on Discord. This period split in terms of word frequency made us think about two different moments in the trend of Reddit's r/Wallstreetbets post, but more could characterize the period. Unfortunately, it is difficult to get an index that

indicates how the users changed the way they post; for this reason, an attempt considering the number of posts made per day could explain the tendency of r/Wallstreetbets better.

	Posts.Volume
Bid Price	0.27
Ask Price	0.27
Spread Price	-0.29
Stock Price	0.27
Bid Size	0.36
Ask Size	0.31
Sperad-size Orders	-0.14
Posts Volume	1.00

By looking at the correlation below, it is possible to see that the posting-volumes of GME's prices have a relation score of 30% with the "Ask Price" and 35% with the "Bid Price". The herd effect, the missing data for almost one day, the split in terms of temporal frequency, the "GME" words in every sentiment frequency sample and the high correlation between price and volume led to think about some sort of relation between the results. This was the new starting point for a new analysis. In the following chapters, this research tries to understand if there is any economic variable behind the GME phenomenon from a quantitative perspective and if any insights can explain how the variable is related.

7 FAMA AND FRENCH

By using Fama e French model, the focus was to understand the price movement and whether it was anyhow related to related an economic reason.

7.1 The three (plus one) factors model.

The three-factor Fama and French model was developed to expand the CAPM model by adding risk and value factors to the market. Under this model, which is based on the historic returns of different asset classes, small-cap and value-based stocks, tends to have returns above the market average.¹⁰

$$return_{GME} = r_f + \beta_1(r_m + r_f) + \beta_2(SMB) + \beta_3(HML) + \beta_4(SENTIMENT)$$

- r = Portfolio expected rate of return
- r_f = Risk-free rate
- β = Factor's coefficient (sensitivity)
- $(r_m - r_f)$ = Market risk premium
- SMB = Small Minus Big
- HML = High Minus Low
- SENTIMENT

The CAPM formula shows that a stock's return equals the risk-free return plus a risk premium based on that stock's beta. The Fama-French model aims to describe returns on stocks through three factors: market risk, the outperformance of small-cap companies and small-cap stocks. The rationale behind the model is that high-value small-cap companies tend to outperform the overall market regularly. Small Minus Big (SMB) is a size effect based on a company's market capitalization. It measures the historical excess of small-cap companies over large-cap companies. Once the SMB is identified, its coefficient is a measure of the sensitivity. This score is a representation of the correlation of a stock or investment portfolio to the movements in the overall market. A statistical measure of risk was obtained by comparing the returns of a single security/portfolio with the overall market returns, which can take positive and negative values. The main reason behind this factor is that small-cap companies tend to achieve higher returns in the long run than large-cap companies. High Minus Low (HML) is an added value. It represents the spread of returns between companies with a high book value ratio to market value (company value) and companies with a low ratio

10. Ziyang Ji et al., "Empirical research on the fama-french three-factor model and a sentiment-related four-factor model in the chinese blockchain industry," *Sustainability (Switzerland)* 12, no. 12 (2020).

of book value to book value.¹¹ As the "Plus One" factor, the sentiment from the previous analysis was considered.

7.1.1 Is GME event related to economic reason?

Multiple linear regressions were ran by changing the "plus one" parameters with different scores for the General Inquiry dictionary; every model performed present but not significant results. This evidence made it plausible to think that there is no economic reason why the analyzed stock had such a shock in prices.

7.2 Results

In this particular case, Fama and French was used to understand if the GME event had any fundamentally economical reason and it has not. The risk resulted from this stock is all idiosyncratic. In other words, there are particular and intrinsic factors that can have a negative effect on the analyzed security. Several things can lead to the devaluation of a single stock without damaging an entire portfolio. For example, a company may experience a strike action that lowers stock prices when consumers worry. Likewise, a company may be subject to major lawsuits, a drop in earnings, or a similar event that causes them to decline in value as investors become less confident. These are all risks that can affect securities at any given time. In this case the exogenous risk could be attributed to the GME trigger. By looking at the price graph over a long period it is possible to notice that something caused the beginning of a swing in term of prices. The concept of dumped oscillator motion made it plausible to think about a specific event as the beginning of everything.¹² As in physics, a damped harmonic oscillator can consist of a mass that moves friction on a vertical plane subjected to an elastic force and a damping force. After a few moments, the motion stops due to the swing of momentum and the weight will stabilize itself (See Appendix C).

11. Abderrazak Dhaoui and Nesrine Bensalah, "Asset valuation impact of investor sentiment: A revised Fama-French five-factor model," *Journal of Asset Management* 18 (2017).

12. Leonidas Sandoval Junior and Italo De Paula Franca, "Shocks in financial markets, price expectation, and damped harmonic oscillators," no. April (2011).



Figure 4: "Motion" over time Game Stop Corp as a mean reversion

This is just a naïve comparison between the harmonic oscillator and the mean reversion analysis. The latter concept refers to the idea that the maximum and minimum price of a security is temporary and that it tends to have an average price over time. Even if the approach presented can have a model to explain the trend of GME, they hardly reveal why the initial peaks took place and the reason why this phenomenon could have caused such a swing on the GME prices. In order to move further with the analysis, it is worthy to keep focus on the r/Wallstreetbets group and finding a relation between posts volume and prices. The aim is to find the phenomenon that caused the GME shock prices.

8 THE TRIGGER

The evidence found led to believe that an external factor could have been the starting point of everything. Could have been the market manipulation the trigger of GME event?

8.1 MARKET MANIPULATION

"Market manipulation fraud—commonly referred to as a "pump and dump"—creates artificial buying pressure for a targeted security, generally a low-trading volume issuer in the over-the-counter securities market largely controlled by the fraud perpetrators. This artificially-increased trading volume has the effect of artificially increasing the price of the targeted security (i.e., the "pump"), which is rapidly sold off into the inflated market for the security by the fraud perpetrators (i.e., the "dump"). This results in illicit gains for the perpetrators and losses for innocent third-party investors. Typically, the increased trading volume is generated by inducing unwitting investors to purchase shares of the targeted security through false or deceptive sales practices and/or public information releases." ¹³ ("Market Manipulation: 'Pump and Dump' Fraud." FBI)

8.2 "Pump and Dump" Fraud

"Pump and Dump" is a simple technique. Scammers using this technique find a secure channel where presenting financial advisory services that are attractive from the point of view of profitability. The organizers of the fraud pre-purchase stocks with low trading volume and influence users to buy, which cause a sharp rise in the stock price. Next, the organizers have to spread the news to make the unwary investor receive the rumor or the offer. The system needs an increase in purchasing, leading to a further increase in the price. As soon as the scammers see the stock price rise, they start selling, leaving the victim with a handful of flies. At this point, the stock deflated as if in a mini speculative bubble guaranteeing the scammers generous profits.

8.2.1 Is the market manipulation the trigger of GME event?

Until now the proportion evidence collected, led the analysis to consider many different periods over the subreddit. In the following chapter the focus will be to establish a connection between the Pump and Dump scheme and r/Wallstreetbets group.

13. Market Manipulation ('Pump and Dump') Fraud." FBI. FBI, June 15, 2016. <https://www.fbi.gov/scams-and-safety/common-scams-and-crimes/market-manipulation-pump-and-dump-fraud>.

9 DISCORD: The use of a social gaming platform to perform fraud

9.1 Discord

Born as a communication app for video game fanatics for over a year, Discord has become quite popular in other areas. It is a platform that works on a server basis. These server are set up with special intentions or emphasis as a way for people to come together, exchange pictures, links and information. These servers offer both voice and text communications. Since it allows to manage large groups of people and to switch fairly smoothly from written chats to calls and video calls. Many people have adopted it to stay in touch with loved ones, organize various activities remotely such as online lessons, communicate with family or groups of friends, book clubs, study groups, and so on. Users may either be invited to servers or find connections to them online. The person who create the server normally establishes specific ground rules for the community that is part of the server. The degree of courtesy and the protocol that should be required. These roles are netiquette, the proper manner of interacting inline. The term netiquette means the set of rules that govern the behavior that people must have on the Internet in relation to others, such as blogs, forums, emails, social media, and much more.¹⁴

9.2 Pump and dump Groups

In Discord, it is easy to take part in a scam server. The most common servers are regarding cryptocurrencies. An individual can just type common words, like "pump and dump" or "crypto pump" in the search bar. Later s/he will join not just as a group but also enter a system. From now on, they can reach out to her/him since the user name will appear in their server. The chat section will be easily spammed by scammers and people who will try to promote Ponzi Schemes or other fake business opportunities in a few days.

Most of the group show a hierarchical pyramid on invitation basis. In other words, they propose promotions and more advantages by the number of invitees that someone can bring. In the Appendix D.1 is possible to see how the rank basis works in a pump and dump group.

The idea is: the closer to the top people are, the better possibility of benefiting from the pump and dump that are interested in. This is because the "detect pushes intervals" are shorter as you climb the group's hierarchy. The further from the managers of the scam people are, lately they know the Pump and Dump target. Being at the bottom of the pyra-

14. Massimo La Morgia et al., "Pump and Dumps in the Bitcoin Era: Real Time Detection of Cryptocurrency Market Manipulations," *Proceedings - International Conference on Computer Communications and Networks, ICCCN 2020-Augus*, no. May (2020).

mid means you get better benefits in term of time over the bottom of the pyramid. The pump and dump organizer will use the Discord channel as the main way to communicate and to "sell" their ideas with their victims. Those groups promote the pump and dump as a technique that guarantee high returns without any effort. The Appendix D.2 shows how they describe the action and the scheme. In the explanation, it is possible to see a paragraph called "*What you can do to help make the pump as successful as possible*" with a description of the "*Social media effect.*" With this terminology the moderators are referring to the fact that social media is capable of spreading the news more rapidly and to a larger audience than any other channel. In order to instruct the naïve users they provide a step by step guide to follow in order to make the phenomenon quickly viral (See Appendix D.3). The first step for setting up a pump and dump is to purchase the stock by the villains when the price is low. The graph in the figure showed the relation between purchase and price

9.3 The GME “CASE”

The GME event, from a social media perspective, looks similar at the modus operandi that was just presented. Even if r/Wallstreetbet was private just on Wednesday for less than one hour, the Pushift database does not have any posts available for almost twenty-four hours. The public reason why a group shut down for one hour was related to some Discord policy violation. The r/Wallstreetbets server in this platform had some hateful and racist content and little to do with GameStop’s recent stock spike. The Discord moderators gave the following reason for the temporal ban: *"To be clear, we did not ban this server due to financial fraud related to GameStop or other stocks, Discord welcomes a broad variety of personal finance discussions, from investment clubs and day traders to college students and professional financial advisors. We are monitoring this situation and in the event there are allegations of illegal activities, we will cooperate with authorities as appropriate."* Sonnemaker, Tyler. “Discord Bans r/Wallstreetbets Server over Hate Speech as the Group Drives GameStop Shares through the Roof.” Business Insider. Business Insider, January 27, 2021.¹⁵ In that period, the Reddit group became private and a message displayed on their group page reporting: *"We are experiencing technical difficulties based on unprecedented scale as a result of the newfound interest in WSB. We are unable to ensure Reddit’s content policy and the WSB rules are enforceable without a technology platform that can support automation of this enforcement. WSB will be back."* In reality, it is plausible to think that during that "blind" period the "Apes Kings"¹⁶ (the people at the top of the pyramid who controlled all the r/Wallstreetbets apes) have organized the pump and dump. With the help of a smaller group of Apes (a second level on the hierarchical scheme) could have spread the news in order to increase the sentiment that made users buy the stock. This caused r/Wallstreetbets to spread like wildfire and become viral.

The first step for setting up a pump and dump is to purchase the stock by the villains when the price is low.¹⁷ Then it is possible to move in the group and manage the crown. In support, from a quantitative perspective it is peculiar to notice the relation between price, volume of posts and ask size of the orders.

Firstly, there will be a focus on the time series evaluating the number of posts and the price of the stock. The two variables followed a similar trend in a particular day (01-28). Both volumes of posts (orange line in Figure 5) and price (black line) grow together in that particular day. The aim is to present that the relationship is an increase of volumes and a correspondent spike in the price.

Next, the ask size is considered. The ask size is the amount of various sell orders. The

15. Sonnemaker, Tyler. “Discord Bans r/Wallstreetbets Server over Hate Speech as the Group Drives GameStop Shares through the Roof.” Business Insider. Business Insider, January 27, 2021

16. Matt Reeves, director. “War for the Planet of the Apes”. Twentieth Century Fox, 2017

17. Clearly, before January 28, it is possible to notice an increasing in the demand of GME stock from the Appendix E.

hypothesis is that organizers are knowledgeable that the price is driven by the sentiment of the people (green dashed line). Also, they know that the stock is overpriced and to capitalize they need to sell a large amount of shares. .

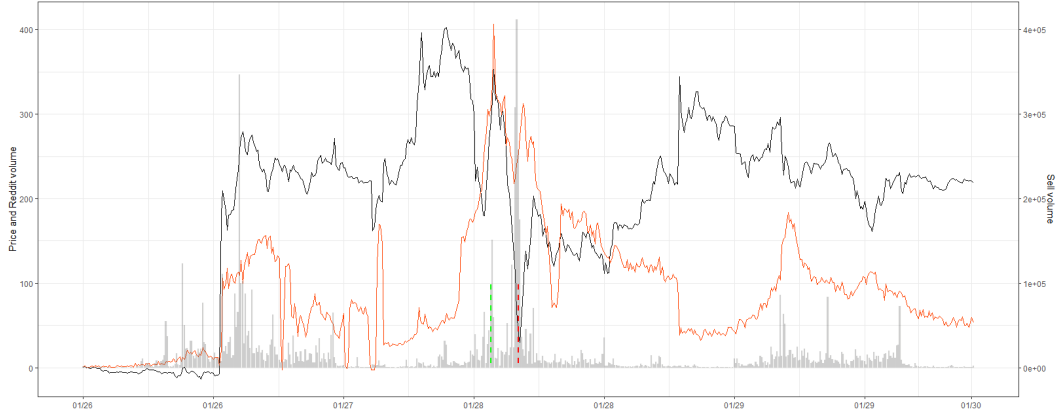


Figure 5: Price and "Ask Orders" VS Reddit (red) volumes.

The second huge spike could indicate a delayed reaction taken by the scam apes shown by the red dashed line. Volume matched the first spike but GME stock price was significantly lower than the first instance.

To give more consistency at the hypothesis, the volume analysis has been conducted also for twitter. The data downloaded of the “GME” post in the same temporal window 01/26 to 01/29 made it possible to conduct the proceed with the analysis. From the Figure 6, is not hard to notice a similar dependence between variable and analogue situation in term of peak of dependence. The peak in volume is bind to a high rise in the stock price: result is a first phase of selling (green dashed line). Once the price drop, another larger number of ask orders took place; this indicated many stock positions sold on 01/28 causing many private investors to lose a lot of money (red dashed line).

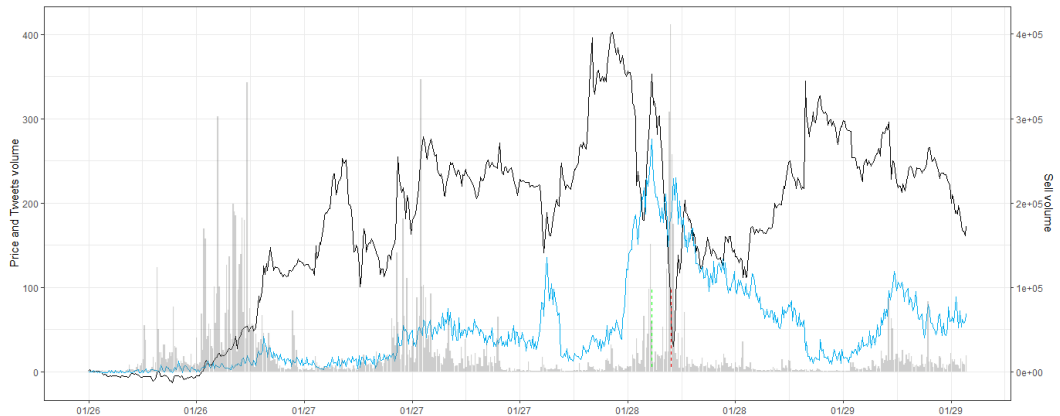


Figure 6: Price and "Ask Orders" VS Tweeter (blue) volumes.

The overall plot shows better how both Twitter and Reddit posts volume are dependent and how their shapes are related to phases presented of a Pump and Dump.

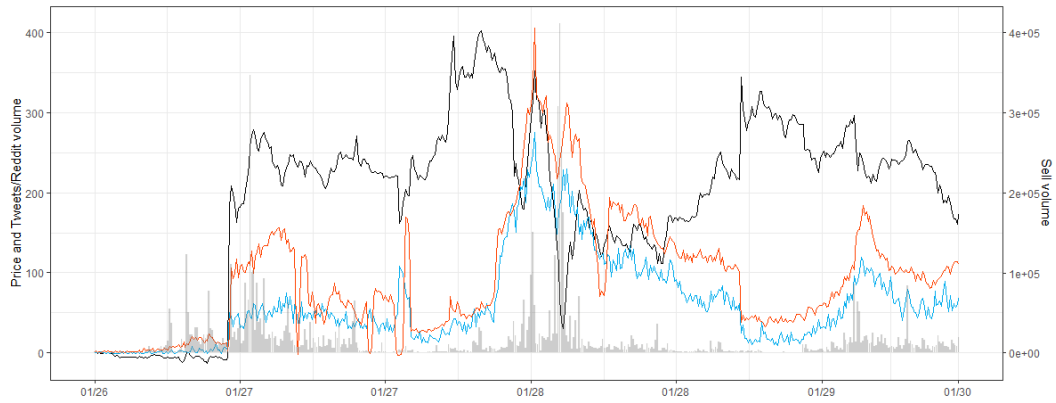


Figure 7: Price and "Ask Orders" VS Reddit (red) and Tweeter (blue) volumes.

10 CONCLUSION

This report aimed to prove that a price manipulation could have taken place in the first price shock "wave" of GME. The straight relationship between volumes of posts and prices could be related to a simple increase of the extraordinary event popularity but with a closer look everything is related. During the "blind temporal window" moderators and a small number of people could have organized the Pump and Dump. By starting to spread the news of "free meal" over the "Apes" the GME event became even more popular and caused people to take a long position in the stock. As shown in sentiment analysis the users could be associated to irrational investors. These types of individuals are characterized by the following mistakes: 1. Following instinct driven by anxiety. 2. Following fear of not making a profit despite an apparent possibility. 3. Being influenced by other peoples' choices to make an investment (herd effect). The mass purchase of the overpriced GME stock incremented a rally on the GME price. This run of the price made pression over the hedge fund and create a the short squeeze that boosted even more the epic run of Game Stop. The combination of those factors drove the price reached unique high levels. It has been found that the "Apes Kings" known in advance what was going to happened at GME time series. Indeed, in the moment the stock reached their target price, they started to sell before the imminent drop taking enormous profits. The evidence reported, clearly identifies an analogy with the pump and dump scheme that affect cryptocurrency. The one lingering question is when the exact moment in which this pump and dump took place was? The analysis identifies the starting point and a possible moment when the financial felonious left their long positions but there is no clear indicator of that. In the future, to answer this question clearer it will be needed to find evidence of the apes organizing premeditated action over GME.

Appendices

A “SentimentAnalysis” RStudio package (Nicolas and Proelochs 2021)

Dictionaries list

- GI: Dictionary with a list of positive and negative words according to the psychological Harvard-IV dictionary as used in the General Inquirer software. This is a general-purpose dictionary developed by the Harvard University.
- HE: Dictionary with a list of positive and negative words according to the Henry’s finance-specific dictionary. This dictionary was first presented in the Journal of Business Communication among one of the early adopters of text analysis in the finance discipline.
- LM: Dictionary with a list of positive, negative and uncertainty words according to the Loughran McDonald finance-specific dictionary. This dictionary was first presented in the Journal of Finance and has been widely used in the finance domain.
- QDAP: Loads polarity terms from the package qdap’s data object; After that, the object is transformed into a structured dictionary object. It is a collection of multiple dictionaries and Word Lists to Accompany that are built for the quantitative sentiment analysis.

B "WordCloud"

B.1 Common words before 01-28



Figure 8: Most common positive and negative words among post on r/Wallstreetbets before 01-28

B.2 Common words after 01-28



Figure 9: Most common positive and negative words among post on r/Wallstreetbets after 01-28

C Damped Harmonic Oscillator motion

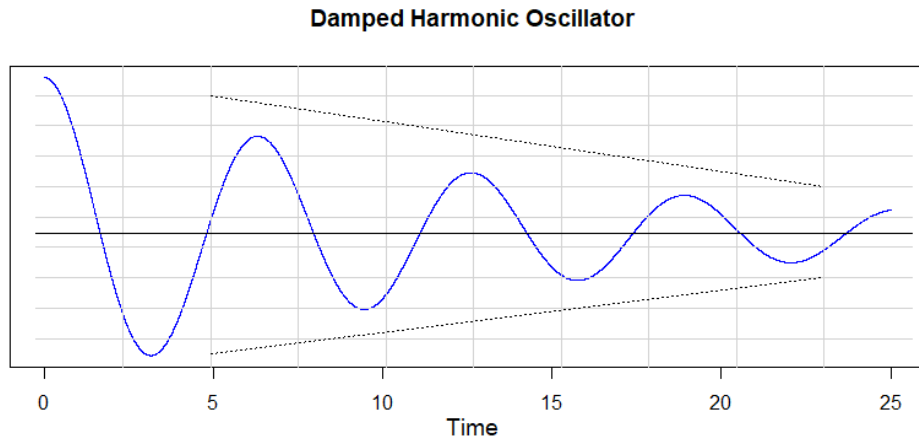


Figure 10: Motion over time of a weight bind to a spring after being dropped.

D Appendix : Discord scam group netiquette

D.1 The Invite Chain

The following extract has been downloaded form the netiquette of a Discord's Pump and Dump scam group:

Freshman (1 invite)

- Access to crypto alerts which will detect pushes within 20 minute intervals

Sophomore (5 invites)

- *Can send photos in chat*
- *Access to crypto bot which will detect pushes within 4 minutes intervals*
- *Access to crypto bot which will detect new coin listings on 32 exchanges*
- *Access to 🏠 post-grads-podium with crypto calls from @Post Grad's*
- *Access to 📊 signal-bar with premium signals from an AI bot*

Junior (15 invites)

- *Can send photos in chat*
- *Can send links in chat*
- *Access to crypto bot which will detect pushes within 60 seconds intervals*
- *Access to crypto bot which will detect new coin listings on 32 exchanges*
- *Access to crypto bot which can detect pre-pumps for pump groups at 60 seconds*
- *Access to 🏠 post-grads-podium with crypto calls from @Post Grad's*
- *Access to 📊 signal-bar with premium signals from an AI bot*
- *Access to 🏠 big-chads-dorm*

Senior (40 invites)

- *Can send photos in chat*
- *Can send links in chat*
- *Can change nickname*
- *Access to crypto bot which will detect pushes within 1 second intervals*
- *Access to crypto bot which will detect new coin listings on 32 exchanges*
- *Access to crypto bot which can detect pre-pumps for pump groups at 1 seconds*
- *Can opt in on instant alerts for pre-pump detections*
- *Access to 🏠 post-grads-podium with crypto calls from @Post Grad's*
- *Access to 📊 signal-bar with premium signals from an AI bot*
- *Access to 🏠 big-chads-dorm*
- *Early access to new projects*
 - *beex: 15% off the Beex Pump Bot!*
 - *cryptshot: 15% off the CryptShot Pump Bot! "*

D.2 PUMP and DUMP details

The following extract has been downloaded from the netiquette of a Discord's Pump and Dump scam group:

"Here is how a pump works:

1. Before the pump, you need to have Binance.com and Discord open.
2. At the moment the countdown ends, the pump will start. There will be a call in the pump-signal channel to announce the coin (example: The coin we picked is \$XXX).
3. Right after we announced the coin, we will buy it collectively and wait for the price to go up.

In the meanwhile, we will ensure that multiple media sites publish articles about the coin to attract outside investors. Everyone in the group can help us by sharing the coin on their social media as well.

What you can do to help make the pump as successful as possible:

- Keep the pump alive by setting buy walls. When the price is going up, set some lower buy walls to create waves. (example: Starting price 100, price 10 minutes after the pump 200, set buy wall at 150)
- Everyone in the group can help us by sharing the coin on their social media, Reddit and Twitter using Hashtags.

This so-called "**Social Media Effect**" does not hit instantly. It will take around 20 minutes to hit. So, buying bottoms during the pump is never a bad idea.

Some basic tips on buying and selling during a pump:

1. When the pump starts the price of the coin can spike pretty fast. If you limit buy a coin advise you to buy 75%. A 100% limit buy does not go through most of the time since the price is up. If you market buy the coin, the order will always go through.
2. After buying the given coin, do not sell 100% of the given coin at once, slowly sell (25% at a time). This way, the market will not drop instantly.
3. DO NOT PANIC SELL. We have seen in the past that a lot of people will panic sell when they're down on their initial investment. Just know pumps come in waves so there should be plenty of times within the pump that you will be able to sell with profit, so do not panic.

If you have any questions, feel free to ask a moderator or the pump team directly. "

D.3 Social media guide

The following extract has been downloaded from the netiquette of a Discord's Pump and Dump scam group:

"What you can do to help make the pump as successful as possible:

How to get the word out on social media about the coin we're pumping. Spreading news or "buy signals" on social media is one of the most important things that must take place in order to produce a good pump. If we succeed and news about the coin goes viral, even more outsiders will FOMO (fear of missing out) in. Besides hugely impacting the initial pump, outsiders massively buying the coin will create a second wave bigger than the initial spike. Thus, creating more profit for you and all the fellow members of the group. Twitter is the source where most crypto traders get their information from, and this is why we advise everyone to at least post on this platform once.

Step 1

At the moment the countdown ends, the pump will start. There will be a call in the pump-signal. For example: The coin we picked is \$XXX. After you've read the coin, enter it on Binance and market buy (75%) OR limit buy the coin.

Step 2

The most important part. Everyone will start posting on social media, This is a crucial factor to attract outsiders. When people outside the group notice that the coin we are pumping is in the Top gainer list of Binance, the usual process for an average trader will be to find explanations or reasons for the rise in price. We have made a list of what you can do to make the coin go viral on social media: -Twitter: tweet using the Coin \$(tag) or hashtag, like, and retweet as much as you can about the coin. -Instagram: react on crypto-related pages with a hashtag of the chosen coin. -Reddit: Post on crypto-related subreddits and react under posts. -Telegram: Head over to groups and start talking about the chosen coin.

Step 3

At this point, the coin will gain a lot of attention. The massive increase in price volume combined with the sheer amount of hype on every social media platform will undoubtedly cause many outsiders to buy into the coin. From here on, you should slowly start selling your bags. Do not sell 100% of the given coin at once, slowly sell (10-25% at a time). This

prevents the market from dropping in price and improves having a steady rise. (Useful groups to post on will be shared closer to the pump in news ews) For this we ask everyone in the group to spread hype as soon as you closed your positions. "

E Bid - Volume

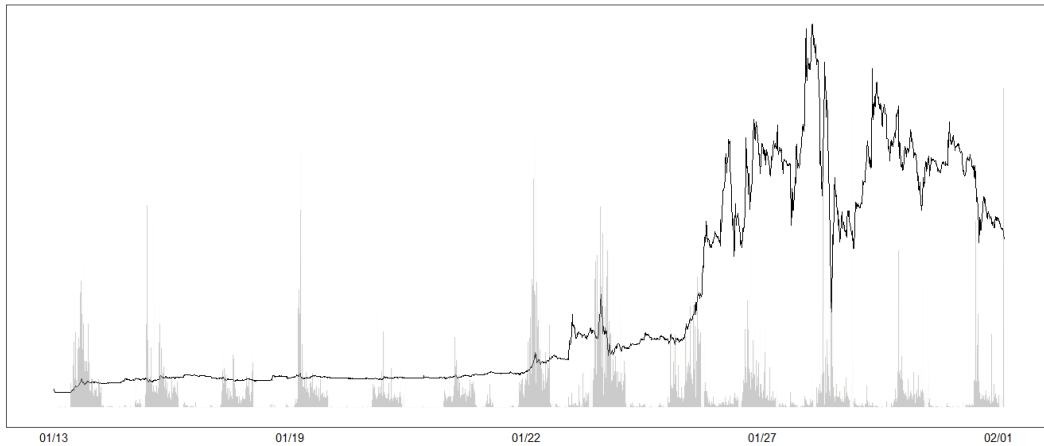


Figure 11: Bid Volume for Game Stop Corp

References

- “An efficient method for computing local cross-correlations of multi-dimensional signals.”
Colorado School of Mines Consortium Project on Seismic Inverse Methods for Complex Structures, 2006.
- Dhaoui, Abderrazak, and Nesrine Bensalah. “Asset valuation impact of investor sentiment: A revised Fama-French five-factor model.” *Journal of Asset Management* 18 (2017).
- Ji, Ziyang, Victor Chang, Hao Lan, Ching Hsien Robert Hsu, and Raul Valverde. “Empirical research on the fama-french three-factor model and a sentiment-related four-factor model in the chinese blockchain industry.” *Sustainability (Switzerland)* 12, no. 12 (2020).
- Junior, Leonidas Sandoval, and Italo De Paula Franca. “Shocks in financial markets, price expectation, and damped harmonic oscillators,” no. April (2011).
- La Morgia, Massimo, Alessandro Mei, Francesco Sassi, and Julinda Stefa. “Pump and Dumps in the Bitcoin Era: Real Time Detection of Cryptocurrency Market Manipulations.” *Proceedings - International Conference on Computer Communications and Networks, ICCCN 2020-Augus*, no. May (2020).
- Mamun, Al, Abu Syeed, and Farida Yasmeen. “Journal of Economic Financial Studies Are investors rational , irrational or normal ?” 03 (2015).
- Nicolas, Author, and Maintainer Nicolas Proellocks. “Package ‘ SentimentAnalysis ’,” 2021.
- Senthilnathan, Samithambe. “Usefulness of Correlation Analysis.” *SSRN Electronic Journal*, no. July (2019).
- Vandekerckhove, Joachim, Dora Matzke, and Eric-Jan Wagenmakers. “Model comparison and the principle of parsimony.” *Oxford Handbook of*, 2014.