

# Power to the Zealots: How Enthusiastic Stock Fans Keep Prices from Fundamental Valuations

David Jackson • September 2023

## Abstract

Threads and comments were scraped over the course of a month from the subreddit r/SuperStonk to compare Gamestop's price action to poster sentiment, as judged by popular sentiment analysis techniques. The sentiment of the subreddit, dedicated to the idea that Gamestop is a strong investment likely to have another shortsqueeze and then strong following fundamentals as a company, did not correlate in a meaningful way to price movements when regressed on one another with simple regression. Further comparison between it and similar equity assets in terms of popular pricing ratios shows signs that GME is overvalued in comparison, and thus leads to conclusions of how these investors effect the price from being more in line with other assets in accordance to their choice of investment habits compared to investors in other popular equities.

## Introduction

In 2021, Gamestop saw a meteoric rise in its stock price from \$5 just the previous winter to well over \$400 at its peak (pre-stock split). Many papers have since been dedicated to the driving force behind a company who saw some fundamental changes to its core business, but largely is believed to have ballooned due to investor sentiment. As discussed in "I just like the stock": The role of Reddit sentiment in the GameStop share rally' (Long et. al), sentiment analysis has been of growing interest in understanding financial markets. With Gamestop specifically, much of retail interest seems to have coalesced on the popular subreddit r/wallstreetbets. However, after the interest for the stock faded from mainstream attention and the users on r/wallstreetbets moved back to more general discussions of the market, a few dedicated subreddits were created to further discuss the Gamestop stock. One of those, the most popular by far, is the focus of this paper's research: r/Superstonk.

The r/Superstonk subreddit is an entire subreddit dedicated to the Gamestop stock (ticker: GME). As seen on their primary "Due Diligence" thread linked to at the top of the subreddit, they believe that despite the massive run up in price in January 2021, that this was not the final short squeeze effect on the company, stating: "[The 2021 squeeze] was just the beginning. We approached an almost complete implosion of the propped up financial system and most brokers removed the ability for retail investors to buy shares of GME. Once the buy button was turned off the thesis is based on the idea that short hedge funds doubled down and shorted the stock even more. We have tons of DD (Due Diligence) explaining this that I will link below. Since then, as more and more retail investors continued to buy GME the price has risen back up again

and has been consistently shorted even more to try and kick the can down the road as long as possible.”

When the subreddit mentions a short squeeze, they are referring to the phenomenon in which a stock that has been heavily shorted, or borrowed and then sold in hopes the price will move down and thus can be rebought at a lower price and returned for profit, sees its price driven rapidly higher as entities that shorted the stock are forced to buy it back to cover their borrowed shares. At the time of the rise of GME initially, short interest was relatively high, and was believed by the community to be the mechanism behind the quick rise of the stock price.

## **Data & Methodology**

### **Data**

This paper first gathered posting history from the r/Superstonk subreddit. Changes to the Reddit API around time of data gathering made backwards looking data collection challenging, as reddit’s policy changes to API rendered previous libraries that could collect vast collections of past posting data too expensive to run, and so many were closed. To work around this, posting data was taken weekly and cleaned of duplicates over the course of a month from June 22<sup>nd</sup>, 2023 until July 20<sup>th</sup>, 2023. Thread titles and comment history in those threads were collected, and in the end, we ended up with 3938 threads and 166,920 comments (excluding the auto-pinned message in every thread by the subreddit, and deleted or removed comments). Figure 1.1 shows the wordcloud for our threads, and Figure 1.2 shows the wordcloud for our comments. Figure 1.3 captures the total posts made each day, showing no clear pattern or decline in postings over the month the data was collected.



Figure 1.1



Figure 1.2

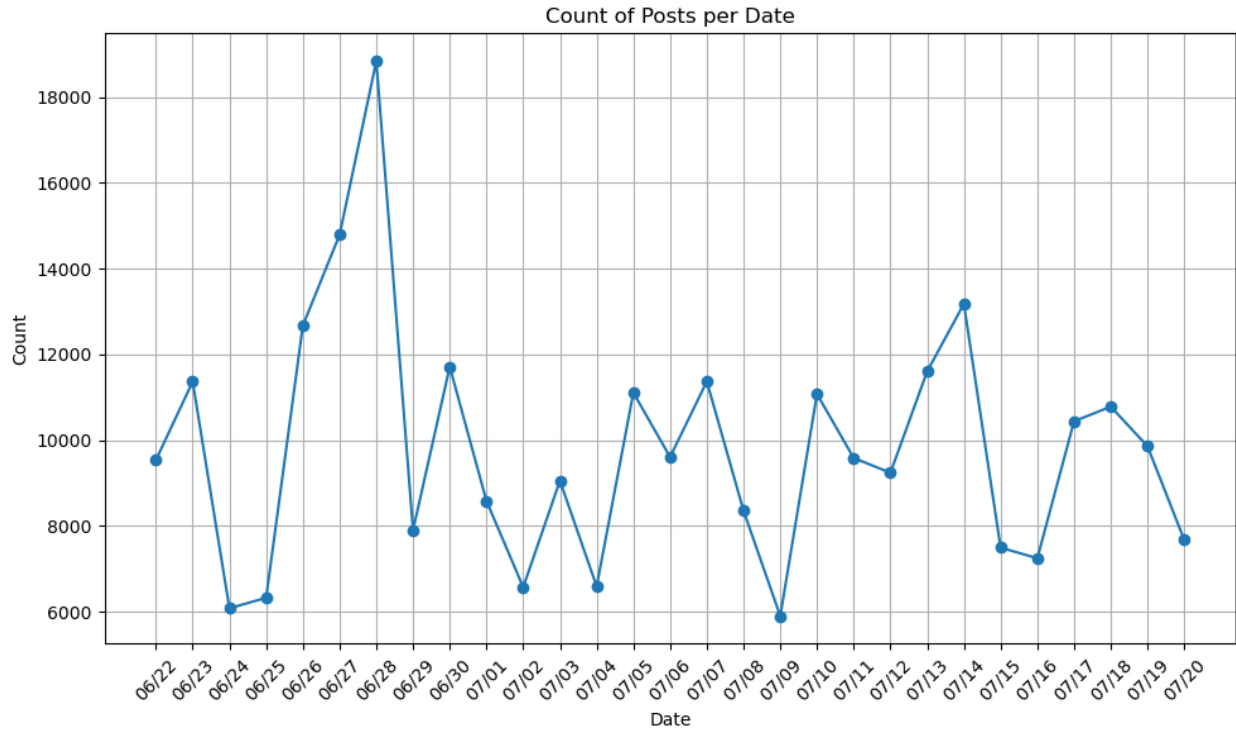


Figure 1.3

This paper found its stock pricing data via the Cboe exchange, which offers hourly open and closes in price of stocks. This was used to collect the hourly open price, hourly close price, and net returns of Gamestop shares during those hours between our dates of interest from the pre-market trading hourly open to the after-hours trading hourly closing. Cboe sees about 25% of the trading volume that the New York Stock Exchange (NYSE) sees, and thus can have minor price discrepancies from that exchange, especially at lower time frames. It was selected over the NYSE pricing because it offers the data free of charge, at levels lower than the daily open and closing prices.

Sentiment was then taken over hourly posting time periods using the nltk library’s Vader Sentiment Analysis tool, a popular Python library tool for measuring the sentiment score of strings (Hutto & Gilbert). In short, this library measures a probability that a given string is of positive, neutral, or negative sentiment, and then creates a compound score based on these in order to more easily bin scores as either positive, neutral, or negative. From their documentation: “The compound score is computed by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to be between -1 (most extreme negative) and +1 (most extreme positive). This is the most useful metric if you want a single unidimensional measure of sentiment for a given sentence. Calling it a ‘normalized, weighted composite score’ is accurate.”

## Methodology

First, given thread titles and comment strings had their sentiment measured. As suggested by the work done through the sentiment library (Hutto & Gilbert), valuations where the compound sentiment score was above a threshold of .5 were given 'Positive Sentiment' and scores below negative .5 given 'Negative Sentiment.' Everything in between was marked as neutral sentiment. These values are then transformed to either a -1 for negative sentiment, a 0 for neutral sentiment, or a 1 for positive sentiment. These totals for each category are listed in Table 1.

Table 1

Positive Sentiment	62,290
Neutral Sentiment	392,636
Negative Sentiment	22,952

We then move to average the sentiment hourly to map to the hourly stock price movements. Our counts came out as displayed in Table 2. Below that, in Figure 2.1, we can see our hourly average sentiment over the course of the data collection.

Table 2

Positive Sentiment Hours	270
Negative Sentiment Hours	47
Positive Return Hours	154
Negative Return Hours	147

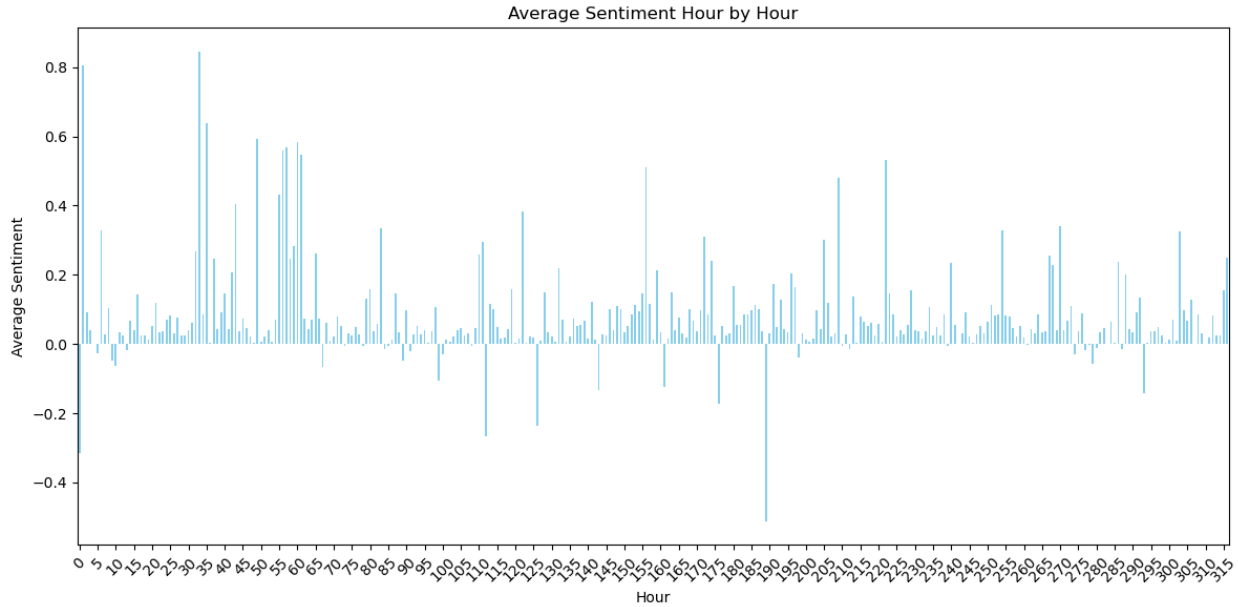


Figure 2.1

Figure 2.2 is the candlestick chart for all hourly prices collected, which demonstrates a total decline in price of 5.56% by the end of the data gathering process.



Figure 2.2

## Results

This jarring disconnects between sentiment, which seems overwhelmingly more positive than negative, and price movements, which seems to evenly move up and down hour to hour, is already quite telling, but a simple regression between the two continues the story.

Figure 2.3 is a histogram of net returns over the hours of our collection. As we can see, the vast majority of hourly returns sit close to 0, with a few outliers going out to +/- 2.5%, and one very large outlier at a little over 20%.

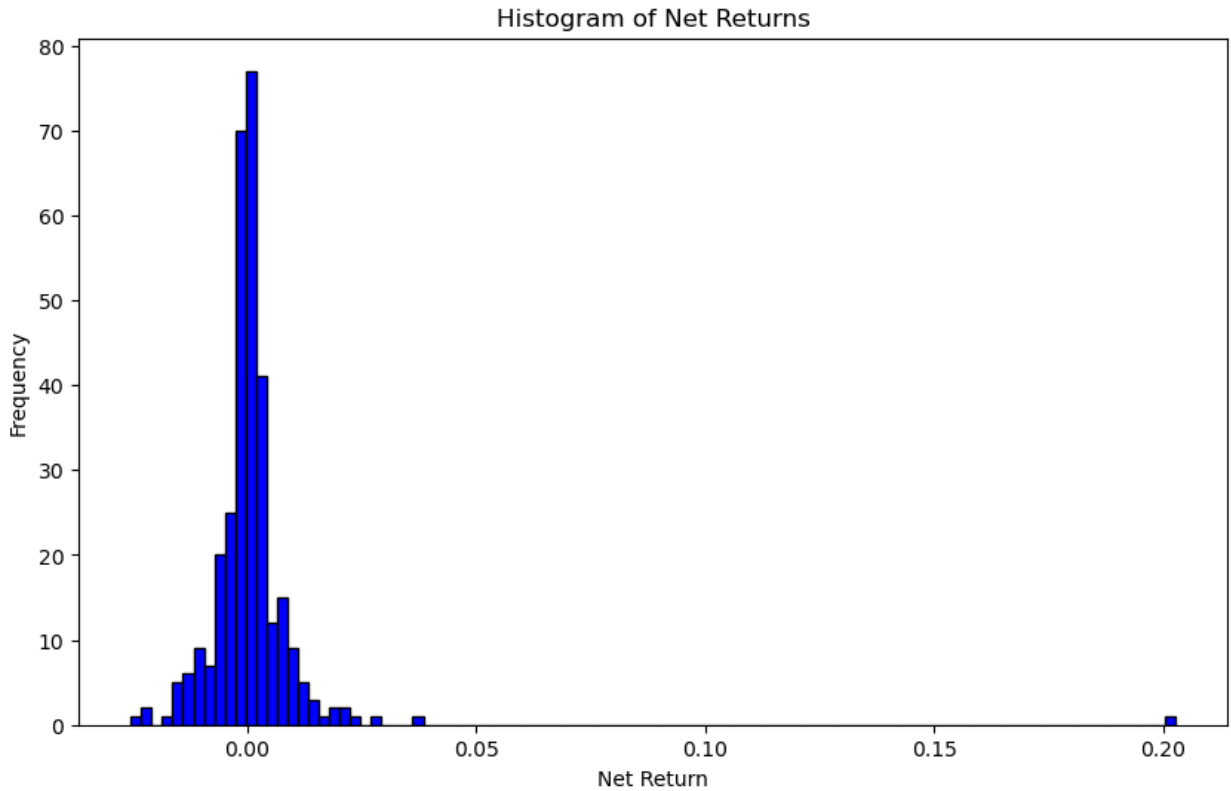


Figure 2.3

Table 3.1 below contains the data on a linear regression of Net Returns hour to hour on hourly sentiment averages.

Table 3.1, Linear Regression of Net Returns on Sentiment

<b>Sentiment</b>	0.0005
	(0.005)
<b>Intercept</b>	.0006
	(.001)
<b>N</b>	317
<b>R<sup>2</sup></b>	0.000

To say the association is weak would be an understatement. This is further compounded when we realize that over the course of the month this data was taken, GME’s price actually declined. Despite this, we see sentiment follow a different path.

Table 3.2 reruns the regression, controlling for the hour of the trading day and day of the week.

*Table 3.2, Multi-Linear Regression of Net Returns on Sentiment, Hour Traded, and Days of the Week*

<b>Sentiment</b>	-0.0029
	(.005)
<b>Hour</b>	-0.003
	(0.000)
<b>Monday</b>	0.0029
	(0.002)
<b>Tuesday</b>	0.0029
	(0.003)
<b>Wednesday</b>	0.0050
	(0.002)
<b>Thursday</b>	0.0012
	(0.002)
<b>Intercept</b>	0.0027
	(0.003)
<b>N</b>	317
<b>R<sup>2</sup></b>	0.026

The only variable that shows any statistical significance past a 95% confidence interval is trading on Wednesdays.

Next, we take a measure of the stock price of Gamestop vs its actual valuation of a company. In order to do this, we will utilize popular value measurements of equity, and compare them to the averages in other companies in Gamestops sector, which in this case is companies under the header of ‘Retail Trade.’

**Total Market Cap:** GME shares have actually taken a significant downturn from the close of the sentiment research (they closed Friday, September 15<sup>th</sup> at 18.22 a share). This gives GME a total market capitalization, or the price of shares times the number of shares, of 5.561 Billion dollars, greater than the median of other retail stocks of 905M.

**P/E Ratio, EV/EBITDA, Free Cash Flow:** We start with the price to earnings ratio, a popular stock ratio comparing a companies stock price to their earnings per share. This is taken simply by taking the price of a given share, and dividing by the companies total earnings over a given period by the total number of shares available. Here, Gamestop has no P/E Ratio, as their company has not brought in a positive profit. While this isn’t unheard of, especially in sectors like tech where prices can reflect future hope for growth or business model changes in the



company, it is less common for Retail Trade company, and puts Gamestops in a minority. Because of its non-profitability, other popular metrics, such as EV/EBITDA (Or Enterprise Value, which is all assets + debt minus cash, and EBITDA, or Earnings Before Interest, Taxes, Depreciation, and Amortization Ratio) and Free Cash Flow (the money the company has left over after paying its operating expenses and capital expenditures) are also non-existent, as again they bring in no profit. To put each in perspective, the median values among other companies in the same sector, 'Retail Trade,' are 16.949, 9.902, and 2.779, respectfully. For further perspective, there are 63 publicly traded companies on the NYSE with a lower total market cap than Gamestop, but with a positive Price to Earnings ratio, 84 with a positive EV/EBITDA ratio, and 46 with positive Free Cash Flow.

**DRS Totals:** When most retail traders or investors buy a stock, their broker is buying it on their behalf, and so the stock is in their name. Brokers can then, in the course of regular trading, do other things with that stock, such as lend it to another party for a short position. Members of the r/Superstonk subreddit seem to be keenly aware of this, and so often discuss 'DRS,' or directly registered shares, in which a share is directly registered to a single person, and thus that person has total control of what is done with those shares. During our data collection, as of June 1<sup>st</sup> 2023, the total was 76.6 million shares. Notably, that June total is around 25% of all outstanding stock of the company. Companies do not have to report their DRS totals, but to put this in perspective, after DDS (Dillard's INC) at 22% and APE (AMC Theaters) also at 22% DRS of total freefloat shares, the next highest company that reports this data, CINF (Cincinnati Financial Corporation), has 0.01% of total free float shares DRS. The DRS for Gamestop is believed, therefore, to likely be the highest of any publicly traded stock. This shows an active interest in the investors of the company not wanting to allow others to short their stock and thus drive price down.

## Conclusion

There was no suggested relationship between price action and sentiment from the users of r/SuperStonk. A casual glance at the front page of the subreddit on any given day will reveal what a deeper dive into all the posts did here: despite GME's slow price drop, sentiment has remained rather positive. In addition, despite the slow bleed out of Gamestop's price, it still remains highly elevated among its peers and to its own price point before its rally, still trading at over a 300% increase from the price it traded at in December 2020. The stock is also priced quite highly compared to companies that are in better financial position via popular metrics for checking the health of the companies' issuing shares.

The thought experiment of zealots of the stock market keeping prices artificially high has been supposed before (Shiller). To paraphrase the thought experiment, we might imagine that buyers and sellers of stock would buy a stock if they believe it is undervalued, and short a stock if it were overvalued, and thus in time drive it to its fair market value. However, if there were

'zealots' in the market, who believed in a stock despite public information about the fundamentals of the company, they could have an outsized effect on the price. The thought experiment imagines if they were to buy up a large position of the company, then after all available shares were shorted by actors who believed the company was overvalued, all that would be left were the shares that were traded among the zealots, thus keeping the price artificially high. Here, however, we have had this effect quite literally happen: the incredibly high DRS of Gamestop shares by investors has removed potential action investors can take against the company by removing shares that can be shorted. Whether or not the DRS of these shares can totally capture the seemingly high valuation of the Gamestop company in comparison to other retail locations is beyond the scope of this paper, as the stock still has a majority of shares trading non-DRS, and indeed the price has still been falling, albeit quite slowly.

# Bibliography

## **“I just like the stock”: The role of Reddit sentiment in the Gamestop share rally**

(by Suwan (Cheng) Long, Brian Lucey, Ying Xie, Larisa Yarovaya)

The Financial Review, Volume 58, Issue 1 Feb 2023

## **VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text**

(by C.J. Hutto and Eric Gilbert)

Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

## **From Efficient Markets Theory to Behavioral Finance**

(by Robert J. Shiller)

Journal of Economic Perspectives, Volume 17, Number 1, Winter 2003 Pages 83-104