Title: (Crosspost) GME Short Indicator - a data driven analysis showing how the shorts haven't closed
Author: LasVegasWasFun
Created 2022-01-31 19:02:01 UTC
Permalink: /r/GME/comments/sh93g9/crosspost_gme_short_indicator_a_data_driven/
Url:
https://old.reddit.com/r/FWFBThinkTank/comments/sgmppd/gme_short_indicator_updated_20220128/

Linked Post Content:
Cheers everybody!

GME Hodlers have had interesting weeks so far ;).

This is a update of my previous [post](https://www.reddit.com/user/SpiffyGrr/comments/s2gz9a/a_visualis
ation_of_short_volume_data_of_gme/) (thank you for all the comments). It's a visualisation of Net Short
Volume Data (FINRA, NYSE, CBOE, Nasdaq). Any constructive comment is appreciated. Feel free to
share and link for feedback.

TLDR;

>It seems that the "market" is still short on GME. The run-ups (after January 2021) of the GME price,
which seems to occur quarterly (every three months), do not affect the reported Short Volume. Even the
massive suppression recently has not affected the reported Short Volume.
>
>It seems that the short positions have not closed yet. The "short situation" seems to have deteriorated.
>
>The current situation compared to January 2021 has not changed. Everyone just has to be cool. I know
you are ;)

https://preview.redd.it/xfe2evnmpwe81.jpg?width=400&format;=pjpg&auto;=webp&s;=f0055e9d4be7bb14
d993981034c255aa7c0fc877

0. Disclaimer

This is not a financial advise. I have no financial background. I just like math, numbers, colourful plots and
crayons. I am holding some of the mentioned stocks and, therefore, my point of view is completely biased.
The presented data could be inaccurate or completely incorrect. Do your own research. Actions and their
consequences taken on the basis of this post are the sole responsibility of the respective executor. English
is not my mother tongue so this post will contain various and several mistakes. My approach could be
completely nonsense. Last but not least, my implementation and data extraction (Python, MATLAB) could
not have worked as intended. The presented data may not belong to the stated stock symbol.

1. Database

* FINRA short data: [https://www.finra.org/finra-data/browse-catalog/short-sale-volume-data/daily-short-sal
e-volume-files](https://www.finra.org/finra-data/browse-catalog/short-sale-volume-data/daily-short-sale-vol
ume-files)
* CBOE short data: [https://www.cboe.com/us/equities/market_statistics/short_sale/](https://www.cboe.co
m/us/equities/market_statistics/short_sale/)
* NYSE short data: https://ftp.nyse.com/ShortData/
* Nasdaq short data:
https://nasdaqtrader.com/Trader.aspx?id=shortsale

I have downloaded all the available data (csv text files) from 2008 (CBOE), 2009 (FINRA, NYSE) and 2010
(Nasdaq BX, PSX), respectively. I also included the free Nasdaq GSM data (prior to August 1, 2010). All
used data are freely available. Current [Nasdaq GSM
data](https://nasdaqtrader.com/Trader.aspx?id=DPUSdata#shortsale) cost $750 per month.

All the data (Short Volume and Total Volume) were joined into a single dataset. The Short Exempt Volume

will not be used since it's included in Short Volume (see [FINRA](https://www.finra.org/sites/default/files/2021-07/DailyShortSaleVolumeFileLayout.pdf)).

Furthermore, I gathered market data (price, outstanding shares, float, short interest etc.) from Yahoo Finance via python and [yfinance · PyPI](https://pypi.org/project/yfinance/) . I also downloaded [FTD data](https://www.sec.gov/data/foiadocsfailsdatahtm).

We have to assume that all these data are inaccurate and manipulated to an unknown extent (imperfect data).

Data verification: I exemplarily verified the plotted data with data from [ChartExchange](https://chartexchange.com/) for GME, AMC, TSLA, AMZN, MSFT, GOOG for a time period from 2019-01-02 to 2021-01-27. The data matches except for 2021-12-07 (across all mentioned stocks). ChartExchange shows, e.g., "FINRA Off Exchange = 307770" for GME. I have checked the "raw" data from [FINRA 2021-12-07](https://cdn.finra.org/equity/regsho/daily/CNMSshvol20211207.txt). It states 431079 for GME. I'm sticking with the "raw" data, as this small difference will not bring any changes to the SHI.

Download: https://drive.google.com/file/d/1WsXTx0HwsRZy1fLtBFjtbMzM0v1WXVHZ/view?usp=sharing

Everbody is invited to verify the data and share the results.

2. Assumption / Method of calculation

First of all, Short Volume is not Short Interest.

From the data (section 1), we know that

>Total Volume T = Short Volume S + Long Volume L.

Let NetShort be a daily "Short Indicator" (in number of traded shares) with the following property

>NetShort := S - L = S - (T - S) = 2S - T.

This idea was already published somewhere (e.g., [Quiver Quantitative](https://www.quiverquant.com/offexchange/GME)), also on Reddit.

NetShort can be positive or negative on a particular day. The approach is to integrate (accumulate) this quantity NetShort over a long time period.

>Short Indicator SHI = Sum of all NetShort over a specific time period

This quantity is called the Short Indicator (SHI) to distinguish daily Net Short volume from cumulative Net Short volume. In the following all presented SHI are integrated from beginning of the available data, i.e., 2008. Therefore, it should be possible to identify a trend or correlation if the market is "short" or "long" on a particular stock.

A long-term point of view of SHI should reveal the trend despite the fact that theses data could be manipulated. And we know that Wall Street firms were "fined" for violation of the short reporting rules.

Since we do not know the initial state, i.e., the number of shorts at the beginning, the SHI are shifted by an unknown offset. Therefore, only the changing of the SHI curve (derivative, delta) is meaningful.

3. Results

The following plot shows the following information:

1. Axis: SHI (red curve) of GME in comparison with the price; in the left corner some ticker Yahoo Finance

information are displayed.
2. Axis: Total volume from Yahoo and the ratio (blue) of Total Volume (FINRA, CBOE, NYSE, Nasdaq as far as available) to Yahoo Total Volume. The ratio is plotted as moving mean (5 days) period to reduce the "noise".
3. Axis: Total and Short Volume from FINRA, CBOE, NYSE and Nasdaq (as far as available). The blue line shows the ratio of the Short and Total Volume within this dataset. The ratio is plotted as moving mean (5 days) to reduce the "noise".
4. Axis: Fails-To-Deliver (updated 2021-12-31).

(mm = million)

https://preview.redd.it/g4qbrckagwe81.png?width=3000&format;=png&auto;=webp&s;=fde075c294b130715c63e0c98e52adbf03c6f93a

* The beginning can be neglected since in 2008 only CBOE data are available.
* The data from Augst 2009 until August 2010 is interesting since it contains data from FINRA, CBOE, NYSE and complete NASDAQ. The ratio of these total volumes to the Yahoo volume is nearly 100%. This means that all major trading volume is handled on these exchanges. Nasdaq data as of August 2010 is not freely available.
* According to [GME Dark Pool Statistics | ChartExchange](https://chartexchange.com/symbol/nyse-gme/stats/) the latest volume on Nasdaq GSM is \~10% (up to max. \~23%). From August 2010 on the total volume from FINRA, CBOE and NYSE is \~80% of Yahoo total volume. I assume that the a large portion of missing volume is traded over Nasdaq GSM. From August 2009 to August 2010, the mean value of Short Volume divided by Total Volume (only Nasdaq GSM) was 55% (min 25%, max 75%). The following histogram (bin width 5%-point) shows the distribution.
* Speculation: I assume that the cumulated net short volume on the Nasdaq GSM is at least not negative and therefore will not lower the SHI. Unfortunately I don't have any data to back it up. Maybe someone has an idea how to verify this.

https://preview.redd.it/5h46ys2qmwe81.png?width=1920&format;=png&auto;=webp&s;=633fae2c169f15dfeadb604edf468b388e98a109

* Starting in 2010 the SHI (cumulated net short volume) is rising more or less linearly.
* In 2017 / 2018 the derivative d(SHI)/dt has changed from linear to exponential or at least faster than linear, i.e., the SHI is increasing significantly faster.
* The SHI approaches a "limit" in 2020. The value has increased by over 1050 million (mm) shares from 2010 until 2020. The starting point is unknown and, therefore, a shift of the curve in y-direction is likely. But the change of 1050 mm shares remains unaffected.
* During the events in January 2021, the SHI dropped by ca. 200 mm. Since then the SHI is rising again and has reach in January 2022 a higher level as in January 2021 (ca. 1150 mm). **The "short situation" seems to have deteriorated.**
* This drop in January 2021 correlates with the SEC report stating that only a small amount of shorts were covered during this event.
* There was a GME share dilution in 2021. And the SHI is still continuously increasing since February 2021.
* The run-ups after January 2021 which seems to occur quarterly does not affect the SHI. The SHI is still increasing. Even the massive suppression recently has not affected the reported Short Volume. SHI is constantly increasing.
* All these observations are indicators that the short positions may not have closed yet.
* The size of the current float and the price must be taken into consideration while assessing the SHI.

The following plot shows a zoom of GME from 2019 to 2022.

https://preview.redd.it/ln1rgypbgwe81.png?width=3000&format;=png&auto;=webp&s;=749f7338edabe9c9540426b1f45ef641a11991da

The SHI plot of GME looks very unusual. Unless you assume that the goal was/is to short GME into oblivion. It seems that the short sellers are not able or willing to close their short positions.

Is this behaviour unique for GME? What about other stocks? What are their trends?

* The SHI of Popcorn Stock has reach a maximum level in December 2020. The delta of SHI from 2004 to 2020 is ca. 420 mm. Since then the SHI dropped by ca. 1800 mm shares. In particular, the drops occur at the two run-ups. This could be an indicator for (partial?) closing of short positions.
* It must be considered that the float of Popcorn Stock is much higher than GME (513 mm vs. 63 mm). Popcorn has a massive share dilution in 2021. Not mentioning the insider selling.

https://preview.redd.it/eart7pj3qwe81.png?width=3000&format;=png&auto;=webp&s;=bb44e118dde814c5a6230ffff8a6d3b079f79b94

In the following, you will find other SHI plots for comparison without commentary.

Be aware that correlation is not the same as causality!

"Look behind you, a three-headed monkey!" $Mic Drop$

https://preview.redd.it/cv9ls45oqwe81.png?width=3000&format;=png&auto;=webp&s;=30f3cf64a0894b76c0da8cbd74e14ef55f96840d

https://preview.redd.it/uqhv9c8pqwe81.png?width=3000&format;=png&auto;=webp&s;=2c355305fd36bf45a0b2073bfa03be68c1fa1fac

A stock split causes the YF ratio (2. axis) to increase (see, e.g., TSLA and AAPL) as Yahoo adjusts its historical volume data while FINRA, CBOE and NYSE do not.

https://preview.redd.it/ji6fhu60rwe81.png?width=3000&format;=png&auto;=webp&s;=ca49ab3d001b546eab8dd8e6b1193f5998cc8df1

https://preview.redd.it/5l9m2g01rwe81.png?width=3000&format;=png&auto;=webp&s;=a58469e878fd2295a08d69e2fae04408e9cada1e

https://preview.redd.it/wwpqmnm2rwe81.png?width=3000&format;=png&auto;=webp&s;=b6ae5461d4e906fec3be20ede50b75cd4396b40e

https://preview.redd.it/7fqy8e35rwe81.png?width=3000&format;=png&auto;=webp&s;=d11b664722cc24cf5c344e3461e2eb66b2a47597