MULTI-DIMENSIONAL ALPHA

LUO'S QES RESEARCH
Quantitative Research
Economics
Portfolio Strategy

January 29, 2021

A BATTLE BETWEEN RETAIL AND INSTITUTIONAL INVESTORS

Quantifying Retail Behaviors using Reddit, Robinhood, and NLP Analysis

- **Retail Red Alert Screen.** To help our readers navigate in this highly unusual time – a battle between retail and institutional investors – we have created a Retail Red Alert Screen that is updated daily before market open and reflects the latest retail buzz (on the Reddit), participation (via retail trading), and a suite of risk profiles (e.g., short interest, hedge fund crowding, options open interest and trading volume).
- **Quantifying Retail Participation using Reddit, Robinhood, Retail Trading, and NLP.** We have developed a massive scale of web scraping, NLP algorithms, and parallel GPU computing to process information on the Reddit platform. In particular, we identify the stocks with the most daily mentions within the r/wallstreetbets community, targeting the thread "What Are Your Moves Tomorrow", where Reddit users potentially share speculative investing ideas for the upcoming trading day. We then supplement our analysis with historical retail popularity metrics from Robinhood and an ongoing data feed from another major retail broker.
- **Retail Investor Behavior Biases.** Individual investors tend to have very different behavior biases from institutional asset managers. Furthermore, retail sentiment from Reddit also appears to have different characteristics from actual retail trading. Specifically, retail investors demonstrate highly speculative behavior, by investing in companies that are heavily mentioned in the news (either positive or negative news) and concentrating their stock picks in the technology and consumer sectors.
- **Predicting Short Squeeze.** The recent retail investing frenzy has caused significant damages to short sellers. Historically, short squeezes are rare events but highly damaging. Stocks that are at risk of short squeezes tend to be retail investors' favorite bets. Our empirical analysis suggests that removing Reddit buzz stocks from the short side of active strategies can boost performance and reduce risk.
- **Retail Forced Hedge Fund De-leveraging.** So far, the retail triggered hedge fund de-leveraging has most impacted the short side, as some hedge fund managers were forced to close short positions that were moving against them to reduce overall leverage and risk. On the long side, our hedge fund crowding signal has been muted in recent days with most of the selling in highly liquid large-cap stocks. However, should retail investor hysteria further accelerate, it could prompt further de-risking with more significant impact on crowded longs.

Source: pixabay.com /

Javed Jussa
JJussa@wolferesearch.com

Sheng Wang
SWang@wolferesearch.com

Yin Luo, CFA, CPA
YLuo@wolferesearch.com

Xiaohui (Victor) Li
XLi@wolferesearch.com

Jason Zhong, PhD
JZhong@wolferesearch.com

Hallie Martin
HMartin@wolferesearch.com

Gaurav Rohal, CFA
GRohal@wolferesearch.com

Shu Liu
SLiu@wolferesearch.com

QES: 1.646.582.9230
Luo.QES@wolferesearch.com

Table of Contents

RETAIL RED ALERT SCREEN

To help our readers navigate in this highly unusual time – a battle between retail and institutional investors, we have created a Retail Red Alert Screen (see Figure 1). Detailed methodology is elaborated in the rest of this research. This screen includes stocks at the center of the retail trading frenzy, generated with our proprietary NLP (Natural Language Processing) analysis of the Reddit platform, data from a major retail brokerage firm, historical data from the dominant retail trading platform – Robinhood, and our suite of machine learning algorithms. We expect continued and heighted volatility for these stocks. Therefore, unless managers have strong convictions, these stocks probably should be carefully watched on both long and short sides. The screen is updated daily, prior to market open, to help investors position for the upcoming trading day. We further incorporate social media posts captured in the pre-market; and therefore managers can be informed about potentially impending market moving events.

Specifically, we construct the Retail Red Alert basket using the following data:

1. **Reddit**: Identifying the stocks with the most daily mentions within r/wallstreetbets community. We specifically target the thread *What Are Your Moves Tomorrow*, where Reddit users potentially share speculative investing ideas for the upcoming trading day. This is a forward-looking idea bank of stock trading ideas. This type of thread can be informative, because comments, posts, votes can be accessed any time (e.g., prior to market open).
2. **Robinhood**: Robinhood is a dominant commission-free trading platform that is popular among individual investors. Although Robinhood discontinued its data access in August 2020, historical data still reveals tremendous insights on retail investors' behavior biases.
3. **Another major retail broker**: We are also able to identify the most actively bought and sold stocks traded through another major retail broker during recent market sessions. This is updated throughout the trading day (8AM to 4PM EST); and therefore captures premarket trading activities.

The top 30 stocks with the most buzz and participation from retail investors are selected in our final stock screen. We also included a few other critical metrics to inform investors about the companies highlighted:

- Company name, ticker, market cap, sector, and industry
- Final ranking score (a high reading indicates high risk)
- Reddit Score
- Retail Broker Score
- Short Interest % (reflecting the extent that a stock is being shorted)
- HF Concentration % (reflecting the extent that a stock is owned by hedge funds on the long side)
- Call Options Open Interest (total number of outstanding call options contracts that have not been settled[1])

[1] This is not delta or gamma adjusted. It is based on the upcoming weekly option expiry date and includes options at various strike prices.

- Put Options Open Interest (total number of outstanding put options contracts that have not been settled[8])
- Call Options Contract Volume (total number of call contracts actively traded during the previous trading day[2])
- Put Options Contract Volume (total number of put contracts actively traded during the previous trading day [9])

Figure 1 Wolfe QES Retail Red Alert Screen

Ticker	Name	Final Average Rank (Higher = WORST)	Market Cap ($Mlns)	Large/Small Cap Index	GICS Sector	GICS Industry	Reddit Score	Retail Broker Score	Short Interest (%)	HF Concentration	Call Open Interest (# of contracts) *	Put Open Interest (# of contracts)*	Call Contract Volume *	Put Contract Volume *
GME	GameStop Corp.	30	$ 13,503	Russell 2000	Consumer Discretionary	Retailing	20710	33.9	68.86	32%	109,269	329,279	46,968	136,938
AMC	AMC Entertainment Holdings Inc	29	$ 2,479	Russell 2000	Communication Services	Media & Entertainment	16329	31.6	63.76	4%	238,895	191,945	106,601	115,211
NOK	Nokia Corp	28	$ 26,517		Information Technology	Technology Hardware & Equipment	10200	3.9	69.80		434,848	90,806	357,938	72,497
SNDL	Sundial Growers Inc	27	$ 916		Health Care	Pharmaceuticals, Biotechnology & Life Sciences	1720	18.7	52.24		59,899	11,006	183,342	11,838
BB	BlackBerry Ltd	26	$ 8,243		Information Technology	Software & Services	7810	3.6	51.57		235,950	114,756	194,327	97,814
SPCE	Virgin Galactic Holdings Inc	25	$ 10,102	Russell 1000	Industrials	Capital Goods	366		75.04	10%	93,574	42,298	83,902	27,084
PLTR	Palantir Technologies Inc	24	$ 52,486		Information Technology	Software & Services	502	1.4	89.36		198,264	130,498	176,894	84,471
RKT	Rocket Cos Inc	23	$ 2,459	Russell 1000	Financials	Banks	86		86.38	27%	75,281	32,621	32,455	9,810
AAL	American Airlines Group Inc	22	$ 10,950	Russell 1000	Industrials	Transportation	651	1.0	56.42	3%	118,470	82,685	147,049	77,732
LMFA	LM Funding America Inc	21	$ 40		Financials	Diversified Financials		0.9	97.68					
TSLA	Tesla Inc	20	$ 791,905	Russell 1000	Consumer Discretionary	Automobiles & Components	935	18.7	14.90	2%	281,942	239,214	372,779	194,133
AMD	Advanced Micro Devices Inc	19	$ 105,261	Russell 1000	Information Technology	Semiconductors & Semiconductor Equipment	121	2.5	9.79	5%	334,261	167,809	166,590	45,565
NAKD	Naked Brand Group Ltd	18	$ 134		Consumer Discretionary	Consumer Durables & Apparel	72	1.5	73.33					
FB	Facebook Inc	17	$ 637,444	Russell 1000	Communication Services	Media & Entertainment	69		0.08	4%	165,849	93,280	232,437	75,898
EXPR	Express Inc	16	$ 305	Russell 2000	Consumer Discretionary	Retailing		1.1	64.78	24%				
BNGO	BioNano Genomics Inc	15	$ 2,862		Health Care	Pharmaceuticals, Biotechnology & Life Sciences		2.1	85.20					
BBBY	Bed Bath & Beyond Inc.	14	$ 4,078	Russell 2000	Consumer Discretionary	Retailing	305		50.58	21%	81,460	57,329	74,790	60,492
STAR	iStar Inc	13	$ 1,158	Russell 2000	Real Estate	Real Estate	707		23.84	7%				
NIO	Nio Inc	12	$ 75,432		Consumer Discretionary	Automobiles & Components	111	1.4	14.53		232,854	110,752	199,035	48,152
AG	First Majestic Silver Corp	11	$ 3,730		Materials	Materials	180	2.2	44.79					
AAPL	Apple Inc	10	$ 2,306,301	Russell 1000	Information Technology	Technology Hardware & Equipment	397	2.4	0.14	1%	807,933	366,046	932,868	283,231
AMZN	Amazon.com Inc	9	$ 1,624,479	Russell 1000	Consumer Discretionary	Retailing	65	1.4	0.14	2%	91,437	73,357	121,156	64,323
RIOT	Riot Blockchain Inc	8	$ 1,346		Information Technology	Software & Services		0.9	70.40					
APPS	Digital Turbine Inc	7	$ 5,418	Russell 2000	Information Technology	Software & Services	473		3.87	11%				
CTRM	Castor Maritime Inc	6	$ 222		Industrials	Transportation		2.1	63.15					
WISH	Contextlogic Inc	5	$ 13,603		Consumer Discretionary	Retailing	381		62.78					
MARA	Marathon Patent Group Inc	4	$ 1,869		Information Technology	Software & Services		0.8	69.19					
ALLY	Ally Financial Inc	3	$ 14,447	Russell 1000	Financials	Diversified Financials	117		1.63	16%				
NVAX	Novavax Inc	2	$ 8,531	Russell 2000	Health Care	Pharmaceuticals, Biotechnology & Life Sciences		0.7	26.55	9%				
TLRY	Tilray Inc	1	$ 2,960		Health Care	Pharmaceuticals, Biotechnology & Life Sciences	64		67.59					

*This is not delta or gamma adjusted. It is based on the upcoming weekly option expiry date and includes options at various strike prices. The volume is cumulated across all strike prices based on the upcoming weekly option expiry date.

Sources: Reddit, Robinhood, Wolfe Research Luo's QES

[2] The volume is cumulated across all strike prices based on the upcoming weekly option expiry date.

A TRADING WAR BETWEEN RETAIL AND INSTITUTIONAL INVESTORS

Since we published a research note on the rise of retail participation in the US equity market using Robinhood data (see *Retail Investor Behavior and Participation Alpha*) in summer 2020, the retail trading frenzy has further accelerated. A weak labor market, working from home policies, and generous government stimulus and unemployment benefits, have encouraged many, especially the younger generation, to become enthusiastic investors since the onset of COVID-19 pandemic in early 2020.

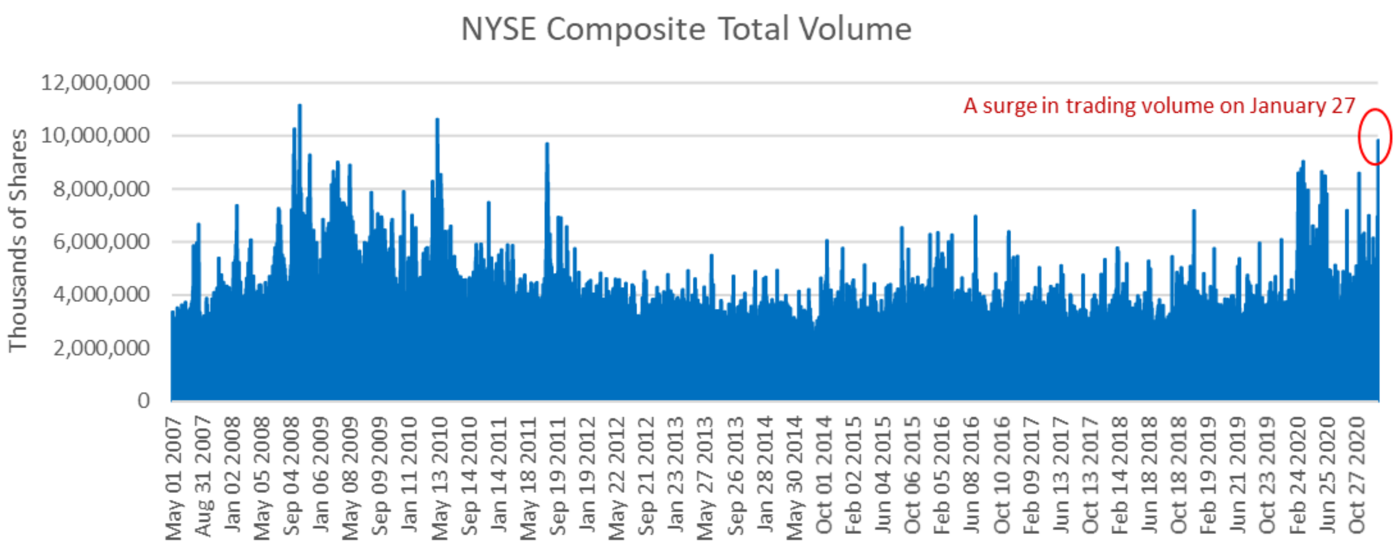
Retail investors are exerting increasing influence on the market. It turns out that they can achieve much more than anyone could imagine, as is evidenced by the recent epic battle centered around GME (GameStop), AMC and other heavily shorted stocks in recent days. An army of individual investors is wielding growing influence across global financial markets, causing ructions in share prices and forcing sophisticated managers such as hedge funds to retreat.

The frenzied battle between YOLO enthusiasts and sophisticated hedge funds has intensified and pushed trading volumes to all-time highs, exceeding the previous peaks observed during the 2008 GFC. As shown in Figure 2, the trading volume in the NYSE and NASDAQ exchanges both surged on January 27, especially the more volatile NASDAQ market.

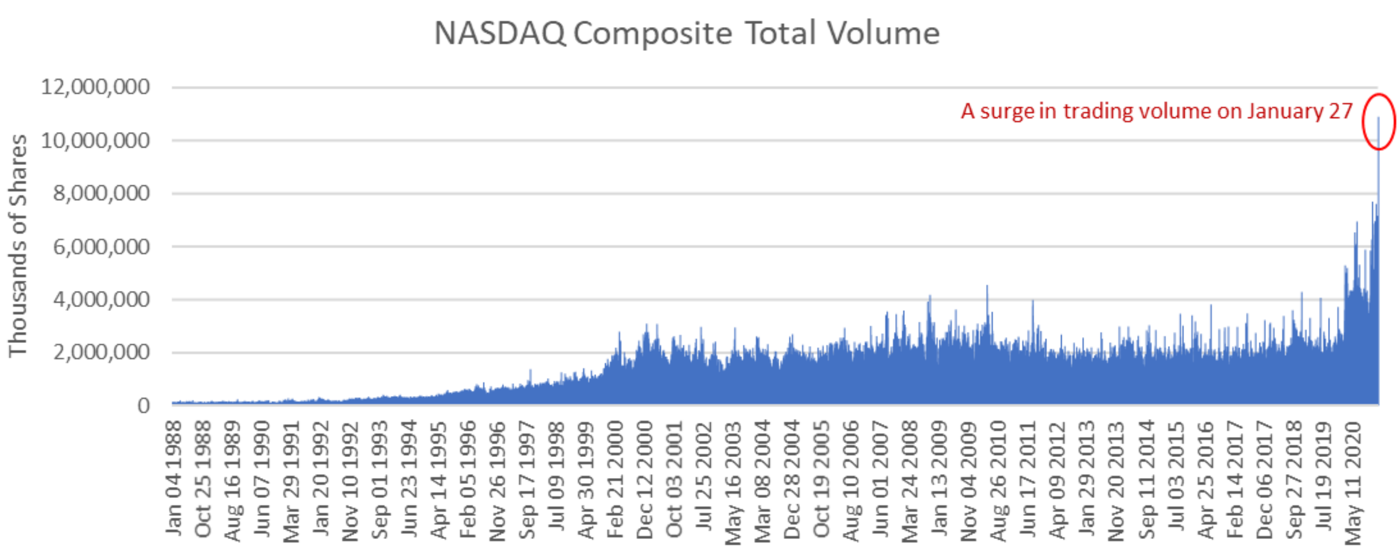
This battle is heating up and situation changes rapidly. On Thursday, Robinhood and other retail brokers restricted trading in a number of companies whose stock prices spiked recently, which draws anger from customers/users, a class action lawsuit against Robinhood, and criticism from politicians. Robinhood cited capital requirements from regulators and clearing houses for its trading restrictions. After tapping into its credit facilities, Robinhood suggests that it may relax some of the trading restrictions.

Figure 2 US Equity Trading Volume on January 27, 2021

A) NYST Trading Volume



B) NASDAQ Trading Volume

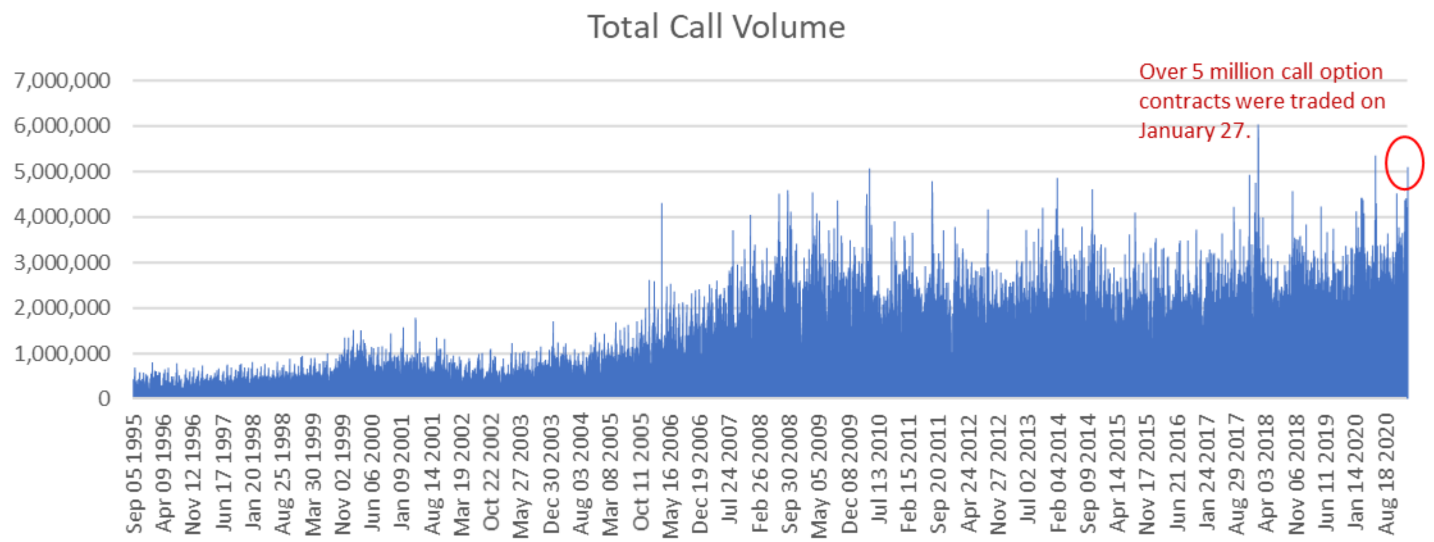


Sources: Wall Street Journal, Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Refinitiv, Thomson Reuters, Wolfe Research Luo's QES

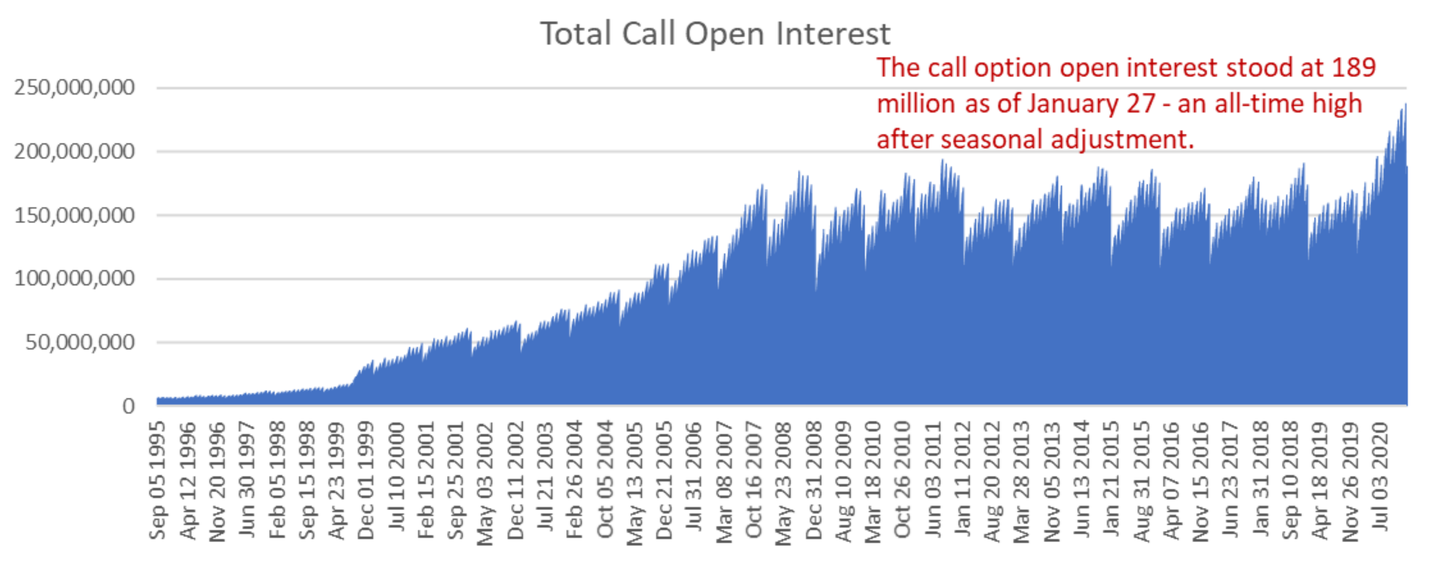
The soaring equity trading volume was matched by record options activities. Based on CBOE's statistics, over five million call option contracts were traded on January 27 alone (see Figure 3A). Total open interest of call options also hit 189 million – a new all-time high after seasonal adjustment[3] (see Figure 3B).

Figure 3 US Option Trading Volume on January 27, 2021

A) Call Option Trading Volume



B) Call Option Open Interest



Sources: CBOE Global Markets, Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Refinitiv, Thomson Reuters, Wolfe Research Luo's QES

[3] There are strong seasonal patterns for option's open interest due to expiration of options. For instance, the open interest would drop significantly in mid-January after a large number of previous year-end options have expired.

Measuring Retail Participation

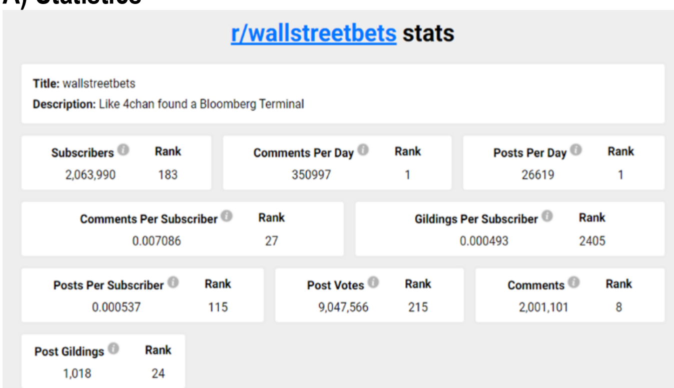
As retail investors become increasingly influential, there is limited data covering individual investor activities and sentiment. In this section, we introduce three interesting data sources – Reddit, Robinhood, and another retail broker. Equipped with cutting edge NLP (Natural Language Processing) and machine learning algorithms, we are able to reveal interesting insight from this investor segment of the market.

Reddit

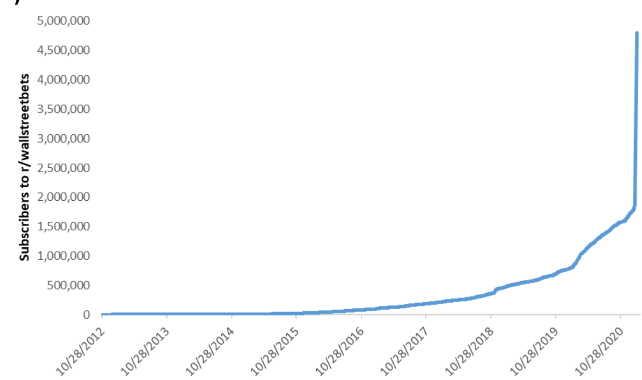
Reddit is a vast social platform and discussion website, comprising user generated content such as text-based posts, video, links, and forums. Registering an account with Reddit is free and does not require an email address. As of the beginning of 2020, Reddit boasts approximately 430 million users[4]. Reddit is also available as a mobile app. It is in fact, the sixth most popular mobile app in the US[5]. The site's content is divided into several communities or "subreddits". There are approximately 2.2 million subreddit communities[6]. The WSB or "r/wallstreetbets" is a popular subreddit community for users who enjoy discussions on stock trading and speculative investing (see Figure 4A). It has over 2 million subscribers on average with over 350,000 comments per day. The number of subscribers nearly triples in less than a month recently (see Figure 4B).

Figure 4 r/wallstreetbets

A) Statistics



B) Number of Subscribers



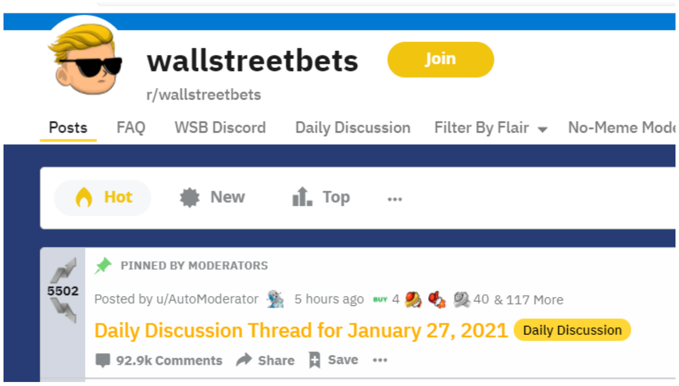
Sources: Reddit, Wolfe Research Luo's QES

Reddit discussion communities can often be peppered with an enormous dose of conversation, debate, dispute, irony, memes, and even expletives. Subreddits are overseen by moderators who are volunteers that manage their community and enforce its guidelines and rules. Content within subreddits can be categorized using flairs. These are essentially tagging systems for posts as well as individual users. It can be used to filter content, unique identifiers, and resources. Figure 5 shows a popular flair called *Daily Discussion,* which is used to filter posts within r/wallstreetbets. The Daily Discussion tag can be used further to filter a thread called *What Are Your Moves Tomorrow* – a daily discussion thread where trading ideas are shared and empowered.

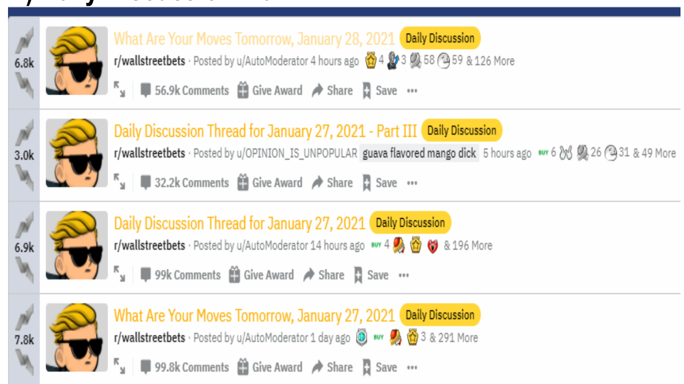
[4] (Reddit, 2019)
[5] (Statista, 2019)
[6] (Reddit Metrics, 2020)

Figure 5 Reddit Screenshots

A) r/wallstreetbets



B) Daily Discussion Flair



Sources: Reddit, Wolfe Research Luo's QES

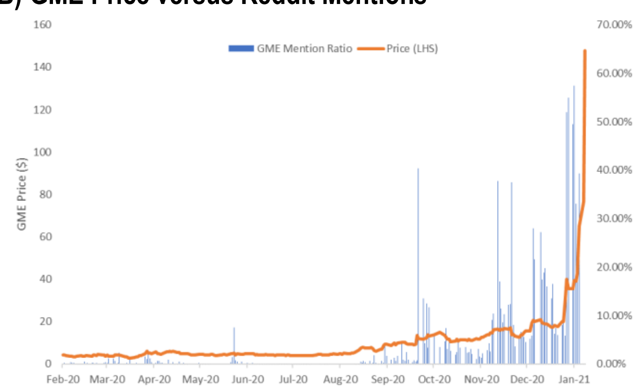
An Example – GME

Empowered trading ideas can quickly multiply on Reddit to astonishing levels. Figure 6(A) shows a word cloud of r/wallstreetbets posts on January 22nd, 2021, with all common expletives removed. As expected, GME is the most mentioned word. The sheer volume of those mentions can dramatically affect share prices. As shown in Figure 6(B), there is a clear positive correlation between GME's share price and its Reddit mentions.

Figure 6 GME

A) World Cloud (January 22nd, 2021)

B) GME Price versus Reddit Mentions



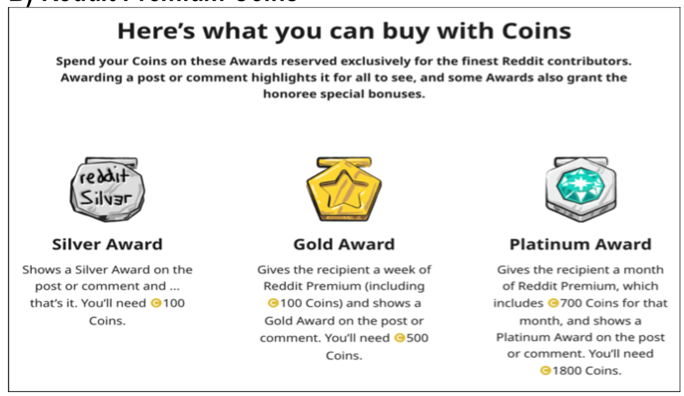
Sources: Reddit, Wolfe Research Luo's QES

A key feature of Reddit is that users can provide both positive and negative feedback on post by upvoting or downvoting. The number of upvotes or downvotes determines the comment's visibility within the site (see Figure 7A). Users can also earn "karma" for their posts and comments, a status that reflects their standing within the community and their contributions to Reddit. Reddit also offers Premium members. It allows users to view the site ad-free. Users may also be gifted coins if another user valued their comment or post, generally due to humorous or high-quality content (see Figure 7B).

Figure 7 Reddit Features

A) Upvoting/Downvoting

B) Reddit Premium Coins

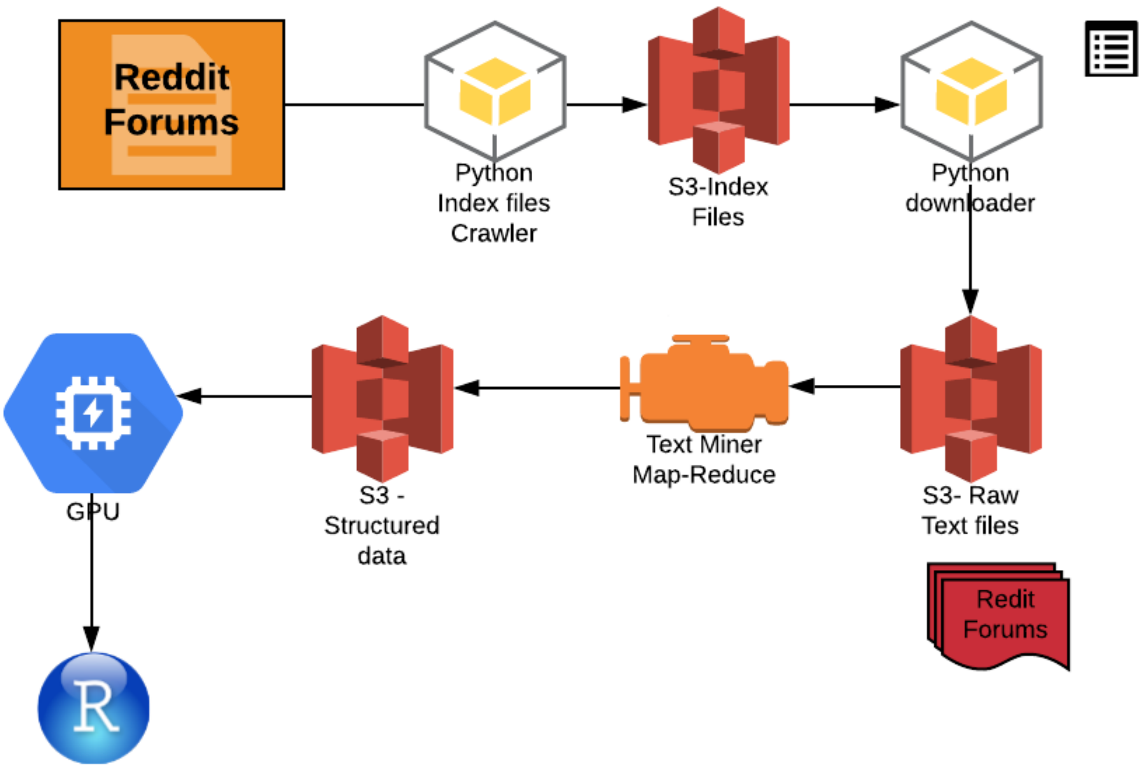


Sources: Reddit, Wolfe Research Luo's QES

Accessing Reddit using our Fault-Tolerant NLP Infrastructure

To optimally process the sheer volume of data within Reddit and other sources, we leverage our cutting edge and fault tolerant NLP framework (see Figure 8). We take advantage of the PySpark programming – a collaboration of Apache Spark and Python. Apache Spark is an open-source cluster-computing framework, built for speed, ease of use, and streaming analytics, while Python is a general-purpose, high-level programming language that is becoming the de facto language of choice for data scientists. The PySpark framework gives us the ability to conduct parallel computing with multiple nodes (machines and instances). We also make heavy usage of the EMR (Elastic Map Reduce) service from AWS. The EMR can process terabytes of data across a Hadoop cluster of virtual servers on Amazon EC2 (Elastic Compute Cloud) and S3 distributed data storage.

Figure 8 QES NLP/ML Computing Framework



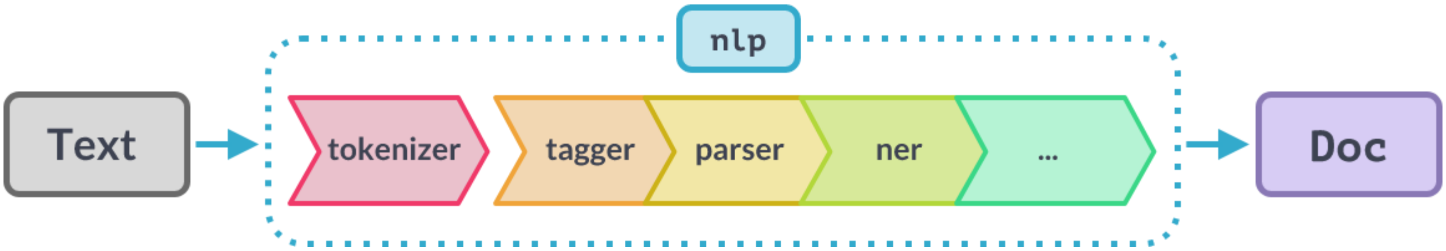
Sources: Wolfe Research Luo's QES

The core of our NLP engine is written in the Python programming language, fully developed by our team over the past decade. In recent years, we have made further transitions from R to Python, especially in the NLP area, to leverage Python's exhaustive set of NLP/ML libraries. To further accelerate our text processing, we exploit GPU (Graphical Processing Unit) computing. A GPU is a specialized microprocessor, optimized to display graphics and images. While a typical CPU (Central Processing Unit) consists of four to eight cores, GPUs have hundreds or even thousands of cores. Although GPUs are designed primarily for special-purpose graph manipulation, we can tune them to perform mathematical operations. Leveraging GPU computing reduces the run time of our NLP/ML models from days to minutes. This enables us to scrape discussion forms in real time as forums get updated continually with users posts (e.g., Reddit, Twitter, Facebook).

The NLP Nuances of Reddit

Comments posted to r/wallstreetbets are notoriously noisy and filled with the unique Reddit-style slang and lingo, which are not only confusing for newbies, but also pose challenges for autonomous NLP algorithms. Consequently, we create a customized NLP processing pipeline to address the unique textual and terminology characteristics of Reddit (see Figure 9). Specifically, we leverage the spaCy NLP toolkit – an industrial strength NLP and deep learning processor. It excels at large-scale and complex information extraction tasks, especially in the area of text preparation for deep learning algorithms.

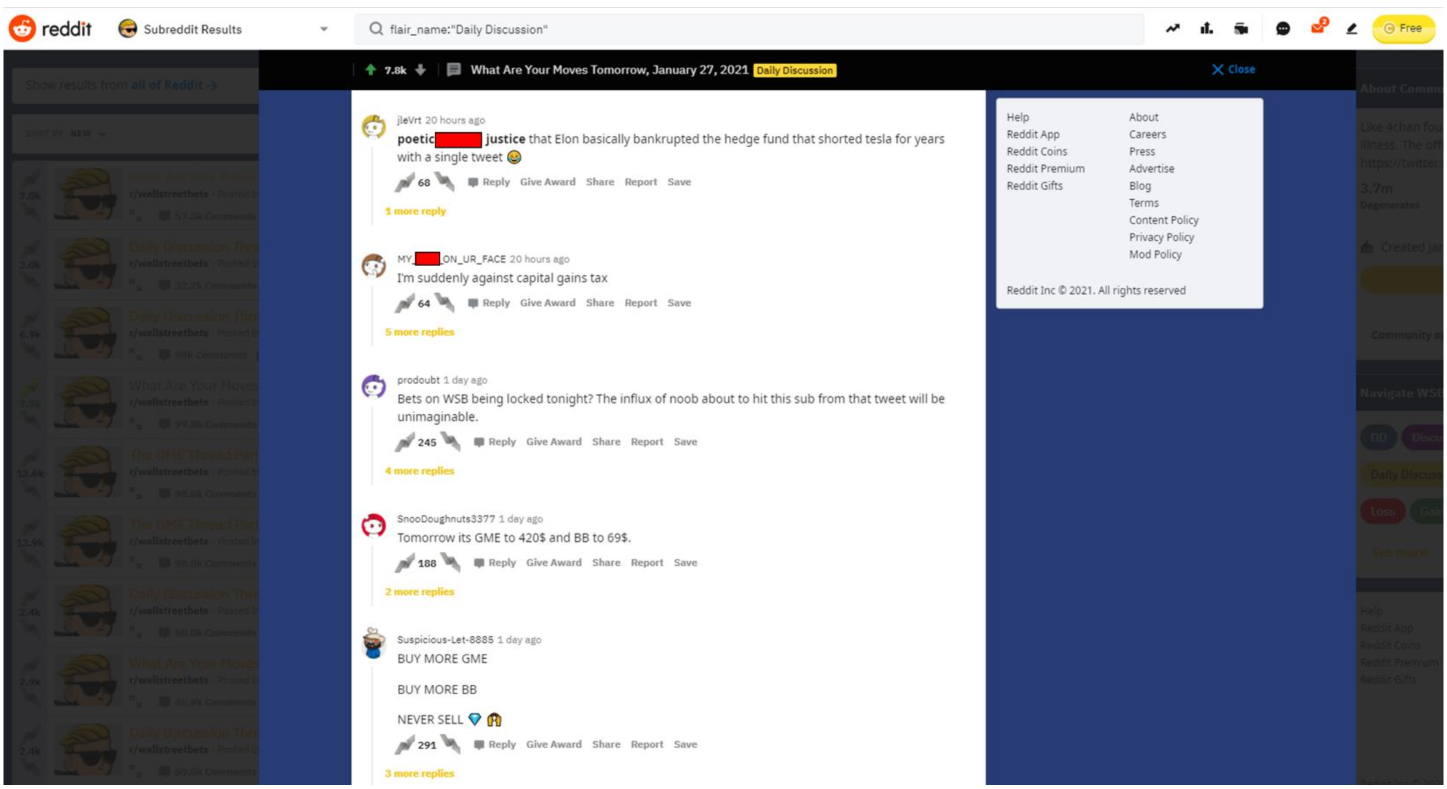
Figure 9 A Simple Illustration of the spaCy Framework



Sources: https://spacy.io/, Wolfe Research Luo's QES

We apply our customized NLP processor to Reddit threads. We focus on the subreddit community r/wallstreetbets and filter using the Daily Discussion flair tag. Next, we focus on a specific thread that speculates on trading ideas for the upcoming trading day. The thread is called What Are Your Moves Tomorrow (see Figure 10). Please note that all expletives have been blurred out. This appears to be a forward-looking idea bank of stock trading ideas for individual investors. This type of thread can be informative, because comments, posts, votes can be accessed any time (e.g., prior to market open).

Figure 10 Reddit Thread: What Are Your Moves Tomorrow



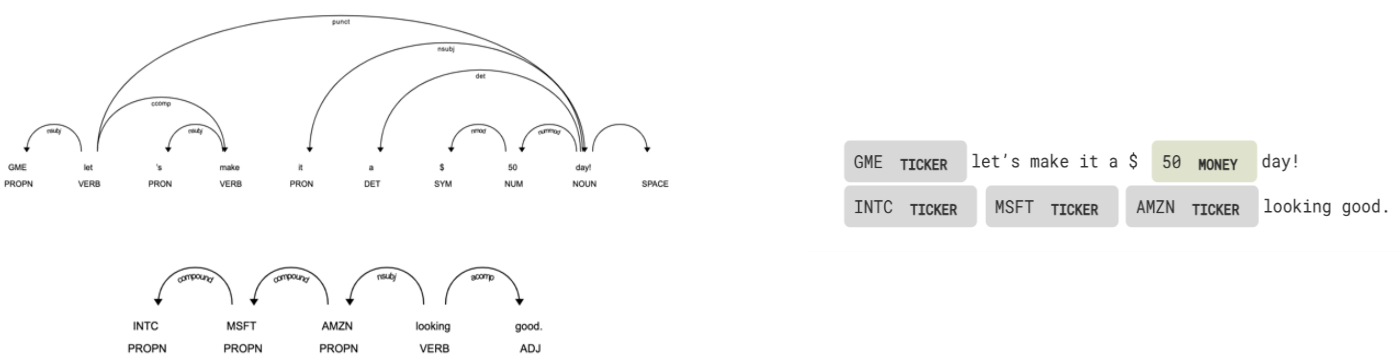
Sources: Reddit, Wolfe Research Luo's QES

Once our machines access the Reddit thread, we need to extract the number of mentions for each stock. To do this, we apply a part-of-speech (POS) tagging algorithm to assign each token using a tense label (see Figure 11). Next, we parsed each token and execute a rule-based ticker matching algorithm using our point-in-time framework. We only recognize the ticker entity that is both a pronoun and in the Russell 3000 universe. This enables a high-quality, point-in-time mapping framework that lays the foundation behind our research.

Reddit users typically leverage low-cost trading platforms like Robinhood to trade stocks and options.

Figure 11 Sector Neutral Long/Short Signal Weighted Portfolio for the Russell 1000, November 9th

A) Part-of-speech tagging **B) Ticker Name Entity Recognition**



Sources: Wolfe Research Luo's QES

ROBINHOOD

Robinhood is one of the pioneers of commission-free investing. The free app allows users to make unlimited commission-free trades in stocks, ETFs, options, and cryptocurrencies. The Robinhood app even enables users to buy/sell fractional shares, build customized and balanced portfolios, as well as trade in real time, with no minimum required account balance. However, short selling is also restricted on the platform. Users can buy inverse ETFs and put options, which allow investors to profit from market declines. Robinhood offers both a mobile and web streamlined trading platforms.

As of May 2020, Robinhood boasts about 13 million users. As a comparison, two of the largest existing retail brokers – TD Ameritrade and E-Trade have approximately 11 and 5 million customers, respectively.

Robintrack

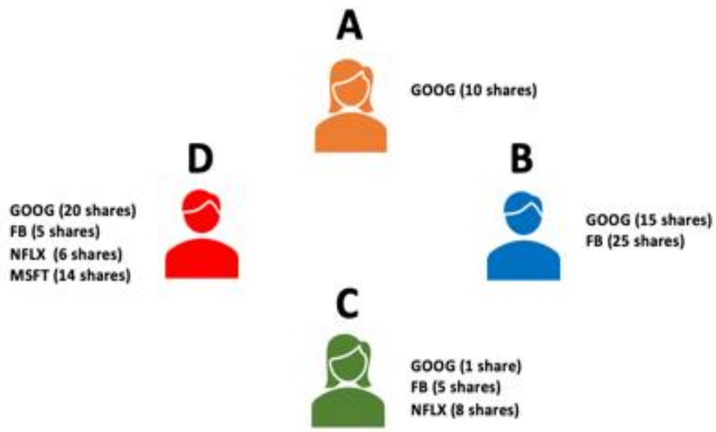
A key dimension of Robinhood data is provided by Robintrack. which offers the retail popularity score i.e., number of Robinhood users owning a specific stock. Robintrack also provides its users the ability to download Robinhood popularity data daily. Robintrack data is updated nightly[7]. Historical data starts from 2018.

Figure 12 shows an example of how the popularity score is computed. In short, popularity is simply the number of Robinhood users holding a stock. In the example above, Google is held by four users and

[7] Users can also download incremental data on an hourly basis.

therefore has a popularity score of four, while Microsoft is owned by only one investor with a popularity score of one.

Figure 12 Popularity Score Example



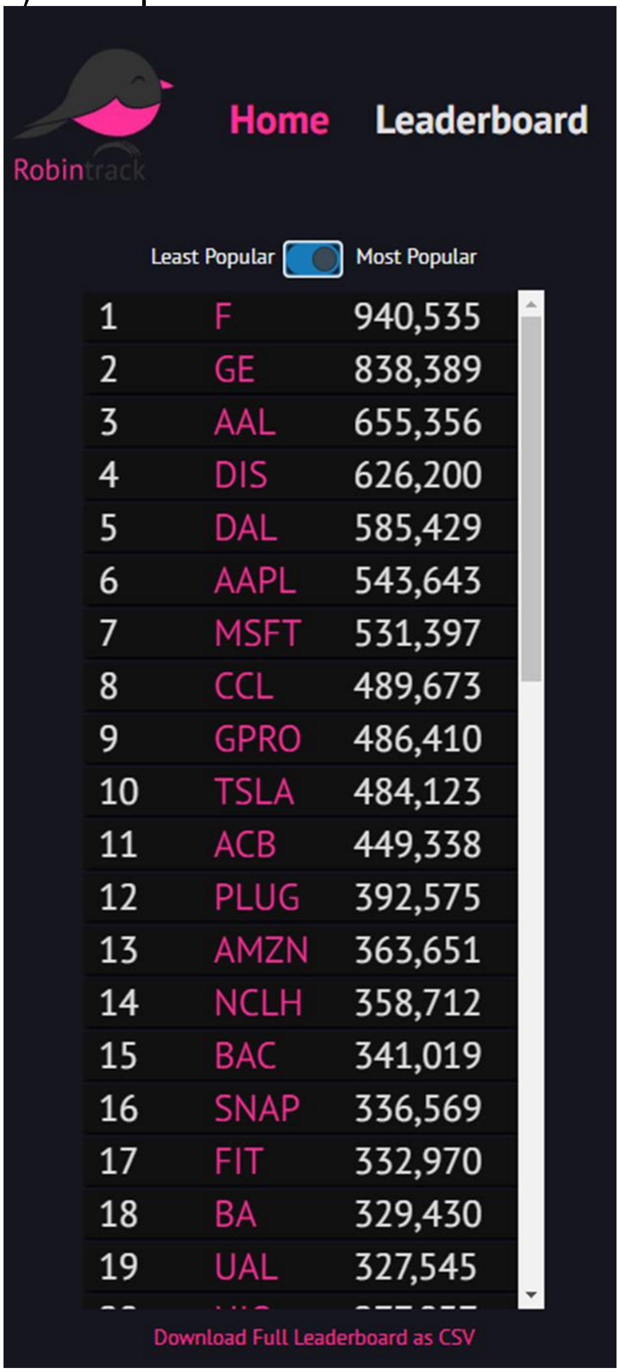
User/Company Shares	GOOG	FB	NFLX	MSFT
User A	10			
User B	15	25		
User C	1	5	8	
User D	20	5	6	14
Popularity	**4**	**3**	**2**	**1**

Sources: Wolfe Research Luo's QES

A full list of the most and least popular stocks can be found on Robintrack's leaderboard (see Figure 13A and Figure 13B for two examples). As of July 15, 2020, Ford, General Electric, American Airlines, and Walt Disney are among the most popular stocks on Robinhood.

Figure 13 Robintrack's Leaderboard

A) Most Popular



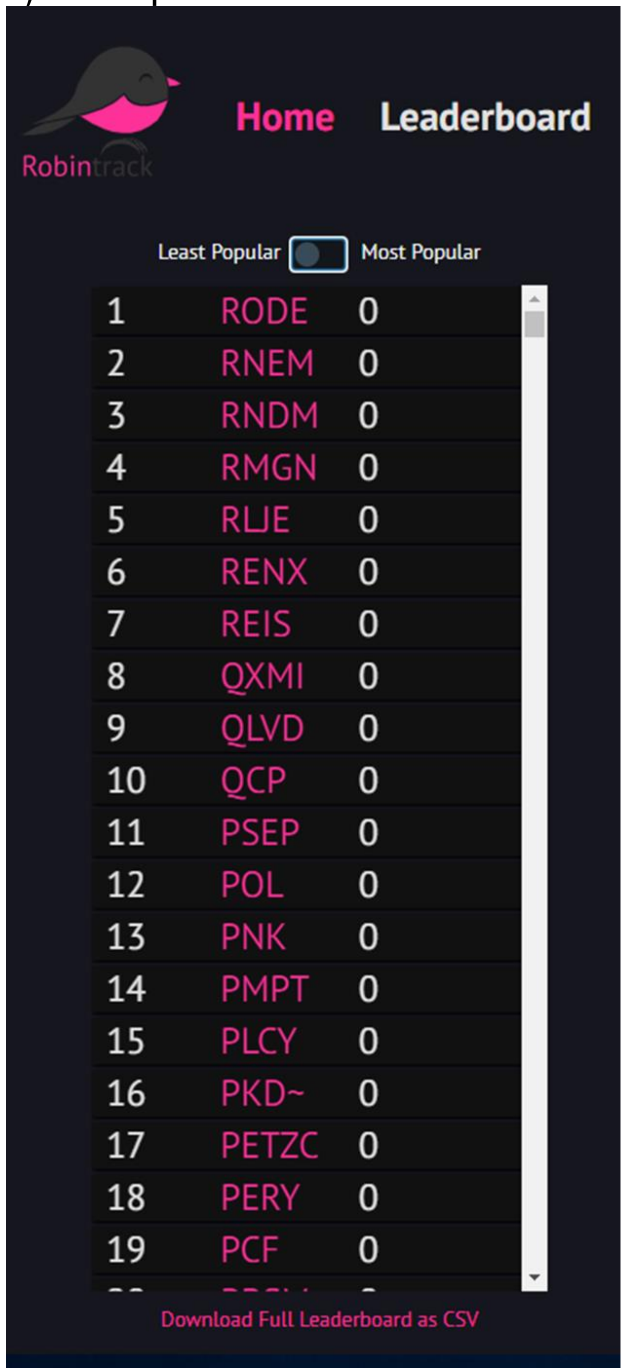
Robintrack | Home | Leaderboard

Least Popular | Most Popular

1	F	940,535
2	GE	838,389
3	AAL	655,356
4	DIS	626,200
5	DAL	585,429
6	AAPL	543,643
7	MSFT	531,397
8	CCL	489,673
9	GPRO	486,410
10	TSLA	484,123
11	ACB	449,338
12	PLUG	392,575
13	AMZN	363,651
14	NCLH	358,712
15	BAC	341,019
16	SNAP	336,569
17	FIT	332,970
18	BA	329,430
19	UAL	327,545

Download Full Leaderboard as CSV

B) Least Popular



Robintrack | Home | Leaderboard

Least Popular | Most Popular

1	RODE	0
2	RNEM	0
3	RNDM	0
4	RMGN	0
5	RLJE	0
6	RENX	0
7	REIS	0
8	QXMI	0
9	QLVD	0
10	QCP	0
11	PSEP	0
12	POL	0
13	PNK	0
14	PMPT	0
15	PLCY	0
16	PKD~	0
17	PETZC	0
18	PERY	0
19	PCF	0

Download Full Leaderboard as CSV

Sources: https://robintrack.net/

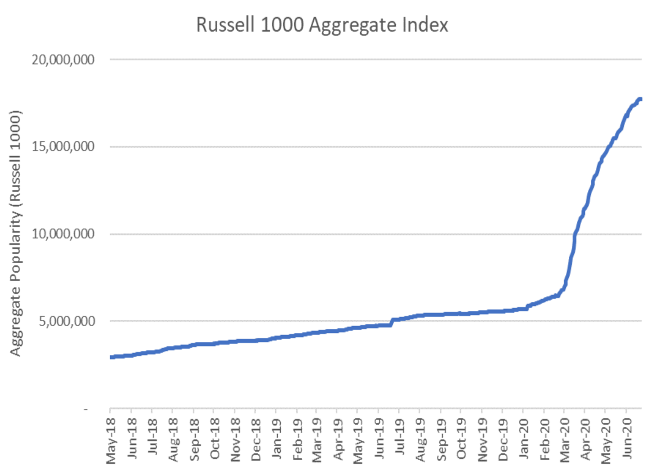
Unfortunately, in August 2020, soon after our publication *Retail Investor Behavior and Participation Alpha*, Robinhood quietly disabled the streaming of its popularity score. However, using statistical algorithms, after combining data from Reddit and another retail broker, we can still reveal significant insights from retail trading activities.

Robinhood's broad user base offers a unique perspective into the trading behavior of retail investors. To better understand the Robinhood data, we proxy the overall user base. We aggregate the number of users/investors of all stocks in the point-in-time Russell 1000 index, i.e., $AggregatePopularity_t = \sum_{i=1}^{1000} Popularity_{i,t}$.

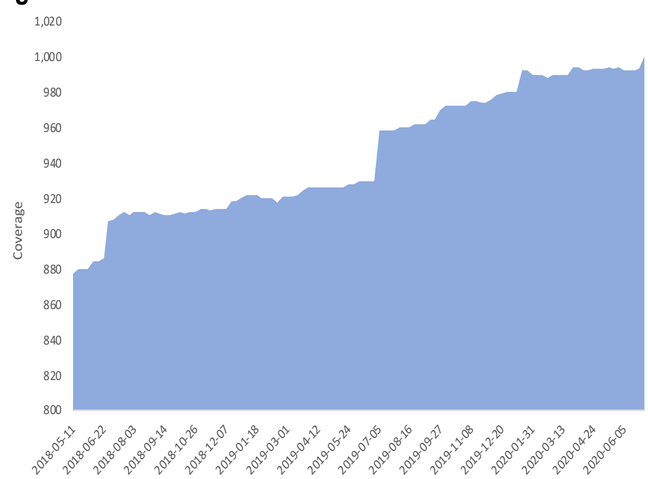
As shown in Figure 14(A), the total number of users/investors has grown since 2018, but there is a massive spike since the onset of the COVID-19 pandemic in February/March 2020. With the stay-home and quarantine policy across the country, it appears that many people take advantage of their free time to conduct online trading. Similarly, the breadth of coverage (i.e., number of stocks in the Russell 1000 index with Robinhood users/investors) has also been rising, with two jumps in mid-2018 and mid-2019, respectively (see Figure 14B).

Figure 14 Retail Investor Behavior Factor Coverage (Russell 1000)

A) Aggregate Popularity



B) Coverage



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Wolfe Research Luo's QES

ANOTHER MAJOR RETAIL BROKERAGE TRADING DATA

In addition to Robinhood, we supplement our study with retail trading data from another major retail broker, using the most actively bought and sold stocks and ETFs facilitated through the company as a proxy for retail interest. It covers some of the most heavily trading stocks and is updated on a daily basis.

Quantifying Social Media Fueled Retail Trading Frenzy

Recall that Reddit is a discussion forum, while Robinhood is a trading venue. Redditors who share speculative trading ideas on the Reddit platform, but they cannot conduct their trade directly. Retail investors can then move on to actual trading platforms such as Robinhood and many other retail brokers.

Reddit versus Robinhood

In this section, we first compare the stock-selection signals derived from Reddit and Robinhood, in terms of their predictive power and unique investing styles. As a reminder, Robinhood data is only available until July 2020. In later sections, we will discuss how to proxy retail participation using Reddit, another major retail broker data, and statistical algorithms.

Let's start from two portfolios:

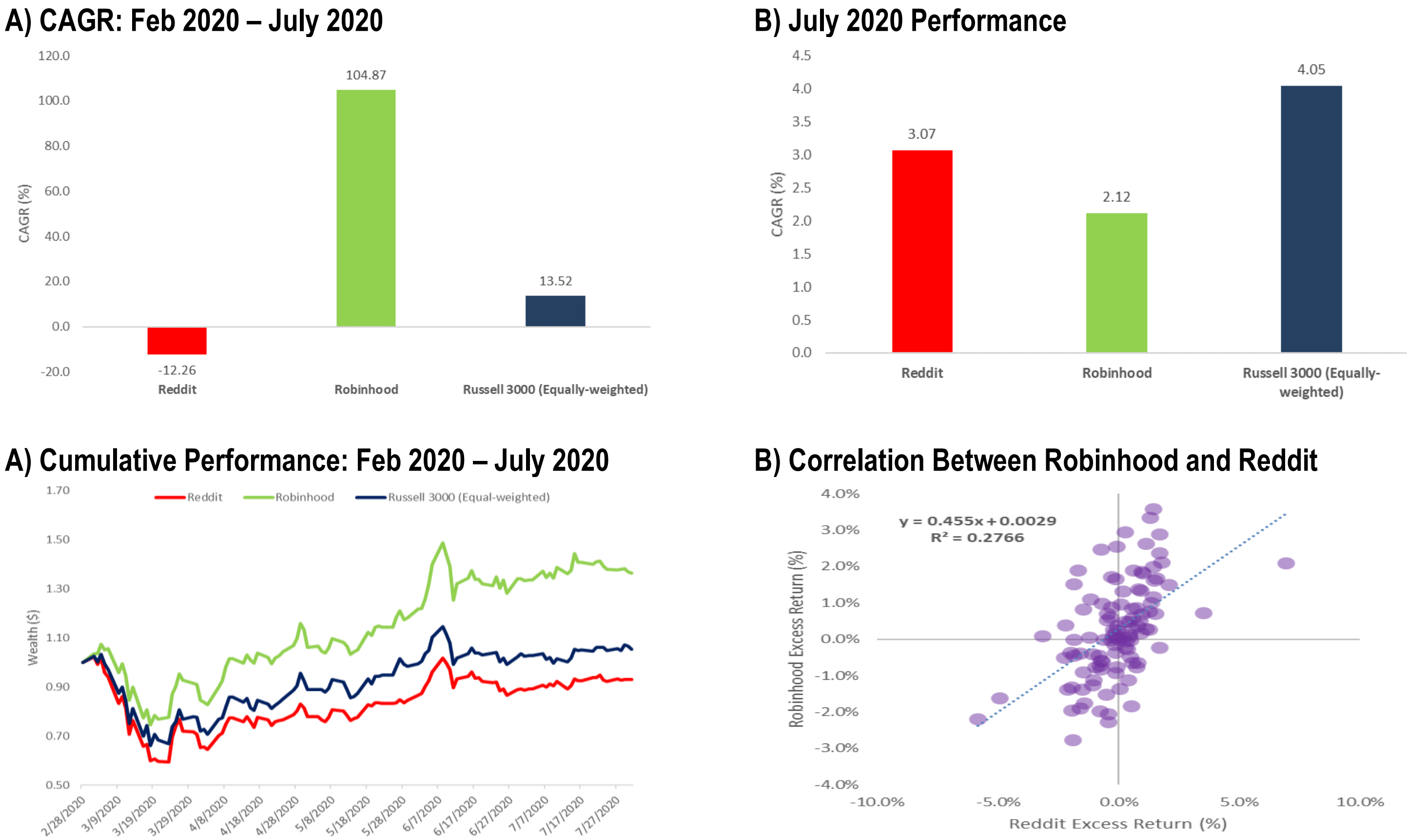
- **Reddit**: Select the top 30 stocks based on daily mentions within r/wallstreetbets community[8] in the previous week. Stocks are equally weighted in the portfolio.
- **Robinhood**: Select the top 30 stocks based on the Robinhood popularity metric. The top 30 stocks are selected based on the average daily popularity during the previous week and then equally weighted in the portfolio.

The top 30 stocks in both portfolios are equally weighted. Both portfolios are long-only and rebalanced every Friday. Our backtesting is conducted using data from February 2020 to July 2020, when both data sources were available to the public.

During the peak of the COVID crisis (i.e., February to July 2020), the Robinhood portfolio significantly outperformed both the Reddit basket and the broad equity market (see Figure 15A). However, as the Reddit user base began to explode, the Reddit-based investment strategy quickly outpaced the Robinhood portfolio in July (see Figure 15B). As expected, the two retail-centric proxies are somewhat correlated with an average correlation nearly 60% (see Figure 15D). Therefore, these two data sources offer consistent but also complementary views of the market.

[8] We further filter by the *Daily Discussion* flair and the thread *What Are Your Moves Tomorrow.*

Figure 15 Reddit Versus Robinhood Performance (Feb 2020 – July 2020)

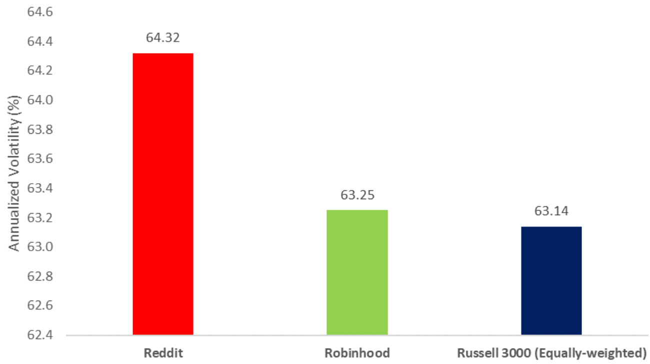


Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Reddit, Wolfe Research Luo's QES

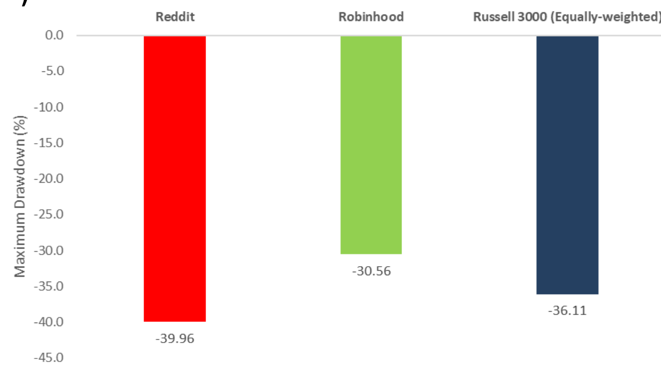
As shown in Figure 16(A), the Reddit portfolio appears to be even more speculative than the Robinhood strategy, with a higher volatility, larger drawdown risk (see Figure 16B), greater turnover (see Figure 16D), and weaker risk adjusted performance (see Figure 16C). The results are as expected, since the Reddit portfolio is based on social media mentions which can reflect more impulsive behavior by retail investors. On the other hand, the Robinhood dataset represents a more consolidated view of how the collective group of retail investors actually invested.

Figure 16 Reddit Versus Robinhood Performance (Feb 2020 – July 2020), Cont'd

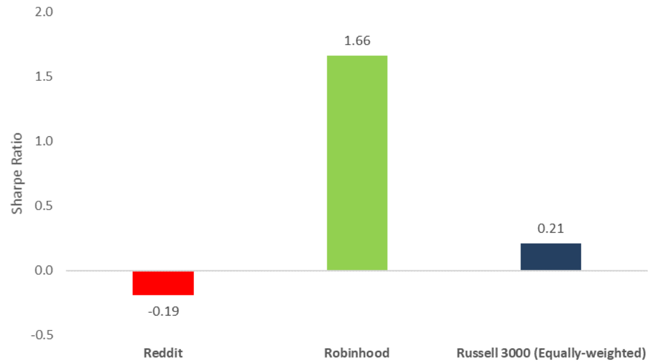
A) Annualized Volatility



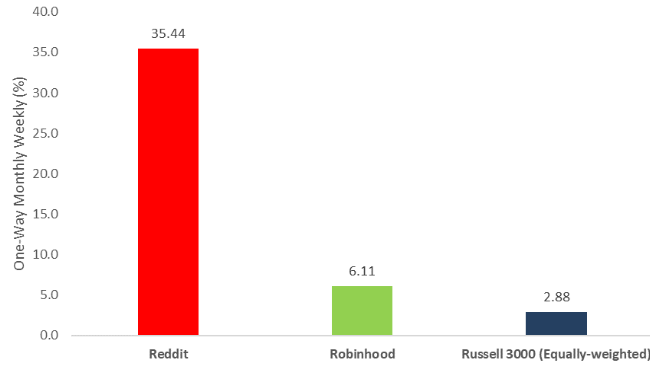
B) Maximum Drawdown



C) Sharpe Ratio



D) One-Way Weekly Turnover



Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Reddit, Wolfe Research Luo's QES

INVESTMENT STYLES COMPARISON

Next, we analyze and compare the unique investing styles and behavior biases of Redditors versus Robinhood users, by leveraging our Port@ Custom Attribution System (see *Port@able Ownership* for more details). Compared to standard off-the-shelf commercial risk models, the Port@ platform provides a more robust, custom view of the sources driving portfolio risk and return. Managers can choose any factors from our 300+ common factor library, 100+ macro sensitivity variables, and 1,000+ alternative data-based signals to analyze their portfolios.

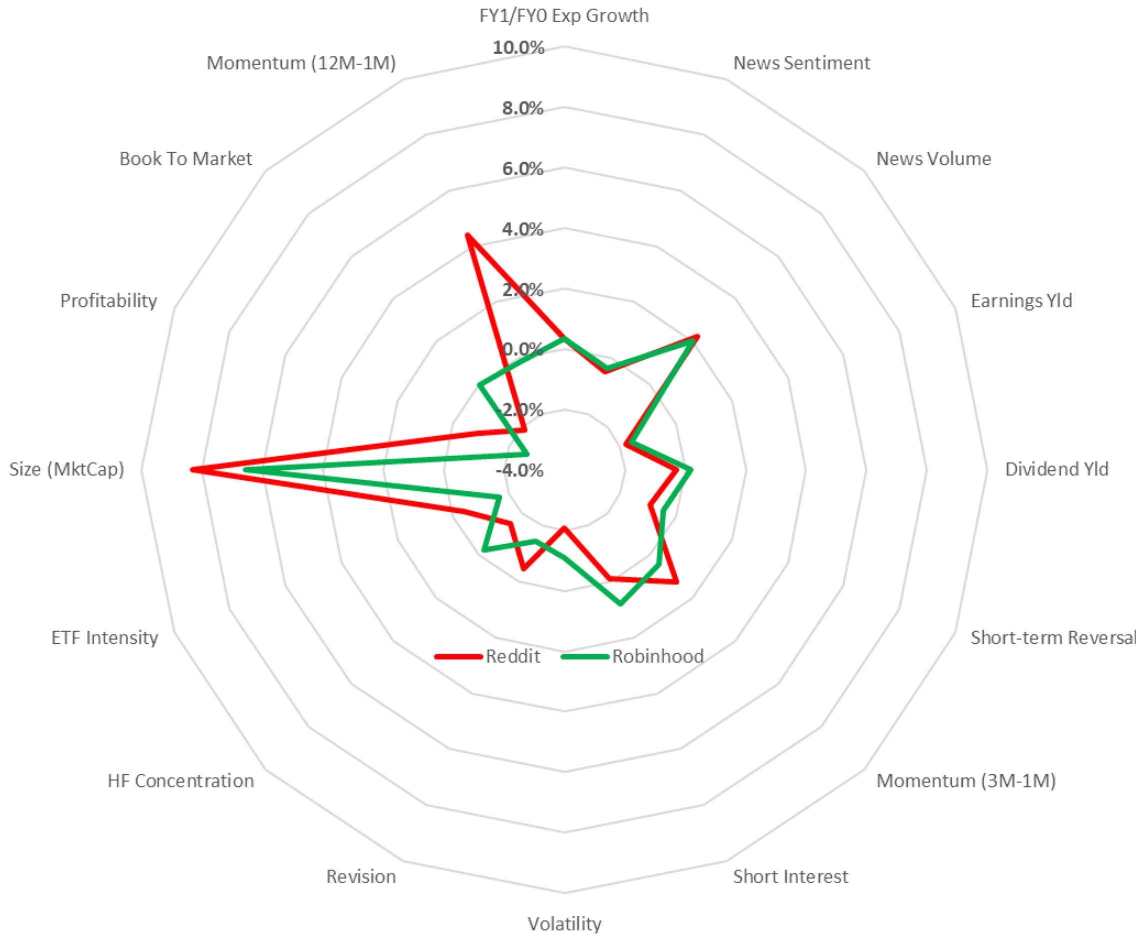
Factor Exposures

We incorporate a host of non-traditional factors in the attribution analysis such as short interest, hedge fund crowding, news volume/sentiment to reveal the unique characteristics behind these two retail participation strategies. Our analysis is conducted using the long-only, Reddit and Robinhood portfolios described above. The benchmark is the Russell 3000 equally weighted index. All analysis is relative to the benchmark.

As shown in Figure 17, retail investors (both Robinhood and Reddit) tend to favor large cap, high growth names that are frequently mentioned in the media (as proxied by news volume). Retail portfolios tend to be negatively correlated to profitability which reflects their appetite for risk taking.

Interestingly, the Reddit portfolio tends to more negatively exposed to hedge fund concentration than the Robinhood portfolio, reflecting Reddit's more rebellious nature. The Reddit portfolio is also more negatively exposed to value than Robinhood. Lastly, unlike Robinhood investors who typically short momentum, the Reddit portfolio takes on a positive momentum bet, indicating a trend following behavior.

Figure 17 Volatility Adjusted Factor Exposure

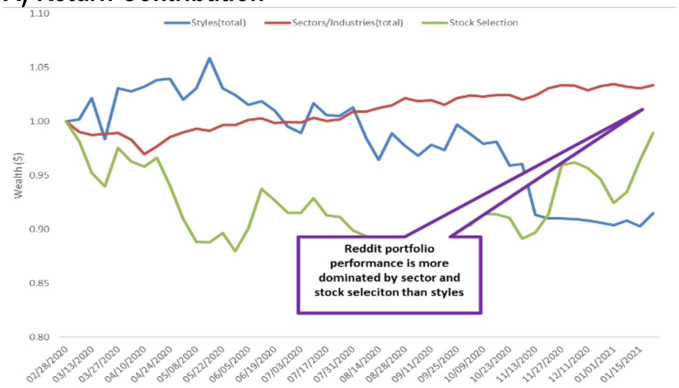


Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Reddit, Wolfe Research Luo's QES

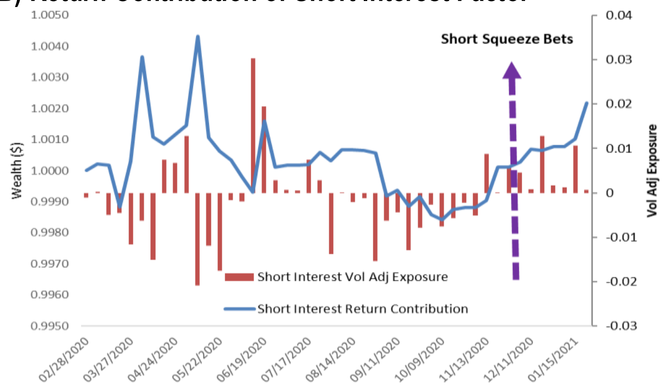
Furthermore, Figure 18(A) decomposes the Reddit portfolio return by style, sector, and stock specific drivers. Interestingly, the performance of the Reddit portfolio is increasingly being driven by sector and stock specific attributes and less so by style factors. This reflects the very idiosyncratic nature of the Reddit community. Given the Reddit's bet against heavily shorted stocks (mostly by hedge funds) in recent days, in Figure 18(B), we show the return contribution by the short interest factor for the Reddit portfolio. We observe that the short interest exposure has increased recently, possibly due to the Reddit trading community's forced short squeezes. The recent heightened exposure to short interest has generated positive returns.

Figure 18 Attribution Analysis

A) Return Contribution



B) Return Contribution of Short Interest Factor

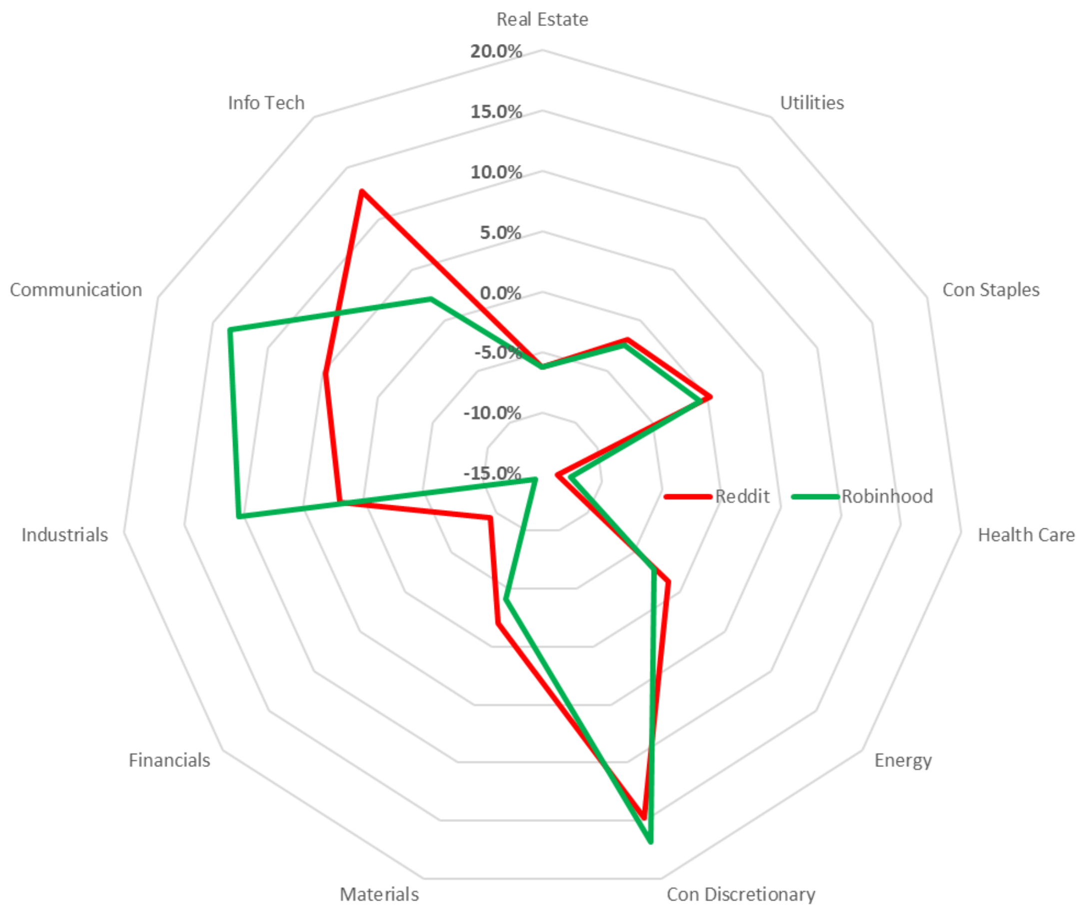


Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Reddit, Wolfe Research Luo's QES

Sector Exposures

Figure 19 reveals the sector decomposition of the Reddit and Robinhood portfolios relative to the benchmark. As expected, retail investors shy away from more defensive and complex sectors such as the REITs, utilities, health care and financials. Conversely, they tend to invest heavily into more famed industries such as the information technology, communication services, and consumer discretionary.

Figure 19 Sector Exposure



Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Reddit, Wolfe Research Luo's QES

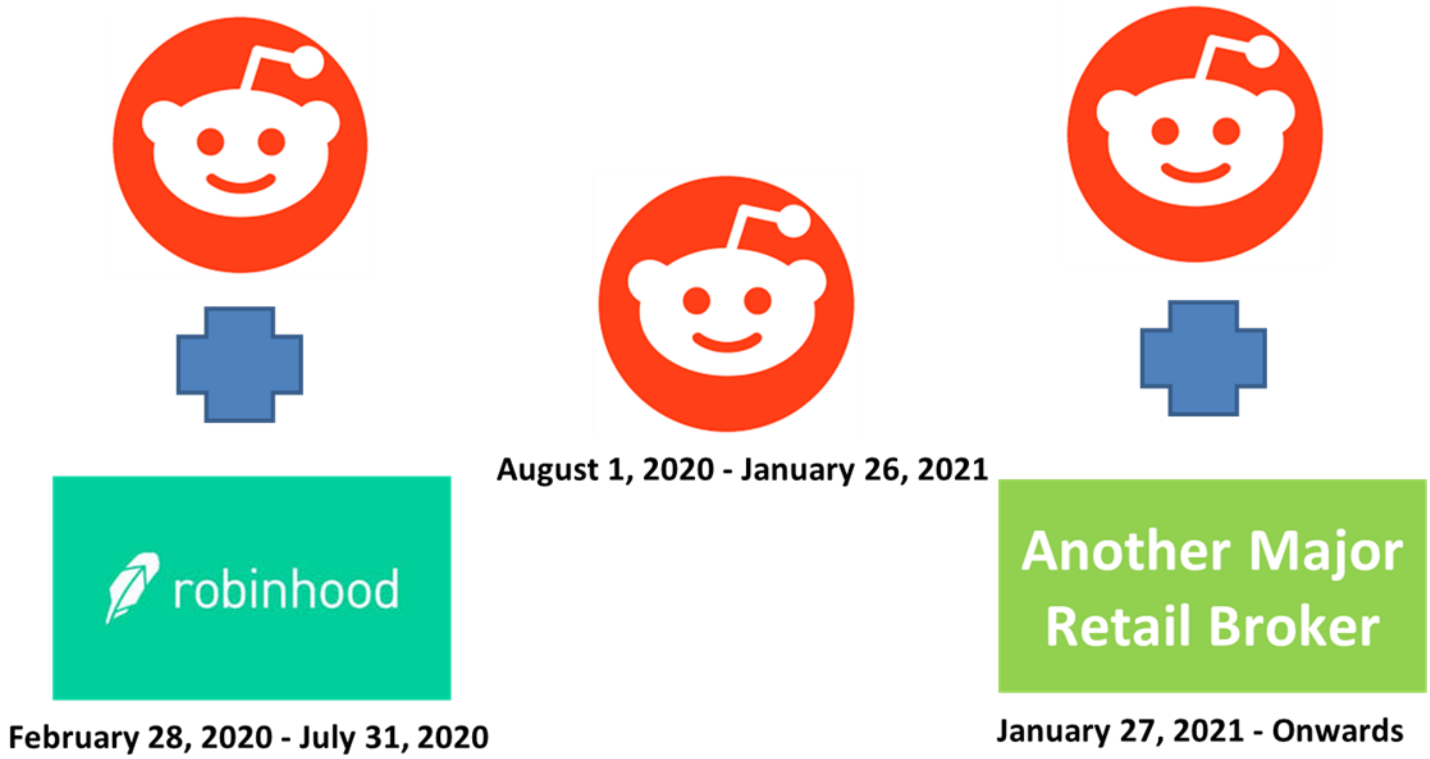
ROBINHOOD + REDDIT + TRADING DATA FROM ANOTHER RETAIL BROKER

Lastly, we form a composite portfolio by combing the top stock picks from Robinhood, Reddit, and another major retail broker. We need to be cognizant that each dataset has different start and end dates, as shown in Figure 20.

In early 2020, both Reddit and Robinhood were available[9], which lead our composite portfolio during this time. The two signals are equally weighted and the top 30 stocks are included. Starting from August 2020, Robinhood discontinued its data access. Therefore, post-August 2020, we rely solely on data from Reddit. More recently, we began accessing another retail broker's most actively traded names; therefore, our composite is now based on both Reddit and another retail broker trading data.

[9] The Robinhood dataset was in fact available in May 2018; but we have not included that in our analysis. However, we will form a more comprehensive retail participation dataset that commences from May 2018 onwards. Please reach out to us for more details.

Figure 20 Composite Retail Signal

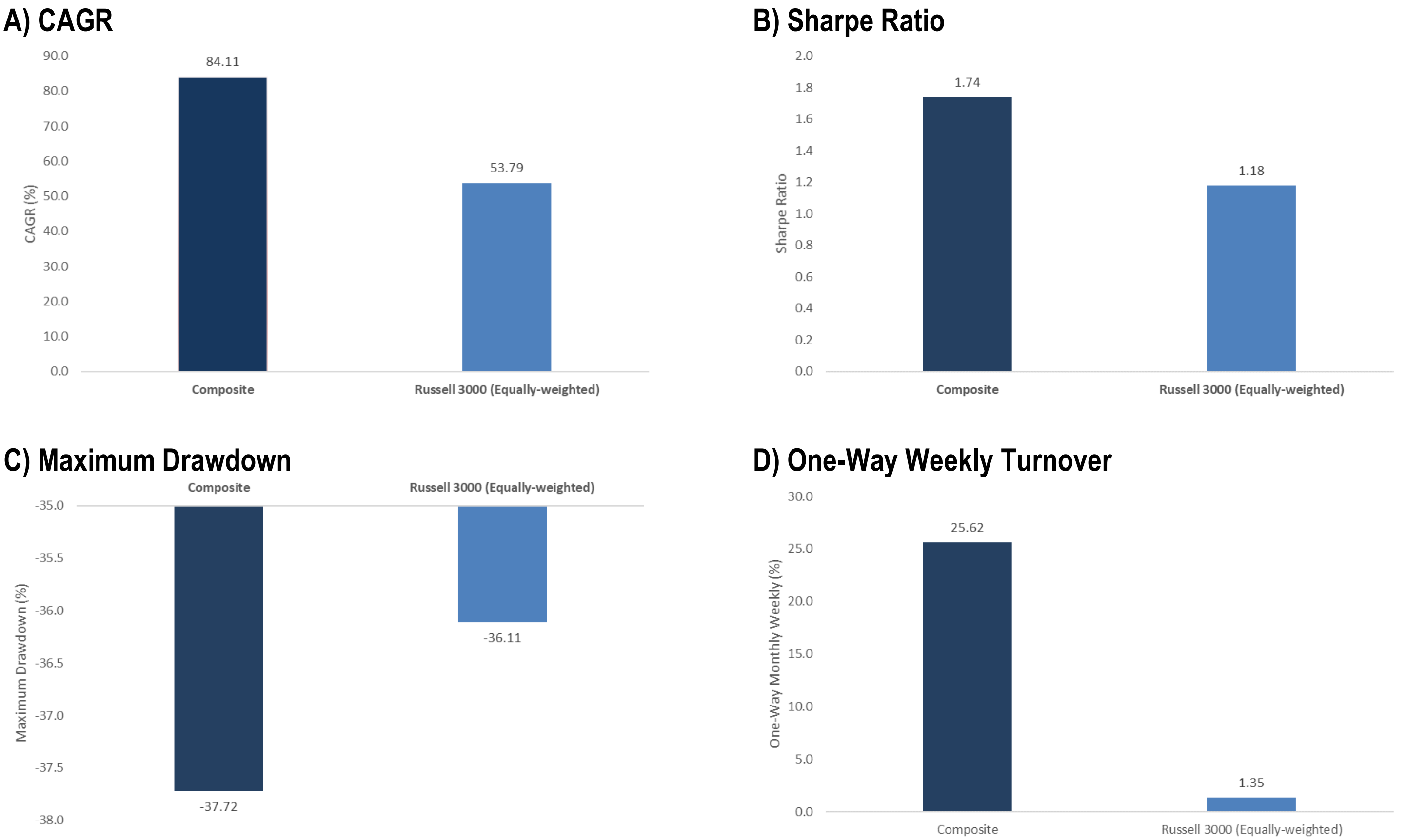


Sources: Reddit, Robinhood, Wolfe Research Luo's QES

To show the performance of retail participation strategies, we backtest our composite signal from February 2020 onwards. Our retail portfolio is long-only and comprises 30 stocks (equally weighted). We rebalance the portfolio on a weekly basis, every Friday.

As shown in Figure 21, despite a highly concentrated portfolio, retail investors demonstrate impressive stock-picking skills. The representative retail portfolio outpaces the market benchmark on an absolute (see Figure 21A) and risk adjusted basis (see Figure 21B). However, it does suffer from a higher downside risk (see Figure 21C) and turnover (see Figure 21D) as compared to the benchmark.

Figure 21 Composite Performance (Feb 2020 – Present)



Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Reddit, Wolfe Research Luo's QES

Lastly, as shown in Figure 22, during the COVID pandemic period, retail investors have demonstrated impressive stock picking skills.

Figure 22 Cumulative Performance

Sources: Reddit, Robinhood, Wolfe Research Luo's QES

Short Squeeze and Retail Popularity

In this section, we conduct an in-depth study on the interaction of retail participation and short squeeze. First, we study the statistical properties of short squeezes. Then, we use our global factor library as tools to identify potential short squeeze candidates. Lastly, data suggests that avoiding going against retail investors, especially on the short side, can boost performance meaningfully. The performance improvement is particularly strong in January this year, as expected.

What is a Short Squeeze?

The SEC defines a "short squeeze" based on two criteria:

- The pressure on short sellers to cover their positions, because of a sharp increase in security prices; and
- A lack of borrow availability

The rush by short sellers to cover their positions can put significant upward pressure on share price, which further exacerbates short squeezes. A short squeeze can also result from short sellers' inability to borrow securities, due to lack of available inventory or excessive borrow fees.

As detailed in *More Bang on the Short*, short squeezes are rare events. However, the damages to short sellers can be significant, as stock prices can go up significantly in short period of time. Unlike long positions, where losses are limited to -100%, there is no theoretical limit on damages from the short side. To study short squeeze, we use the short interest utilization factor, i.e., # of shorted shares over total lendable inventory (see *New Insights in Short Interest – A Global Perspective* for more details) as a proxy for difficult in borrow securities.

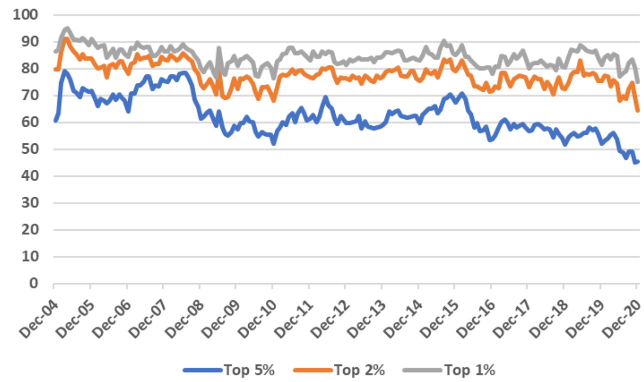
As shown in Figure 23(A), the top 5% most heavily shorted stocks (based on the utilization factor) have an average short interest utilization ratio of 50%, while the proportion jumps to over 80% for the top one percentile. Therefore, the most intensively shorted 1% companies utilize over 80% of available inventories, leaving little room for additional short borrowing. Even worse, as shown in Figure 23(B), the skewness of these stocks – based on their forward one-month returns is considerably higher than the broad market. A high and positive skewness measure suggests that surprises in stock returns are more likely to be on the positive side – more troublesome for more densely shorted stocks.

Similarly, Figure 23(C) illustrates that on average stocks in the broad market (Russell 3000 index) roughly follow a normal distribution, with a median return slightly positive (because the stock market tends to go up in the long run) and some limited outliers. On the other hand, as shown in Figure 23(D), stocks with high short interest suffer from a negative median return, but with large number of positive outliers. On average, in the long term, heavily shorted stocks tend to underperform the market. However, we see a much heavy tail on the positive side, meaning a few stocks are likely to generate large positive returns in the subsequent month – even over 100%.

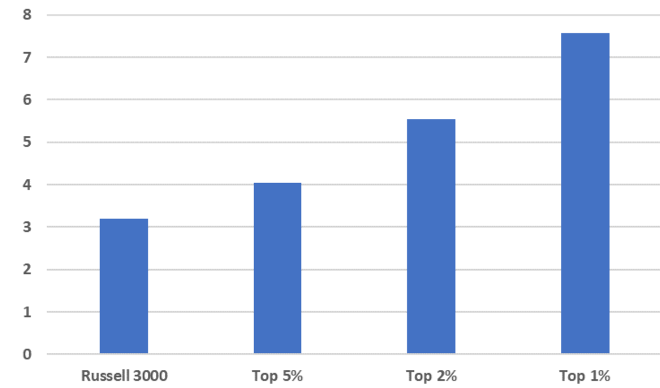
For the rest of this section, we use the top 5% most heavily shorted stock universe as our research target, because most short squeezes occur in this universe.

Figure 23 Statistical Properties of Heavily Shorted Stocks

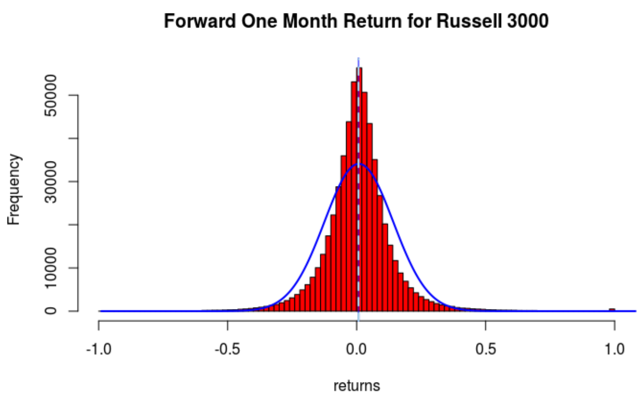
A) Short Interest Utilization Level in Russell 3000



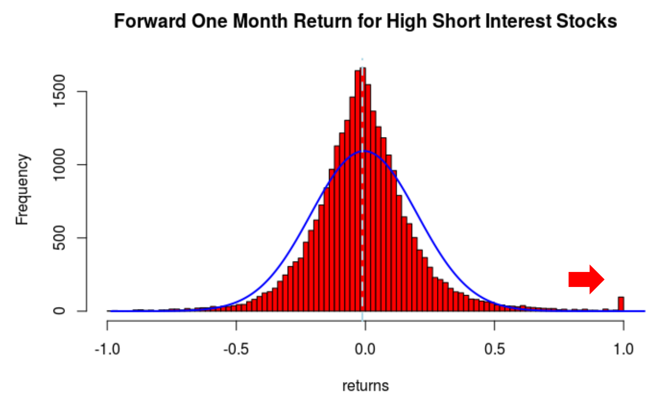
B) Skewness of the Next Month Return



C) Return Distribution, all Russell 3000 Stocks



D) Return Distribution, Top 5% Most Shorted Stocks

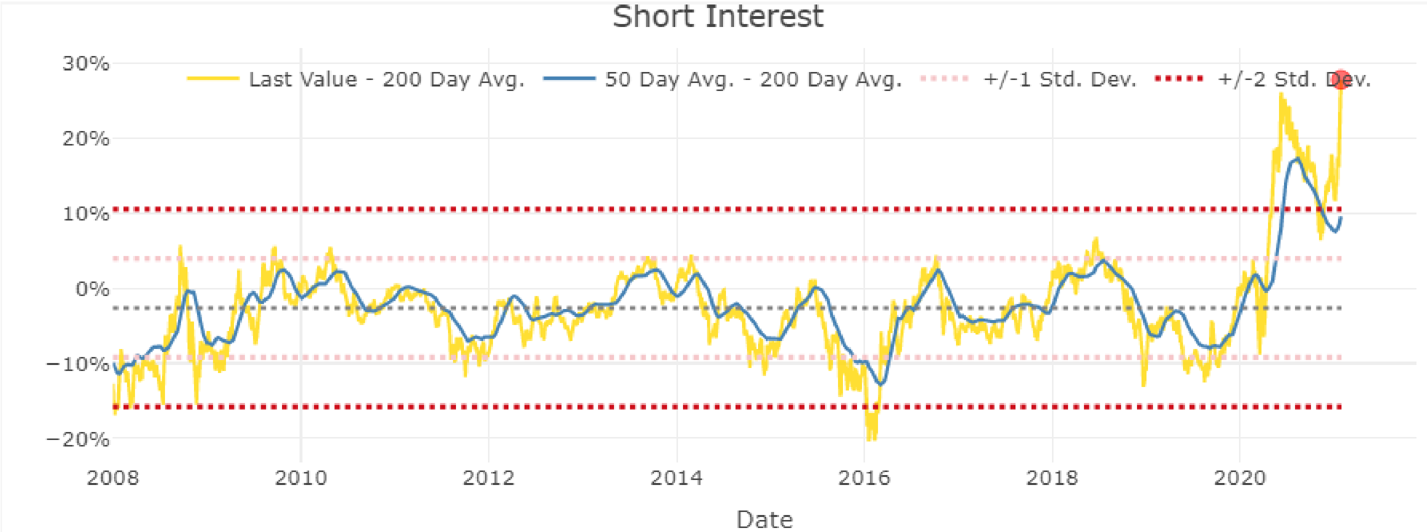


Sources: Bloomberg Finance LLP, MSCI, FTSE Russell, Markit, RavenPack, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

IDENTIFYING SHORT SQUEEZES

As shown in Figure 24, currently, the short interest factor is overstretched at a historical record, above three standard deviation (see our Factor Awareness Dashboard), suggesting that heavily shorted stocks have outperformed their lightly shorted peers considerably.

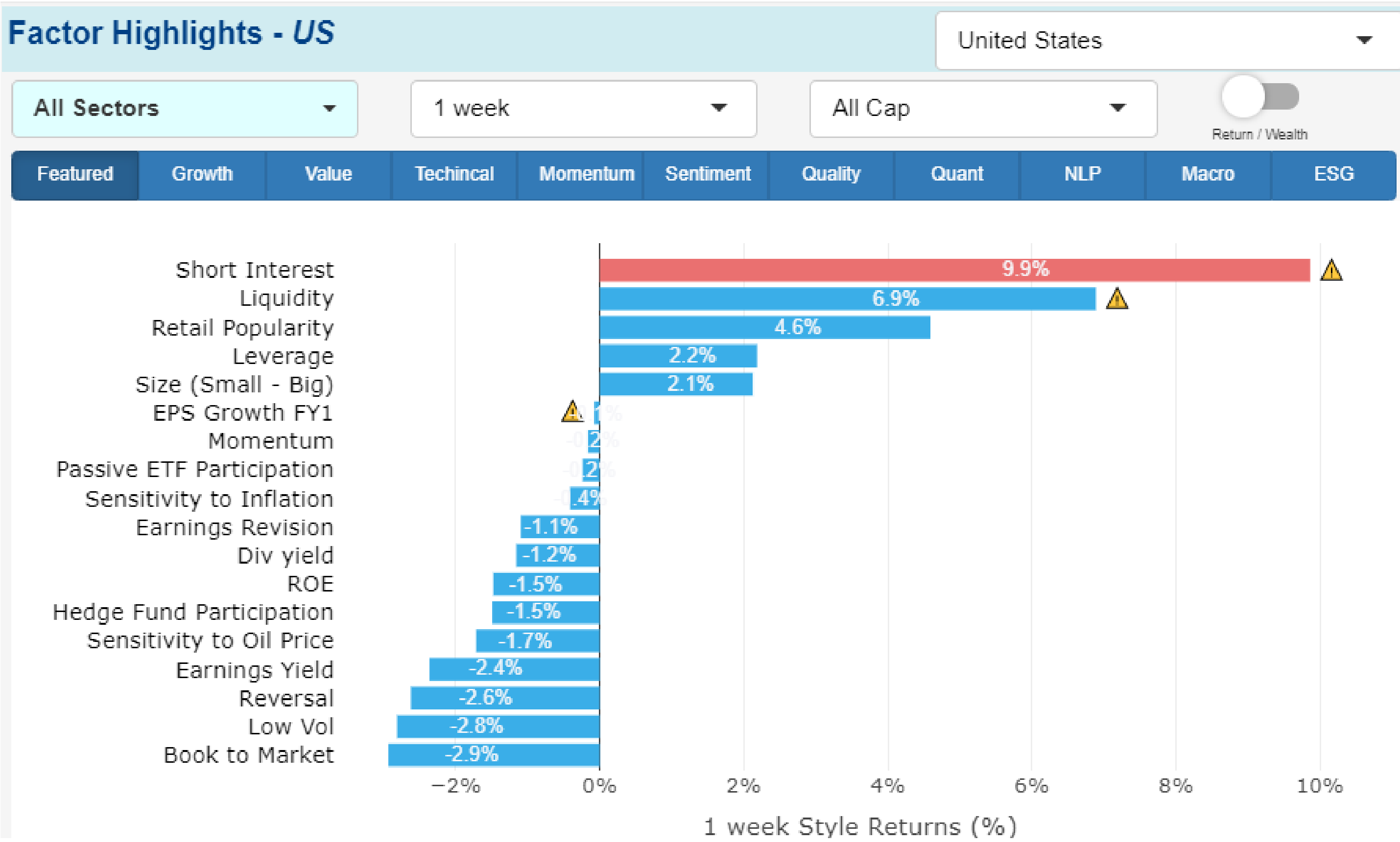
Figure 24 Short Interest Distortion



Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Wolfe Research Luo's QES

The short interest factor generated the largest return in the past week among ~60 common factors – a 9.9% return based on the signal weighted portfolio in the US (see Figure 25), due to the significant short squeezes and the resulting hedge fund de-leveraging activities as discussed in the previous sections.

Figure 25 Factor Returns from QES Factor Awareness Dashboard, as of January 28, 2021



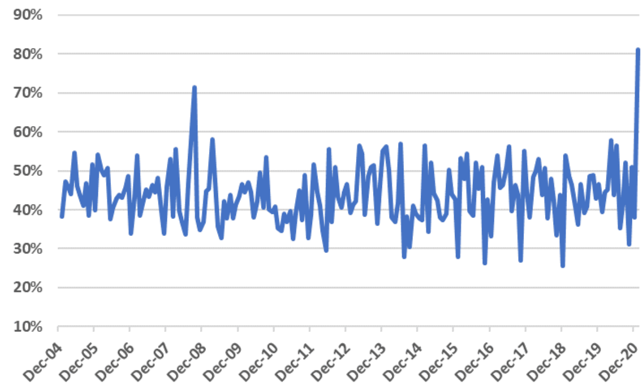
Sources: Bloomberg Finance LLP, FTSE Russell, Haver, S&P Capital IQ, Thomson Reuters, Robinhood, Robintrack, Wolfe Research Luo's QES

Figure 26(A) shows the percentage of heavily shorted stocks (in the top 5%) that have outperformed the market in the past 16 years. The ratio typically hovers between 40% and 50%, but surged to 80% in the past few days, surpassing the previous peak by more than 10%.

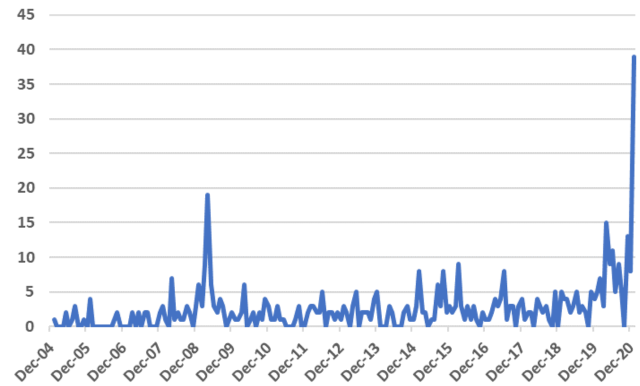
To understand short squeezes, we further limit our universe to those stocks with excess returns (i.e., returns above the market) more than 50% in the subsequent month. As shown in Figure 26(B), in a typical month, fewer than five stocks fall into this category. However, currently, nearly 40 intensively shorted stocks outperformed the market by 50%. We use this universe as our proxy for short squeezes. As expected, there is no short squeeze more than a quarter of the time. In January 2021, among the 150 heavily shorted names, almost 40 stocks outperformed market by over 50% and therefore, are defined as short squeezes – twice as many as the previous record during the 2009 market rally post GFC.

Figure 26 High Short Interest Stocks (Top 5%) in the Russell 3000 Universe

A) % of Stocks Outperformed the Market



B) # of Stocks with >50% Excess Return



Sources: Bloomberg Finance LLP, MSCI, FTSE Russell, Markit, RavenPack, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

To understand the characteristics of stocks that had suffered from short squeezes, we compute the difference in median factor percentile, between our short squeeze universe and the rest of the high short interest universe. To avoid potential sector bias, factor scores are sector-adjusted.

Figure 27(A) shows the median factor percentile for some of the most differentiating factors, prior to short squeezes. In Figure 27(B), we repeat our analysis, but only during those months when the market was down. The factor percentile spread is even wider during bear market months. For example, the abnormal volume factor is reasonably predictive of future short squeezes, and its forecasting power increases substantially during bear market. The intuition to separate bear markets is that, during those months when the market had positive returns, the significant outperformance of heavily shorted stocks might simple due to risk rally rather than short squeeze.

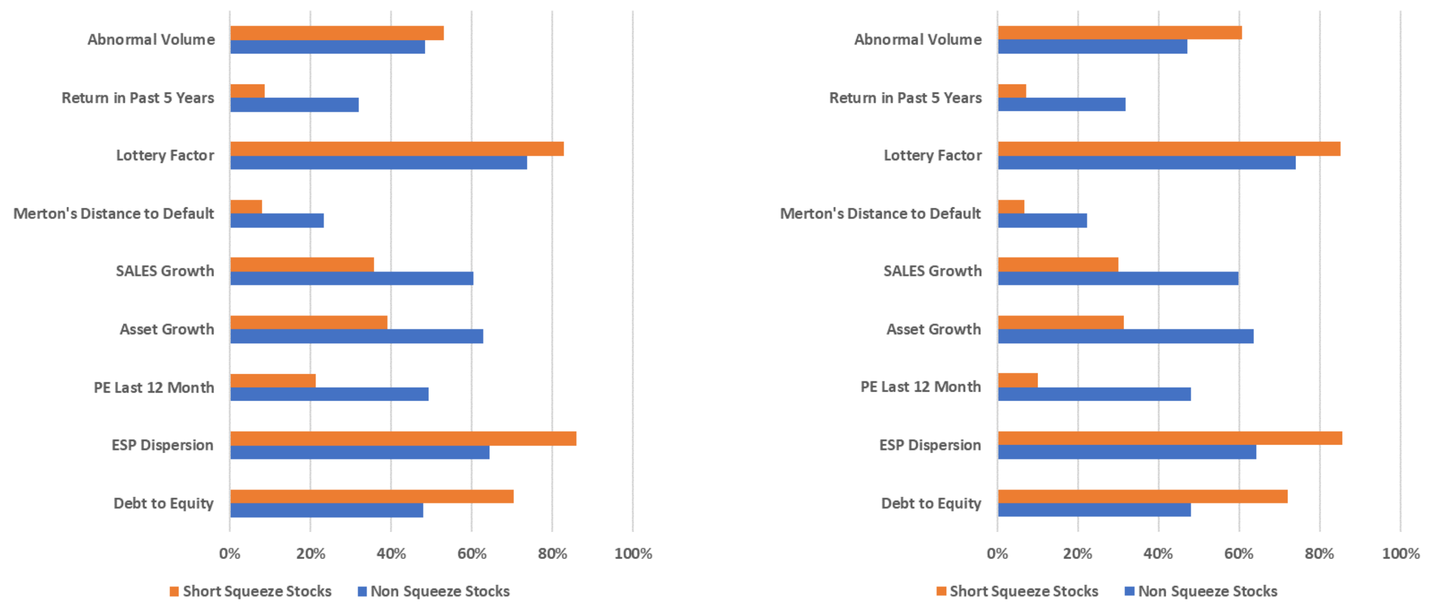
Most of the factors in Figure 27 are quite intuitive. In summary, short squeezes stocks tend to have:

- Poor long-term momentum (defined as past five-year return)
- Low quality, i.e., high default probability and high financial leverage
- Negative growth profiles (as measured by both sales and asset)
- High uncertainty (i.e., high EPS dispersion)
- Cheap valuation (based on PE ratio)
- High abnormal volume and lottery-type of return profile (defined as the highest daily return in the past month) factor[10]

[10] Interestingly, although the short-term reversal factor is conceptually related to the lottery signal, it is not useful in differentiating short squeezes.

Figure 27 Median Factor Percentile, Short Squeeze Stocks versus Other Heavily Shorted Stocks

A) All Months   B) Bear Market Months



Sources: Bloomberg Finance LLP, MSCI, FTSE Russell, Markit, RavenPack, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

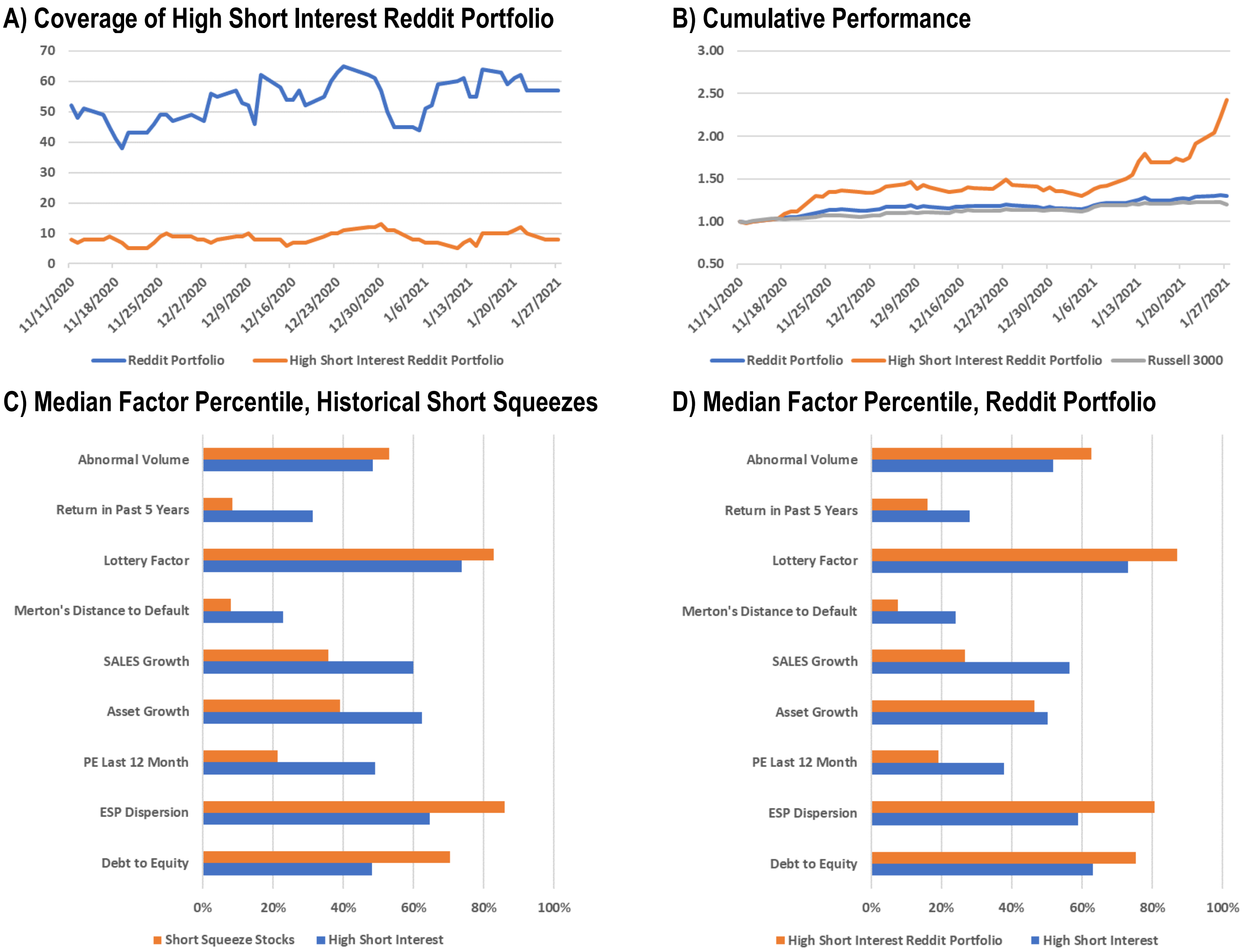
When Retail Popularity Meets Short Squeezes

Given the recent short squeezes are largely due to retail investor participation (as detailed in the previous sections), we want to understand how retail frenzy and short squeeze interacts. Leveraging our Reddit retail buzz signal, we can study those heavily shorted stocks (i.e., top 5% based on short interest) that are also in the Reddit buzz portfolio.

As shown in Figure 28(A), among the top 50 to 60 retail investors' favorite stocks, around 10 stocks also have high short interest. As explained in *Retail Investor Behavior and Participation Alpha*, given the large difference in retail behavior from institutional investors, it is not surprising to see that individual investors can take on significant bet against sophisticated hedge fund managers. As shown in Figure 28(B), on average, retail investors show only modest stock picking still. However, their ability to push up the prices of heavily shorted stocks has been impressive, especially in January. The equally weighted high short interest Reddit portfolio delivered a 150% return in the past two months, with two-thirds of that outperformance generated in January.

The high short interest Reddit portfolio shows a similar factor pattern as the broad short squeeze universe (see Figure 28C and D), which confirms that the recent rally in heavily shorted stocks is mainly driven by retail frenzy.

Figure 28 Short Squeeze and Retail Popularity (Reddit Buzz)



Sources: Bloomberg Finance LLP, MSCI, FTSE Russell, Markit, RavenPack, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

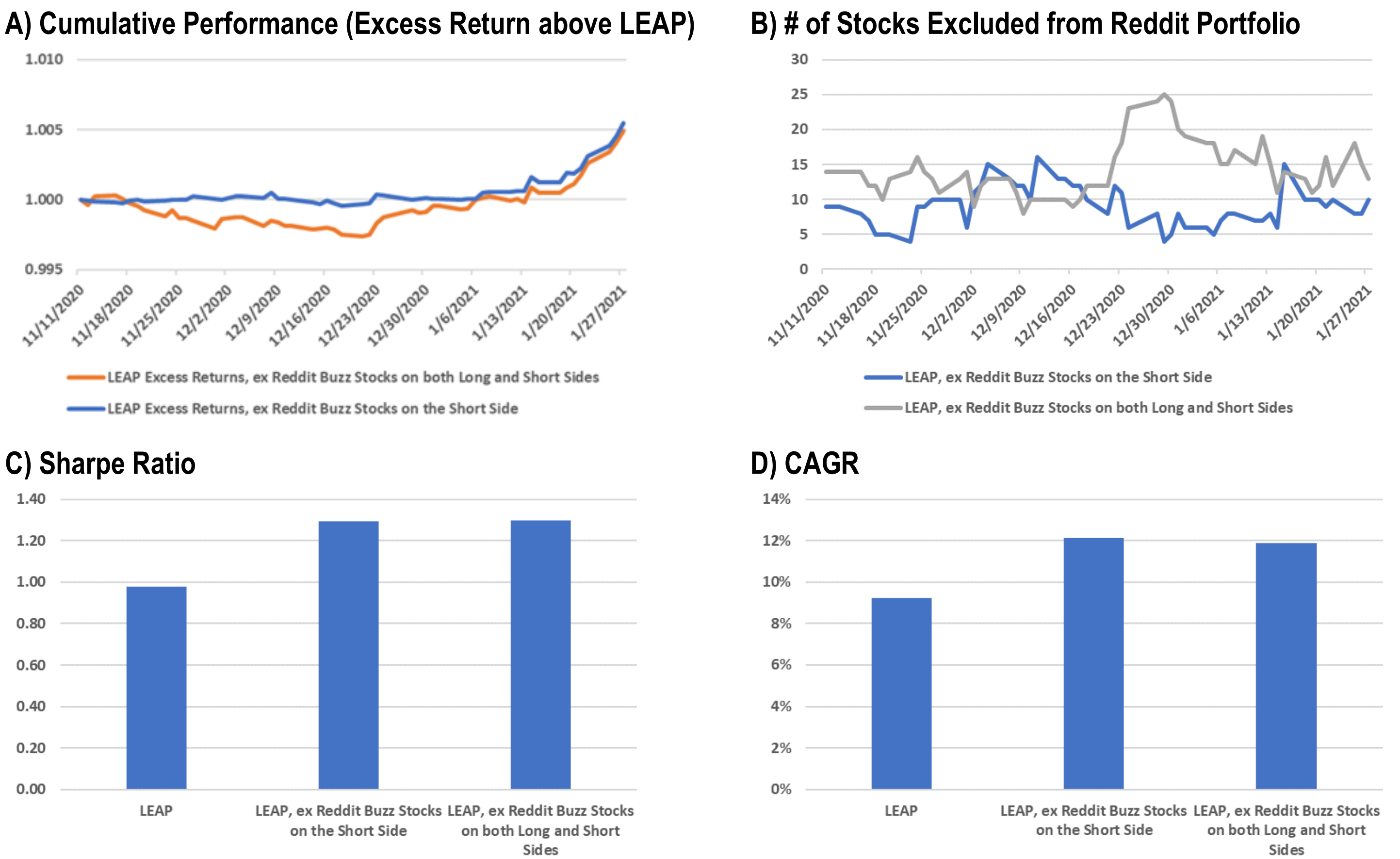
Don't Fight with the Crowd!

In investing, we have all learned the lesson, i.e., "don't go against the flow". Given the collective power by a large number of retail investors, it is probably wise not to short their favorite stocks. Using one of our flagship global stock-selection model, LEAD (see *Style Rotation, Machine Learning, and the Quantum LEAP*) as an example, we will show the impact of excluding Reddit buzz names from the short side. As illustrated in Figure 29(A), excluding Reddit buzz stocks from the short quintile portfolio improves performance meaningfully, especially given that we have only eliminated fewer than 15 stocks (the short quintile includes around 600 stocks). On average, there are around 10 and 15 Reddit buzz stocks on the short and long sides of the LEAP portfolio, respectively (see Figure 29B). Not betting against retail flows boosts Sharpe ratio and CAGR by 32% and 28%, respectively (see Figure 29C and

D). Given the strong performance of the Reddit buzz signal, not participating on the long side hurts our performance slightly, especially before January 2021.

Given the highly speculative nature of the Reddit buzz signal and retail investor behavior in general, we recommend using retail participation as a risk factor instead of an alpha signal. We will discuss the risk implications in more details in the next section.

Figure 29 Performance Improvement by excluding Reddit Buzz Stocks



Sources: Bloomberg Finance LLP, MSCI, FTSE Russell, Markit, RavenPack, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

RETAIL FORCED HEDGE FUND DE-LEVERAGING

Hedge fund de-leveraging events typically are preceded by some exogenous event. This week, the retail buying frenzy caused a massive surge in heavily shorted stocks, which forced some sophisticated managers to cover these positions *en masse,* with the one-way demand from retail exacerbating the market impact. Despite the wide media speculation that hedge funds subsequently de-leveraged also on the long side, limited downside in our hedge fund crowding factor suggests this has not yet been as widespread and damaging as the de-risking that occurred in March 2020. We do see evidence that crowded mega/large-cap stocks with significant total hedge fund allocations and popularity across many hedge funds have underperformed the market in recent days. But given the highly liquid nature of the crowded longs that were likely sold, the market impact of exiting these positions was less severe. Should the retail investor hysteria further accelerate, it could prompt further de-risking with more significant impact on crowded longs.

De-Leveraging events typically follow a predictable pattern

Hedge fund de-leveraging events typically follow a similar pattern. After some exogenous event causes PnL volatility across equity long/short portfolios, some hedge funds start to hit their VaR limits and prime brokers increase the cost of leverage offered to these accounts. Then portfolio managers are forced to exit some of their positions to bring down total risk and leverage. When this happens to many funds across the industry at the same time, we typically see a de-leveraging event characterized by a spike in the short Interest factor[11] and drawdown in the HF crowding signal[12] that happens very quickly over a few days or maybe weeks. In our Wolfe QES US Equity risk model suite, we include the HF crowding and short interest factors to measure the risk and return associated with the equity long/short positioning, which is typically elevated during these de-leveraging episodes.

Most equity L/S managers are exposed to Short Interest (-) and HF Crowding (+)

Most equity L/S managers, almost by definition, run positive exposure to the HF crowding factor from their long positions and negative exposure to the short interest factor on their short side, due to the well-documented herding behavioral bias. In a de-leveraging event, these two factors move in the opposite directions as what many managers hope of, i.e., rallies in the short interest factor and plunges in the HF crowding side.

To gauge the impact of such de-leveraging episodes, we can leverage our two tradable portfolios that proxy the long and the short sides of popular hedge fund bets, using the HF crowding and short interest signals, respectively. These two portfolios are sector-neutral and signal-weighted, by investing the top 150 longs and top 150 shorts in the S&P 1500 index. These portfolios are tradable as custom basket swaps, excluding hard-to-borrow securities and adjusted for liquidity, with Bloomberg tickers **NMWCRWDL** (most crowded HF longs) and **NMWSINTL** (most crowded shorts). Our proxy of the aggregate equity long/short book is a market neutral portfolio that purchases NMWCRWDL and shorts NMWSINTL.

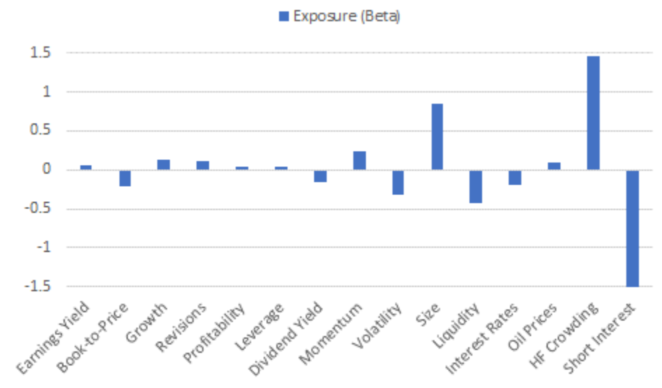
We then conduct an attribution analysis of our equity L/S proxy using our US equity risk model. As shown in Figure 30(A), this portfolio would generate a loss of -1.5% on a +1% move in the short interest

[11] In that heavily shorted stocks outperform the market.
[12] Such that the popular long names fall sharply.

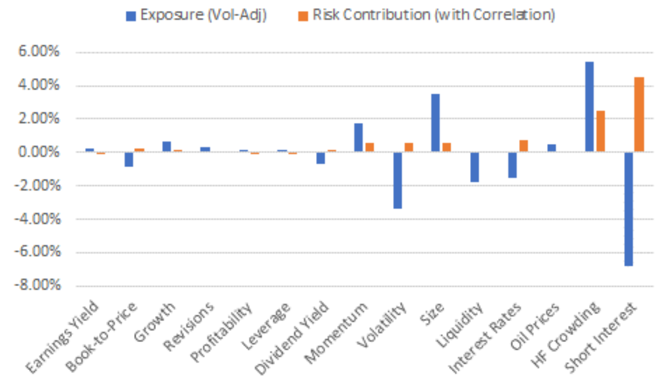
FMP (Factor Mimicking Portfolio), all else held constant. Given that our short interest factor rallied 5% over the past five trading days, it contributes a -7.5% return to the portfolio.

Figure 30 The representative equity L/S portfolio is positively exposed to the HF crowding signal and negatively exposed to the short interest factor

A) Style Exposures



B) Style Vol-Adj Exposures & Risk Contributions



Sources: Bloomberg Finance LLP, EDGAR, FTSE Russell, Haver, I/B/E/S, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

In our risk models, factors are normalized to z-scores to construct FMPs[13] and then further adjusted with statistical sensitivity of stock returns to FMPs returns. Consequently, our factor exposures reflect not only the cross-sectional ranking of stocks to a particular signal, but also whether these stocks have recently been moving with a particular factor. You can think of the FMP as a "pure" factor with no exposure to other systematic factors in the risk model. However, if a factor is "missing" from the model, some of that systematic risk may be picked up by other included factors. For example, in recent days, retail behavior has been picked up by the short interest factor in our US equity risk model.

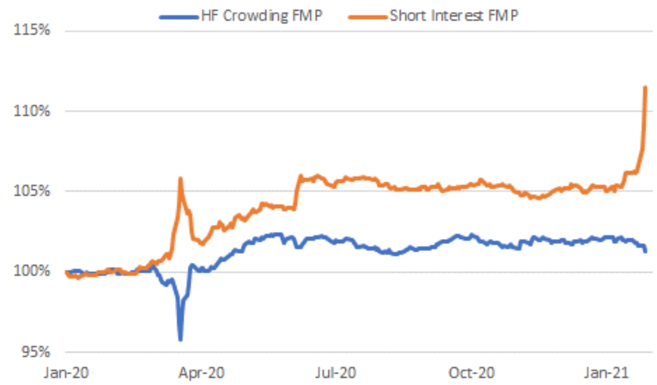
Short interest FMP rallied +5%, but HF Crowding FMP was only down -57bps this week

While the short interest factor rallied violently in the past four trading days, the HF crowding signal was relatively flat (see Figure 31A), down by only -57bps with most of that downside on Wednesday. We see a similar dynamic across our US sector risk models, where the spike in the short interest FMPs have been quite significant, the damage to the HF crowding factors have not been as nearly severe (see Figure 31B). The short rally has been more substantial in the consumers, health care, and TMT sectors, which happen to be the battleground between retail investors and institutional investors recently, even after controlling for other systematic risks.

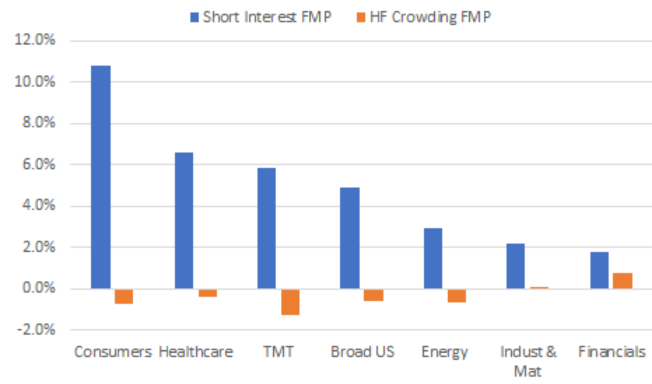
[13] Using a risk model allows us to dis-entangle the sources of risk and return across different factors. Each factor in the model is represented by a FMP with the following characteristics: (1) z-score of 1 to the target factor, (2) z-score of 0 to all other systematic factors (styles, industries, countries) in the risk model, and (3) minimum idiosyncratic risk across the multiple portfolios that satisfy (1) and (2).

Figure 31 Short interest FMP spiked in our broad US and sector risk models, most notably in consumers, health care, TMT

A) Recent FMP Returns – Broad US Market



B) FMP Return in Sector Risk Models (Jan 22nd-27th)



Sources: Bloomberg Finance LLP, EDGAR, FTSE Russell, Haver, I/B/E/S, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Therefore, while we have heard anecdotal accounts of HF de-leveraging on both the long side and short sides; and media speculations of "crowded" longs selling off, our HF crowding factor does not appear to support this view. Rather, the damage is so far mostly on the short side. Given the fast-moving nature of the market in recent days, we are mindful about the timeliness of our data sources.

Short interest signals in our risk model are updated daily

Our short interest factor is based on IHS Markit's global securities lending database. IHS Markit sources its data directly from prime brokers, custodian banks, and hedge funds. As a result, our short interest data is updated daily, with a lag of T+2 days, which is far superior to the bi-weekly data provided with a longer lag by exchanges.

HF crowding data is primarily based on 13F filings

Unlike short interest, our HF crowding data is primarily powered by our collections of 13F filings, which are of quarterly frequency, with a lag of 45 days. The most recent filings reflect hedge fund holdings as of September 30, 2020. The next batch update for Q4/2020 portfolios and will only become available on or around February 15, 2021.

Our empirical analysis (see *Port@ble Ownership*) suggests that ~70% of the hedge fund assets has a holding period of six months or longer. Consequently, our lagged data should reflect much of the long holdings. However, hedge fund turnover may have been higher in Q4/2020 than usual, due to Presidential election, COVID vaccine development, and significant factor rotations. Additionally, our risk model methodology adjusts for the statistical sensitivity of a stock to a factor, so a stock that is no longer crowded should in theory have a smaller beta coefficient to the HF crowding factor; and therefore lower exposure in our risk models.

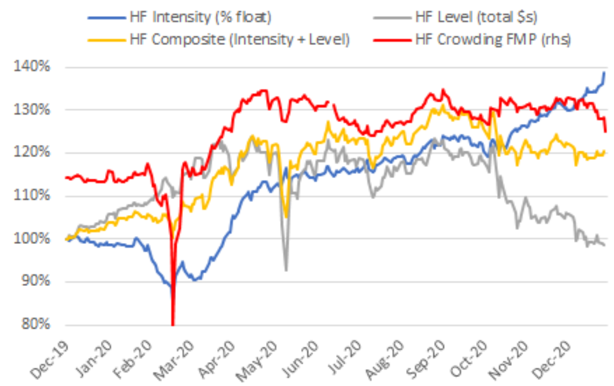
Evidence shows a divergence between HF crowding signals

To investigate further, we compare the performance of our HF crowding FMP with signal-weighted portfolios constructed from three additional HF signals based on the level (aggregate HF holdings), intensity (aggregate HF holdings relative to market cap), and popularity (# of HF owning a given stock). Note that the level factor is highly correlated (about 95%) with the popularity signal, whereas intensity represents stocks heavily owned by hedge funds relative to other types of shareholders. Our HF crowding FMP blends the level and intensity factors. Interested readers should refer our white paper *Port@able Ownership* for details.

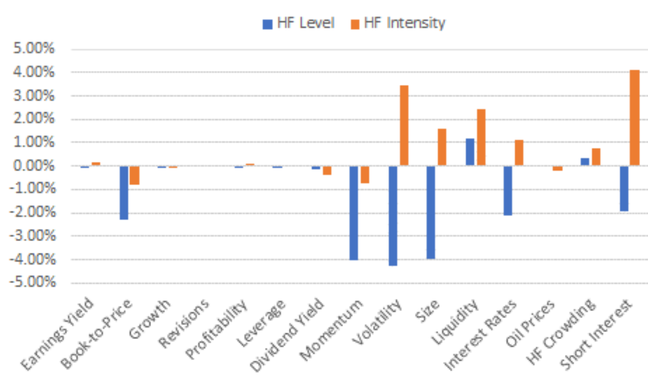
As shown in Figure 32(A), we observe a significant divergence between the level and intensity factors since early November's Presidential election. As shown in Figure 32(B), most of the return for the level and intensity factors has been driven by their exposures to technical factors in the past three months. The intensity factor's impressive rally has been driven by its exposures to short interest(+), volatility(+), and liquidity[14] (+), while the sell-off of the level signal is due to its exposures to volatility(-), momentum(+), and size(+). The positive exposure of the intensity factor to short interest suggests that there is some overlap between heavily shorted stocks and high participation long names, i.e., a two-way demand from hedge funds for some securities.

Figure 32 Hedge Fund intensity and level factors diverged significantly post-election and COVID vaccine approval

A) Signal-Weight HF Factors and FMP



B) Return Contributions from Style Factors



Sources: Bloomberg Finance LLP, EDGAR, FTSE Russell, Haver, I/B/E/S, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

De-risking on the long side may have been concentrated in mega/large-cap stocks

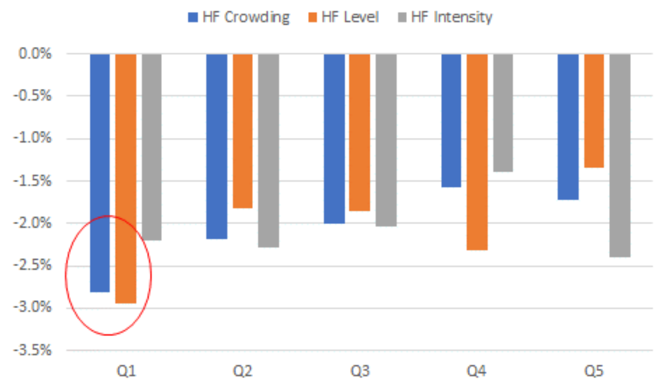
As shown in Figure 33(A) and (B), the HF level factor (i.e., stocks with the highest $ HF ownership) underperformed in recent days. This suggests that de-risking on the long side was likely concentrated in mega/large-cap stocks that tend to have significant allocations across a large number of hedge funds. Given the recent underperformance of small-cap, less liquid stocks, selling these crowded large-cap stocks to bring down total leverage appears to be a logical choice. Hedge fund selling of crowded large-caps was less likely to have significant market impact due to the highly liquid nature of these stocks.

Lastly, we sort all stocks into five quintile portfolios, by the three HF factors, respectively, where Quintile 1 contains the most heavily shorted stocks. The relatively mild sell-off in the "pure" HF crowding FMP and level factor suggests that there was some de-risking on Wednesday, when the retail frenzy last peaked. Should the retail investor insurgence continue and intensity, it is certainly possible that hedge funds would be forced to further de-leverage, which may trigger a larger scale sell-off of popular long names with more significant market impact.

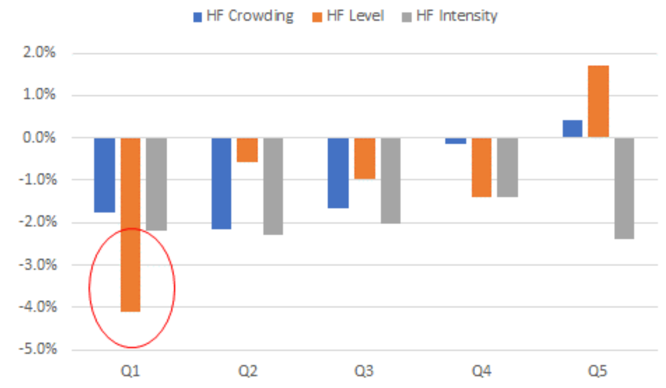
[14] The liquidity signal in our risk model is controlled for size. Therefore, stocks with positive (negative) liquidity exposure have seen higher (lower) volumes relative to their market-cap.

Figure 33 High HF level stocks underperformed in recent days, suggesting some de-risking on the long side

A) Quintile Portfolio Return on Jan 27th



B) Quintile Portfolio Return during Jan 22nd-27th



Sources: Bloomberg Finance LLP, EDGAR, FTSE Russell, Haver, I/B/E/S, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

DISCLOSURE SECTION

Analyst Certification:
The analyst of Wolfe Research primarily responsible for this research report whose name appears first on the front page of this research report hereby certifies that (i) the recommendations and opinions expressed in this research report accurately reflect the research analysts' personal views about the subject securities or issuers and (ii) no part of the research analysts' compensation was, is or will be directly or indirectly related to the specific recommendations or views contained in this report.

Other Disclosures:
Wolfe Research, LLC does not assign ratings of Buy, Hold or Sell to the stocks it covers. Outperform, Peer Perform and Underperform are not the respective equivalents of Buy, Hold and Sell but represent relative weightings as defined above. To satisfy regulatory requirements, Outperform has been designated to correspond with Buy, Peer Perform has been designated to correspond with Hold and Underperform has been designated to correspond with Sell.

Wolfe Research Securities and Wolfe Research, LLC have adopted the use of Wolfe Research as brand names. Wolfe Research Securities, a member of FINRA (www.finra.org) is the broker-dealer affiliate of Wolfe Research, LLC and is responsible for the contents of this material. Any analysts publishing these reports are dually employed by Wolfe Research, LLC and Wolfe Research Securities.

The content of this report is to be used solely for informational purposes and should not be regarded as an offer, or a solicitation of an offer, to buy or sell a security, financial instrument or service discussed herein. Opinions in this communication constitute the current judgment of the author as of the date and time of this report and are subject to change without notice. Information herein is believed to be reliable but Wolfe Research and its affiliates, including but not limited to Wolfe Research Securities, makes no representation that it is complete or accurate. The information provided in this communication is not designed to replace a recipient's own decision-making processes for assessing a proposed transaction or investment involving a financial instrument discussed herein. Recipients are encouraged to seek financial advice from their financial advisor regarding the appropriateness of investing in a security or financial instrument referred to in this report and should understand that statements regarding the future performance of the financial instruments or the securities referenced herein may not be realized. Past performance is not indicative of future results. This report is not intended for distribution to, or use by, any person or entity in any location where such distribution or use would be contrary to applicable law, or which would subject Wolfe Research, LLC or any affiliate to any registration requirement within such location. For additional important disclosures, please see www.wolferesearch.com/disclosures.

The views expressed in Wolfe Research, LLC research reports with regards to sectors and/or specific companies may from time to time be inconsistent with the views implied by inclusion of those sectors and companies in other Wolfe Research, LLC analysts' research reports and modeling screens. Wolfe Research communicates with clients across a variety of mediums of the clients' choosing including emails, voice blasts and electronic publication to our proprietary website.

This report is limited for the sole use of clients of Wolfe Research. Authorized users have received an encryption decoder which legislates and monitors the access to Wolfe Research, LLC content. Any distribution of the content produced by Wolfe Research, LLC will violate the understanding of the terms of our relationship.