Did the Game Stop for Hedge Funds?

Jun Chen, Byoung-Hyoun Hwang, and Melvyn Teo[☆]

Abstract

Can retail investors on social media platforms effectively target hedge fund short positions? We show that the disclosure of hedge fund short positions drives social media activity on WallStreetBets, which in turn precipitates price increases for heavily shorted stocks. The resultant short squeezes hurt hedge funds, which respond by shorting less aggressively, leading to prolonged overpricing in the stock market. In line with a causal interpretation, we find that the impact of social media on stock returns manifests around the publication dates for short sales, but not around the settlement dates, and attenuates during the trading restrictions imposed by Robinhood.

[☆]Chen is at the School of Business, Renmin University. E-mail: chen.jun@rmbs.ruc.edu.cn. Hwang is at the Nanyang Business School, Nanyang Technological University. E-mail: bh.hwang@ntu.edu.sg. Teo (corresponding author) is at the Lee Kong Chian School of Business, Singapore Management University. Address: 50 Stamford Road, Singapore 178899. E-mail: melvynteo@smu.edu.sg. Tel: +65-6828-0735. Fax: +65-6828-0427.

1. Introduction

Retail trading accounts for an increasing share of US equity trading volume. The 2022 Bloomberg Intelligence report shows that retail investors' share of equity trading volume approached 23 percentage points in 2021, up from 10 percentage points just ten years ago.[1] Survey evidence indicates that these investors frequently draw from social media when deciding which stocks to buy and sell.[2] The reliance on social media can lead to more coordinated actions either because these retail investors draw from a common information source or because they use social media to actively encourage each other to pile into certain stocks for ideological or strategic considerations. The growing prevalence of retail investors and their increasingly coordinated actions imply that retail investors are becoming more important in setting prices. To the extent that retail investors are informed, this development can lead to more efficient markets. To the extent that they are less informed or trade for ideological or strategic reasons, it may destabilize financial markets and pose a new source of risk to institutional investors such as hedge funds.

Perhaps the clearest indication of the latter possibility is the short squeeze on GameStop. In January 2021, the dramatic rally in GameStop shares, ostensibly fueled by social media users on Reddit's WallStreetBets, led to the capitulation of the multi-billion dollar hedge fund Melvin Capital. Anecdotal evidence suggests that by placing a massive short bet on GameStop via put options, which have to be disclosed in regulatory Form 13F filings every quarter, Melvin Capital effectively placed itself in the crosshairs of retail investors on WallStreetBets.[3] Market observers postulate that the high short interest in GameStop compounded the problem for Melvin Capital as it increased the likelihood of a classic short

[1] "Stock market gamification unlikely to end soon or draw new rules," Bloomberg Intelligence, 19 February 2021.

[2] "Social media is the most popular source of investment ideas for young investors, CNBC survey finds," CNBC, 26 August 2021.

[3] Melvin Capital's regulatory filings indicate that it owned put options on 3.4 million GameStop shares by the end of the second quarter of 2020 and 5.4 million shares by the end of the third quarter of the same year. See "Melvin Capital, GameStop, and the road to disaster," The Financial Times, 6 February 2021. According to the Financial Times, Melvin Capital lost 53% in January 2021. The rally on GameStop was allegedly sparked by a Reddit user Stonksflyingup who posted a video on WallStreetBets entitled "GME squeeze and the demise of Melvin Capital" on 27 October 2021.

squeeze whereby following a rise in GameStop shares, multiple short sellers rush to cover their short positions simultaneously, driving up the price of GameStop further.[4]

Motivated by the coordinated actions of retail investors on social media which lead to the short squeeze in GameStop, we assess three research questions: (1) Do retail investors on social media platforms target hedge funds' short positions? (2) On balance, do retail investors succeed in pushing up the prices of stocks shorted by hedge funds? Relatedly, how much do hedge funds lose when they are targeted by retail investors? (3) How do hedge funds respond to the new threat posed by retail investors on social media platforms, and what are the broader implications for asset prices and market efficiency?

To test whether retail investors on social media platforms target hedge funds' short positions, we turn to Reddit's WallStreetBets and examine how social media activity for a specific stock changes with the publication of that stock's short interest level. Our analysis comprises all US stocks from January 2020 through March 2022, a period characterized by intense social media activity on WallStreetBets. To measure social media activity on a stock, we evaluate the number of posts, comments to posts, unique posters writing posts, emojis used in posts, and meme stock lingos employed in posts that reference the stock. Our five-dimensional measure of social media activity comprehensively captures both the level and the intensity of the online discussion activity for a stock. If retail investors target hedge funds' short positions, we should observe a spike in social media activity in response to high levels of short interest.

To show that the levels of social media activity and short interest are not simultaneously determined by fundamental news and, instead, causally tied, we exploit the following institutional feature. Short sellers need to report their short positions to the Financial Industry Regulatory Authority (FINRA) by the 15th of each month or the preceding business day if the 15th is not a business day (settlement date). FINRA compiles the short interest data

[4]See "Melvin Capital, GameStop, and the road to disaster," The Financial Times, February 6, 2021. In that article, a prominent short seller was quoted as saying, "I don't get why Melvin were there, I just don't get it. We get really uncomfortable if one of our shorts has a 10 per cent short interest ratio. The higher the short interest, the higher the risk, since if everyone rushed to exit their positions at once, a sudden surge in demand to buy back stock would push the price up further – a classic short squeeze. That's the part where the retail people got it right, to their credit."

and discloses the data to the public eight business days after the settlement date (publication date). If social media activity and short interest are simultaneously determined by fundamental news, we should observe a positive correlation around the settlement date. In contrast, if social media users react to the publication of high levels of short interest, we should observe a positive correlation between social media activity and short interest around the publication date and not around the settlement date.

We find that consistent with the view that social media users are reacting to the publication of short interest, social media activity on a stock increases following the publication date of short interest but not after the settlement date. Specifically, a one-standard deviation increase in short interest (i.e., a 7.5% increase) precipitates a meaningful 22.8% increase (t-statistic = 4.47) in the number of posts referencing the stock the day after the publication date. In additional tests, we take advantage of the regulatory feature whereby, unlike stock positions that are borrowed for shorting, hedge fund put option positions have to be disclosed on a quarterly basis on their Form 13F with a disclosure lag. We show that social media attention in a stock increases following the disclosure of hedge fund put option positions in the stock but not after the purchase of those put options. These results suggest that retail investors are indeed targeting hedge fund short positions once they are revealed to the investment public.

To address the impact of social media activity on stock prices, we examine whether the spikes in social media activity are followed by noticeable stock price appreciations for heavily shorted stocks. We find that an increase in social media activity is associated with higher future stock returns for high-short interest stocks but not for other stocks.[5] By our estimates, a one-standard deviation increase in the number of posts on a stock relative to the mean precipitates an economically meaningful 2.29% increase (t-statistic = 3.55) in stock returns the next day for high-short interest stocks relative to a similar increase in the number of posts for other stocks.

Next, we show that increases in social media activity engender greater stock returns

[5]We classify as high-short interest stocks those in the top one percentile based on short interest. Our results are robust when we reclassify as high-short interest stocks those in the top two percentile based on short interest.

around the publication dates for short interest data but not around the settlement dates for short sales. The close correspondence of social media activity and stock returns, as well as the non-result around settlement dates suggest that the spikes in social media activity and the ensuing stock price increases are causally tied. Retail investors encourage each other on social media to pile into stocks that hedge funds currently short either for ideological reasons or in an attempt to engineer a short squeeze. The buying pressure from retail investors, perhaps exacerbated by the ensuing squeeze with hedge funds also buying to cover their short positions, cause stock prices to rise.

To understand whether the price appreciations of stocks with high short interest are indeed driven by retail investors targeting hedge funds' short positions, we test whether the effects of social media activity on stock returns are amplified when social media users allude to short sellers or hedge funds in their online discussions about the stock. We find that the impact of social media on stock returns is stronger when the words "short seller", "shorts", "squeeze", and "Melvin" are referenced in social media posts about the stock. Moreover, social media activity is a better harbinger of higher stock returns when hedge funds publicly disclose put options on the stock or when the names of hedge funds with publicly disclosed put option positions on a stock are mentioned in social media posts about that stock.

To further investigate the causal relation between social media activity and stock prices, we leverage the temporary trading restrictions imposed on several stocks in January 2021 by Robinhood, a popular trading platform favored by retail investors. Robinhood states that it curtailed trading in these stocks to meet its capital obligations and clearinghouse deposit requirements. The trading restrictions surprised many retail investors and precipitated at least one class action suit, which alleges that Robinhood "deprived their customers of the ability to use their service" as well as the potential gains from trading for "no legitimate reason." If the stock price increases in high-short interest stocks are indeed driven by social media activity, the trading restrictions imposed by Robinhood should weaken the positive relation between social media activity and subsequent stock returns. This is indeed what we find. For stocks that were affected by the trading restrictions, social media activity no longer had an impact on future stock returns during the period when those trading restrictions were

in effect. These results lend credence to the view that social media activity and stock prices are causally related.

How much do hedge funds lose when their short positions are targeted by retail investors? We show that hedge fund monthly returns and Fung and Hsieh (2004) seven-factor alphas negatively relate to prior month's social media activity on the heavily shorted stocks that they sold short. Specifically, a one-standard deviation increase (relative to the mean) in the monthly number of posts on the heavily shorted stocks stocks short sold by a fund engenders a 0.44% decrease in fund returns and a 0.65% reduction in fund Fung and Hsieh (2004) seven-factor alpha the next month. Conversely, hedge fund performance is unrelated to social media activity on the non-heavily shorted stocks that a fund sold short. Our findings validate concerns raised by practitioners that it has become increasingly risky to short sell heavily shorted stocks as social media platforms allow retail investors to mount coordinated attacks against short sellers.

To understand the response of hedge funds to the new threat posed by retail investors on social media platforms, we examine changes in hedge fund short positions after the first quarter of 2021. We focus on changes around the first quarter of 2021 as predatory attacks on hedge fund short positions by social media users were particularly salient in January 2021, a period during which the prices of GameStop and other meme stocks surged multiple times.[6] We find that after the first quarter of 2021, hedge funds reduced both the dollar value of and the equivalent number of shares outstanding associated with their publicly disclosed short positions on high-short interest stocks. Relative to the four-quarter period before the first quarter of 2021, over the next four quarters, hedge funds reduced the value of their publicly disclosed short positions in high-short interest stocks by US$345.44 million or 56.78%. During the same time, they also publicly short sold 22.68 million or 67.47% fewer shares in high-short interest stocks. Concomitantly, we neither observe a reduction in hedge funds' publicly disclosed short positions on other stocks nor a reduction in their non-publicly disclosed short positions on high-short interest stocks.

[6] According to the Wall Street Journal, "At the worst point in January 2021, Melvin Capital Management was losing more than $1 billion a day as individual investors on online forums such as Reddit banded together to push up prices of stocks Melvin was betting against." See "Hedge fund Melvin lost $6.8 billion in a month. Winning it back is taking a lot longer," Wall Street Journal, 28 January 2022.

We show further that the reduction in hedge fund publicly disclosed short selling activity is driven by high-short interest stocks that appreciated in price while garnering significant social media attention. Relative to other comparable hedge fund management companies, those that publicly shorted high-short interest stocks that subsequently experienced above-median increases in posts on WallStreetBets in the first quarter of 2021 are 28.5 percentage points less likely to short sell high-short interest stocks via put options after that quarter. Similarly, compared to other similar hedge fund management companies, those that publicly shorted high-short interest stocks that subsequently experienced above-median price appreciations in the first quarter of 2021 are 32.6 percentage points less likely to short sell high-short interest stocks via put options thereafter. These results support the notion that the reduction in hedge fund publicly disclosed short selling activity following the first quarter of 2021 is driven by the intense social media activity in the stocks that they publicly sold short and the resultant mark-to-market losses in their put options on those stocks.

Do retail investors push prices away from fundamentals and destabilize financial markets when attacking hedge fund short positions? We find that in the process of squeezing hedge fund short positions in high-short interest stocks, social media users move stock prices above fundamental values. Specifically, heavily shorted stocks that experience high social media traffic, and therefore are more likely to appreciate in price, are also less likely to announce positive cash flow news. Moreover, we find that following the imposition of trading restrictions by Robinhood, the prices of heavily shorted stocks with high social media activity fall. These results are consistent with the view that social media activity pushes stock prices temporarily above fundamental values through retail trading activity. The trading restrictions imposed by Robinhood substantially reduce such activity, thereby allowing stock prices to revert back to fundamental values.

Our study adds to the growing literature on word-of-mouth effects in financial markets. One of the first to consider the transmission of financial information through social interactions, Shiller and Pound (1989) show that in general investors do not derive investment ideas by themselves, but are instead drawn to stocks through conversations with their peers. Subsequent work analyzing the decisions of investors residing in a common locale (Kaustia

and Knüpfer, 2012) or sharing a common workplace (Hvide and Östberg, 2015) arrive at the same conclusion. With the advent of digital technologies, an increasing share of our interactions occurs virtually. Early research on word-of-mouth via social media shows that views expressed on online investment platforms can predict future returns and earnings surprises (Chen, De, Hu, and Hwang, 2014; Avery, Chevalier, and Zeckhauser, 2016; Jame, Johnston, Markov, and Wolfe, 2016). More recent work highlights the negative implications of social media platforms, including how they encourage the formation of echo chambers (Cookson, Engelberg, and Mullins, 2022) and can contain fake news (Kogan, Moskowitz, and Niessner, 2023), which lead to extreme and erroneous beliefs. Our study complements the aforementioned perspectives by studying coordination by retail investors on social media, and showing that such coordination attempts are often not driven by fundamental news, can alter other market participants' behavior, and create temporary price dislocations.

2. Data and methodology

2.1. Social media and stock data

We collect social media activity data from the largest Reddit investing subreddit – WallStreetBets (also known as r/wallstreetbets), which has 13.5 million subscribers as of the end of 2022. We collate all submission data from WallStreetBets between January 1, 2020 and March 31, 2022 using the Pushshift Application Programming Interface, which is designed to provide enhanced functionality and search capabilities for searching Reddit comments and submissions (i.e., posts).[7] Our sample period spans four quarters before and four quarters after the first quarter of 2021, a time characterized by intense social media activity in several high-short interest stocks, including GameStop.

For each submission, we identify the stock referenced by the submission. To do so, we first compile a list of tickers mentioned on WallStreetBets during our sample period. Since users on WallStreetBets frequently use the term "$TICKER" to refer to a stock with ticker

[7] https://files.pushshift.io/reddit/

symbol "TICKER", we find all capitalized words or symbols that follow a dollar sign. After retrieving a list of potential tickers, we clean them via the following method. First, we restrict the potential tickers to the list of tickers belonging to stocks that were publicly traded during our sample period. Second, we manually remove potential tickers such as "THE", "WSB", "DD", "ON", "ALL", "IN", "FOR" and "GO" as they are unlikely to refer to the ticker of a company. After cleaning, we narrow down the list to 1,861 valid tickers. Equipped with this list of tickers, we proceed to tag all WallStreetBets submissions. Specifically, if the title or the main text of a submission contains the ticker of a company, we assign the submission to that company on the date when the submission was posted. However, a submission could include multiple tickers. Since such multiple-ticker submissions allude to more than one stock in the same post, it makes it difficult for us to assign social media activity to the appropriate stock (Chen et al., 2014). Therefore, we focus our analysis on single-ticker submissions, which comprise about 72.7% of all submissions on WallStreetBets that contain at least one ticker. In addition to data on submissions, we also download data on comments that are posted in response to the submissions on WallStreetBets. We assign a comment to a stock on the date when the comment is posted if the comment is in response to a submission assigned to that stock.

To gauge social media activity, for each stock, we analyze the number of posts, the number of unique posters creating posts, the number of comments to posts, the number of emoji's used in posts, as well as the number of meme stock lingos used in posts that reference the stock. Since emojis are used by posters to convey emotion and lingos allow users to communicate nuances of meaning and emotion better than regular language, they help measure the intensity of social media discussions.

We merge social media data with daily stock return data from CRSP using ticker symbols. Following prior literature, we focus on common stocks with CRSP share codes equal to 10 or 11 and exclude securities such as warrants, preferred shares, American Depositary Receipts, closed-end funds, and REITs. We also supplement our stock-level return data with short interest data that we retrieve from the Compustat Short Interest File, which reports mid-month and month-end short interest data for stocks listed on the NYSE, AMEX, and

NASDAQ.

We classify as top shorted (or heavily shorted/high-short interest) stocks those in the top one percentile relative to all other stocks based on short interest. During our sample period, at any point in time there are 27.41 top shorted stocks. We note that a total of 138 stocks were classified as top shorted stocks at some point during the sample period.[8] This constitutes a much larger set than the sample of 13 high-short interest stocks investigated by Allen et al. (2022).

To ameliorate the concern that some firms may not attract a minimum level of social media activity during our sample period because they do not particularly interest users on WallStreetBets, we restrict the sample to firms that have been mentioned on WallStreetBets at least five times during our sample period.[9] We conduct our stock-level regression analyses on a final sample of 953 firms.

2.2. Hedge fund data

We study the relation between social media activity and hedge fund performance using monthly net-of-fee returns and assets under management (henceforth AUM) data of live and dead hedge funds from the Morningstar database from January 2020 to March 2022. In our fund universe, we have a total of 2,551 US hedge funds, comprising 2,027 live funds and 524 dead funds.

Since hedge funds are not required to list on commercial databases, hedge fund data are susceptible to self-selection biases. For example, hedge funds often include returns prior to fund listing dates onto the databases. Because funds that have good track records are more likely to list on databases so as to attract investment capital, the backfilled returns tend to be higher than non-backfilled returns, which leads to a backfill bias (Liang, 2000; Fung and Hsieh, 2009; Bhardwaj, Gorton, and Rouwenhorst, 2014). To alleviate backfill bias,

[8] Our baseline results are robust to reclassifying as top shorted stocks, those stocks in the top two percentile based on short interest.

[9] We obtain similar baseline results when we do not impose this restriction or when we restrict the sample of firms to those that are mentioned on WallStreetBets at least ten times during our sample period.

throughout this paper, we study hedge fund returns reported post fund database listing date. Since Morningstar does not provide listing date information, we rely on the Jorion and Schwarz (2019) algorithm to infer fund database listing dates.

We estimate hedge fund performance relative to the Fung and Hsieh (2004) seven factors. These factors are S&P 500 return minus the risk-free rate (SNPMRF), Russell 2000 return minus the S&P 500 return (SCMLC), change in the constant maturity yield of the 10 year U.S. Treasury bond appropriately adjusted for the duration (BD10RET), change in the spread of Moody's BAA bond over 10-year Treasury bond appropriately adjusted for duration (BAAMTSY), bond PTFS (PTFSBD), currency PTFS (PTFSFX), and commodity PTFS (PTFSCOM), where PTFS is primitive trend following strategy. Fung and Hsieh (2004) show that their model captures a substantial 84% of the variation in hedge fund index returns.

We obtain hedge funds' stock holding data from Form 13F filings. Under the Securities Exchange Act of 1934, all institutional investors, including hedge fund management companies, with investment discretion over $100 million are required to make quarterly disclosures of portfolio holdings to the SEC on Form 13F within 45 days of the quarter-end. The types of securities that must be reported on Form 13F include stocks and equity options.

To study hedge funds' publicly disclosed short positions, we extract institutional investors' put option data including CUSIP, fair value, and shares of the securities underlying the options. As the option data are not included in standard commercial databases, e.g., Thomson Reuters, as per Aragon, Martin, and Shi (2019) we access the data directly through the SEC Electronic Data Gathering, Analysis, and Retrieval system (henceforth EDGAR).

To identify managers of hedge funds from the universe of 13F filers, we match the 13F filings data to the sample of hedge funds from the Morningstar database via fund management company name. We are able to match the 13F filing data to 267 hedge fund firms that operate 1,506 hedge funds and hold put options on 1,719 unique stocks during our sample period. Of these funds, 235 hedge funds operated by 70 hedge fund firms report returns during our sample period.

[Insert Table 1 here]

Panels A and B of Table 1 report the summary statistics for the stock- and hedge fund-level attributes used in the paper. Panel C of Table 1 reports the difference in attributes between the top shorted stocks and the non-top shorted stocks. Panel A reveals that the social media activity for the average stock in the sample is low. On average, stocks garner 0.3610 posts and 9.52 comments a day. However, there is significant heterogeneity in social media attention across stocks. Panel C reveals that top-shorted stocks attract significantly greater social media attention than do non-top shorted stocks. For example, top shorted stocks attract 6.25 posts and 248.51 comments a day while non-top shorted stocks only elicit 0.18 posts and 2.31 comments a day. We observe similar results with other measures of social media activity such as the number of unique posters, the number of emojis used, and the number of lingos used.[10] Therefore high-short interest stocks are likely to be synonymous with meme stocks or stocks with substantial social media activity.[11]

3. Empirical results

3.1. Social media response to hedge fund short positions

Do retail investors on social media platforms increasingly target hedge funds' short positions? To address our first research question, we first explore the relation between short interest and social media activity. To show that the levels of social media activity and short interest are not simultaneously determined by fundamental news and, instead, causally tied, we exploit the following institutional feature. Short sellers are required to report to FINRA their short positions as of settlement on the 15th of each month or the preceding business day if the 15th is not a business day, and as of settlement on the last business day of the month. FINRA then compiles the short interest data and provides it for publication on the 8th

[10] We use the set of lingos listed under the "A meme stock glossary" in the investopedia page https://www.investopedia.com/meme-stock-5206762. These lingos include "Apes", "BTFD", "Diamond hands", "FOMO", "Hold the line", "Paper hands", "Stonks", "Tendies", "To the moon", and "YOLO".

[11] Meme stocks are typically defined as stocks with elevated social media activity (https://www.investopedia.com/meme-stock-5206762). However, some practitioners classify as meme stocks those with high short interest *and* elevated social media activity. See, for example, the definition of meme stocks used in the construction of the meme stock ETF (https://www.roundhillinvestments.com/etf/meme/).

business day after the reporting date.[12] This creates a lag of at least eight business days between the settlement of short sales and the publication of short sales data. If investors on WallStreetBets are responding to short sellers then social media activity should increase following the publication of short interest as opposed to the settlement of short sales.

To test whether social media activity responds to the publication of short interest data or to the settlement of short sales, we estimate the following multivariate regression on changes in social media activity:

$$\begin{aligned}\Delta Social\ Media\ Activity_{it+1} = \alpha + \beta_1 Short\ Interest_{it} \\ + \sum_i \beta_2^i Firm\ Dummy_i \\ + \sum_t \beta_3^t Settlement\ Cycle\ Dummy_t + \epsilon_{it},\end{aligned} \tag{1}$$

where $\Delta Social\ Media\ Activity$ is a placeholder for one of five measures for changes in social media activity on WallStreetBets, $ShortInterest$ is short interest in the stock, $FirmDummy$ is firm dummy, and $Settlement\,Cycle\ Dummy$ is settlement cycle dummy. The measures of changes in social media activity include the change in the number of posts ($\#Posts$), the change in the number of comments to posts ($\#Comments$), the change in the number of posters writing posts ($\#Posters$), the change in the number of emojis in posts ($\#Emojis$), the change in the number of meme stock lingos used in posts ($\#MemeLingos$)) that reference the stock. Change in social media activity is measured on day $t+1$ relative to the average social media activity from day $t-3$ to day $t-1$, where day t is either the publication date or the settlement date for short sales.[13] Statistical inferences are based on White (1980) robust standard errors clustered by firm and settlement cycle.

If traders are indeed responding to the publication of short interest as opposed to the settlement of short sales, we should expect the coefficient estimates on short interest to be positive and statistically significant in the regressions on the social media variables post

[12] See https://www.nasdaqtrader.com/trader.aspx?id=shortintpubsch

[13] Our results are qualitatively similar when we evaluate the percentage change in social media activity on day $t+1$ relative to average social media activity from day $t-5$ to day $t-1$.

publication, and to be economically modest and statistically indistinguishable from zero in the regressions on the social media variables post settlement. This is precisely what we find. The coefficient estimates on *Short Interest* in the regressions on *#Posts*, *#Comments*, *#Posters*, *#Emojis*, and *#Meme Lingos* reported in Panel A of Table 2 are all positive and statistically significant at the 5% or 1% level. Conversely, the coefficient estimates on *Short Interest* in the regressions on *#Posts*, *#Comments*, *#Posters*, *#Emojis*, and *#Meme Lingos* reported in Panel B of Table 2 are all statistically indistinguishable from zero at the 10% level. The coefficient estimate on *Short Interest* in the *#Posts* regression reported in Panel A indicates that a one-standard deviation increase in short interest (i.e., a 7.5% increase) precipitates a meaningful 22.8% (or 0.11 standard deviation) increase in the number of posts referencing the stock on WallStreetBets the day after the publication date.

[Insert Table 2 and Figure 1 here]

Figure 1 illustrates the daily percentage change in the number of posts referencing a stock on WallStreetBets around the publication of short interest and the settlement of short sales for the stock. It corroborates the results from the regressions and suggests that investors on WallStreetBets respond to the publication of short sales data as opposed to the actual settlement of short sales. These results support the view that short sales drive social media activity on WallStreetBets.

Next, to further test whether retail investors on social media respond to hedge fund short positions, we take advantage of the regulatory feature whereby, unlike stock that hedge funds borrow for shorting, hedge fund put option positions have to be disclosed on a quarterly basis on their Form 13F with a disclosure lag. We test whether the disclosure of hedge fund put option positions precipitates social media activity in the stock by estimating regressions analogous to those in Eq. (1). The dependent variables include the changes in the social media activity measures featured in Eq. (1). Change in social media activity is evaluated on $t+1$ relative to the average social media activity from $t-3$ to $t-1$, where day t is the 13F filing deadline. The independent variable of interest is the size of publicly disclosed hedge fund short positions on the stock scaled by short interest (*HF Short Position*).

The results reported in Panel A of Table 3 indicate social media activity increases after the disclosure of hedge fund short positions. The analysis in Panel A assumes that most of the 13F disclosures are reported on the filing deadline. However, only 68.09% of the put option positions on Form 13F are disclosed on day t itself. Conversely, 97.36% are disclosed between $t-2$ and t inclusive. Therefore, a three-day filing period may be more appropriate. Panel B shows that our findings are robust when we to accommodate a three-day filing window and study percentage change in social media activity that is measured on $t+1$ relative to the average social media activity from $t-5$ to $t-3$. Panel C reveals that we do not observe a similar effect when we conduct a placebo test and evaluate social media activity around the reporting date, i.e., the last day of the reporting quarter.

[Insert Table 3 here]

3.2. Social media and stock returns

Do retail investors succeed in their campaigns against hedge fund short positions? To address our second research question, we estimate the following multivariate regression on stock returns:

$$\begin{aligned} Return_{it} = \alpha &+ \beta_1 High\ Short\ Interest_{it-1} + \beta_2 \ln(1 + Social\ Media\ Activity_{it-1}) \\ &+ \beta_3 \ln(1 + Social\ Media\ Activity_{it-1}) * High\ Short\ Interest_{it-1} \\ &+ \beta_4 Sentiment_{it-1} + \beta_5 Dow\ Jones_{it-1} \\ &+ \beta_6 Analyst\ Upgrades_{it-1} + \beta_9 Return_{it-1} \\ &+ \beta_{10} Return_{it-5,t-2} + \beta_{11} Return_{it-60,t-6} \\ &+ \sum_i \beta_{12}^i Firm\ Dummy_i + \sum_t \beta_{13}^t Time\ Dummy_t + \epsilon_{it}, \end{aligned} \tag{2}$$

where *Return* is daily stock return, *High Short Interest* is an indicator variable that takes a value of one if the stock is in the top one percentile of stocks based on short interest, *Social Media Activity* is a placeholder for one of the five measures for social media activity on WallStreetBets, *Sentiment* is Dow Jones Newswire daily sentiment, *Dow Jones* is

an indicator variable that takes a value of one if any Dow Jones Newswire mentions that stock that day, *Analyst Upgrades* is the number of analyst upgrades on the stock that day, *Firm Dummy* is firm dummy, and *Time Dummy* is year-month-day dummy. Our measures for social media activity include the number of posts ($\#Posts$), the number of comments to posts ($\#Comments$), the number of unique posters writing posts ($\#Posters$), the number of emojis used in posts ($\#Emojis$) and the number of meme stock lingos used in posts ($\#Meme\ Lingos$) that reference the stock. We also estimate analogous regressions on Daniel, Grinblatt, Titman, and Wermers (1997) DGTW-adjusted returns, Carhart (1997) four-factor adjusted returns, and Fama and French (2015) five-factor adjusted returns. Statistical inferences are based on White (1980) robust standard errors clustered by firm and day.

[Insert Table 4 here]

The coefficient estimates on the interaction variables reported in Table 4 support the view that social media activity influences stock prices. They are all positive and statistically significant at the 1% level regardless of whether we estimate regressions on stock returns, DGTW-adjusted returns, Carhart (1997) four-factor adjusted returns or Fama and French (2015) five-factor adjusted returns. The coefficient estimate on the interaction between $\ln(1+\#Posts)$ and $High\ Short\ Interest$ indicates that a one-standard deviation increase in $\#Posts$ relative to the mean precipitates a meaningful 2.29% increase in stock returns the next day for high-short interest stocks relative to other stocks. The coefficient estimates on the stock control variables accord with the extant literature. Stock returns positively relate to stock sentiment, stock mentions by media articles, and the number of analysts issuing upgrades on the stock.

We conduct a series of robustness tests to verify the strength of our baseline regression results. First, we exclude GameStop Corp. (GME) from the sample and redo the baseline tests. Second, we remove AMC Entertainment Holdings (AMC) from the sample and reestimate the baseline regressions. These are the two most commonly discussed meme stocks on WallStreetBets during our sample period. Third, we focus on posts that reference the

respective stocks in the title as opposed to those that either reference the stock in the title or in the body. Fourth, we expand the set of high-short interest stocks to include those in the top two percentile of stocks based on short interest. Fifth, we remove posts on WallStreetBets with due diligence reports from the sample due to concerns that such posts could contain new fundamental information about the stock (Bradley et al., 2021). Table IA1 of the Internet Appendix reveals that our results remain qualitatively unchanged with these adjustments although as expected the findings weaken but are still statistically significant at the 10% or 5% level after removing GameStop from the sample.

Next, to test whether social media activity and stock prices are causally related, we examine the impact of social media activity on stock prices around the publication dates for short interest data and around the settlement dates for short sales. To do so, we estimate the following multivariate regression on stock returns:

$$\begin{aligned} Return_{it} = \alpha &+ \beta_1 High\ Short\ Interest_{it-1} + \beta_2 \ln(1 + Social\ Media\ Activity_{it-1}) \\ &+ \beta_3 \ln(1 + Social\ Media\ Activity_{it-1}) * High\ Short\ Interest_{it-1} \\ &+ \beta_4 Sentiment_{it-1} + \beta_5 Dow\ Jones_{it-1} \\ &+ \beta_6 Analyst\ Upgrades_{it-1} + \beta_9 Return_{it-1} \\ &+ \beta_{10} Return_{it-5,t-2} + \beta_{11} Return_{it-60,t-6} \\ &+ \sum_i \beta_{12}^i Firm\ Dummy_i + \sum_t \beta_{13}^t Settlement\ Cycle\ Dummy_t + \epsilon_{it}, \end{aligned} \tag{3}$$

where the variables are as per Eq. (1) and (2). Statistical inferences are based on White (1980) robust standard errors clustered by firm and settlement cycle. We also estimate regressions on Fama and French (2015) five-factor adjusted returns.

If traders' response to the publication of short interest (as opposed to the settlement of short sales) drives the positive relation between social media activity and future stock returns, we should expect the coefficient estimates on the interactions between *High Short Interest* and post publication social media variables to be positive and statistically significant, and those on the interactions between *High Short Interest* and the post settlement social media

variables to be economically modest and statistically indistinguishable from zero. The results reported in Table 5 confirm this. The coefficient estimates on the interactions between *High Short Interest* and the social media activity variables post publication date are all positive and statistically significant at the 5% or 1% level. Conversely, the coefficient estimates on the interactions between *High Short Interest* and the social media activity variables post settlement date are all statistically indistinguishable from zero at the 10% level.

[Insert Table 5 here]

To test whether the price appreciations of stocks with high social media activity are indeed driven by retail investors targeting hedge funds' short positions, we ask whether the effects of social media activity on stock returns are amplified when social media users allude to short sellers or hedge funds in their online discussions about the stock. To test, we include in the Eq. (2) regressions interactions between our five measures of social media activity, $HighShortInterest$, and the natural logarithm of one plus proxies for social media mentions of short selling activity. These proxies include the number of posts referencing the stock that mentions "short seller" (*#Short Seller*), "shorts" (*#Shorts*), "squeeze" (*#Squeeze*), and "Melvin" (*#Melvin*). We study the last term as Melvin Capital, a prominent hedge fund that incurred significant losses from its short position in GameStop, could be used by retail investors on WallStreetBets to rally other investors against short sellers in general.

The coefficient estimates on the triple interaction terms reported in Table 6 are all positive and statistically significant at the 1% or 5% level. They indicate that the effects of social media activity on stock returns are indeed stronger when short sellers are mentioned in social media posts on the stock. These results support the view that social media users on WallStreetBets are motivated in part by a desire to challenge short sellers.

[Insert Tables 6 and 7 here]

Next, we test whether the impact of social media activity on stock returns should be strengthened when hedge funds publicly disclose short positions on the stock or when hedge

fund short selling activity is more salient for social media users. In that effort, we construct several measures of hedge fund short selling activity or the salience of such activity for social media users. Our main measure of hedge fund short selling activity is the size of the publicly disclosed hedge fund short positions on the stock scaled by short interest (*HF Short Position*). We supplement that with additional proxies such as the number of hedge funds that publicly shorted the stock (*#HF Short*), the number of hedge funds with large publicly disclosed short positions in the stock (*#HF Large Short*), the natural logarithm of one plus the number of posts on a stock that mention the names of hedge funds that shorted the stock (ln(1+*#HF Short Posts*)), and the natural logarithm of one plus the number of posts on a stock that mention the names of hedge funds with large short positions in the stock (ln(1+*#HF Large Short Posts*)). The last two variables measure the salience of hedge fund short selling activity for social media users. Large publicly disclosed short positions are those in the top 20th percentile relative to all hedge fund publicly disclosed short positions that quarter. Next, we reestimate the Eq. (2) regressions after including the interactions between the natural logarithm of one plus our measures of social media activity, $High\,Short\,Interest$, and these proxies for hedge fund publicly disclosed short selling activity.

The coefficient estimates on the triple interaction terms reported in Table 7 are, with few exceptions, all positive and statistically significant at the 1% or 5% level. They indicate that the impact of social media activity on a stock's returns is indeed stronger when hedge funds publicly disclose significant short positions in the stock or when social media posts about the stock mention the names of hedge funds that publicly shorted the stock. These results are consistent with the notion that social media users purchase heavily shorted stocks to squeeze hedge funds that are short those stocks.

To further investigate the causal relation between social media activity and stock prices, we exploit the trading restrictions imposed on several stocks in January 2021 by Robinhood, a trading platform favored by many retail investors. These stocks include AAL, AMC, BB, BBBY, CTRM, EXPR, GME, KOSS, NAKD, NOK, SNDL, TR, and TRVG. Robinhood raised margin requirements on these securities, restricted buy transactions by clients, and closed out some positions automatically if clients were assessed to be at risk of not having

the necessary collateral.[14] According to Robinhood, it curtailed trading in these stocks to meet its capital obligations and clearinghouse deposit requirements. The trading restrictions caught many retail investors off guard and precipitated at least one class action suit from individual investors who alleged that Robinhood "deprived their customers of the ability to use their service" as well as the potential gains from trading for "no legitimate reason."

To test whether the trading restrictions affect the association between social media activity and stock returns, we include in the Eq. (2) regressions interactions between the natural logarithm of one plus the measures of social media activity, *High Short Interest*, and *Restriction*, an indicator variable that takes a value of one if Robinhood imposed trading restrictions on that stock that day. If the stock price increases in high-short interest stocks are indeed driven by social media activity, the trading restrictions imposed by Robinhood should weaken the positive relation between social media activity and subsequent stock returns, since the retail investors on WallStreetBets often traded via Robinhood. This is indeed what we find. The coefficient estimates on the triple interaction terms reported in Table 8 are all negative and statistically significant at the 5% or 1% level, consistent with the view that social media activity engenders stock price increases for heavily shorted stocks.

[Insert Table 8 here]

3.3. Social media and hedge fund performance

How much do hedge funds lose when their short positions are targeted by retail investors? Since social media activity is positively related to stock returns, we should find that a hedge fund's performance deteriorates when it short sells top-shorted stocks that subsequently attract significant social media traffic. To test, we estimate the following multivariate re-

[14] See, "Robinhood restricts trading in GameStop, other names involved in frenzy," CNBC, 28 January 2021 and "Robinhood, other brokerages restrict trading on GameStop, AMC," Wall Street Journal, 28 January 2021

gression on hedge fund monthly performance:

$$\begin{aligned} HF\ Alpha_{im} = {} & a + b\ln(1 + Social\ Media\ Activity\ High\ Short\ Interest_{im-1}) \\ & + c\ln(1 + Social\ Media\ Activity\ Low\ Short\ Interest_{im-1}) + d\ln(Size_{im-1}) \\ & + \sum_j e^j Fund\ Dummy_i^j + \sum_n f^n Year\ Month\ Dummy_m^n + \epsilon_{im}, \end{aligned} \tag{4}$$

where *HF Alpha* is fund alpha, *Social Media Activity High Short Interest* is a placeholder for one of five monthly social media activity measures aggregated over all top-shorted stocks that are short sold by the fund, *Social Media Activity Low Short Interest* is a placeholder for one of five monthly social media activity measures aggregated over all non-top-shorted stocks that are short sold by the fund, *Size* is fund AUM in US$ millions, *Fund Dummy* is the fund dummy, and *Year Month Dummy* is the year-month dummy. Fund alpha is monthly abnormal return from the Fung and Hsieh (2004) model, with the factor loadings estimated over the prior 24 months.[15] Top-shorted stocks are the top one percentile of stocks based on short interest. The social media activity measures include $\#Posts$, $\#Comments$, $\#Posters$, $\#Emojis$, and $\#Meme\ Lingos$ as per defined in Eq. (1) and Eq. (2). We base statistical inferences on robust standard errors that are clustered by hedge fund management company and year-quarter since hedge fund short positions are reported at the fund management company level and on a quarterly basis.

[Insert Table 9 here]

The results reported in Table 9 support the view that social media activity hurts the risk-adjusted performance of hedge funds that shorted high-short interest stocks. The coefficient estimates on the social media variables that relate to top-shorted stocks are negative, economically meaningful, and statistically significant at the 1% level in the regression on fund alpha. Conversely, we do not observe a similar negative relation between the social media variables that pertain to non-top-shorted stocks and fund alpha. The coefficient estimates on ln(1+*Social Media Activity High Short Interest*) reported in column 1 indicate that a

[15] Inferences do not change when we use factor loadings estimated over the past 36 months instead.

one-standard deviation increase (relative to the mean) in the monthly number of posts on top-shorted stocks sold short by a fund relative foreshadows a 0.44% decrease in fund return and a 0.65% reduction in fund seven-factor alpha the next month. We obtain qualitatively similar, albeit economically weaker, results for the regression on fund returns.[16]

3.4. Hedge funds' response to social media

How do hedge funds respond to the new threat posed by retail investors on social media platforms? To address our third research question, we study the change in hedge fund put option positions around the first quarter of 2021. The results in this paper suggest that social media activity has made it increasingly risky for hedge funds to short stocks via put options as hedge fund put option positions must be reported on a quarterly basis as part of hedge funds' mandatory 13F disclosures. The resultant disclosures leave hedge funds vulnerable to predatory attacks from retail investors on social media platforms. These attacks were particularly salient in January 2021 during which the price of GameStop and other high-short interest stocks surged multiple times. A natural question to ask is whether hedge funds curtailed their use of put options on heavily shorted stocks after the first quarter of 2021.

Figure 2 graphs the dollar value and size of hedge funds' aggregate publicly disclosed short positions over the sample period. It provides prima facie evidence that hedge funds in general reduced both the dollar value and the equivalent number of shares outstanding for the put option positions that they report on their 13F form disclosures after the first quarter of 2021. Relative to the four quarters before Q1 2021, over the next four quarters, hedge funds reduced the value of their aggregate publicly disclosed short positions in top-shorted stocks by US$345.44 million or 56.78%. During the same period they also publicly short sold 22.68 million or 67.47% fewer shares in top-shorted stocks. These results are consistent with the view that hedge funds respond to the heightened risk of short selling created by social media platforms by curtailing the amount and number of the short positions they express in the form of put options.

[16]While Melvin Capital is included in the analysis on hedge fund 13F filings, it is not included in the hedge fund performance analysis as Melvin did not report returns to the Morningstar database.

[Insert Figure 2 here]

To understand the drivers of the reduction in hedge fund publicly disclosed short selling activity, we estimate the following two difference-in-differences models:

$$\begin{aligned} HF\ Short_{iq} = a + bTreatment1_i * After\ Q12021_{iq} \\ + \sum_j e^j Fund\ Management\ Company\ Dummy_i^j \\ + \sum_n f^n Year\ Quarter\ Dummy_q^n + \epsilon_{iq}, \end{aligned} \tag{5}$$

$$\begin{aligned} HF\ Short_{iq} = a + bTreatment2_i * After\ Q12021_{iq} \\ + \sum_j e^j Fund\ Management\ Company\ Dummy_i^j \\ + \sum_n f^n Year\ Quarter\ Dummy_q^n + \epsilon_{iq}, \end{aligned} \tag{6}$$

where *HF Short* is a placeholder for one of three measures of hedge fund short selling activity, *Treatment1* is an indicator variable that takes a value of one if a hedge fund management company shorted top-shorted stocks at the end of Q4 2020 that subsequently received above-median percentage increases in number of posts on WallStreetBets in Q1 2021, *Treatment2* is an indicator variable that takes a value of one if a hedge fund management company shorted top-shorted stocks at the end of Q4 2020 that subsequently experienced above-median percentage increases in stock price in Q1 2021, *After Q12021* is an indicator variable that takes a value of one if the quarter is equal to or occurs after Q1 2021, *Fund Management Company Dummy* is the fund management company dummy, and *Year Quarter Dummy* is year-quarter dummy. Since *Treatment1* and *Treatment2* are subsumed by the fund management company fixed effects and *After Q12021* is subsumed by the year-quarter fixed effects, we omit them as standalone independent variables in the regression. The three measures of hedge fund short selling activity include an indicator variable that takes a value of one if a hedge fund management company publicly shorts a top-shorted stock at the end of that quarter (*Shorted High Short Interest*), the number of top-shorted stocks publicly shorted by the hedge fund

management company at the end of that quarter (*High Short Interest #Shorts*), and the natural logarithm of one plus the number of shares of top-shorted stocks shorted by the hedge fund management company at the end of that quarter (*High Short Interest #Shares*). Top-shorted stocks are stocks in the top one percentile based on short interest that quarter. We base statistical inferences on robust standard errors that are clustered by hedge fund management company and year-quarter.

[Insert Table 10 here]

The results reported in Panel A of Table 10 are consistent with the view that hedge funds that incur losses due to social media activity in the stocks that they shorted via put options in the first quarter of 2021 become more reluctant to short sell stocks with high short interest. The coefficient estimates on the interaction terms in the regressions are, with one exception, all statistically significant at the 5% or 1% level. The coefficient estimate on the interaction term reported in column 1 indicates that a hedge fund management company that publicly shorted a top-shorted stock at the end of Q4 2020 that subsequently garnered above-median increases in posts on WallStreetBets in Q1 2021 are 28.5 percentage points less likely to publicly short sell top-shorted stocks after that quarter. Similarly, the coefficient estimate on the interaction term reported in column 4 reveals that hedge fund management companies that publicly shorted a top-shorted stock at the end of Q4 2020 that subsequently experienced above-median increases in price in Q1 2021 are 32.6 percentage points less likely to publicly short sell top-shorted stocks after that quarter. Panel B of Table 10 reveals that our results remain qualitatively unchanged when we match treatment fund management companies to control fund management companies based on the size of fund management company publicly disclosed short positions. These results support the notion that the reduction in hedge fund short selling activity post the first quarter of 2021 is driven by the losses that they suffered as a result of social media activity in the high-short interest stocks that they publicly shorted.

3.5. Social media and market efficiency

Do retail investors push prices away from fundamentals and destablize financial markets? We postulate that social media users may be so focused on attacking hedge funds' short positions that they neglect fundamental information and move prices away from fundamental values. To test, we estimate regressions analogous to Eq. (2) but with a measure of cash flow news as the dependent variable. Our measure of cash flow news *Earnings Surprise* takes a value of one if a firm announces a positive earnings surprise that day, a value of negative one if a firm announces a negative earnings surprise that day, and a value of zero otherwise.

The results reported in columns 1 to 5 of Table 11 indicate that conditional on a firm being in the top one percentile of stocks based on short interest, social media activity negatively predicts cash flow news about the firm.[17] To understand whether this is driven by positive or negative cash flow news about the firm, we estimate analogous regressions on *Positive Earnings Surprise* a variable that takes a value of one if a firm announces a positive earnings surprise that day and a value of zero otherwise, and *Negative Earnings Surprise* a variable that takes a value of negative one if a firm announces a negative earnings surprise that day and a value of zero otherwise. The results in columns 6 to 15 indicate that our findings are driven by positive earnings surprises. Relative to high-short interest stocks with low social media activity, high-short interest stocks with high social media activity are less likely to announce positive earnings surprises. These results, in combination with those from Table 4, suggest that social media activity in high-short interest stocks moves prices above fundamental value.

[Insert Table 11 here]

[17] While the coefficient estimates on *High Short Interest* suggest that stocks in the top one percentile based on short interest are more likely to announce positive earnings surprises and/or are less likely to announce negative earnings surprises, we find that this finding is sensitive to the regression specification used. When we omit firm fixed effects from the regression, we find that the coefficient estimates on *High Short Interest* are neither positive nor statistically distinguishable from zero at the 10% level. We note that our main finding, i.e., that top-shorted stocks with high social media activity are less likely to announce positive cash flow news, is robust to omitting firm fixed effects.

4. Conclusion

Despite the intense discussion around how social media users on WallStreetBets triggered the massive short squeeze on GameStop that ultimately led to the collapse of Melvin Capital, it is not clear whether the GameStop episode was simply an isolated, one-off event or whether it has broader ramifications for financial markets. We shed light on this issue by investigating (i) whether social media users on WallStreetBets target hedge funds' short positions, (ii) their ability to trigger price increases in heavily shorted stocks, and (iii) the response of hedge funds to the short squeezes triggered by social media users.

We find that the disclosure of hedge fund short positions triggers social media activity on WallStreetBets. Increases in short interest on a stock leads to heightened social media activity on the stock. The fact that this occurs around the publication date for short interest data but not around the settlement date for short sales suggests that short interest and social media activity are causally tied. Moreover, we show that social media attention in a stock increases following the disclosure of hedge fund put options in the stock but not immediately following the quarter-end reporting date for those put options. These results suggest that retail investors are indeed targeting hedge fund short positions once they are revealed to the investment public.

Social media activity is in turn associated with higher future stock prices for heavily shorted stocks. Consistent with a causal relation between social media attention and stock returns, increases in social media activity engender greater stock returns around the publication dates for short interest data but not around the settlement dates for short sales. In line with a causal interpretation, the trading restrictions imposed by Robinhood in January 2021 attenuated the relation between social media attention and stock prices for stocks affected by the trading restrictions relative to other heavily shorted stocks. Social media users appear motivated to target hedge funds and other short sellers. The relation between social media activity and stock performance is amplified when short sellers in general or when the hedge funds that publicly shorted the stock are alluded to in social media posts about the stock.

The resultant short squeezes by social media users led to substantial losses for hedge funds and affected the way those hedge funds conduct short selling operations. Hedge funds that shorted high-short interest stocks that subsequently attracted intense social media attention last month experience meaningful losses the next month. Hedge funds respond to the increased short selling risk brought about by social media platforms by reducing both the dollar value and shares outstanding of their publicly disclosed short positions. The reduction in short positions is driven by hedge funds that experienced elevated social media activity in the high-short interest stocks that they publicly shorted.

Our results suggest that the social media-induced short squeeze in GameStop, which lead to the collapse of Melvin Capital, was not a one-off event. By acting as coordinating devices, online investment forums allow retail investors to credibly challenge hedge fund short positions. In response to the increased short sales risk for high-short interest stocks and to avoid appearing in the crosshairs of social media users, hedge funds tactically avoid expressing their bearish opinions on heavily shorted stocks via put options. Therefore, while the game of short selling has not stopped for hedge funds, it has fundamentally evolved to reflect the growing power of social media.

References

Agarwal, V., Jiang, W., Tang, Y., Yang, B., 2013. Uncovering hedge fund skill from the portfolio holdings they hide. Journal of Finance 68, 739–783.

Allen, F., Haas, M., Nowak, E., Pirovano, M., Tengulov, A., 2022. Squeezing shorts through social media platforms. Unpublished working paper, Imperial College.

Aragon, G.O., Martin, J.S., Shi, Z., 2019. Who benefits in a crisis? Evidence from hedge fund stock and option holdings. Journal of Financial Economics 131, 345–361.

Avery, C.N., Chevalier, J.A., Zeckhauser, R.J., 2016. The "CAPS" prediction system and stock market returns. Review of Finance 20, 1363–1381.

Barber, B.M., Huang, X., Jorion, P., Odean, T., Schwarz, C., 2022. A (sub)penny for your thoughts: tracking retail investor activity in TAQ. Unpublished working paper, University of California at Irvine.

Bhardwaj, G., Gorton, G., Rouwenhorst K.G., 2014. Fooling some of the people all of the time: the inefficient performance and persistence of commodity trading advisors. The

Review of Financial Studies 27, 3099–3132.

Bradley, D., Hanousek, J., Jame, R., Xiao, Z., 2021. Place your bets? The market consequences of investment research on Reddit's Wallstreetbets. Unpublished working paper, University of Southern Florida.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Chen, H., Du, P., Hu, J., Hwang, B.H., 2014. Wisdom of the crowds: the value of stock opinions transmitted through social media. Review of Financial Studies 27, 1367–1403.

Cookson, J.A., Engelberg, J., Mullins, W., 2023. Echo chambers. Review of Financial Studies 36, 450–500.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Farrell, M., Green, T.C., Jame, R., Markov, S., 2022. The democratization of investment research and the informativeness of retail investor trading. Journal of Financial Economics 145, 616–641.

Fung, W., Hsieh, D., 2004. Hedge fund benchmarks: a risk-based approach. Financial Analysts Journal 60, 65–80.

Fung, W., Hsieh, D., 2009. Measurement biases in hedge fund performance data: an update. Financial Analysts Journal 65, 36–38.

Hu, D., Jones, C.M., Zhang, V., Zhang, X., 2021. The rise of Reddit: how social media affects retail investors and short sellers' roles in price discovery. Unpublished working paper, Peking University.

Hvide, H., Östberg, P., 2015. Social interaction at work. Journal of Financial Economics 117, 628–652.

Jame, R., Johnston, R., Markov, S., Wolfe, M.C., 2016. The value of crowdsourced earnings forecasts. Journal of Accounting Research 54, 1077–1110.

Jorion, P., Schwarz, C., 2019. The fix is in: properly backing out backfill bias. Review of Financial Studies 32, 5048–5099.

Kaustia, M., Knüpfer, S., 2012. Peer performance and stock market entry. Journal of Financial Economics 104, 321–338.

Kogan, S., Moskowitz, T.J., Niessner, M., 2023. Social media and financial news manipulation. Unpublished working paper, University of Pennsylvania.

Liang, B., 2000. Hedge funds: the living and the dead. Journal of Financial and Quantitative Analysis 35, 309–326.

Lou, X., Ozik, G., Sadka, R., Shen, S., 2022. Innocuous noise? Social media and asset prices. Unpublished working paper, Boston College.

Shiller, R.J., Pound, J., 1989. Survey evidence on diffusion of interest and information among investors. Journal of Economic Behavior and Organization 12, 47–66.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

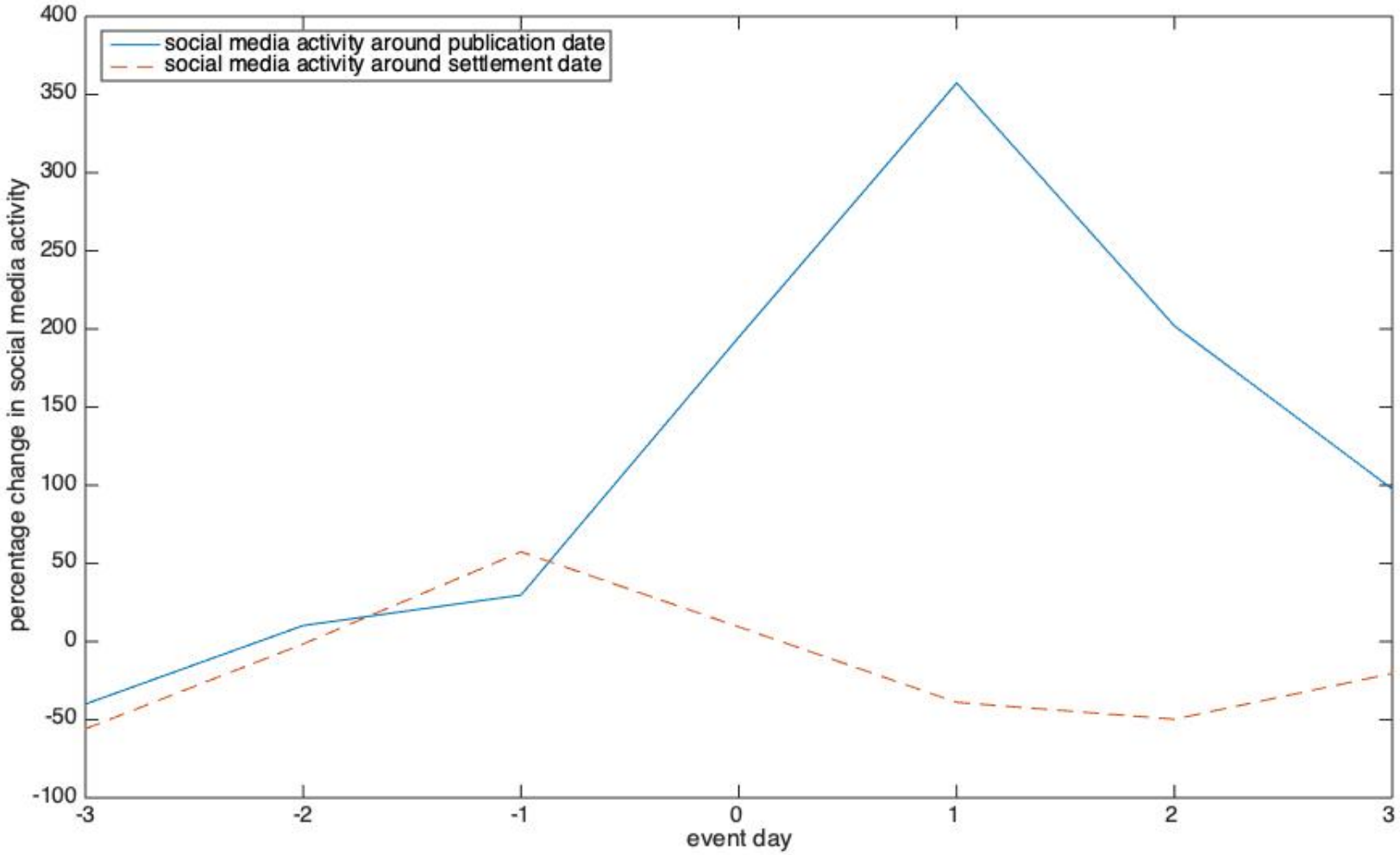

Figure 1: Social media activity around the publication and the settlement of short sales. Social media activity is the number of posts referencing the stock on day t on WallStreetBets. Short sellers are required to report to FINRA their short positions as of settlement on the 15th of each month or the preceding business day if the 15th is not a business day, and as of settlement on the last business day of the month. FINRA then compiles the short interest data and provides it for publication on the 8th business day after the reporting settlement date. Therefore, there is a lag of at least eight business days between short sales settlement and publication. Percentage change in social media activity is measured on day t relative to the average social media activity from day -3 to -1 inclusive, where day 0 is either the publication date or the settlement date. The solid line denotes social media activity around the publication of short sales data. The dashed line denotes social media activity around the settlement of short sales. The sample period is from January 2020 to March 2022.

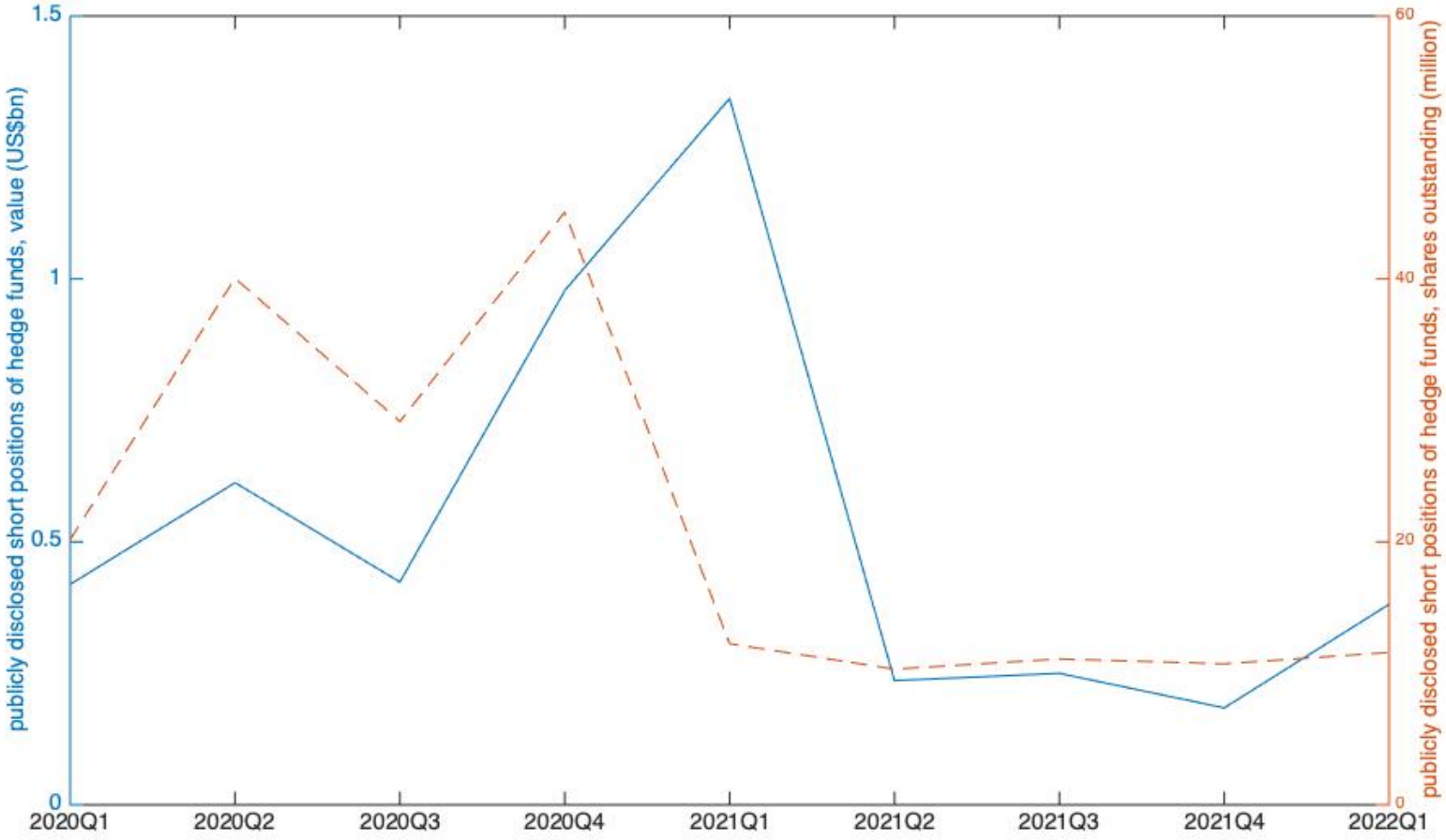

Figure 2: Hedge fund aggregate publicly disclosed short selling activity in top-shorted stocks before and after the first quarter of 2021. Publicly disclosed short positions are put option positions reported by hedge fund management companies as part of their mandatory 13F disclosures. Top-shorted stocks are stocks in the top one percentile based on short interest. The solid line denotes the aggregate publicly disclosed short positions of hedge funds in top-shorted stocks in millions of shares outstanding (y-axis on the right). The dashed line denotes the aggregate publicly disclosed short positions of hedge funds in top-shorted stocks in US$ billions (y-axis on the left). The sample period is from January 2020 to March 2022.

Table 1: **Summary statistics**

This table reports summary statistics for the key variables used in the paper. Social media activity data are from WallStreetBets. Hedge fund publicly disclosed short positions are hedge fund put option positions that have to be reported on Form 13F. Panel A reports summary statistics for stock-level attributes. Panel B reports summary statistics for hedge fund-level attributes. Panel C reports the differences in attributes between top-shorted and non top-shorted stocks. Top shorted stocks are those in the top one percentile based on short interest. The sample period is from January 2020 to March 2022.

Panel A: Summary statistics for stock-level attributes

Stock attributes	Number of observations	Mean	Std dev	5-%ile	25-%ile	50-%ile	75-%ile	95-%ile
Daily stock returns (%)	463999	0.144	6.396	-6.600	-2.024	-0.012	1.895	7.077
DGTW-adjusted returns (%)	463999	0.044	6.001	-5.787	-1.751	-0.120	1.456	5.958
Carhart (1997) 4-factor alpha (%)	449557	0.044	6.338	-6.236	-1.810	-0.116	1.519	6.424
Fama and French (2015) five-factor alpha (%)	449557	0.077	6.373	-6.238	-1.784	-0.095	1.549	6.527
Short interest	463999	0.065	0.071	0.006	0.018	0.041	0.088	0.194
Number of posts on the stock	463999	0.361	35.451	0.000	0.000	0.000	0.000	0.000
Number of comments to posts on the stock	463999	9.520	1187.362	0.000	0.000	0.000	0.000	0.000
Number of unique posters writing posts on the stock	463999	0.294	28.782	0.000	0.000	0.000	0.000	0.000
Number of emojis used in posts on the stock	463999	0.334	40.604	0.000	0.000	0.000	0.000	0.000
Number of meme stock lingos used in posts on the stock	463999	0.048	3.313	0.000	0.000	0.000	0.000	0.000
Number of mentions of "short seller" in posts on the stock	463999	0.001	0.106	0.000	0.000	0.000	0.000	0.000
Number of mentions of "shorts" in posts on the stock	463999	0.007	0.612	0.000	0.000	0.000	0.000	0.000
Number of mentions of "squeeze" in posts on the stock	463999	0.015	1.051	0.000	0.000	0.000	0.000	0.000
Number of mentions of "Melvin" in posts on the stock	463999	0.003	0.523	0.000	0.000	0.000	0.000	0.000
Dow Jones newswire sentiment on the stock	463999	0.135	2.815	0.000	0.000	0.000	0.000	4.000
Stock has a Dow Jones newswire	463999	0.214	0.410	0.000	0.000	0.000	0.000	1.000
Number of analysts upgrading the stock	463999	-0.000	0.115	0.000	0.000	0.000	0.000	0.000
Size of hedge fund publicly disclosed short positions scaled by total number of shares shorted	463048	0.096	0.198	0.000	0.000	0.007	0.101	0.495
Number of hedge funds that publicly shorted the stock	463999	4.251	5.980	0.000	0.000	1.000	7.000	16.000
Number of mentions of hedge funds that publicly shorted the stock in posts on the stock	463999	0.001	0.050	0.000	0.000	0.000	0.000	0.000

Panel B: Summary statistics for hedge fund-level attributes

Hedge fund attributes	Number of observations	Mean	Std dev	5-%ile	25-%ile	50-%ile	75-%ile	95-%ile
Monthly hedge fund returns (%)	2034	0.673	3.811	-4.931	-0.918	0.672	2.295	6.110
Fung and Hsieh (2004) seven-factor alpha (%)	2017	0.303	3.390	-4.241	-0.792	0.275	1.358	5.530
Hedge fund assets under management (US$m)	2034	270.610	760.678	0.029	4.130	50.000	158.120	1291.210

Panel C: Differences in attributes between top-shorted and non top-shorted stocks

	Top-shorted stocks		Non top-shorted stocks			
Stock attributes	Mean	Std dev	Mean	Std dev	Difference	p-value
Daily stock returns (%)	0.20	7.42	0.14	6.36	0.06	0.332
DGTW-adjusted returns (%)	0.08	6.72	0.04	5.98	0.04	0.475
Carhart (1997) four-factor alpha (%)	0.12	7.27	0.04	6.31	0.08	0.211
Fama and French (2015) five-factor alpha (%)	0.18	7.17	0.07	6.35	0.11*	0.090
Short interest	0.29	0.14	0.06	0.06	0.24***	0.000
Number of posts on the stock	6.25	192.11	0.18	13.44	6.07***	0.000
Number of comments to posts on the stock	248.51	6889.32	2.31	138.75	246.20***	0.000
Number of unique posters writing posts on the stock	4.99	155.16	0.15	11.26	4.84***	0.000
Number of emojis used in posts on the stock	7.00	220.56	0.13	15.17	6.87***	0.000
Number of meme stock lingos used in posts on the stock	0.80	17.93	0.03	1.26	0.77***	0.000
Number of mentions of "short seller" in posts on the stock	0.02	0.57	0.00	0.04	0.02***	0.000
Number of mentions of "shorts" in posts on the stock	0.14	3.46	0.00	0.16	0.14***	0.000
Number of mentions of "squeeze" in posts on the stock	0.31	5.97	0.01	0.24	0.30***	0.000
Number of mentions of "Melvin" in posts on the stock	0.09	3.05	0.00	0.03	0.09***	0.001
Dow Jones newswire sentiment on the stock	-0.02	2.98	0.14	2.81	-0.16***	0.000
Stock has a Dow Jones newswire	0.20	0.40	0.21	0.41	-0.02***	0.000
Number of analysts upgrading the stock	0.00	0.12	0.00	0.12	0.00***	0.000
Size of hedge fund publicly disclosed short positions scaled by total number of shares shorted	0.06	0.08	0.10	0.20	-0.03***	0.000
Number of hedge funds that publicly shorted the stock	4.66	4.50	4.24	6.02	0.43***	0.000
Number of mentions of hedge funds that publicly shorted the stock in posts on the stock	0.01	0.18	0.00	0.04	0.01***	0.000

Table 2: **Change in social media activity after the publication of short interest data**
Short sellers are required to report their short positions to the Financial Industry Regulatory Authority (FINRA) as of the 15th of each month or the preceding business day if the 15th is not a business day ("settlement date"). The short positions are reported to the public eight business days later ("publication date"). This table reports results from multivariate regressions on changes in social media activity after the publication date of short interest data and after the settlement date for short sales. The dependent variables include the change in the number of posts referencing the stock (*#Posts*), the change in the number of comments to posts referencing the stock (*#Comments*), the change in the number of unique posters writing posts that reference the stock (*#Posters*), the change in the number of emojis in posts referencing the stock (*#Emojis*), the change in the number of meme stock lingos used in posts referencing the stock (*#Meme Lingos*). Change in social media activity is measured on day $t + 1$ relative to the average social media activity from day $t - 3$ to day $t - 1$, where day t is either the publication date (Panels A and B) or the settlement date (Panels C and D). The independent variable of interest is either short interest in the stock (*Short Interest*) or an indicator variable that takes a value of one for the top one percentile of stocks based on short interest the previous day (*High Short Interest*). The control variables include dummy variables for firm and settlement cycle. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and settlement cycle. Panels A and C report regressions with *Short Interest*. Panels B and D report regressions with *High Short Interest*. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Change in social media activity after the publication date for short interest

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Short Interest	3.036***	2.736**	2.537***	2.337***	1.147*
	(4.47)	(2.24)	(4.35)	(4.41)	(1.84)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.316	0.092	0.339	0.272	0.214
#Observations	4073	3153	4057	1111	899

Panel B: Change in social media activity after the publication date for short interest

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
High Short Interest	0.895***	0.897*	0.693***	1.263***	0.959
	(3.19)	(1.76)	(2.94)	(2.87)	(1.49)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.090	0.055	0.119	-0.113	0.163
#Observations	4016	3110	4000	1094	883

Panel C: Change in social media activity after the settlement date for short sales [Placebo]

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Short Interest	0.055	-0.159	0.081	0.237	-0.023
	(0.19)	(-0.26)	(0.30)	(0.78)	(-0.07)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.054	0.005	0.057	0.051	0.059
#Observations	5603	4146	5542	1668	1269

Panel D: Change in social media activity after the settlement date for short sales [Placebo]

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
High Short Interest	0.051	-0.084	0.094	-0.555	-0.030
	(0.37)	(-0.11)	(0.93)	(-1.66)	(-0.42)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	-0.001	-0.002	0.021	0.075	0.046
#Observations	5514	4066	5453	1644	1244

Table 3: **Change in social media activity after the disclosure/reporting dates for hedge fund short positions**

U.S. institutional investment managers exercising investment discretion over $100 million are required to report their long-equity and long-option-positions via Form 13F to the U.S. Securities and Exchange Commission (SEC) within 45 days of each quarter. The public can observe the form immediately after it is filed (Agarwal, Jiang, Tang, and Yang, 2013). This table reports results from multivariate regressions on changes in social media activity after the disclosure date or reporting date for hedge fund short positions (via put options) on Form 13F. The dependent variables include the change in the number of posts referencing the stock (*#Posts*), the change in the number of comments to posts referencing the stock (*#Comments*), the change in the number of unique posters writing posts that reference the stock (*#Posters*), the change in the number of emojis in posts referencing the stock (*#Emojis*), the change in the number of meme stock lingos used in posts referencing the stock (*#Meme Lingos*). The independent variable of interest is the size of publicly disclosed hedge fund short positions on the stock scaled by short interest (*HF Short Position*). The control variables include dummy variables for firm and 13F disclosure cycle. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and 13F disclosure cycle. In Panel A, change in social media activity is measured on $t+1$ relative to the average social media activity from $t-3$ to $t-1$ where day t is the 13F filing deadline. In Panel B, change in social media activity is measured on $t+1$ relative to the average social media activity from $t-5$ to $t-3$ to accommodate a three-day window for the filing of 13F reports. In Panel C, change in social media activity is measured on $t+1$ relative to the average social media activity from $t-3$ to $t-1$ where day t is the 13F reporting date (i.e., the last day of the reporting quarter). The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Change in social media activity after disclosure date for HF short positions, assuming a one-day 13F filing window

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
HF Short Position	0.433**	1.094*	0.209**	0.262**	0.402***
	(2.93)	(2.16)	(2.58)	(2.37)	(3.65)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Disclosure Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.346	0.379	0.365	0.518	0.497
#Observations	645	421	638	150	171

Panel B: Change in social media activity after disclosure date for HF short positions, assuming a three-day 13F filing window

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
HF Short Position	0.422**	1.195***	0.451**	3.761*	0.743**
	(2.94)	(3.41)	(3.26)	(2.10)	(3.05)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Disclosure Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.128	0.012	0.177	-0.481	0.089
#Observations	311	233	311	67	70

Panel C: Change in social media activity after reporting date for HF short positions [Placebo]

	Dependent variable				
Independent variable	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
HF Short Position	-0.459	-3.580	-0.487	0.798	-0.874
	(-1.25)	(-1.18)	(-1.61)	(1.27)	(-0.74)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Disclosure Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.340	0.254	0.369	0.573	0.473
#Observations	430	362	430	106	110

Table 4: **Social media activity and stock returns**
This table reports results from multivariate regressions on daily stock returns and risk-adjusted returns. The dependent variables include daily stock returns (Panel A), Daniel, Grinblatt, Titman, and Wermers (1997) DGTW-adjusted returns (Panel B), Carhart (1997) four-factor adjusted returns (Panel C), and Fama and French (2015) five-factor adjusted returns (Panel D). The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day and their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*). The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t\text{-}1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t\text{-}5,t\text{-}2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t\text{-}60,\ t\text{-}6}$) as well as dummy variables for firm and year-month-day. The coefficient estimates for these firm control variables are omitted for brevity. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and day. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel A Regressions on daily stock returns					
High Short Interest	-0.045	-0.022	-0.041	0.003	0.009
	(-0.47)	(-0.23)	(-0.43)	(0.03)	(0.09)
ln(1+*Social Media Activity*)	0.063	0.024	0.079	0.026	0.165
	(0.41)	(0.54)	(0.44)	(0.12)	(0.60)
ln(1+*Social Media Activity*)**High Short Interest*	0.693***	0.349***	0.737***	0.811***	1.194***
	(3.55)	(2.95)	(3.67)	(4.46)	(5.51)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.109	0.109	0.109	0.109	0.109
#Observations	463999	463999	463999	463999	463999
Panel B: Regressions on Daniel et al. (1997) DGTW-adjusted returns					
High Short Interest	-0.052	-0.030	-0.048	-0.009	-0.003
	(-0.60)	(-0.34)	(-0.56)	(-0.10)	(-0.03)
ln(1+*Social Media Activity*)	0.069	0.031	0.085	0.021	0.159
	(0.47)	(0.74)	(0.49)	(0.10)	(0.61)
ln(1+*Social Media Activity*)**High Short Interest*	0.644***	0.322**	0.687***	0.771***	1.135***
	(3.36)	(2.58)	(3.48)	(4.37)	(5.43)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.015	0.015	0.015	0.015	0.015
#Observations	463999	463999	463999	463999	463999

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel C: Regressions on Carhart (1997) four-factor adjusted returns					
High Short Interest	-0.046	-0.021	-0.042	0.008	0.014
	(-0.47)	(-0.21)	(-0.43)	(0.08)	(0.15)
ln(1+*Social Media Activity*)	0.148	0.068	0.166	0.119	0.320
	(0.92)	(1.42)	(0.89)	(0.53)	(1.14)
ln(1+*Social Media Activity*)**High Short Interest*	0.757***	0.388***	0.800***	0.834***	1.238***
	(4.49)	(4.30)	(4.66)	(5.03)	(6.06)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.018	0.018	0.018	0.018	0.018
#Observations	449557	449557	449557	449557	449557
Panel D: Regressions on Fama and French (2015) five-factor adjusted returns					
High Short Interest	0.030	0.049	0.035	0.081	0.087
	(0.28)	(0.43)	(0.32)	(0.72)	(0.77)
ln(1+*Social Media Activity*)	0.154	0.052	0.175	0.104	0.277
	(0.94)	(1.11)	(0.92)	(0.44)	(0.99)
ln(1+*Social Media Activity*)**High Short Interest*	0.719***	0.387***	0.757***	0.778***	1.138***
	(4.11)	(3.92)	(4.45)	(6.52)	(9.05)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.016	0.016	0.016	0.016	0.016
#Observations	449557	449557	449557	449557	449557

Table 5: **Social media activity around the publication of short interest data and stock returns** This table reports results from multivariate regressions on daily stock returns and risk-adjusted returns. The dependent variables include daily stock returns (Panels A and C) and Fama and French (2015) five-factor adjusted returns (Panels B and D). The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day and their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*). The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t-1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t-5,t-2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t-60,\ t-6}$) as well as dummy variables for firm and settlement cycle. The coefficient estimates for these firm control variables are omitted for brevity. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and settlement cycle. Panels A and B report results for social media activity post publication date for short interest. Panels C and D report results for social media activity post settlement date for short sales. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel A Regressions on daily stock returns after publication date for short interest					
High Short Interest	0.194	0.150	0.119	4.196	0.106
	(0.17)	(0.14)	(0.10)	(1.08)	(0.04)
ln(1+*Social Media Activity*)	0.074	-0.009	0.087	-0.071	-0.568**
	(0.65)	(-0.33)	(0.68)	(-0.26)	(-2.39)
ln(1+*Social Media Activity*)**High Short Interest*	0.602**	0.247**	0.688**	1.634*	2.265**
	(2.08)	(2.46)	(2.06)	(1.92)	(2.22)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.219	0.236	0.217	0.115	0.221
#Observations	3569	2604	3542	825	737
Panel B: Regressions on five-factor adjusted returns after publication date for short interest					
High Short Interest	0.240	0.129	0.195	4.045	-0.488
	(0.23)	(0.12)	(0.19)	(1.33)	(-0.19)
ln(1+*Social Media Activity*)	0.045	-0.010	0.066	-0.015	-0.823**
	(0.38)	(-0.26)	(0.50)	(-0.05)	(-2.67)
ln(1+*Social Media Activity*)**High Short Interest*	0.540**	0.244**	0.598**	0.847	2.650***
	(2.31)	(2.32)	(2.20)	(0.97)	(3.06)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.035	0.048	0.035	-0.003	0.062
#Observations	3447	2500	3420	793	716

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel C: Regressions on daily stock returns after settlement date for short sales [Placebo]					
High Short Interest	-0.471	-1.200	-0.388	4.083*	2.359
	(-0.56)	(-0.98)	(-0.46)	(1.74)	(1.39)
ln(1+*Social Media Activity*)	0.014	-0.001	-0.003	-0.686	0.482*
	(0.27)	(-0.49)	(-0.04)	(-1.39)	(1.81)
ln(1+*Social Media Activity*)**High Short Interest*	0.246	0.013	0.299	1.421	2.275
	(0.96)	(0.50)	(1.09)	(0.73)	(1.21)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.183	0.184	0.178	0.203	0.180
#Observations	1874	1565	1845	387	414
Panel D: Regressions on five-factor adjusted returns after settlement date for short sales [Placebo]					
High Short Interest	-0.543	-2.429*	-0.427	3.097	-0.468
	(-0.48)	(-1.77)	(-0.39)	(1.15)	(-0.20)
ln(1+*Social Media Activity*)	0.080	-0.002**	0.071	0.051	0.033
	(1.30)	(-2.35)	(0.95)	(0.09)	(0.07)
ln(1+*Social Media Activity*)**High Short Interest*	0.160	0.029	0.262	1.610	0.273
	(0.75)	(1.21)	(1.21)	(1.11)	(0.10)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Settlement Cycle Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.080	0.137	0.087	0.011	0.308
#Observations	1874	1565	1845	387	414

Table 6: **Social media activity, mentions of short sellers, and stock returns**
This table reports results from multivariate regressions on daily stock returns and risk-adjusted returns. The dependent variables include daily stock returns and Fama and French (2015) five-factor adjusted returns. The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day as well as their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*) and the natural logarithm of one plus measures of short seller mentions on social media. The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The measures of short seller mentions include the number of posts referencing the stock that mentions "short seller" (*#Short Seller*), the number of posts referencing the stock that mentions "shorts" (*#Shorts*), the number of posts referencing the stock that mentions "squeeze" (*#Squeeze*), and the number of posts referencing the stock that mentions "Melvin" (*#Melvin*). The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t-1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t-5,t-2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t-60,\ t-6}$) as well as dummy variables for firm and year-month-day. The coefficient estimates for these firm control variables are omitted for brevity. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and year-month-day. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Regressions on daily stock returns					Regressions on five-factor adjusted returns				
Social media activity									
#Posts	*Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*	*#Posts*	*Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and ln(1+*#Short Seller*)									
0.615***	0.943***	0.748***	0.946***	0.723***	0.616***	0.692***	0.721***	0.926***	0.658***
(3.84)	(9.74)	(4.85)	(7.65)	(3.50)	(4.06)	(7.50)	(4.79)	(7.38)	(3.43)
Panel B: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and ln(1+*#Shorts*)									
0.468***	0.517***	0.470***	0.341***	0.772***	0.381***	0.460**	0.381***	0.342***	0.670**
(8.09)	(4.98)	(7.19)	(6.82)	(6.77)	(3.82)	(2.03)	(4.85)	(2.93)	(2.41)
Panel C: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and ln(1+*#Squeeze*)									
0.241***	0.309***	0.225***	0.115**	0.422***	0.192***	0.271***	0.168***	0.142***	0.415***
(4.07)	(8.60)	(3.80)	(2.03)	(6.09)	(6.06)	(4.00)	(5.13)	(4.15)	(3.71)
Panel D: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and ln(1+*#Melvin*)									
0.465***	0.751***	0.567***	0.715***	0.512***	0.771***	0.875***	0.859***	1.019***	0.876***
(4.56)	(14.89)	(5.50)	(8.70)	(3.98)	(12.31)	(26.06)	(13.29)	(17.25)	(10.79)

Table 7: **Social media activity, hedge fund shorting activity, and stock returns**
This table reports results from multivariate regressions on daily stock returns and risk-adjusted returns. The dependent variables include daily stock returns and Fama and French (2015) five-factor adjusted returns. The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day as well as their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*) and measures of hedge fund publicly disclosed short selling activity. The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The measures of hedge fund shorting activity include the size of the publicly disclosed hedge fund short positions on the stock scaled by short interest (*HF Short Position*), the number of hedge funds that publicly shorted the stock (*#HF Short*), the number of hedge funds with large publicly disclosed short positions in the stock (*#HF Large Short*), the number of posts on the stock that reference hedge funds that publicly shorted the stock (*#HF Short Posts*), and the number of posts on the stock that reference hedge funds with large publicly disclosed short positions in the stock (*#HF Large Short Posts*). The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t-1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t-5,t-2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t-60,\ t-6}$) as well as dummy variables for firm and year-month-day. The coefficient estimates for these firm control variables are omitted for brevity. Large publicly disclosed short positions are those in the top 20th percentile relative to all hedge fund publicly disclosed short positions that quarter. The *t*-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and year-month-day. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Regressions on daily stock returns					Regressions on five-factor adjusted returns				
Social media activity									
#Posts	*Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*	*#Posts*	*Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and *HF Short Position*									
3.008***	2.227***	3.156***	3.306***	3.455**	2.711***	1.842***	2.817***	3.203***	3.313**
(3.86)	(4.98)	(3.77)	(4.28)	(2.07)	(3.92)	(4.43)	(3.75)	(4.37)	(2.08)
Panel B: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and *#HF Short*									
0.043***	0.037***	0.045***	0.050***	0.045*	0.030**	0.026***	0.031*	0.046***	0.028
(3.42)	(4.25)	(3.30)	(4.24)	(1.78)	(2.04)	(3.76)	(1.88)	(3.12)	(0.98)
Panel C: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and *#HF Large Short*									
0.208***	0.170***	0.217***	0.204***	0.201*	0.148**	0.120***	0.147**	0.189**	0.127
(3.91)	(4.38)	(3.76)	(2.98)	(1.84)	(2.26)	(4.64)	(2.08)	(2.45)	(1.10)
Panel D: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and ln(1+*#HF Short Posts*)									
0.871***	1.520***	0.969***	1.029***	0.456	1.002***	1.568***	1.131***	1.189***	0.704**
(4.17)	(7.16)	(4.24)	(3.63)	(1.10)	(6.90)	(7.75)	(6.89)	(5.58)	(2.33)
Panel E: Interaction between ln(1+*Social Media Activity*), *High Short Interest*, and ln(1+*#HF Large Short Posts*)									
2.074***	3.092***	2.228***	1.833***	1.428	2.297***	3.184***	2.480***	2.057***	1.716*
(5.47)	(3.07)	(5.58)	(3.62)	(1.61)	(7.91)	(2.83)	(7.61)	(4.78)	(1.90)

Table 8: **Social media activity, Robinhood trade restrictions, and stock returns**
This table reports results from multivariate regressions on daily stock returns and risk-adjusted returns. The dependent variables include daily stock returns (Panel A) and Fama and French (2015) five-factor adjusted returns (Panel B). The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day as well as their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*) and an indicator variable for the days when trading restrictions were imposed by Robinhood on the stock (*Restriction*). The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t-1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t-5,t-2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t-60,\ t-6}$) as well as dummy variables for firm and year-month-day. The coefficient estimates on the firm controls are omitted for brevity. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and year-month-day. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel A Regressions on daily stock returns					
High Short Interest	-0.080	-0.042	-0.076	-0.014	-0.009
	(-0.83)	(-0.42)	(-0.79)	(-0.15)	(-0.10)
*High Short Interest*Restriction*	0.069	1.042	-0.384	1.915	-1.144
	(0.01)	(0.19)	(-0.05)	(0.37)	(-0.27)
ln(1+*Social Media Activity*)	0.094	0.031	0.117	0.073	0.283
	(0.61)	(0.69)	(0.64)	(0.33)	(1.10)
ln(1+*Social Media Activity*)**High Short Interest*	0.967***	0.457**	1.040***	1.167***	1.772***
	(2.78)	(2.48)	(2.88)	(3.27)	(4.06)
ln(1+*Social Media Activity*)**Restriction*	0.435	0.415	0.560	0.285	-0.740
	(0.30)	(0.43)	(0.36)	(0.23)	(-0.53)
ln(1+*Social Media Activity*)**High Short Interest***Restriction*	-1.971***	-1.441***	-2.045***	-2.522***	-2.573***
	(-3.86)	(-4.59)	(-3.71)	(-5.08)	(-3.24)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.110	0.109	0.110	0.110	0.110
#Observations	463999	463999	463999	463999	463999
Panel B: Regressions on five-factor adjusted returns					
High Short Interest	-0.003	0.035	0.001	0.068	0.073
	(-0.03)	(0.32)	(0.01)	(0.64)	(0.69)
*High Short Interest*Restriction*	-2.311	-1.659	-2.763	-0.718	-3.663
	(-0.31)	(-0.29)	(-0.36)	(-0.13)	(-0.81)
ln(1+*Social Media Activity*)	0.142	0.056	0.162	0.117	0.370
	(0.90)	(1.22)	(0.88)	(0.52)	(1.42)
ln(1+*Social Media Activity*)**High Short Interest*	1.026***	0.499***	1.101***	1.156***	1.765***
	(3.50)	(3.30)	(3.70)	(3.87)	(4.99)
ln(1+*Social Media Activity*)**Restriction*	0.592	0.487	0.722	0.439	-0.616
	(0.42)	(0.53)	(0.49)	(0.37)	(-0.47)
ln(1+*Social Media Activity*)**High Short Interest***Restriction*	-1.979***	-1.390***	-2.050***	-2.425***	-2.440***
	(-3.96)	(-4.53)	(-3.79)	(-4.98)	(-3.04)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.018	0.018	0.018	0.018	0.018
#Observations	449557	449557	449557	449557	449557

Table 9: **Social media activity in hedge fund short positions and hedge fund performance**
This table reports multivariate regressions on hedge fund performance. The dependent variables include hedge fund monthly return (Panel A) and Fung and Hsieh (2004) seven-factor monthly alpha (Panel B), where factor loadings are estimated over the last 24 months. The primary independent variable of interest is the natural logarithm of one plus social media activity in the top-shorted stocks publicly shorted by the hedge fund last month (ln(1+*Social Media Activity High Short Interest*)). The measures of social media activity include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The regressions control for the natural logarithm of one plus social media activity in the non top-shorted stocks publicly shorted by the hedge fund last month (ln(1+*Social Media Activity Low Short Interest*)). Top-shorted stocks are the top one percentile of stocks based on short interest. The other independent variables include the natural logarithm of last month's fund AUM in US$m (ln(*Size*)) as well as fixed effects for fund and year-month. The *t*-statistics, in parentheses, are derived from robust standard errors that are clustered by fund management company and year-quarter. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel A Regressions on hedge fund monthly returns					
ln(1+*Social Media Activity High Short Interest*)	-0.277**	-0.203**	-0.280**	-0.202*	-0.232*
	(-2.45)	(-2.46)	(-2.41)	(-2.09)	(-2.26)
ln(1+*Social Media Activity Low Short Interest*)	0.097	0.005	0.101	0.096	0.080
	(0.82)	(0.05)	(0.76)	(0.57)	(0.78)
log(*Size*)	-0.904**	-0.874**	-0.902**	-0.909**	-0.898**
	(-2.53)	(-2.65)	(-2.53)	(-2.50)	(-2.51)
F-test: ln(1+*Social Media Activity High Short Interest*) - ln(1+*Social Media Activity Low Short Interest*) = 0	16.48***	5.28*	13.19***	2.84	8.17**
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year-month Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.374	0.374	0.374	0.373	0.372
#Observations	2034	2034	2034	2034	2034
Panel B Regressions on hedge fund monthly seven-factor alpha					
ln(1+*Social Media Activity High Short Interest*)	-0.410**	-0.278**	-0.413**	-0.317**	-0.430***
	(-3.19)	(-3.14)	(-3.21)	(-2.96)	(-3.93)
ln(1+*Social Media Activity Low Short Interest*)	0.276**	0.129*	0.277**	0.158	0.317
	(2.73)	(2.03)	(2.40)	(1.02)	(1.84)
log(*Size*)	-0.327	-0.295	-0.324	-0.332	-0.321
	(-1.25)	(-1.25)	(-1.25)	(-1.27)	(-1.23)
F-test: ln(1+*Social Media Activity High Short Interest*) - ln(1+*Social Media Activity Low Short Interest*) = 0	42.93***	26.72***	41.08***	7.63**	13.72***
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year-month Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.066	0.065	0.066	0.063	0.065
#Observations	2017	2017	2017	2017	2017

Table 10: **Social media activity and hedge fund publicly disclosed short selling activity**
This table reports multivariate regressions on hedge fund publicly disclosed short selling activity. The dependent variables include an indicator variable that takes a value of one if a hedge fund management company publicly shorts a top-shorted stock at the end of that quarter (*Shorted High Short Interest*), the number of top-shorted stocks publicly shorted by the hedge fund management company at the end of that quarter (*High Short Interest #Shorts*), and the natural logarithm of one plus the number of shares of top-shorted stocks shorted by the hedge fund management company at the end of that quarter (*High Short Interest #Shares*). Top-shorted stocks are stocks in the top one percentile based on short interest that quarter. Publicly disclosed short sales denote hedge fund management company put option positions, which have to be reported in Form 13F as part of hedge fund mandatory disclosures. The first independent variable of interest is the interaction between an indicator variable that takes a value of one if a hedge fund management company shorted top-shorted stocks at the end of Q4 2020 that subsequently received above-median percentage increase in number of posts in Q1 2021 (*Treatment1*) and an indicator variable that takes a value of one if the quarter is equal to or occurs after Q1 2021 (*After Q12021*). The second independent variable of interest is the interaction between an indicator variable that takes a value of one if a hedge fund management company shorted top-shorted stocks at the end of Q4 2020 that subsequently experienced above-median percentage increases in stock price in Q1 2021 (*Treatment2*) and *After Q12021*. The regressions include fixed effects for fund management company and year-quarter. The *t*-statistics, in parentheses, are derived from robust standard errors that are clustered by fund management company and quarter. Panel A reports results when treatment fund management companies are not matched to control fund management companies. Panel B reports results when treatment fund management companies are matched to control fund management companies based on fund management company publicly disclosed short positions. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Dependent variable					
Independent variable	*Shorted High Short Interest* (1)	*High Short Interest #Shorts* (2)	ln(1+*High Short Interest #Shares*) (3)	*Shorted High Short Interest* (4)	*High Short Interest #Shorts* (5)	ln(1+*High Short Interest #Shares*) (6)
Panel A: Treatment fund management companies not matched to control fund management companies						
Treatment1 x *After Q12021*	-0.285**	-1.148**	-1.789*			
	(-2.96)	(-2.91)	(-1.83)			
Treatment2 x *After Q12021*				-0.326***	-0.991**	-2.274**
				(-3.52)	(-3.13)	(-2.70)
Fund Management Company Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.560	0.812	0.606	0.571	0.809	0.611
#Observations	1494	1494	1494	1494	1494	1494
Panel B: Treatment fund management companies matched to control fund management companies based on publicly disclosed short positions						
Treatment1 x *After Q12021*	-0.542***	-0.948***	-3.587*			
	(-3.45)	(-3.38)	(-1.98)			
Treatment2 x *After Q12021*				-0.502***	-0.841***	-3.425**
				(-3.41)	(-3.38)	(-2.61)
Fund Management Company Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.476	0.639	0.464	0.546	0.609	0.463
#Observations	231	231	231	258	258	258

Table 11: **Social media activity and stock cash flow news**
This table reports results from multivariate regressions on daily stock cash flow news. The dependent variables includes a variable that takes a value of one if the firm announces a positive earnings surprise that day, a value of negative one if the firm announces a negative earnings surprise that day, and a value of zero otherwise (*Earnings Surprise*), a variable that takes a value of one if the firm announces a positive earnings surprise that day and a value of zero otherwise (*Positive Earnings Surprise*), and a variable that takes a value of negative one if the firm announces a negative earnings surprise that day and a value of zero otherwise (*Negative Earnings Surprise*). The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day and their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*). The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t-1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t-5,t-2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t-60,\ t-6}$) as well as dummy variables for firm and year-month-day. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and day. Panel A reports results from regressions on *Earnings Surprise*. Panel B reports results from regressions on *Positive Earnings Surprise*. Panel C reports results from regressions on *Negative Earnings Surprise*. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel A: Regressions on *Earnings Surprise*					
High Short Interest	0.004***	0.004**	0.004***	0.003**	0.003**
	(2.80)	(2.58)	(2.82)	(2.10)	(2.36)
ln(1+*Social Media Activity*)	0.012***	0.007***	0.012***	0.002	0.011***
	(5.05)	(6.04)	(4.82)	(1.08)	(2.75)
ln(1+*Social Media Activity*)**High Short Interest*	-0.010***	-0.005**	-0.011***	-0.003	-0.011**
	(-2.79)	(-2.25)	(-2.71)	(-1.39)	(-2.05)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.011	0.011	0.011	0.010	0.010
#Observations	463999	463999	463999	463999	463999
Panel B: Regressions on *Positive Earnings Surprise*					
High Short Interest	0.003**	0.002*	0.003**	0.001	0.002*
	(2.25)	(1.90)	(2.30)	(1.39)	(1.79)
ln(1+*Social Media Activity*)	0.017***	0.010***	0.019***	0.005***	0.016***
	(7.09)	(8.61)	(6.74)	(3.08)	(3.94)
ln(1+*Social Media Activity*)**High Short Interest*	-0.012***	-0.006***	-0.014***	-0.004***	-0.015***
	(-5.58)	(-4.09)	(-5.55)	(-3.52)	(-3.93)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.030	0.031	0.030	0.029	0.029
#Observations	463999	463999	463999	463999	463999

	Social media activity				
	#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
Independent variable	(1)	(2)	(3)	(4)	(5)
Panel C: Regressions on *Negative Earnings Surprise*					
High Short Interest	0.001*	0.002**	0.001*	0.002**	0.002**
	(1.84)	(2.04)	(1.71)	(2.38)	(2.31)
ln(1+*Social Media Activity*)	-0.005***	-0.003***	-0.006***	-0.003***	-0.005***
	(-5.82)	(-6.50)	(-5.85)	(-3.67)	(-3.56)
ln(1+*Social Media Activity*)**High Short Interest*	0.003	0.001	0.003	0.001	0.004
	(1.31)	(1.02)	(1.56)	(1.12)	(1.27)
Firm Controls	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.017	0.017	0.017	0.017	0.017
#Observations	463999	463999	463999	463999	463999

Internet Appendix:
Did the Game Stop for Hedge Funds?

Table IA1: **Social media activity and stock returns, robustness**

This table reports results from multivariate regressions on daily stock returns and risk-adjusted returns. The dependent variables include daily stock returns and Fama and French (2015) five-factor adjusted returns. The primary independent variables of interest are the natural logarithm of one plus measures of social media activity on WallStreetBets in the previous day and their interactions with an indicator variable for the top one percentile of stocks based on short interest the previous day (*High Short Interest*). The measures of social media activity (*Social Media Activity*) include the number of posts (*#Posts*), the number of comments to posts (*#Comments*), and the number of unique posters writing posts (*#Posters*), the number of emojis used in posts (*#Emojis*) and the number of meme stock lingos used in posts (*#Meme Lingos*) that reference the stock. The other independent variables include Dow Jones Newswire sentiment that day (*Sentiment*), an indicator variable for a Dow Jones Newswire that mentions the firm that day (*Dow Jones*), number of analysts upgrading the stock that day (*Analyst Upgrades*), prior day's stock return ($Returns_{t-1}$), cumulative stock return from $t-5$ to $t-2$ ($Returns_{t-5,t-2}$), cumulative stock return from $t-60$ to $t-6$ ($Returns_{t-60,\ t-6}$) as well as dummy variables for firm and year-month-day. The coefficient estimates for these firm control variables are omitted for brevity. The t-statistics, in parentheses, are derived from robust standard errors that are clustered by firm and day. The sample period is from January 2020 to March 2022. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Regressions on daily stock returns					Regressions on five-factor adjusted returns				
Interaction between ln(1+*Social Media Activity*) and *High Short Interest*									
Social media activity									
#Posts	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*	*#Posts*	*#Comments*	*#Posters*	*#Emojis*	*#Meme Lingos*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Excluding GME									
0.491**	0.176	0.507*	0.509**	1.055**	0.548**	0.243**	0.576**	0.511*	1.100**
(2.05)	(1.49)	(1.89)	(1.98)	(2.15)	(2.16)	(2.10)	(2.04)	(1.87)	(2.14)
Panel B: Excluding AMC									
0.818**	0.413**	0.870***	0.974***	1.355***	0.824***	0.444***	0.879***	0.945***	1.358***
(2.56)	(2.38)	(2.61)	(3.14)	(3.74)	(2.99)	(3.21)	(3.12)	(3.60)	(4.34)
Panel C: Stocks referenced in titles of posts									
0.750***	0.382***	0.795***	0.909***	1.485***	0.721***	0.385***	0.757***	0.841***	1.348***
(4.11)	(3.67)	(4.26)	(4.78)	(6.47)	(4.32)	(3.80)	(4.74)	(5.44)	(8.74)
Panel D: *High Short Interest* = top two percentile of stocks based on short interest									
0.920**	0.426*	1.003**	1.187**	1.623***	0.848**	0.431**	0.914**	1.054**	1.425***
(2.17)	(1.93)	(2.19)	(2.54)	(2.80)	(2.23)	(2.19)	(2.24)	(2.51)	(2.77)
Panel E: Excluding WallStreetBets posts with due diligence reports									
0.747***	0.383***	0.793***	0.861***	1.236***	0.701***	0.389***	0.738***	0.795***	1.129***
(4.08)	(3.59)	(4.24)	(5.50)	(5.78)	(4.33)	(3.98)	(4.79)	(6.95)	(7.64)