Place your bets? The market consequences of investment research on Reddit's Wallstreetbets[*]

Daniel Bradley[a], Jan Hanousek Jr.[b, c], Russell Jame[d], and Zicheng Xiao[e]

[a]Department of Finance, University of South Florida, Tampa, FL 33620, danbradley@usf.edu

[b]Department of Finance, University of South Florida, Tampa, FL 33620, janhanousek@usf.edu

[c]Faculty of Business and Economics, Department of Finance, Mendel University in Brno, Czech Republic

[d]Gatton College of Business, University of Kentucky, Lexington, KY 40515, 859.218.1793, russell.jame@uky.edu

[e]Department of Finance, University of South Florida, Tampa, FL 33620, zicheng@usf.edu

March 2022

Abstract

We examine the market consequences of due diligence (DD) reports on *Reddit's Wallstreetbets* (*WSB*) platform. Over the 2018-2020 sample, we find that DD recommendations are significant predictors of one-month ahead returns, earnings forecast revisions, and earnings surprises, and user comments are incrementally informative. However, these benefits fully reverse in the first half of 2021. During this period, the fraction of reports emphasizing price-pressure strategies (rather than fundamentals) quadruples, and the decline in informativeness is concentrated in "price-pressure" reports. Our findings are consistent with the successful GME short squeeze causing WSB users to place too much emphasis on coordinated trading strategies.

Keywords: Reddit, Wallstreetbets, *WSB*, retail trading, social media, Gamestop

JEL classifications: G20, G23

Place your bets? The market consequences of investment research on Reddit's Wallstreetbets

1. Introduction

On February 18, 2021 the CEOs of Reddit and Robinhood along with a Reddit user testified before Congress for their role in the well-publicized Gamestop (GME) short squeeze that sent shares to almost $500 before plummeting to around $50 a few days later.[1] With the explosion of social media platforms devoted to investment research in recent years, it is not surprising that regulators expressed concerns about the impact of social media on stock market efficiency and retail investor welfare.[2] Existing academic research on the topic is mixed. For example, some studies find that social media contains value-relevant information (Chen, Du, Hu, and Hwang, 2014) and results in more informative retail trading (Farrell, Green, Jame and Markov, 2021), but other recent work suggest that social media can induce cognitive biases that harm investors and impede price discovery (Cookson, Engelberg, and Mullins, 2022; Jia, Redigolo, Shu, and Zhao, 2020).

In this paper, we focus on the investment research provided on Reddit's *Wallstreetbets (WSB)*, the forum targeted in the recent Congressional probe. *WSB* is a forum (called a subreddit) where users post investment analysis and the community comments on the idea. It is by far the most popular finance-related subreddit experiencing explosive growth with currently over 10 million subscribers (10x increase year-over-year). However, several unique features of *WSB* suggest that it may be less informative than many other prominent social media sites that have been studied more extensively in the academic literature (e.g., Seeking Alpha). For example, in contrast to Seeking Alpha, *WSB* users are easily able to post and comment anonymously and there is virtually no editorial review. Posts also tend to focus on highly speculative strategies that emphasize small probabilities of large gains, possibly at the expense of lower expected returns.[3] Finally, many *WSB* users were attracted to the site based on the extraordinary success of the Gamestop short squeeze. Given investors tendency to extrapolate from salient recent events, new users may flood the site with research reports that emphasize strategies unrelated to firm-fundamentals (e.g., short-squeezes). To the extent that users overestimate the

---

[1] Representatives from hedge funds Melvin Capital and Citadel also testified. For the transcript of the testimony, see https://www.c-span.org/video/?508545-2/GameStop-hearing-part-2

[2] While the GameStop event may be the most publicized recent example, policy makers have long been concerned about social media and financial markets. For instance, see https://www.nytimes.com/2013/04/29/business/media/social-medias-effects-on-markets-concern-regulators.html.

[3] For example, strategies may emphasize stocks with high volatility or skewness which have been shown to earn lower expected returns (e.g., Ang, Hodrick, Xing, and Zhang, 2006; and Bali, Cakici, and Whitelaw, 2011)

effectiveness of such strategies, *WSB* research in the post-GME period may be particularly uninformative.[4]

The purpose of our paper is to offer a first look at the determinants and consequences of the investment research provided on *WSB* in the periods both before and after the Gamestop event. We focus primarily on single firm 'Due Diligence" (DD) reports, which are reports identified by the poster (and verified by moderators) as containing some type of analysis and a clear buy or sell signal.[5] Our sample includes 5,050 DD reports issued between July 2018 and June 2021, of which roughly half are issued in the post-GME period (January-June of 2021).

Consistent with the view that *WSB* research emphasizes high-risk investments, we find that DD reports tilt towards young, volatile stocks with high skewness, and high short interest. *WSB* preference towards speculative investments also increases substantially in 2021. For example, *WSB* tendency to cover more volatile stocks nearly triples in 2021, while *WSB* coverage of stocks with heavy short interest increases by more than 500%. The time-series patterns are consistent with the Gamestop event attracting even more risk-seeking users.

The informativeness of DD reports also exhibits considerable time-series variation. In the pre-GME period, we find that DD reports are significant predictors of future returns. For example, an incremental DD buy recommendation is associated with a 6.04% increase in one-month ahead returns for the full sample and a 2.32% increase after excluding GME and AMC.[6] However, the one-month return predictability fully reverses in the post-GME period. As a benchmark, we also consider the informativeness of Seeking Alpha (SA) research reports over the sample period. Consistent with prior work (e.g., Chen et al., 2014), we find that SA research report recommendations predict future returns. However, we find no evidence that SA research reports exhibit any decline in informativeness in the post-GME period, which alleviates the concern that broad economic forces resulted in a deterioration in investment research across all social media platforms.

---

[4] Prior work suggests that strategies based on extrapolative expectations often earn lower expected returns (see, e.g., Greenwood and Shleifer (2014), Barberis, Greenwood, Jin, and Shleifer (2015), Cassella and Gulen (2018), and Da, Huang, and Jin (2021)).

[5] We focus on DD reports because they contain clear investment recommendations and are most comparable to other forms of social-media investment research studied in the past (e.g., Seeking Alpha reports). We consider other *WSB* posts in supplemental tests (see Section 3.5.1).

[6] AMC, along with GME, are widely publicized as two of the original meme stocks targeted by *WSB* users (https://www.cnbc.com/2021/07/14/amc-share-price-cut-in-half-as-reality-sets-in-for-meme-stock-investors.html). Consistent with this view, these two stocks represent close to 25% of our 2021 sample (top 2 in our data). We have also considered excluding all meme stocks, defined as the 50 stocks for which Robinhood imposed a trading halt (https://www.theverge.com/2021/1/29/22256419/robinhood-limits-wall-street-bets-stock-buys). Excluding other meme stocks apart from GME and AMC has a negligible impact on the results (see Panel B of Table 4).

WSB research reports could provide value by skillfully interpreting major news events (i.e., information processing) or by producing new information. To distinguish between these channels, we partition DD reports into those that coincide with abnormal media coverage, earnings announcements, or another DD report (*confounded reports*) versus all other reports (*non-confounded*). We find that *WSB* reports are similarly informative for both types of reports in the pre-GME period, suggesting that the informativeness of *WSB* research is not limited to reports that piggyback on major information events.[7]

A natural concern is that the predictive ability of *WSB* research is not driven by reports containing value-relevant information (*information*), but rather because DD reports induce uninformed demand shocks that push prices beyond fundamentals (*price pressure*). Several pieces of evidence favor the *information* hypothesis. First, we find no evidence that the returns patterns reverse over longer-horizons (up to 12-weeks following the report), which is inconsistent with the joint hypothesis of price pressure and short-term downward sloping demand curves.[8] Second, if *WSB* posts either always incite demand shocks or simply reflect small investor sentiment, then we might expect that even uninformative *WSB* posts predict returns. However, we find no evidence that non-research related *WSB* posts, i.e., posts that likely contain no useful information such as posting a meme or bragging about recent trading gains or losses in a stock, are predictive of one-month ahead returns in the pre-GME period. Finally, we find DD reports issued during the pre-GME period positively forecast cash-flow news, as measured by either media sentiment, earnings surprises, or analyst earnings forecast revisions. However, these effects again fully reverse to zero (and sometimes become significantly negatively) in the post-GME period.

The decline in report informativeness starting in 2021 is consistent with the GME event contributing to this decline. The remarkable success of the GME short squeeze may have caused users to place too much emphasis on coordinated trading strategies, possibly at the expense of analyzing firm fundamentals. To test this prediction, we develop a dictionary of words that measure the reports' emphasis on short squeezes or other forms of price pressure (e.g., short interest, squeeze, gamma, hedge, etc.) versus its emphasis on fundamentals (e.g., earnings, revenue, growth rate, store visits, etc.).

---

[7] We also emphasize that in all of our test, we skip at least one full day between the issuance of the DD report and the future returns. For example, for a report issued on a Monday, future returns are measured assuming that investor purchases at the close of trading on Tuesday. This conservative approach ensures that the returns associated with following a DD report are not contaminated by the immediate market reaction to any news that was released on the same day of the report.
[8] While testing for subsequent reversals is a commonly used to distinguish information from price pressure (see, e.g., Harris and Gurel, 1986), we acknowledge that recent events, most notably the persistently elevated price of GameStop and other meme stocks, suggests that price pressure can persist over longer horizons.

We classify a report as price-pressure focused (*PP Report*) if the number of price pressure words exceeds the number of fundamental words. We find that the fraction of *PP Reports* increases from 8% in the pre-GME period to 31% in the post-GME period. The differences are highly significant and are robust to excluding GME and AMC from the sample. Moreover, consistent with price pressure strategies contributing to the deterioration in report informativeness, we find that the decline in return predictability of DD reports in the post-GME period is significantly stronger among *PP Reports*.

Our final set of tests examine whether consumers of *WSB* research are able to discern report quality. We contrast DD reports where the comments following the DD report are consistent with the report itself (e.g., bullish comments following a purchase recommendation) with DD reports where commenters disagree with the content of the report. We find that high comment agreement is associated with significantly greater report informativeness in the pre-GME period, but not the post-GME period. Thus, both report informativeness and the quality of user comments declines in the more recent sample period.

As a broader, albeit noisier, test of whether users are able to discern report quality, we also contrast DD reports where the order imbalances following the DD report is consistent with the report itself (e.g., net buying following a purchase recommendation) with DD reports where trading is inconsistent with the content of the report. We separately examine three groups of traders with varying levels of sophistication (and presumably reliable on *WSB*): institutional investors, large retail investors (as proxied by volume-based measures of retail order imbalance), and small retail investors (as proxied by trade-based measures of retail order imbalance). We find that DD report informativeness is significantly greater in the pre-GME period when small retail traders are trading in the same direction of the report., but this effect is fully eliminated in the post-GME period. This finding echoes the previous evidence that *WSB* comments contain incremental investment value, but only in the pre-GME period. We find no evidence that order imbalances of large retail investors or institutional investors contain any incremental predictive power in either period, which is consistent with *WSB* primarily influencing small retail trader order imbalances.[9]

Our study adds to the literature that examines the value of investment research provided on social media. While some papers find a significant positive relation between investment opinions on social finance sites and future stock returns (e.g., Chen et al., 2014; Jame, Johnston, Markov, and

---

[9] Consistent with this view, we find that institutional order imbalances are uncorrelated with *WSB* investment recommendations, large retail investor order imbalances exhibit a weak correlation, and small retail traders exhibit a substantial correlation.

Wolfe, 2016; Crawford, Gray, Johnson, and Price, 2018; Bartov, Faurel, and Mohanram, 2018), others do not (e.g.., Tumarkin and Whitelaw, 2001; Kim and Kim, 2014; Giannini, Irvine, and Shu, 2018; Ammann and Schaub, 2021). We offer a first look at the investment value of DD reports on *WSB*, which has recently become the most influential social finance site by many metrics and offers many unique features relative to most existing social media sites. We find that *WSB* research is significantly informative prior to the GME event, which highlights that social media can be informative even in settings with complete user anonymity and almost no editorial board oversight. At the same time, the dramatic shift in strategy and the significant decline in informativeness in the post-GME period suggests that this structure can also promote excessive trend-chasing that can have detrimental effects on the usefulness of the site. We also emphasize that the patterns we uncover for *WSB* research and *SA* research often differ substantially. Our finding of significant differences across social media sites echoes recent literature that uncovers significant heterogeneity in the trading behavior across different types of retail investors (e.g., Jones, Shi, Zhang, Zhang, 2020; Eaton, Green, Roseman, and Wu, 2021).

We also contribute to the literature on retail trading. Early work finds that retail traders are uninformed 'dumb money' (e.g., Hvidkjaer, 2008; Frazzini and Lamont, 2008, Barber, Odean, and Zhu, 2009). However, more recent evidence suggests that retail traders are informed (e.g., Kaniel, Liu, Saar, and Titman, 2012; Kelley and Tetlock, 2013; Boehmer, Jones, Zhang, and Zhang, 2021) and are skilled at processing public information (Farrell et al., 2022). Our evidence that WSB research is more informative when WSB commenters and small traders agree with the report recommendation is consistent with the more recent literature. However, the reversal of this relation in the post-GME period is consistent with attention-grabbing events, such as the GME short squeeze, attracting less sophisticated investors (Da, Engelberg, and Gao 2011; Barber, Xuang, Odean, and Schwarz, 2021).

Finally, our study contributes to the nascent literature that explores the growing importance of *WSB* and its impact on financial markets. Several contemporaneous papers focus on the dynamics between *WSB* activity and one-day ahead returns, trading volume, short interest, volatility, and market quality (e.g., Aharon, Kizy, Umar, and Zaremba, 2021; Semenova and Winker, 2021; Hu, Jones, Zhang, and Zhang, 2021; Long, Lucey, and Yarovaya, 2021; Eaton, Green, Roseman, and Wu, 2021; and Allen, Haas, Nowak, Pirovano, and Tengulov, 2021). In contrast, our emphasis is on the informativeness of *WSB* research reports (i.e., due diligence reports), how the informativeness changes following the GME event, and whether users of *WSB* research are adept at discerning report quality.

2. Data and Descriptive Statistics

2.1 Reddit and Wallstreetbets – Background

Reddit is a social media platform founded in June 2005. Like many other social media websites, contributors post content, and users can add comments in response to the original post. The Reddit community is a collection of forums, where each forum is dedicated to a particular topic called a subreddit. Each subreddit is then organized into several pages based on users' ranking criteria. For instance, the default page is the 'Hot Page,' which lists the currently most viewed posts or posts with the most active commentators. 'New Posts' lists posts based on the listing timestamp, and 'Top posts' lists the posts with the most likes (upvotes) and comments for a specified period. When a new post is written, it is only visible in the new post category. The post can then move up to the hot page if it reaches sufficient traffic.

Wallstreetbets (*WSB*) is one of many subreddits within the Reddit community. It was created on January 31, 2012, with a particular emphasis on highly speculative trading strategies. While this is not the only subreddit dedicated to investing strategies (i.e., r/Investing, r/Personalfinance, r/Stocks, etc.), we focus on this particular subreddit for two primary reasons. First, with over 10 million subscribers, it is by far larger than other finance-related subreddits. Second, it is the subreddit that has recently attracted significant media and regulatory attention from its role in the GameStop short squeeze. The conventional view is that this forums' userbase is predominantly unsophisticated retail investors who are more interested in gambling than investing. There has also been significant concern that the "research" on *WSB* is at best uninformative, and at worse, a force that destabilizes stock prices and contributes to significant retail trading losses.

2.2. A Comparison of WSB to Seeking Alpha (SA)

WSB shares important similarities with other social finance platforms such as Seeking Alpha (*SA*). Both sites allow non-professional investors to share their investment research, and both sites allow for readers to provide comments on the report and engage in discussions with other users. Prior work on *SA* suggest that a large fraction of *SA* reports and contributors are skilled (see, e.g., Chen et al., 2014; Farrell, Jame, and Qiu, 2020). However, there are several prominent differences between *WSB* and *SA* that suggests that prior work on *SA* may not apply to *WSB*.

First, while *SA* employs an editorial team to review all research reports to ensure quality there is very limited quality control on *WSB*. Relatedly, *WSB* allows users greater anonymity than *SA*.[10] Greater anonymity reduces the incentives to develop a strong reputation and potentially allows users with more nefarious motives (e.g., pump and dump schemes) to switch identities without accountability. *WSB* reports also tend to be considerably less in-depth than the average *SA* report, and the userbase of *WSB* is likely to have significantly less financial sophistication.[11] Anecdotal evidence suggests that *WSB* also places a larger emphasis on highly speculative trading strategies. As a result, investment research on the site may gravitate towards strategies that tend to earn lower expected returns such as buying stocks with high volatility (Ang, Hodrick, Xing, and Zhang, 2006) or stocks with lottery-like features (Bali, Cakici, and Whitelaw, 2011).

Finally, in contrast to Seeking Alpha which has steadily grown over the past 15 years, much of *WSB* growth is attributable to the GME short squeeze event. For example, Figure 1 shows the forum grew from roughly 500,000 users in July of 2018 to roughly 10.7 million users as of June 2021, with a clear spike during the GameStop short squeeze in January of 2021. One concern is the dramatic increase in new users, most of whom were attracted to the site by the extraordinary price increases in Gamestop, can have a profound shift on the culture of the site. For example, given investors tendency to forecast expected returns based on recent performance (e.g., Greenwood and Shleifer, 2014), new users may overestimate the effectiveness of strategies that are unrelated to fundamentals, such as price-pressure induced short squeezes. Indeed, ample anecdotal evidence suggest that new members of *WSB* tend to emphasize coordinated buy-and-hold strategies for a handful of meme stocks with little regard to the company's fundamentals.[12] As a result, many different spinoff subreddits were started by original *WSB* users to attempt to recreate *WSB* prior to the GME event. Given the significant shift in both the userbase and culture of *WSB* following the GME event, our analysis will separately examine *WSB* reports issued in the pre-GME period (July 2018-December 2020) and the post-GME period (January 2021- June 2021).

---

[10] Seeking Alpha allows users to contribute using pseudonyms, but they still require private disclosure of their identities to Seeking Alpha and they do not allow the same user to post under multiple pseudonyms (https://seekingalpha.com/page/policy_anonymous_contributors).

[11] With respect to article depth, we find that the average *WSB* report in our sample is 352 words, which is roughly half of the length of a typical SA report (675 words), as reported in Chen et al., 2014. With respect to investor sophistication, the average Seeking Alpha user has a household income of $321,000 and roughly $1.5 million in investable assets (see https://static.seekingalpha.com/uploads/pdf_income/sa_media_kit_04_2020.pdf). While these figures are unknown for *WSB* users, anecdotal evidence suggests that these estimates would be substantially smaller.

[12] For a summary of these competing views see: https://www.insider.com/wallstreetbets-reddit-forum-divided-as-new-users-flood-subreddit-2021-2

2.3 The WSB Sample

We scrape all posts on *WSB* from July of 2018 through June of 2021 using the Pushshift API, which collects new posts and comments in almost real-time.[13] Posts can be deleted by the original author, moderator of the subreddit, or an "automod" (which is a spam filter robot operated and constructed by moderators). Deletions by the automod typically occur in less than a minute. Deletions of posts by moderators take longer (usually up to a day) if the post breaks the rules of the subreddit and was not already captured by the automod. Lastly, a post can be deleted by the author at any time. Importantly, the API retains posts deleted by the authors, and these posts are included in our sample.

WSB contains more than 100,000 different posts spanning several different categories including: *News* (links to news stories WSB users found interesting), *Discussion* (open-ended discussions, frequently on macroeconomic forces such as proposed regulations, supply chain disruptions, etc.), *YOLO* (posts reporting large upcoming bets), *Gains/Losses* (posts highlighting major investment successes and failures), *Sh$tposts* (ironic investment theses that are meant to entertain rather than inform), and *Due Diligence* reports (posts that contain investment analysis and a clear investment recommendation).

Our analysis focuses on Due Diligence (DD) reports. These reports are vetted by the moderators as containing information where 1) at least some analysis has been performed and 2) the author provides a clear investment recommendation (long or short). We chose to focus on research-related posts because these posts are most similar to other forms of social-finance investment research (e.g., Seeking Alpha research reports) that have been studied in the prior literature, and they are the most likely to contain value-relevant information. In Section 3.5.1, we contrast DD reports with other non-research posts on *WSB*.

For each DD report, we manually review the report to identify the investment recommendation and ticker. Although the author's investment recommendation is clear to anyone reading the report, there is no standardized format for listing the recommendations which necessitates a manual review of each report. The manual review of tickers is also needed for two reasons. First, users may place special characters before or after a ticker symbol that a program would misclassify. Second, users sometimes intentionally report a wrong ticker to misdirect hedge funds and other institutional investors that monitor message boards using algorithms.[14]

---

[13] There is a period between April 13th and August 4th of 2020 where DD reports are missing. This is likely due to an issue with Reddit's API.

[14] For an example of WSB users attempting to mislead hedge funds, see:
https://www.reddit.com/r/wallstreetbets/comments/ly0d4m/how_to_beat_hedge_fund_algorithms_on_wsb_a/

We limit the sample to DD reports focused on a single ticker (e.g., we eliminate DD reports that focus on market-wide or industry analysis) and to common stocks (CRSP share codes 10 and 11) with available data in the CRSP-Compustat merged database. Appendix A provides an example of a DD report in our sample. A manual reading of the report indicates that the author is recommending a "Buy" for "BYND". The header of the report also includes the username and the timestamp of the report. For DD reports that occur outside of trading hours, we set the date of the report equal to the date on which an investor could have first traded on the report.[15]

For each report, we also collect all the comments that are posted in response to the report. We limit the sample to comments that are posted between the publication of the report and the start of the subsequent trading day. This filter helps eliminate comments that are potentially influenced by ex-post returns (i.e., comments praising the author for a buy recommendation that earned subsequently high returns) while still retaining the majority of comments. We also develop a measure to quantify the extent to which commenters agree with the DD recommendation. The language of *WSB* users is very different from typical financial market participants (e.g., greater use of sarcasm, slang, jokes, emojis, and in many cases vulgarity), making traditional measures of text analysis (e.g., Loughran and McDonald, 2011) not well suited for measuring *WSB* user sentiment. Instead, we develop our own data dictionary based on common *WSB* expressions which we describe in greater detail in Appendix B. Using this data dictionary, we define *Commenter Agreement* equal to one if the number of keywords in agreement with the DD report are greater than or equal to the number of keywords that disagree with the DD report, and zero otherwise.

When examining whether DD reports contain value-relevant information, one concern is that reports may simply repeat major information announcements (hereafter: confounded reports). We classify a report as *confounded* if the firm announced earnings or had abnormally high media coverage (as defined in Appendix D) on the day prior to the DD report or the day of the DD report. Since many users may also "piggyback" off other DD reports, we also classify a report as *confounded* if there was a DD report issued on the previous day.[16]

Panel A of Table 1 provides summary statistics. The sample includes 5,050 DD reports covering 3,811 firm days and 909 different firms. The overwhelming majority of DD reports (88%)

---

[15] For example, if a report was issued at 5 pm on Wednesday January 6, we would classify the date of the report as Thursday, January 7, and we would define the [1,5] day return as the return from Friday January 8 through Thursday January, 14. We exclude the Day [0] return to reduce the impact of potentially confounding news that could influence both the DD report and the Day [0] return.

[16] DD reports issued on the same day as an existing DD report are not problematic since the unit of observation in the main analysis is the firm-day.

are buy recommendations. The average report receives 65 comments between the time of the report and the start of the subsequent trading day. Roughly 30% of all DD reports are classified as *Confounded*. We also find the average contributor, as measured by the username on the DD report, issues only 1.32 reports during our sample period. These estimates likely significantly understate the number of reports per person since users often get temporary bans for violating moderator rules and circumvent the ban by joining the forum with a different username. Nevertheless, the lack of repeated posts by the same username suggests that *WSB* users are not especially concerned with developing and maintaining a reputation.

We also partition the sample into the pre-GME period (July 2018-December 2020) and the post-GME period (January 2021-June 2021). Although the post-GME period is substantially shorter in calendar time, it accounts for a slight majority (~54%) of all DD reports. DD reports in the post-GME period attract more comments (81 versus 45), they are more likely to be *Confounded* (35% versus 26%), and they are more likely to recommend a long position (95% versus 81%). The substantial differences in report characteristics in the pre- and post-GME period are consistent with the GME event resulting in a significant shift in the culture of *WSB*.

Given the extreme returns of GME and AMC, we will explore whether our central conclusions are robust to including/excluding GME and AMC. Panel B of Table 1 reports the summary statistics after excluding GME and AMC, while Panel C reports the results for the GME/AMC subsample. Reports on GME and AMC account for 12% of all reports (609/5,050), but they are far more prevalent in the post-GME period (19% of all reports) relative to the pre-GME period (3% of all reports). GME and AMC reports also garner significantly greater attention (154 comments versus 52 comments) and are more likely to be *Confounded*.

2.4 Other Variable Construction

We collect Seeking Alpha research reports over the same sample period (July 2018-June 2021). For each report, we collect the following information: a report ID assigned by Seeking Alpha, the date and time of the publication, the ticker (or tickers) assigned to each report, the author of the report, the number of comments issued within the first day of the report, and the authors stated recommendation (e.g., "bullish", "neutral" or "bearish"). To parallel the *WSB* sample, we convert SA recommendations to a "buy" indicator, which equals one if the recommendation is "very bullish" or "bullish" and zero otherwise. We exclude reports that do not have a stated recommendation, and we limit the sample to reports that are focused on single-ticker articles on US common stocks.

Panel D of Table 1 reports summary statistics for the Seeking Alpha sample. The sample of *SA* reports is much larger than DD reports (21,177 versus 5,050), and the SA sample spans a larger cross-section of firms. *SA* research also tends to be overwhelming bullish, with 85% of all reports being classified as buy recommendations. The average *SA* contributor writes 11.1 reports compared to only 1.3 reports for *WSB*. This finding is consistent with anecdotal evidence that contributors on SA are more interested in using the site to build a reputation.[17]

We combine the data on social media research from *WSB* and *SA* with several additional data sources. We obtain financial statement data, including book value of equity, book value of debt, book value of assets, short interest, and total common shareholders from Compustat. We obtain financial market data, including daily data on share price, shares outstanding, volume, and stock returns from CRSP. Earnings announcement dates and sell-side analyst earnings forecast data are from the I/B/E/S unadjusted US detail history file and sell-side analyst recommendations are from the I/B/E/S detail recommendations file. We collect the number of shares held by institutions from the Thomson Reuters Institutional Holdings database, and media coverage is collected from Bloomberg.

We identify retail trading from TAQ data using the approach of Boehmer, Jones, Zhang, and Zhang (BJZZ, 2021). Specifically, we classify trades with TAQ exchange code "D" and prices just below a round penny (fraction of a cent between 0.6 and one) as retail purchases, while trades with exchange code "D" and prices just above a round penny (fraction of a cent between zero and 0.4) are classified as retail sales. This classification is conservative in the sense that it has a low type 1 error (i.e., trades classified as retail are very likely to be retail). However, this classification omits retail trades that occur on exchanges as well as limit orders that are not immediately executable.

2.5 Determinants of WSB Coverage

In this section, we examine the determinants of the intensity of WSB coverage. We expect that many of the firm characteristics that influence research coverage on other social finance sites (e.g., Seeking Alpha) are likely to be relevant on *WSB* as well. However, relative to SA, we expect that *WSB* users will tend to issue reports on more speculative stocks, and we expect that such effects may be amplified in the post-GME period.

We examine the determinants of coverage by estimating the following panel regression:

$$Coverage_{it} = \alpha + \beta_1 Chars_{it-1} + \beta_2 Chars_{it-1} \times 2021 + Month_t + \varepsilon_{it.} \qquad (1)$$

---

[17] For example, many of the testimonials of Seeking Alpha contributors emphasize the important reputational benefits associated with being a regular contributor on Seeking Alpha (https://seekingalpha.com/page/testimonials).

The dependent variable, *Coverage*, is either equal to *WSB Coverage*, defined as the natural log of 1 plus the total number of DD reports issued for firm i during month t, *SA Coverage*, defined as the natural log of 1 plus the total number of *SA* reports issued for firm i during month t, or the difference between *WSB Coverage* and *SA Coverage* (*WSB – SA Coverage*). *Chars* contains the vector of firm characteristics used in Farrell et al (2021) to explain *SA* coverage, namely the percentage of the firm's shares held by institutional investors at the end of the prior year (*Inst. Ownership*), the number of common shareholders (*Breadth of Ownership*), market capitalization (*Size*), book to market (*BM*), return volatility (*Volatility*), share turnover (*Turnover*), past one-month returns (*Return$_{m-1}$*), past returns over the prior two to twelve months (*Ret$_{m-2, m-12}$*), the number of unique media articles mentioning the firm the prior year (*Media Coverage*), and the number of sell-side analysts issuing a forecast for the firm in the prior year (*IBES Coverage*). In addition, given the ample anecdotal evidence that *WSB* users target stocks with lottery like features, stocks with heavy short interest, and stocks that recently went public, we add indicator variables equal to one if the firm is in the top quintile of the maximum daily return in the previous month (*High Max*), the top quintile of short interest in the previous month (*Heavy Short*), or if the firm went public in the past six months *(Recent IPO)*. We allow the coefficient on all the firm characteristics to vary in the pre- and post-GME period by interacting the firm characteristics with *2021*, an indicator equal to one in the post-GME period (January – June of 2021) and zero otherwise. We include month fixed effects and cluster standard errors by firm and month. We log all continuous variables other than *Return*, and we standardize all continuous variables, including the dependent variables, to have zero mean and unit variance.

Specification 1 of Table 2 reports the results. Consistent with prior work on the determinants of *SA* coverage, we find that *WSB* coverage in the pre-GME period is increasing in firm size, turnover, volatility, and media coverage, and decreasing with institutional ownership. Consistent with anecdotal evidence, we confirm that *WSB* coverage is also significantly greater for lottery-like stocks (i.e., *High Max)*, stocks with high short interest, and stocks that recently went public. The coefficients on the post-GME interaction terms indicate that *WSB* coverage of speculative stocks, including stocks with higher volatility, higher max returns, higher shorter interest, and recent IPO stocks, are significantly greater in the post-GME sample. The magnitudes are economically large. For example, *WSB* coverage of stocks with high max returns increases by 800% in the post GME period (from 0.03 to 0.24), and their coverage of stocks with high short interest increases by roughly 650% (from 0.07 to 0.45). In Table IA.1 of the Internet Appendix, we repeat the analysis after excluding GME and AMC and

continue to find similar results. This dramatic increase in *WSB* coverage of stocks with greater volatility, skewness, and short interest is consistent with the extreme GME returns attracting even more speculative investors.

As a benchmark, we also examine the determinants of *SA* coverage over the same sample period. In the pre-GME period, we find that the determinants of *SA* coverage are typically similar to the determinants of *WSB* coverage. For example, we find that *SA* also has a strong preference for lottery-like stocks, stocks with heavy short-interest, and stocks that recently went public. However, in sharp contrast to *WSB*, *SA* coverage of speculative stocks did not dramatically increase in the post-GME period. Specification 3, which tests whether the determinants of *WSB Coverage* are significantly different from *SA Coverage*, confirms that the increase in the coverage of speculative stocks, as measured by either *Volatility*, *High Max, Heavy Short*, or *Recent IPO*, was significantly larger on *WSB* relative to *SA*. This conclusion is robust to excluding GME and AMC (see Table IA.1). The findings suggest that the GME event had a particularly pronounced impact on *WSB* coverage decisions.

3. The Informativeness of *WSB* Research

3.1 WSB Research and the Cross-Section of Stock Returns

In this section, we examine whether DD report recommendations forecast future stock returns. We estimate the following panel regression:

$$R_{it+1,t+x} = \beta_1 Net\,DD_{it} + \beta_1 Net\,DD \times 2021_{it} + Controls_{it} + Day_t + \varepsilon_{it.} \qquad (2)$$

The dependent variable is the stock return measured over the subsequent week (i.e., $x = 5$ trading days) or the subsequent month ($x = 21$ trading days). *Net DD* is the number of buy DD recommendations for stock i on day t less the number of sell DD recommendations for stock i on day t. We also include *Net DD× 2021,* which interacts *Net DD* with an indicator equal to one for the post-GME period (January-June 2021) and zero otherwise. Thus, *Net DD* captures the average predictive ability of DD reports over the July 2018 – December 2020 period, and *Net DD × 2021* captures the incremental predictive ability of DD reports in the post-GME period. Following Kelley and Tetlock (2013), the controls include *Size, Book-to-Market*, returns measured from days [0], [-5, -1], and [-26, -6] and media sentiment measured from [0], [-5, -1], and [-26, -6]. See Appendix D for detailed definitions.

Day denotes calendar-day fixed effects. To account for the overlapping holding periods, we cluster standard errors by both firm and month.[18]

Specifications 1 and 2 of Table 3 report the results for one-week and one-month holding periods for the full sample, and Specifications 3 and 4 report analogous results after excluding GME and AMC. Across all four specifications, the coefficient on *Net DD* is positive and at least marginally significant ($p < 0.10$). The economic magnitudes are also sizeable. For example, after excluding the extreme returns of GME and AMC, an incremental buy DD report issued over the 2018-2020 period is associated with a 0.92% increase in one-week ahead returns and a 2.32% increase in one-month ahead returns. While this finding is consistent with prior work that suggests that crowdsourced investment research can be informative, it is perhaps surprising that this relation continues to hold in a setting with complete user anonymity, minimal board oversight, and limited reputational incentives.

In contrast, the coefficient on *Net DD × 2021* is significantly negative at the one-month holding period. In particular, in the full sample, DD reports predictive ability declines by 5.21% (to 0.83%), and in the sample that excludes GME and AMC the estimate declines by -3.83% (to -1.51%). Both the 0.83% and the -1.51% estimate are not significantly different from zero, suggesting that DD reports in the post-GME sample are uninformative.

3.2 WSB Research and the Cross-Section of Stock Returns – Robustness

In Table 4, we examine whether the findings reported in Table 3 are robust to different research design choices. For reference, Panel A of Table 4 tabulates our baseline results from Table 3 for the full sample of stocks (Columns 1 and 2) and for the sample that excludes GME and AMC (specifications 3 and 4).

Hu et al. (2021) find that *WSB* is particularly influential among "Robinhood 50" stocks, defined as the 50 stocks that Robinhood imposed trading restriction on beginning on January 28th, 2021 and ending February 5th, 2021. While we already consider specifications that exclude the two most prominent stocks on the list (GME and AMC), in Panel B we exclude the remaining 48 stocks from the sample. We find that the main results are similar which alleviates the concern that the results are driven by a small subset of meme stocks.

---

[18] A similar approach that avoids overlapping holding periods is to examine daily returns and measure *Net DD posts* over different horizons (e.g., the previous five or 21 trading days). The results of this analysis, reported in Table IA.2, are qualitatively similar.

In Panel C, we augment the model by interacting the 2021 indicator with all the control variables. The results are again similar. Thus, while *WSB* users did significantly shift the types of firms they covered in the post-GME period (see Table 2), this shift does not account for the differences in report informativeness.

The pre-GME sample spans from July 2018 through December 2020 and thus includes both the pre- and post-COVID era. Several papers document that retail trading dramatically increased during the pandemic (e.g., Odzik, Sadka, and Shen, 2021) which raises the concern that pandemic induced trading, rather than the GME-event, is driving the deterioration in performance in the post-GME period. To explore this possibility, in Panel D we repeat the analysis after excluding the pre-pandemic period (July 2018-December 2019). The results, if anything, are slightly stronger, which is inconsistent with the decline in WSB report informativeness being driven by the pandemic. To better understand the time-series dynamics of the decline in the predictive ability of DD reports, in Figures 2A and 2B, we estimate Specifications 2 and 4 of Table 3 for each quarter over the 2020-2021 sample period and for the pre-2020 sample. We combine the pre-2020 sample because there are relatively small number of DD reports (606) prior to 2020. After excluding GME and AMC (Figure 3B), we see that the predictive ability of DD reports was stable across all four quarters of 2020, with point estimates ranging from 2.08% to 3.68%.[19] We also observe a sharp decline in informativeness in Q1 of 2021 (-0.71%), which further deteriorated in Q2 of 2021 (-1.42%).

For the sample of firm-days where *Net DD* is not equal to zero, the majority (~75%) of firm-days have a *Net DD* equal to one (e.g., one buy recommendation). However, roughly 13% of firm-days have multiple buy recommendations (i.e., *Net DD* >1), and 12% of all firm-days have a net sell recommendation (i.e., *Net DD* <0). To examine whether either multiple buy recommendations or sell recommendations contain incremental information, in Panel E we report the results after replacing *Net DD* with three separate variables: *Heavy Buy,* an indicator equal to one if *Net DD* is greater than or equal to 2, *Light Buy,* an indicator equal to one if *Net DD* is equal to one; and *Sell*, an indicator equal to one if *Net DD* is negative. The point estimates for *Heavy Buy* are always larger than the estimates on *Light Buy*, which is consistent with multiple buy recommendations being a stronger signal than a single buy recommendation in the pre-GME period. Similarly, over the one-month holding period, the reversals in the post-GME period are larger for *Heavy Buys* than *Light Buys*. However, due to the relatively small sample size of *Heavy Buy* recommendations, the estimates on *Heavy Buy* and *Light Buy,*

---

[19] Including GME and AMC (Figure 3A) generally yields similar results except that the return predictability in Q4 of 2020 is substantially larger (12.64%) due to the very large returns of GME in January of 2021.

in both the pre- and post-GME period, are not statistically different from each other (untabulated). We find no evidence that the relatively small sample of *Sell* recommendations are informative in the pre-GME period, nor do we find any evidence of reversals in the post-GME period. In fact, after excluding GME and AMC there is weak evidence that sell-recommendations become more informative. This is consequence of sell recommendations often being issued in response to a previous buy recommendation, which as we have shown, performed poorly in the post-GME period.

Finally, we consider the relation between *Net DD* and stock returns over longer horizons. We estimate Equation (2) for horizons ranging from one-week (i.e., $x = 5$) through 12 weeks (i.e., $x = 60$). Figures 3A and 3B report the results for the full sample and the sample that excludes GME and AMC. For both samples, we see that the predictive ability of *WSB* reports in the pre-GME period does not reverse over longer horizons. The absence of a return reversal is inconsistent with short-term price-pressure being the primary driver of the informativeness of WSB research in the pre-GME period (additional price pressure tests in Section 3.5 reinforce this finding). In addition, the decline in the predictive ability of DD reports in the post-GME period remains sizeable over longer horizons. For example, at the end of 12 weeks, the coefficient on *Net DD × 2021* is -5.94% for the full sample and -5.00% for the sample that excludes GME and AMC.

3.3 SA Research and the Cross-Section of Stock Returns

The decline in return predictability beginning in Q1 of 2021 points to the possibility that the rapid growth of the *WSB* userbase following the GME short squeeze contributed to the deterioration in the quality of *WSB* research. An alternative view is that the macroeconomic environment in 2021, for whatever reason, made it more difficult to conduct high-quality investment research. To explore this possibility, we repeat the analysis in Table 3 for *SA* research. Specifically, we modify equation (2) by adding two additional variables: *Net SA*, defined as the number of SA reports issuing a buy recommendation for stock *i* on day *t* less the number of *SA* reports issuing a sell recommendation for stock *i* on day *t*, and *Net SA × 2021*, which interacts *Net SA* with an indicator equal to one for the post-GME period (January-June 2021) and zero otherwise.

The results are reported in Table 5. Four findings are noteworthy. First, the coefficient on *Net DD* and *Net DD × 2021* are very similar to Table 3, indicating that the predictive ability of *WSB* research is not subsumed by *SA* research. Second, the coefficient on the *Net SA* is positive and significant. This finding is consistent with prior work suggesting that *SA* research is informative (e.g., Chen et al., 2014; Dim 2021). Third, the estimates on *Net SA* are economically smaller than the

estimates on *Net DD*, although the differences in the coefficients are generally not reliably different from each other. Nevertheless, at a minimum, the findings suggest that *WSB* research was at least as informative as *SA* research in the pre-GME period. Lastly, we find no evidence that the informativeness of *SA* research declined in the post-GME period. This finding is inconsistent with the view that broad macroeconomic forces contributed to widespread decline in the informativeness of investment research across all social media platforms.

3.4 WSB Research and the Cross-Section of Stock Returns – Confounding Information Events

In this section, we explore the economic channel underlying the investment value of *WSB* reports in the pre-GME period. If the return predictability following DD reports is primarily a consequence of DD reports piggybacking off of other news events (e.g., Altinkilic and Hansen, 2009) or skillfully interpreting public news (e.g., Engelberg, Reed, and Ringgenberg, 2012), then we would expect the results to be significantly stronger for reports that coincide with major information events (i.e., *Confounded Reports*). On the other hand, if *WSB* users are independently producing novel information, then the return predictability results may be stronger for reports not issued during major news events (*Non-Confounded Sample*). While both channels are potentially valuable to users who rely on *WSB* for investment research, distinguishing these explanations provides insight into the source of *WSB* investment value in the pre-GME period.[20]

Table 6 reports the results from Equation (2) after partitioning *Net DD* into *Net DD Processing*, defined as the number of confounded buy DD recommendations for stock *i* on day *t* less the number of confounded sell DD recommendations for stock *i* on day *t*, and *Net DD Production*, defined as the number of non-confounded buy DD recommendations for stock *i* on day t less the number of non-confounded sell DD recommendations for stock *i* on day *t*. In the pre-GME period, we find that the coefficients on both *Net DD Processing* and *Net DD Production* are always positive and the estimates are both significantly different from zero for the one-month horizon. The evidence suggests that both information processing and information production contribute to the predictive ability of *WSB* in the pre-GME period. We also find that both components reverse in the post-GME period, and there is some evidence that the reversal is larger for information processing reports.

---

[20] As emphasized in footnote 14, our analysis excludes the Day [0] return of the DD report. Thus, while *WSB* users may piggyback off other information events, if markets efficiently incorporate major news announcements into prices on the day of its release, piggybacking should not be associated with abnormal returns. If the market does not immediately incorporate this information, then *WSB* reports that bring this news to investors' attention are still providing value to investors.

3.5 WSB Research and Returns: Information versus Price Pressure

The results from the prior section are consistent with DD reports issued in the pre-GME period containing value-relevant information that is subsequently impounded into prices over the subsequent month (*information*). However, an alternative view is that DD reports cause (or are correlated with) uninformed demand shocks that induce significant price pressure over the subsequent month (*price pressure*). The lack of reversal over the 12-week holding period is inconsistent with the temporary price pressure explanation, but it is still possible that *WSB* induces price pressure that persists for even longer holding periods. In this section, we consider two complementary tests aimed at testing the price-pressure explanation.

3.51 WSB Non-Research Posts

As a first test, we examine non-research related *WSB* posts. If *WSB* posts either always incite demand shocks (e.g., because naïve investors simply purchase stocks that catch their attention) or if *WSB* posts are correlated with naïve investor demand (e.g., investors tend to post on companies that they are most excited about) then we might expect that even uninformative *WSB* posts predict returns. In contrast, if DD reports predict returns only because they contain new value-relevant information, then *WSB* posts that clearly contain no new information should not predict returns.

We define a *WSB* post as non-research related if it belongs to one of the following categories: *News*, *Losses*, *Gains*, *Charts*, and *Shi$posts*. We focus on these posts categories, because they are not designed to contain any new value-relevant information.[21] For example, *News* include links to articles and does not provide any analysis or interpretation of the news, *Gains* and *Losses* report previous successful and unsuccessful investments, *Charts* are typically graphs of past returns that could be found on any website, and *Shi$posts* are joke investment recommendations that are not intended to be taken seriously.

Since the price pressure explanation is likely to be most relevant for salient postings, we limit the sample to posting that explicitly mention the ticker in the title of the tread. We capture all potential tickers that either have dollar sign before the ticker (e.g., $GME) or that are in parentheses (GME).

---

[21] We exclude *YOLO* and *Discussion* posts. Although *YOLO* posts, which simply detail large bets, do not provide any investment research, if *WSB* users are informed, then their trades could still provide value-relevant signals. Similarly, although *Discussion* posts do not provide a clear investment recommendation, they could provide useful contextual information. In the Internet Appendix (Table IA.3), we find that classifying *YOLO* and *Discussion* posts as non-research related posts yields similar results.

After capturing all potential tickers, we manually review the post and remove tickers that are obviously fake (e.g., $*WSB*, $CUCK, $ROPE, etc.) or do not correspond to a US common stock. Our final sample includes 3,049 firm-days with at least one non-research related posting.

To explore this impact of non-research related *WSB* posts, we repeat Equation 2 after adding two additional variables: *Non-Research Posts* and *Non-Research Posts × 2021*, where *Non-Research Posts* is the number of non-research related postings on *WSB* for stock i on day t and *Non-Research Posts × 2021* interacts *Non-Research Posts* with an indicator equal to one for the post-GME period (January-June 2021) and zero otherwise. Table 7 reports the results. The coefficient on *Non-Research Posts* is statistically insignificant, indicating that non-research *WSB* posts did not forecast returns in the pre-GME period. We also confirm that for the one-month horizon the coefficient on *Net DD* is significantly larger than the coefficient on *Non-Research Post*. These findings alleviate the concern that any WSB post, independent of its content, can predict returns simply because it correlated with naïve investor demand.

3.5.2 WSB Research and Future Cash Flow News

To more directly test whether DD reports contain value-relevant information, we next examine whether DD reports forecast cash flow news. Specifically, we estimate the following panel regression:

$$Y_{it+1,t+x} = \beta_1 Net\ DD_{it} + \beta_1 Net\ DD_{it} \times 2021 + Controls_{it} + Day_t + \varepsilon_{it.} \qquad (3)$$

The dependent variable is a measure of cash flow news measured over the subsequent week (i.e., $x = 5$ trading days) or subsequent month ($x = 21$ trading days). We consider three proxies for cash flows news. The first is *Media Sentiment* obtained from *Bloomberg*. Specifically, for each firm day, *Bloomberg* assigns a sentiment score ranging from -1 (very negative news) to 1 (very positive news), with a median value of 0 (neutral articles). We assign firms with no media coverage a value of 0, and we sum the daily media sentiment over the five-day or 21-day holding period. Our second measure is *Positive Forecast Error*, which equals one if realized earnings exceed the median quarterly forecast across all I/B/E/S analysts as of day t, and zero otherwise. The five-day (21-day) sample is limited to firms that will announce earnings within five (21) trading days of day t, and we also require that the firm have at least one I/B/E/S earnings forecast. While *Positive Forecast Error* is a common proxy for cash flow news (e.g., Kelley and Tetlock, 2013), one limitation is that it restricts the sample to firms that will shortly announce earnings. As a broader measure of earnings-related news, we also compute *Positive Forecast*

Revision, which equals the total number of upward revisions scaled by the total number of revisions. In computing this measure, we consider both quarterly and annual earnings forecast revisions. We exclude firms with zero I/B/E/S coverage, and we set *Positive Forecast Revision* to 50%, the median value across the sample, for firms with I/B/E/S coverage but no forecast revisions over the holding period.[22] *Controls* and *Day* are defined as in Equation (2), and standard errors are clustered by firm and month.

Table 8 reports the results for the full sample, and Table IA.4 reports the results after excluding GME and AMC.[23] In all six specifications, the estimates on *Net DD* are positive and at least marginally significant ($p < 0.10$).[24] The economic magnitudes are also sizeable. For example, the estimate in Specification 3 indicates that an incremental buy DD recommendation issued within 5 days of the earnings announcement is associated with a 5.4% percentage points increase in the likelihood of beating the sell-side consensus forecast, which corresponds to roughly a 10% increase relative to the sample mean of 60%. On the other hand, the estimates on *Net DD* × 2021 are always significantly negative, indicating the ability of *WSB* DD reports to predict fundamentals declines significantly in the post-GME period. These findings, coupled with the return predictability evidence in Table 3, suggest that in the pre-GME period, DD reports contained value-relevant information that could potentially enhance market efficiency, but the informativeness of DD reports is completely eliminated in the post-GME period. Further, in the case of *Media Sentiment* and *Positive Forecast Error*, the post-GME estimate (i.e., *Net DD* + *Net DD* × 2021) is significantly less than zero suggesting that *WSB* reports in the post-GME period are negative predictors of fundamentals.

3.6 Price Pressure Reports and the Decline in WSB Report Informativeness in the Post-GME Period

The existing evidence is consistent with the GME event altering the culture of the site and contributing to a decline in the informativeness of *WSB* DD reports. While the impact of the GME event on the culture of *WSB* is likely far-reaching and multifaceted, anecdotal evidence suggests that a particularly important change was that the site became more focused on identifying potential profit opportunities due to short-squeezes and other forms of coordinated price pressure strategies, possibly because the massive (and salient) success of the GME short-squeeze resulted in upwardly biased

---

[22] The results are robust to excluding all firm with zero forecast revisions.
[23] Neither GME nor AMC have extreme measures of cash flow news, so excluding them from the analysis has a negligible impact on the results.
[24] In Table IA.5 of the Internet Appendix, we also examine whether non-research related *WSB* posts (as defined in Section 3.5.1) can predict cash flow news. In all six specifications, we find an insignificant relation between non-research related *WSB* posts and cash-flow news which reinforces the view of DD reports being distinct from other forms of WSB postings.

expectations of the profitability of this strategy.[25] In this section, we explore where there is an increase in *WSB* reports emphasizing price pressure following the GME event and whether this change at least partially contributes to the decline in the informativeness of *WSB* reports.

We conduct textual analysis to identify whether the report focuses on price pressure-related strategies. We develop a list of price pressure words, and as a benchmark, we also create a list of words related to fundamentals. Both lists are available in Appendix C. We define a report as focusing on price pressure if the number of price pressure words exceeds the number of fundamental words (*PP Report*).[26]

Figure 4A plots the fraction of *PP Reports* by quarter for the full sample of firms. We find that the fraction of *PP Reports* never exceeds 10% for any quarter in the pre-GME period. However, the estimates jump to 32% and 30% in Q1 and Q2 of 2021, respectively. In unreported analysis, we confirm that the difference between the pre-GME mean of 8% and the post-GME mean of 31% is highly significant ($t = 6.09$) based on standard errors double-clustered by firm and month. The differences after excluding GME and AMC (Figure 4B) are less dramatic, but still economically large, and the difference between the pre-GME mean of 7% and the post-GME mean of 24% remains significant ($t = 4.38$).

We next examine whether the increase in *PP Reports* in the post-GME period contributes to the decline in report informativeness. We repeat Equation (2) after partitioning *Net DD* into *Net DD PP*, defined as the number of buy DD recommendations for stock i on day t less the number of sell DD recommendations for stock i on day t computed over the subset of *PP Reports*, and *Net DD Non-PP*, defined as the number buy DD recommendations for stock i on day t less the number of sell DD recommendations for stock i on day t computed over all reports that are not classified as *PP Reports*.

Specifications 1 and 2 of Table 9 report the results for the five-day and 21-day holding period for the full sample of stocks, and Specifications 3 and 4 report analogous results after excluding GME and AMC. In the pre-GME period, the coefficients on *Net DD PP* and *Net DD Non-PP* are always positive (albeit not always statistically significant) and the estimates are not significantly different from each other. In other words, there is no evidence that *PP Reports* are less informative than other reports prior to the GME event. In contrast, we find the decline in report informativeness in the post-GME period is significantly larger for *PP Reports* relative to *Non-PP Reports*. For example, Specification 4

---

[25] For example, one user laments about the increasing frequency of posts discussing short squeezes here: https://www.reddit.com/r/wallstreetbets/comments/nujffg/not_every_stock_is_a_short_squeeze/

[26] As a robustness check, we also classify a report as focusing on price pressure if there is at least one price pressure word in the report (*PP Report2*). The results using this alternative classification are very similar (see Table IA.6).

indicates that the relation between *Net DD* and one-month ahead returns declined by 8.62% for *PP Reports* in the post-GME period compared to a 3.34% decline for *Non-PP Reports*; and the difference between the two estimates (5.28%) is significant. The findings are consistent with the GME event resulting in a significant increase in the number of uninformative price-pressure reports, which contributed to the reduced return predictability of DD reports in the post-GME period.

4. Can *WSB* Users Discern Report Quality?

A concern among regulators is that *WSB* induces uninformed trading that is potentially harmful to investors, particularly less-sophisticated investors. Two primary factors that likely influence the impact of *WSB* research on retail trade informativeness. The first factor, and the primary focus of the previous section, is the average informativeness of *WSB* reports. The second factor, and the focus of this section, is the extent to which consumers of *WSB* research are able to identify differences in quality across reports.

4.1 WSB Research and the Cross-Section of Stock Returns – Comment Agreement

A unique aspect of social finance research, relative to professional research (e.g., sell-side analyst recommendations), is that users on the platform can immediately provide comments in response to the report. As shown in Table 1, the average *WSB* research report induces a sizeable number of comments (65) within the first day of the report. These comments offer a direct lens into how the users of *WSB* view the quality of the report and thus provide a nice setting for testing whether users of *WSB* are able to identify more informative reports. Relatedly, the comments themselves may help users evaluate report quality. Even if the average commenter is relatively uninformed, aggregating the opinions of many diverse commenters may contain independently useful information (i.e., the wisdom of crowds). To explore this possibility, we recompute *Net DD* after partitioning the sample into DD reports where user comments agree with the DD reports (i.e., *Comment Agreement* equals one) and all other DD reports.

There are at least two ways in which comments can add incremental value to DD reports. First, comment agreement could decline in the 2021 period as users recognize the deterioration in research quality in the post-GME period (*time-series skill*). Second, within the pre-GME or post-GME period, commenters may be able to identify higher-quality reports (*cross-sectional skill*). To account for both cross-sectional and time-series skill, we explore the impact of comment agreement over the full time-series (i.e., *Net DD Agree* and *Net DD Disagree).* To isolate cross-sectional skill, we also separately

22

estimate the value of comment agreement in the pre- and post-GME period by including specifications that interact *Net DD Agree* and *Net DD Disagree* with the post-GME indicator. We limit the analysis to the one-month holding period.

Specification 1 of Table 10 reports the results for the full time-series. We find that the coefficient on *Net DD Agree* (6.59%) is marginally significant, while the coefficient on *Net DD Disagree* (1.10%) is insignificant. In addition, the difference between *Net DD Agree* and *Net DD Disagree* is also marginally significant ($t = 1.87$). Further, excluding GME and AMC from the sample (Specification 2) results in a much more reliable statistical difference ($t = 3.32$). Collectively, this evidence is consistent with *WSB* commenters being able to identify higher-quality reports.

In Specifications 3 and 4 we add *Net DD Agree × 2021* and *Net DD Disagree × 2021*. We find that comment agreement is incrementally informative in the pre-GME period. For example, after excluding GME and AMC, we find that pre-GME DD reports with comment agreement outperform pre-GME DD reports without comment agreement by 1.44% over the subsequent month. However, we find no evidence that comments are useful in the post-GME period. In fact, the point estimates are in the wrong direction. Thus, in addition to DD report quality declining in the post-GME period, the value of user comments also declines.

4.2 WSB Research and the Cross-Section of Stock Returns – Trader Agreement

While user comments are well-suited for studying how readers of *WSB* perceive report quality, only a relatively small fraction of investors who are influenced by *WSB* research issue comments in response to DD reports. In this section, we consider a complimentary approach based on trader agreement. In particular, we define traders as in agreement with the DD report if the order imbalances of a group of investors on the day of the report release are consistent with the direction of the *WSB* report recommendation. For example, net buying on the day of the buy recommendation would be classified as trader agreement, while net selling on the day of a buy recommendation would be classified as trader disagreement. In using this measure, we acknowledge that the overwhelming majority of traders' decisions are influenced by factors apart from the quality of the *WSB* recommendation. Nevertheless, order imbalances are a useful measure of investors revealed beliefs about the stock, and *WSB* research that is consistent with investor beliefs is likely to be more influential than research that challenges those beliefs.[27]

---

[27] This prediction would be consistent with both confirmation bias (Nickerson, 1998) and selective exposure theory (Knoboch-Westerwick and Meng, 2009).

We measure trader agreement for three subsets of investors: small retail investors, large retail investors, and institutional investors. We proxy for small retail investors by equally weighting retail trades, which tends to be dominated by relatively smaller traders. We proxy for large retail investors by examining retail share volume, which is heavily influenced by large trades. Finally, any trade not classified as retail is classified as an institutional trade. We sign retail trades using the algorithm of Boehmer, Jones, Zhang, and Zhang (2021), and we sign institutional trades using the Lee and Ready (1991) algorithm.

We compute order imbalances for each firm i, day t, and investor group g. For example, *Inst Vol OIB* is defined as institutional buy share volume less institutional sell share volume scaled by total institutional share volume. *Retail Vol OIB* and *Retail Trade OIB* are defined analogously after replacing *Institutional Share Volume* with *Retail Share Volume* and *Retail Number of Trades,* respectively. For each investor group, we define *Trader Agreement* as an indicator equal to one if the *OIB* measure on the day of the report release is in the same direction as the *WSB* report recommendation (e.g., net buying following a buy recommendation), and zero if it is the opposite direction. We then repeat the analysis in Table 10 after replacing *Comment Agreement* with *Trader Agreement.* To control for the fact that net buying may have information about future returns independent of the *WSB* recommendation, we also include *Buy*, an indicator equal to one if the investor group order imbalance is positive, and we allow the coefficient to vary in the post-GME period by including *Buy × 2021*.

Table 11 reports the results. We find that the coefficient on *Net DD Trader Agree* and *Net DD Trader Disagree*, based on institutional trading, are not significantly different from each other over the full sample (Specification 1) or in either the pre-GME or post-GME period. This is perhaps not surprising since institutional trading is unlikely to be heavily influenced or heavily correlated with the investment recommendations on *WSB*.[28] We also do not find very strong differences for large retail traders (Specifications 3 and 4). However, we do observe significant differences for smaller retail traders. In particular, in the pre-GME period, *WSB* reports with small retail trader agreement significantly outperform *WSB* reports with small retail trader disagreement. However, this effect is completely eliminated in the post-GME period. We also confirm that this conclusion is robust to excluding GME and AMC (see Table IA.8 of the Internet Appendix). This finding echoes the comment agreement results in Table 10, which show that commenters were able to discern report

---

[28] Consistent with this view, in Table IA.7 of the Internet Appendix we find that institutional order imbalances are uncorrelated with *WSB* investment recommendations. Large retail investor order imbalances exhibit a modest correlation with WSB recommendations, but the magnitude is roughly five times smaller than the corresponding estimate for smaller retail investors.

quality in the pre-GME period, but not the post-GME period. One potential explanation for both findings is that the GME event attracted a large influx of new *WSB* users and small retail traders who were generally less sophisticated.

5. Conclusion

Wallstreetbets (*WSB*) has become an increasingly prominent source of investment research, particularly for risk-seeking retail investors. This paper offers a first look at the investment value of *WSB* due-diligence (DD) reports. We find that prior to the GME short squeeze event, *WSB* was a source of valuable investment research. In particular, over July 2018 – December 2020, *WSB* DD reports positively forecasted one-month ahead returns. *WSB* research also positively forecasted media sentiment, earnings surprises, and earnings forecast revisions suggesting that *WSB* research contained useful information about future cash flows news. In addition, the informativeness of *WSB* reports was greater when user comments and smaller retail trading was in agreement with the DD report recommendation. However, all of the above finding fully reverse in the post-GME sample period of January 2021 – June 2021. We find that one factor that contributed to the decline in informativeness following the GME event was the dramatic increase in reports placing a greater emphasis on price-pressure rather than fundamentals. Collectively, the evidence is consistent with the surge in new users stemming from the GME short squeeze event significantly altering the content of reports, deteriorating the informativeness of *WSB* research, and consequently, its potential benefits to less sophisticated investors.

Our findings should be of relevance to both regulators and investors. From a regulatory perspective, we believe the collective evidence suggests that the negative impact of *WSB* research on financial markets is likely to be relatively modest. For example, despite regulators' concern that *WSB* research is harming small investors, we find little evidence to suggest that DD reports are negatively forecasting returns even in the post-GME period. On the other hand, the declining informativeness of *WSB* research in the post-GME period should provide caution to the 10 million *WSB* subscribers who turn to *WSB* for investment research. Indeed, our evidence cast doubt on the view that simply following all DD report recommendations will generate significant abnormal returns going forward. However, *WSB* may still be a useful source of information for investors who are adept enough to discern between higher and lower quality *WSB* research. Our findings suggest that users should be particularly cautious of reports that focus on price-pressure strategies. Identifying additional attributes

of *WSB* reports that are associated with better performance, particularly in the post-GME period, is a potentially interesting area for future research.

References

Aharon, D. Y., Kizys, R., Umar, Z., and Zaremba, A., 2021. Did David Win a Battle or the War Against Goliath? Dynamic Return and Volatility Connectedness between the GameStop Stock and the High Short Interest Indices. Working paper.

Allen, F., Haas, M., Nowak, E., Pirovano, M., & Tengulov, A., 2021. Squeezing shorts through social media platforms. Working paper.

Altınkılıç, O. and Hansen, R.S., 2009. On the information role of stock recommendation revisions. *Journal of Accounting and Economics 48*(1), 17-36.

Ammann, M., & Schaub, N. 2021. Do Individual Investors Trade on Investment-related Internet Postings? *Management Science* 67(9), 5679-5702.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61(1), 259-299.

Bali, T. G., Cakici, N., and Whitelaw, R. F, 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99(2), 427-446.

Barber, B. M., Odean, T., & Zhu, N., 2009. Do retail traders move markets? *Review of Financial Studies* 22(1), 151-186.

Barber, B., Huang, X., Odean, T., & Schwarz, C., 2021. Attention Induced Trading and Returns: Evidence from Robinhood Users. *Journal of Finance*, forthcoming.

Barberis, N.,Greenwood, R., Jin, L., & Schleifer, A., 2015. X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115(1) 1-24.

Bartov, E., Faurel, L., and Mohanram, P. S., 2018. Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review* 93(3), 25-57.

Boehmer, E., Jones, C.M., Zhang, X. and Zhang, X., 2021. Tracking retail investor activity. *Journal of Finance* 76(5), 2249-2305.

Cassella, S., and Gulen, H., 2018. Extrapolation bias and the predictability of stock returns by price-scaled variable. *Review of Financial Studies* 31 (11), 4345-4397.

Chen, H., De, P., Hu, J., and Hwang, B.H., 2014. Wisdom of the crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27 (5), 1367-1403.

Cookson, J. A., Engelberg, J. E., and Mullins, W., 2022. Echo chambers. *Review of Financial Studies*, forthcoming.

Crawford, S., Gray, W., Johnson, B., and Price, R., 2018. What motivates buy-side analysts to share recommendation online? *Management Science* 64 (6), 2473-2972.

Da, Z., Engelberg, J., and Gao, P., 2011. In search of attention. *Journal of Finance* 65(5), 1461-1469.

Da, Z., Huang, X., and Jin, L. 2021. Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics* 140(1), 175-196.

Dim, C., 2021. Should retail investors listen to social media analysts? Evidence from text-implied beliefs. Working paper.

Eaton, G.W., Green, T.C., Roseman, B. and Wu, Y., 2021. Retail trader sophistication and stock market quality: Evidence from brokerage outages. Working paper.

Engelberg, J. E., Reed, A. V., & Ringgenberg, M. C., 2012. How are shorts informed? Short sellers, news, and information processing. *Journal of Financial Economics* 105(2), 260-278.

Farrell, M., Green, T.C., Jame, R. and Markov, S., 2021. The democratization of investment research and the informativeness of retail investor trading. *Journal of Financial Economics*, forthcoming.

Farrell, M., Jame, R. and Qiu, T., 2020. The cross-section of non-professional analyst skill. Working paper.

Frazzini, A., and Lamont, O. A., 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics* 88(2), 299-322.

Giannini, R., Irvine, P., and Shu, T., 2018. Nonlocal disadvantage. An examination of social media sentiment. *Review of Asset Pricing Studies* 8(2), 293-336.

Greenwood, R., and Shleifer, A., 2014. Expectations of returns and expected returns. *Review of Financial Studies* 27(3), 714-746.

Harris, L., and Gurel, E., 1986. Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures. *Journal of Finance* 41(4), 815-829.

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. Working paper.

Hvidkjaer, S., 2008. Small trades and the cross-section of stock returns. *Review of Financial Studies*, 21(3), 1123-1151.

Jame, R., Johnston, R., Markov, S., and Wolfe, M., 2016. The value of crowdsourced earnings forecasts. *Journal of Accounting Research* 54(4), 1077-1110.

Jones, C. M., Shi, D., Zhang, X., & Zhang, X., 2020. Heterogeneity in retail investors: Evidence from comprehensive account-level trading and holdings data. Working paper.

Jia, W., Redigolo, G., Shu, S. and Zhao, J., 2020. Can social media distort price discovery? Evidence from merger rumors. *Journal of Accounting and Economics* 70(1), 101334.

Kaniel, R., Liu, S., Saar, G., & Titman, S., 2012. Individual investor trading and return patterns around earnings announcements. *Journal of Finance* 67(2), 639-680.

Kelley, E.K. and Tetlock, P.C., 2013. How wise are crowds? Insights from retail orders and stock returns. *Journal of Finance* 68(3), 1229-1265.

Kim, S.H., and Kim, D., 2014. Investor sentiment from internet message positives and the predictability of stock returns. *Journal of Economic Behavior and Organization* 107, 728-729.

Knobloch-Westerwick S., and Meng, J., 2009. Looking the other way: Selective exposure to attitude-consistent and counterattitudinal political information. *Communication Research* 36(3), 426-448.

Lee, C. M., & Ready, M. J., 1991. Inferring trade direction from intraday data. *Journal of Finance* 46(2), 733-746.

Long, C., Lucey, B. M., & Yarovaya, L., 2021. " I just like the stock" versus "fear and loathing on main street": The role of reddit sentiment in the GameStop short squeeze. Working paper.

Loughran, T., & McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66(1), 35-65.

Nickerson, R.S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2(2), 175-220.

Ozdik, G., Sadka, R., and Shen., S., 2021. Flattening the illiquidity curve: Retail trading during the COVID-19 Lockdown. *Journal of Financial and Quantitative Analysis*, forthcoming.

Semenova, V., and Winkler, J., 2021. Reddit's self-organized bull runs: Social contagion and asset prices. Working paper

Tumarkin, R. and Whitelaw, R.F., 2001. News or noise? Internet postings and stock price, *Financial Analyst Journal* 57, 41-51.

Appendix A: Sample report

Posted by u/ MikeThePutz
Post time: Wednesday, Jul 10, 2019, 02:55:11 PM EST.

BYND is at Costco DD

.

DD

Costco is now carrying Beyond Burgers. I don't see this in any press releases by either Costco or BYND. There was a vegan blog that mentioned this information (it was brought to my attention by **relevant_pet_bug**) and I called the stores to confirm it. The item # is 1338620. It is in approximately 15 stores nationwide. A store in San Diego has it for sure, 2345 Fenton Parkway, and the Costco on 1890 S University Drive in Davie, Florida has it. I don't know the other 13 stores. When I called a purchasing manager in the regional area to find out why the stores only carried a limited supply of the Beyond Burgers, he said that Costco would buy as many Beyond Burgers as they could get their hands on, but that BYND only sold a limited amount because they couldn't keep up with demand. He said that Beyond Burgers are selling really well and they are selling out in "just a few days". Apparently, BYND will be installing new manufacturing lines by the end of 2019 to increase supply and supposedly they will be able to provide Costco nationwide with Beyond Burgers by next year.

I asked why Costco would sell Beyond Burgers when they already sell Morning Star and Don Lee Farm's Veggie Burgers and the purchasing manager said that Beyond Burgers just taste different and customers want them. AGAIN: This man is responsible for buying items for Costco in a large region of the US and he said that they would buy as many BYND Burgers as they could, but that supply was limited and that they will stock Beyond Burgers nationwide once BYND can meet demand in early 2020. This is hugely positive news and I don't see any news reports about it or analyst reports mentioning it (please let me know if I am wrong!)

Photos taken from other groups about it: https://imgur.com/a/FcalTI9

If any reporters want my sources for this story feel free to PM me.

I want to thank u/**relevant_pet_bug** for pointing me in the right direction and bringing the vegan blog post, where this was first mentioned, to my attention.

Edit: fixed an error with the years. I legitimately forgot we are living in 2019.

153 comments
95% Upvoted

Appendix B: List of Keywords in Sentiment Analysis

This table reports the list of keywords assigned as bullish and bearish. Words in red are substitutes for more vulgar expressions typically used on the *WSB* forum. Small spelling differences are not included in the list but are counted when conducting keyword analysis. For example, "calls are gonna print", "calls are going to print", and 'calls gonna print" are all classified as a bullish keyword.   For buy recommendations, we define comment agreement equal to one if the number of bullish words in the comments exceeds the number of bearish words, and zero otherwise. For sell recommendations, comment agreement equals one if the number of bearish words in the comments exceeds the number of bullish words, and zero otherwise.

Bullish Words	Bearish Words
Calls are gonna print	Puts are gonna print
Buy calls	Buy puts
Buy more calls	Buy more puts
Mentions Call Option Contract	Mentions Put Option Contract
Bulls emoji	Bear emoji
Bears are in trouble	Bulls are in trouble
Moon	Crash
Pluto	Fraud
Get in	Pump and Dump
Undervalued	Hold bags
Rocket emoji	Drill Team
Can't go belly up	

Appendix C: List of Keywords in Price Pressure Analysis

This table reports the list of keywords assigned as "price pressure" words or "fundamental" words.

Price Pressure Words	Fundamental Words
Squeeze	Earnings
Short Interest	EPS
Short Sellers	Revenue
Short volume	Sales
Gamma	Growth Rate
Float	Cash Flow
Hedge Funds (HFs)	Net Income
Hedge	Customers
Melvin	Competitors
Robinhood (RH)	Market Share
Dealers	Store Visits
"HODL"[1]	P/S Ratio
	P/E Ratio
	Guidance
	Analysts

---

[1] HODL originated as misspelling of "Hold" in a 2013 WSB post, and it has become a popular inside joke on the site. Many users now also view HODL as an acronym for Hold On for Dear Life.

Appendix D: Variable Definitions

D.1 Outcome Variables

- *WSB Coverage (Table 2)* – the natural log of 1 plus the total number of Wallstreetbets (*WSB*) due diligence (DD) reports written for a firm during the calendar month. (Source: WSB).
- *SA Coverage (Table 2)* – the natural log of 1 plus the total number of Seeking Alpha *(SA)* reports written for a firm during the calendar month. (Source: Seeking Alpha).
- *$Ret_{t+1,t+x}$ (Tables 3,4,5,6,8,9,10)* – the buy and hold return for the DD report recommendation starting on the day after the report and ending on day x, where x typically equals five or 21 trading days. We define the day of the report as the first trading day in which an investor could have traded on the report.
- *$News\ Sentiment_{t+1,t+x}$ (Table 7)*- the sum of a daily sentiment score starting on the day after the report and ending on day x. The sentiment scores are obtained from Bloomberg and range from -1 (very negative news) to 1 (very positive news), with a median value of 0 (neutral articles). We assign firms with no media coverage a value of 0. (Source: Bloomberg).
- *$Positive\ Forecast\ Error_{t+1,\ t+x}$ (Table 7)* – An indicator equal to one if the realized quarterly earnings reported within x days of the DD report exceed the median forecast across all I/B/E/S analysts. The value is set missing for firms that do not have I/B/E/S coverage or for firms that will not announce earnings over the forecast horizon being analyzed (i.e. five or 21 trading days). (Source: I/B/E/S).
- *$Positive\ Forecast\ Revision_{t+1,\ t+x}$ (Table 7)* – the total number of upward revisions issued within x days of the DD reports scaled by the total number of revisions issued over the same period. In computing this measure, we consider both quarterly and annual earnings forecast revision. This value is set to missing for firms that do not have I/B/E/S coverage, and the value is set to 50%, the median value across the sample, for firms with IBES coverage but no forecast revisions over the holding period. (Source: I/B/E/S).

D.2 Other Variable

- *Net DD* – the total number of WSB due diligence (DD) reports that recommend buying the firm over a time period (e.g., one day) less the total number of DD reports that recommend selling the firm during the time period. (Source: WSB).
- *D2021* – an indicator equal to one for the January 2021 June 2021 sample period and zero otherwise.
- *# Comments* – the total number of comments issued in response to a DD report. The sample is limited to comments that are posted between the publication of the report and the start of the next trading day. (Source: WSB).
- *Comment Agreement* – an indicator equal to one if the number of agreement keywords in the comments is at least as large as the number of disagreement keywords in the comments. The analysis is limited to comments that are posted between the publication of the report and the start of the next trading day. The list of agree/disagree keywords are reported in Appendix B. (Source: WSB).
- *Net DD Comment Agree* – The Net DD measure computed using only for the subset of reports where Comment Agreement =1. (Source: WSB).
- *Net DD Comment Disagree* – The Net DD measure computed using only for the subset of reports where Comment Agreement =0. (Source: WSB).

- *Trader Agreement* – an indicator equal to one if group of investors trades in the same direction as the DD report recommendation on the day of the DD report. We measure trade direction for institutional trades, large retail trades, and small retail trades separately using the following approaches:
    - *Inst. Vol OIB* – institutional buy share volume less institutional sell share volume scaled by total institutional share volume. Institutional trades are assigned as buys or sells based on the Lee and Ready (1991) algorithm. (Source: TAQ Intraday Indicators).
    - *Retail Vol OIB* – retail buy share volume less retail sell share volume scaled by total retail share volume. Retail trades are assigned as buys or sells based on the Boehmer, Jones, Zhang, and Zhang (2021) algorithm. (Source: TAQ Intraday Indicators).
    - *Retail Trade OIB* – retail buy trades less retail sell trades scaled by total retail share trades. Retail trades are assigned as buys or sells based on the Boehmer, Jones, Zhang, and Zhang (2021) algorithm. (Source: TAQ Intraday Indicators).
- *Net DD Trader Agree* – The *Net DD* measure computed using only for the subset of reports where *Trader Agreement* =1. (Source: WSB and TAQ).
- *Net DD Trader Disagree* – The *Net DD* measure computed using only for the subset of reports where *Comment Agreement* =0. (Source: WSB and TAQ).
- *Confounded Report* – an indicator equal to one if the report is issued around a confounding information event, defined as a DD report issued on the previous day (i.e., -1), an earnings announcement issued on the previous or current day (-1, 0) or abnormal media coverage on the previous or current day (-1, 0).
    - *Earning Report* – a quarterly or annual earnings announcement *(*Source: I/B/E/S).
    - *Abnormal Media Coverage* – an indicator equal to one if the number of articles on the firm, as reported by Bloomberg, is in the top 20% relative to the firm's typical media coverage over the previous 60 days [-60, -1]. (Source: Bloomberg).
- *Net DD Processing* – *Net DD* computed using only the subset of reports where *Confounded Report* =1. (Source: WSB and Bloomberg).
- *Net DD Production* – The *Net DD* computed using only the subset of reports where *Confounded Report* =0. (Source: WSB and Bloomberg).
- *PP Report* – an indicator equal to one if the number of *price pressure* words in the report exceeds the number of *fundamental* words in the report. The list of *price pressure* and *fundamental* words are available in Appendix C. (Source: WSB).
- *Net DD PP* – *Net DD* computed using only the subset of reports where *PP Report = 1*. (Source: WSB).
- *Net DD Non-PP* – *Net DD* computed using only the subset of reports where *PP Report = 0*. (Source: WSB).
- *Net SA* – the total number of Seeking Alpha research reports that recommend buying the firm over a time period (e.g., one day) less the total number of Seeking Alpha reports that recommend selling the firm during the time period. (Source: SA).
- *Non-Research Posts* – the total number of unique non-research posts where a ticker is mentioned in the title over a time period (e.g., one day). We classify posts in the following WSB categories as non-research related: *News, Gains, Losses, Charts,* and *Shi$posts.* (Source: WSB).
- *Size* – the market capitalization computed as share prices times total shares outstanding at the end of the year. (Source: CRSP).

- *Book-to-Market (BM)* – the book-to-market ratio computed as the book value of equity during the calendar year scaled by the market capitalization at the end of the calendar year. Positive values are winsorized at the 1st and 99th percentile. Negative value and missing values are set equal to zero and we include a corresponding "Missing BM" indicator. (Source: CRSP/Compustat).
- *Volatility* – the standard deviation of daily returns during the month (Source: CRSP).
- *Turnover* – the average daily turnover (i.e., share volume scaled by shares outstanding) during the month.
- *Ret [0]* – the buy-and-hold return on the current day. (Source: CRSP).
  - *Ret [-5, -1]* - the buy-and-hold return on five trading days.
  - *Ret [-26, -6]* - the buy-and-hold return over the previous six to 26 trading days.
  - *Return (m-1)* – the buy-and-hold return in the previous calendar month. (Source: CRSP).
  - *Return (m-2, m-12)* – the buy-and-hold return over the previous two to twelve calendar months. (Source: CRSP).
- *Sentiment [0]* – The average sentiment scores across all news articles on the current day, where the score ranges from -1 (very negative news) to 1 (very positive news), with a median value of 0 (neutral articles). Firms with no media coverage are assigned a sentiment score of 0.
  - *Sentiment [-5, -1]* – the sum of the sentiment score over the previous 1 to 5 trading days prior to the report release.
  - *Sentiment [-26, -6]* – the sum of the sentiment score over the previous six to 26 trading days prior to the report release.
- *Institutional Ownership* – the percentage of the firm's shares held by institutions at year end. (Source: Thomson Reuters Institutional Holdings S34).
- *Breadth of Ownership* – the total number of common shareholders (Source: Compustat).
- *IBES Coverage* – the number of unique brokerage houses issuing earnings forecast for a firm during the calendar year. (Source: I/B/E/S).
- *Media Coverage* – the total number of media articles about a firm during the calendar year. (Source: Bloomberg).
- *High Max* – an indicator equal to one if the firm maximum daily return in the prior month was in the top quintile of the distribution (Source: CRSP).
- *Heavy Short* – an indicator equal to one if the firm is in the top quintile of short interest, defined as the number of shares that have been sold short scaled by shares outstanding. (Source: Compustat).
- *Recent IPO* – an indicator equal to one if the firm went public in the past six months. (Source: CRSP).
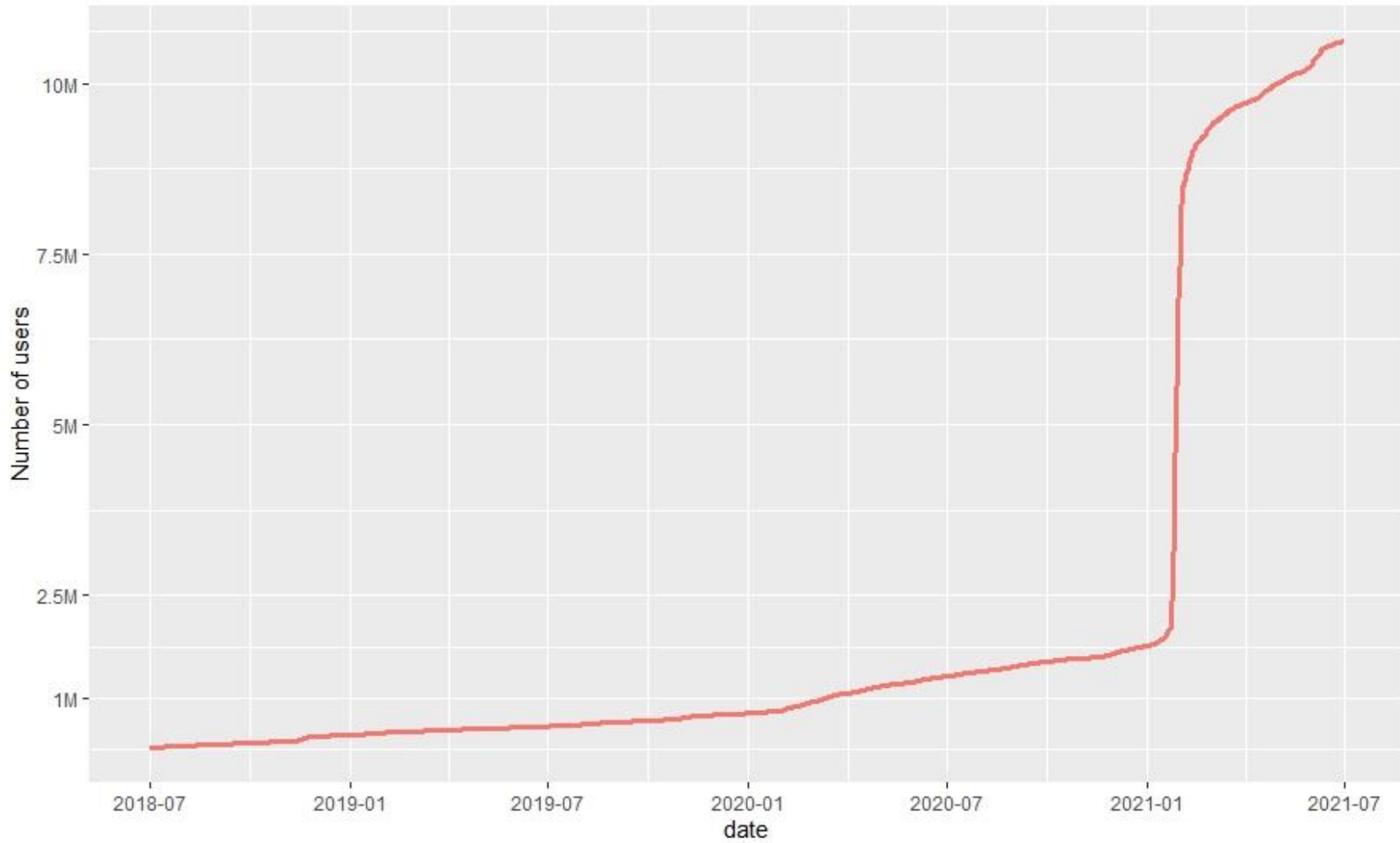
Figure 1: Growth in Reddit's Wallstreetbets (WSB)
This figure plots the total number of users on WSB from July 2018 through June 2021. This data can be found at https://subredditstats.com/r/wallstreetbets.

Figure 2A: WSB Reports and Returns by Quarter (Full Sample)

Figure 2B: WSB Reports and Returns by Quarter (Exclude GME and AMC)
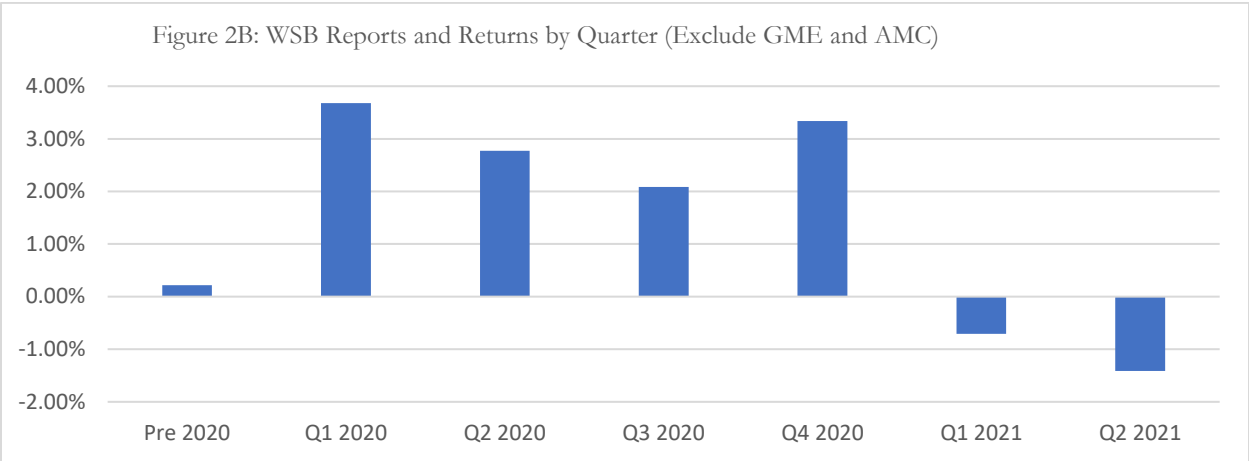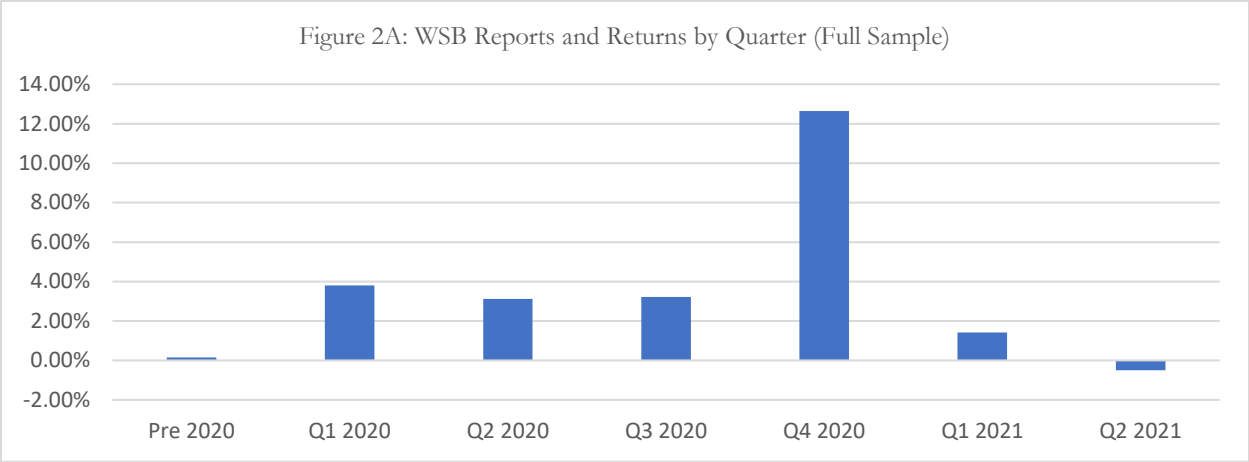
Figure 2: WSB Reports and Future Returns – Quarterly Estimates
This figure reports the estimates on *Net DD* from Table 3 for each quarter from Q1 of 2020 through Q2 of 2021. It also reports the estimates for all reports prior to 2020 (*Pre 2020*). Figure 2A reports the results for one-month holding period and the full sample of firms (i.e., Specification 2 of Table 3), and Figure aB reports the results for the one-month holding period and the sample that excludes GME and AMC (i.e., Specification 4 of Table 3).

Figure 3A: WSB Reports and Longer-Horizon Returns (Full Sample)

Figure 3B: WSB Reports and Longer-Horizon Returns (Exclude GME & AMC)
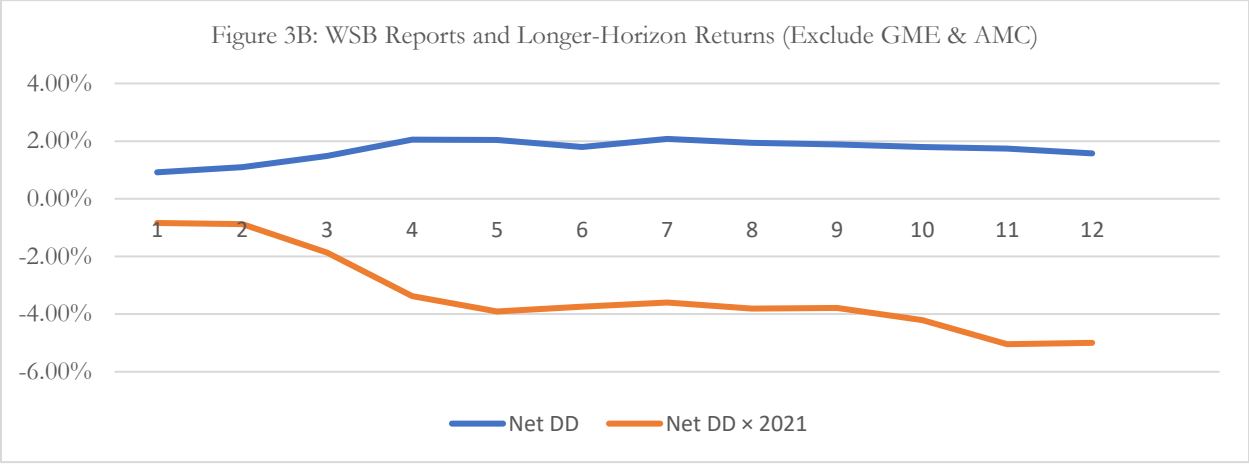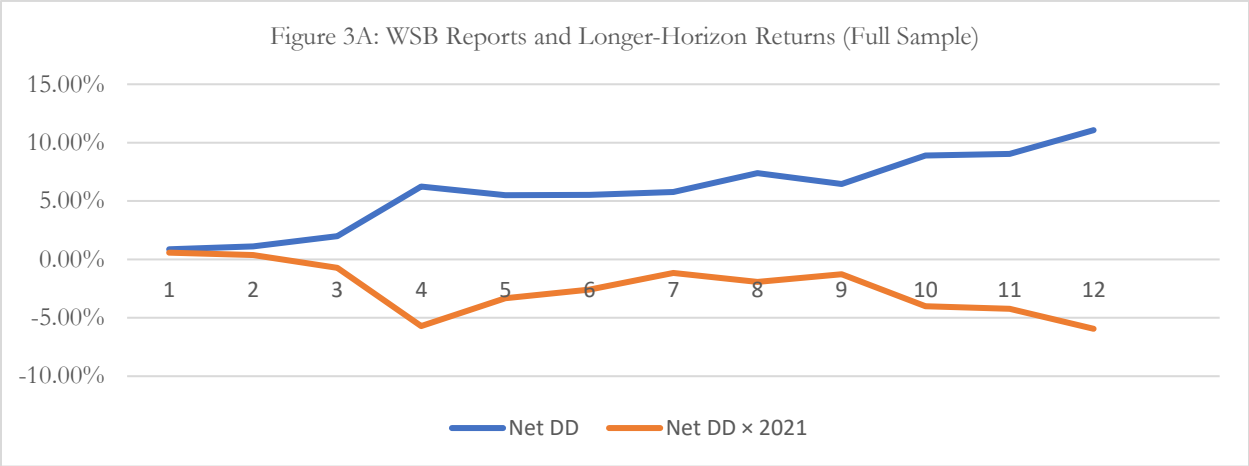
Figure 3: WSB Reports and Future Returns – Longer Horizons
This table repeats the estimates from Table 3 for horizons ranging from one-week (i.e., x =5) through 12 weeks (i.e., x = 60). We report the coefficient estimates on *Net DD* and *Net DD × 2021* for each horizon. Figures 3A and 3B report the results for the full sample and the sample that excludes GME and AMC.

Figure 4A: Percentage of PP Reports - Full Sample

Figure 4B: Percentage of PP Reports - Exclude GME and AMC
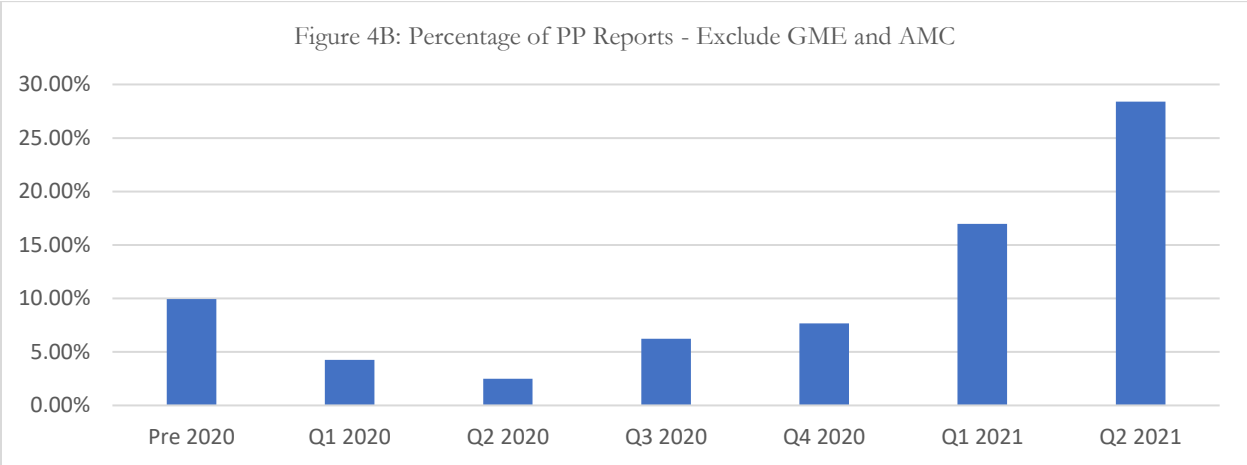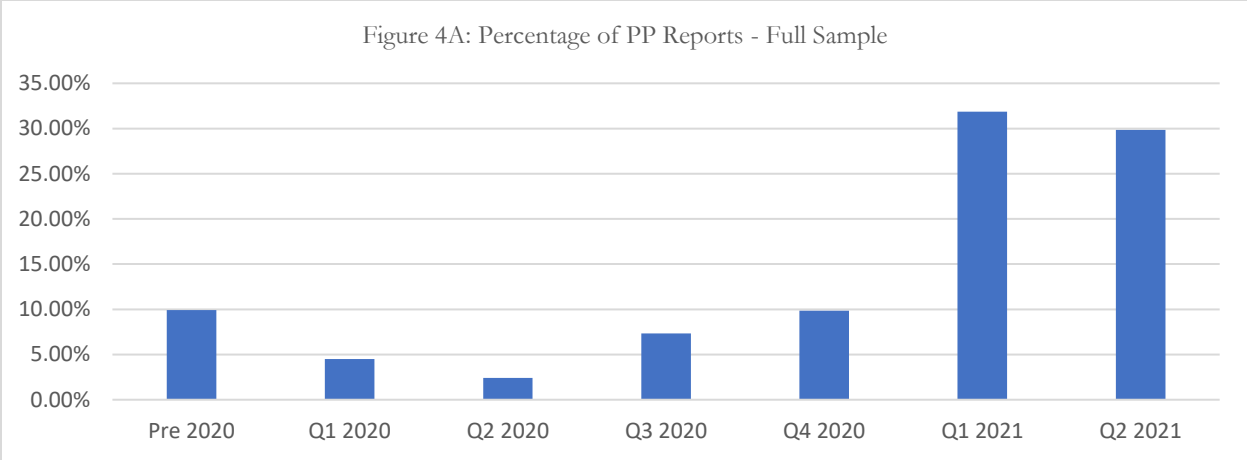
Figure 4: WSB and Price Pressure Reports – Quarterly Estimates
This figure reports the percentage of reports where the number of "price pressure" words exceed the number of "fundamental" words (see Appendix C for the list of "price pressure" and "fundamental" words) for each quarter from Q1 of 2020 through Q2 of 2021. It also reports the estimates for all reports prior to 2020 (*Pre 2020*). Figure 4A reports the results for the full sample, and Figure 4B reports the results for the sample that excludes GME and AMC.

Table 1: Descriptive Statistics

This table reports summary statistics on the sample of Due Diligence (DD) reports on Reddit's Wallstreetbets (*WSB*). DD reports are reports identified by the poster (and verified by the moderator) as containing some analysis and offering a clear buy or sell signal. We report the number of DD reports for the full sample (July 2018-June 2021), the July 2018 – December 2020 sample (*pre-GME*), and the January 2021-June 2021 sample (*post-GME*). We also report the number of firms-days and firms with at least one DD report, the percentage of reports recommending a long position (*Buys*), the total number of comments issued between the DD report and the subsequent trading day (*# Comments*), the percentage of the DD reports that coincide with a confounding information event (*Confounded*), defined as an earnings announcement or abnormal media coverage over the [-1,0] window or another DD report on day *-1,* and the average number of DD reports issued by each username (*Posts per Contributor*). We limit the sample to DD reports that focus on a single common stock ticker. Panel A tabulates the results for the full sample. Panel B excludes GME and AMC from the sample, and Panel C limits to the sample to just GME and AMC. Panel D reports analogous results for the sample of research reports provided by Seeking Alpha.

Panel A: WSB DD Reports - Full Sample

	DD Reports	Firm-Days	Firms	% Buys	# Comments	Confounded	Posts per Contributor
Full Sample	5050	3811	909	88%	65	30.69%	1.32
July 2018-2020	2333	2016	617	81%	45	26.15%	1.34
Jan–June 2021	2717	1795	527	95%	81	34.60%	1.25

Panel B: WSB DD Reports - Exclude GME & AMC

	DD Reports	Firm-Days	Firms	% Buys	# Comments	Confounded	Posts per Contributor
Full Sample	4441	3642	907	88%	52	26.88%	1.34
July 2018-2020	2252	1956	615	80%	43	25.71%	1.33
Jan–June 2021	2189	1686	525	95%	62	28.10%	1.27

Panel C: WSB DD Reports - GME & AMC Only

	DD Reports	Firm-Days	Firms	% Buys	# Comments	Confounded	Posts per Contributor
Full Sample	609	169	2	93%	154	58.46%	1.13
July 2018-2020	81	60	2	89%	108	38.27%	1.37
Jan–June 2021	528	109	2	94%	161	61.55%	1.08

Panel D: Seeking Alpha Reports – Full Sample

	SA Reports	Firm-Days	Firms	% Buys	# Comments	Confounded	Posts per Contributor
Full Sample	23,177	23,871	2,953	85%	45	25.58%	11.08
July 2018-2020	19,853	18,614	2,638	85%	45	23.80%	10.37
Jan–June 2021	4,018	4,018	1,689	88%	45	34.54%	5.22

Table 2; Determinants of WSB Coverage

This table presents the estimates from Equation (1):

$$Coverage_{it} = \alpha + \beta_1 Chars_{it-1} + \beta_2 Chars_{it-1} \times 2021 + Month_t + \varepsilon_{it.}$$

The dependent variable, *Coverage,* is either *WSB Coverage* defined as Log (1 + DD Reports) for firm *i* during month *t, SA coverage* defined as Log (1 + SA Reports) for firm *i* during month *t,* or the difference between *WSB and SA Coverage*. *Chars* include the following firm characteristics: the percentage of the firm's shares held by institutional investors at the end of the prior year (*Inst. Ownership*), the number of common shareholders (*Breadth of Ownership*), market capitalization (*Size*), book to market (*BM),* return volatility (*Volatility*), share turnover (*Turnover*), returns over the prior month (*Ret_{m-1}*), returns over the prior two to twelve months (*Ret_{m-2, m-12}*), the number of media articles mentioning the firm in the prior year (*Media Coverage*), the number of sell-side analysts issuing a forecast for the firm in the prior year (*IBES Coverage*), an indicator equal to one if the firms maximum daily return in the prior month with in the top quintile (*High Max),* an indicator equal to one if the firm is in the top quintile of short interest (*Heavy Short*), and an indicator equal to one if the firm went public in the past six months *(Recent IPO)*. We allow the loading on firm characteristics to vary in the pre-GME and post-GME period by interacting the firm characteristics with *2021,* an indicator equal to one for the 2021 sample period and zero otherwise, and we also include calendar-month fixed effects. All continuous variables are standardized to have mean zero and unit variance. More detailed variable definitions are in Appendix D. Standard errors are clustered by firm and month, and *t*-statistics are reported in parentheses.

	WSB Coverage	SA Coverage	WSB - SA Coverage
	[1]	[2]	[3]
Inst Ownership	-0.03	-0.06	0.04
	(-3.83)	(-5.85)	(3.80)
Inst Ownership × 2021	-0.03	0.04	-0.07
	(-1.03)	(3.84)	(-2.49)
Log (Breadth of Ownership)	0.01	0.03	-0.03
	(0.69)	(2.07)	(-2.10)
Log (Breadth of Ownership) × 2021	-0.01	-0.01	0.00
	(-0.63)	(-0.84)	(0.03)
Log (Size)	0.24	0.80	-0.56
	(4.42)	(12.20)	(-8.57)
Log (Size) × 2021	0.41	-0.11	0.52
	(3.67)	(-1.68)	(4.57)
Log (BM) × 2021	-0.01	-0.03	0.01
	(-1.95)	(-2.31)	(1.44)
Log (BM)	0.04	0.00	0.04
	(2.22)	(-0.13)	(1.70)
Negative BM	0.02	0.02	-0.01
	(0.90)	(0.53)	(-0.13)
Negative BM × 2021	-0.09	-0.01	-0.08
	(-0.84)	(-0.24)	(-0.86)
Log (Vol)	0.10	0.45	-0.35
	(4.50)	(9.45)	(-8.17)
Log (Vol) × 2021	0.14	-0.24	0.38
	(2.42)	(-4.43)	(6.07)
Log (Turn)	0.01	-0.03	0.04
	(0.78)	(-1.79)	(2.87)
Log (Turn) × 2021	0.13	0.08	0.05
	(3.45)	(5.11)	(1.59)
Return (m-1)	0.01	0.01	0.00
	(2.05)	(2.11)	(0.51)
Return (m-1) × 2021	0.01	-0.01	0.02
	(0.22)	(-0.54)	(0.73)
Return (m-2, m-12)	0.28	0.21	0.07
	(2.98)	(4.90)	(0.91)
Return (m-2, m-12) × 2021	-0.24	-0.20	-0.04
	(-2.66)	(-4.70)	(-0.58)
Log (Media Coverage)	0.18	0.39	-0.21
	(3.76)	(6.35)	(-3.81)

Log (Media Coverage) × 2021	0.11	-0.19	0.30
	(1.40)	(-3.75)	(3.21)
Log (IBES Coverage)	-0.01	0.04	-0.05
	(-0.70)	(2.88)	(-4.45)
Log (IBES Coverage) × 2021	-0.08	-0.04	-0.04
	(-2.57)	(-2.91)	(-1.33)
High Max	0.03	0.06	-0.03
	(3.05)	(3.72)	(-1.77)
High Max × 2021	0.21	0.03	0.17
	(5.62)	(1.32)	(6.81)
Heavy Short	0.07	0.16	-0.09
	(2.82)	(5.68)	(-3.10)
Heavy Short × 2021	0.38	-0.07	0.44
	(4.12)	(-1.85)	(4.71)
Recent IPO	0.57	0.97	-0.41
	(3.82)	(6.86)	(-2.56)
Recent IPO × 2021	1.20	0.39	0.81
	(3.93)	(3.05)	(2.63)
Obs. (Firm-Months)	117,519	117,519	117,519
Fixed Effects	Month	Month	Month
R-square	4.65%	15.15%	6.07%
Sample	All Firms	All Firms	All Firms

Table 3: WSB Reports and Future Returns

This table reports results from the estimation of Equation (2):

$$R_{it+1,t+x} = \beta_1 Net\ DD_{it} + \beta_2 Net\ DD_{it} \times 2021 + Controls_{it} + Day_t + \varepsilon_{it.}$$

The dependent variable, *R*, is the stock return measured over the subsequent week (i.e., x = 5 trading days) or the subsequent month (x = 21 trading days). *Net DD,* is the number of buy DD recommendations for stock *i* on day *t* less the number of sell DD recommendations for stock *i* on day *t* and *Net DD × 2021,* interacts *Net DD* with *2021,* an indicator equal to one for the 2021 sample period and zero otherwise. *Controls* includes market capitalization (*Size*), book-to-market (*BM*), prior returns and prior media sentiment measured on the day of the DD report [0], the five days prior to the DD report [-5,-1], and the 6 to 26 days prior to the DD reports [-26,-6]. *Day* denotes date fixed effects. More detail variable definitions are available in Appendix D. Specifications 1 and 2 report the results for the full sample for five-day and 21-day returns, respectively. Specifications 3 and 4 report analogous results after excluding GME and AMC. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether the *Net DD + Net DD × 2021* is significantly different from zero.

| | Ret [1,5] | Ret [1,21] | Ret [1,5] | Ret [1,21] |
	[1]	[2]	[3]	[4]
Net DD	0.86%	6.04%	0.92%	2.32%
	(1.80)	(2.19)	(1.92)	(2.21)
Net DD × 2021	0.58%	-5.21%	-0.84%	-3.83%
	(0.55)	(-2.61)	(-1.39)	(-2.54)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.58)	(-1.26)	(-1.58)	(-1.27)
Log (BM)	-0.07%	-0.25%	-0.08%	-0.26%
	(-0.95)	(-0.84)	(-0.97)	(-0.86)
Ret [0]	-7.22%	-9.35%	-7.28%	-9.30%
	(-5.15)	(-4.67)	(-5.23)	(-4.54)
Ret [-5, -1]	-2.50%	-3.32%	-2.50%	-3.38%
	(-2.27)	(-2.30)	(-2.27)	(-2.36)
Ret [-26, -6]	-0.38%	-0.81%	-0.38%	-0.82%
	(-1.48)	(-0.85)	(-1.48)	(-0.85)
News Sentiment [0]	0.08%	0.08%	0.06%	0.06%
	(2.74)	(0.82)	(1.69)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.22)	(0.14)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.07%
	(0.47)	(0.84)	(0.73)	(1.09)
Net DD + Net DD × 2021	1.45%	0.83%	0.08%	-1.51%
	(1.49)	(0.92)	(0.18)	(-1.31)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table 4: WSB Reports and Future Returns – Robustness

This table examine the sensitivity of the baseline estimates in Table 3 (tabulated for convenience in Panel A) to alterative research design choices. In Panel B, we exclude the Robinhood 50 stocks, defined as the 50 stocks that Robinhood imposed trading restriction on beginning on January 28th,2021 and ending February 5th, 2021 (except Columns 1 and 2 continue to include GME and AMC). In Panel C, we allow the coefficient on the firm characteristics to vary in the pre- and post-GME period by interacting the 2021 indicator with all the control variables. Panel D limits the sample period to January 2020 through June of 2021. Panel E partitions *Net DD* into three separate variables: *Heavy Buy,* an indicator equal to one if *Net DD* is greater than or equal to 2, *Light Buy,* an indicator equal to one if *Net DD* is equal to one; and *Sell*, an indicator equal to one if *Net DD* is negative. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate.

	All Firms		Exclude GME & AMC	
	Ret [1,5]	Ret [1,21]	Ret [1,5]	Ret [1,21]
	[1]	[2]	[3]	[4]
Panel A: Baseline Results				
NET DD	0.86%	6.04%	0.92%	2.32%
	(1.80)	(2.19)	(1.92)	(2.21)
NET DD ×2021	0.58%	-5.21%	-0.84%	-3.83%
	(0.55)	(-2.61)	(-1.39)	(-2.54)
Panel B: Drop "Robinhood 50" Stocks				
NET DD	0.88%	6.31%	0.94%	2.19%
	(1.78)	(2.10)	(1.90)	(1.84)
NET DD × 2021	0.74%	-5.20%	-0.72%	-3.46%
	(0.71)	(-2.17)	(-1.16)	(-2.24)
Panel C: Interact controls with Post-GME Indicator				
NET DD	0.86%	6.04%	0.91%	2.25%
	(1.79)	(2.14)	(1.91)	(2.18)
NET DD × 2021	0.64%	-5.08%	-0.71%	-3.39%
	(0.58)	(-2.25)	(-1.05)	(-1.93)
Panel D: Limit Sample to 2020 -2021				
NET DD	1.17%	8.12%	1.25%	3.30%
	(1.81)	(2.48)	(1.93)	(2.34)
NET DD × 2021	0.32%	-7.16%	-1.07%	-4.45%
	(0.27)	(-2.86)	(-1.45)	(-2.48)
Panel E: Partition Net DD into Sells, Light Buys, Heavy Buys				
Heavy Buy	3.17%	26.75%	3.67%	5.39%
	(4.25)	(3.00)	(5.72)	(1.21)
Light Buy	0.81%	4.78%	0.76%	3.48%
	(1.32)	(3.33)	(1.24)	(2.77)
Sell	0.49%	1.60%	0.44%	1.95%
	(0.43)	(0.99)	(0.39)	(1.18)
Heavy Buy × 2021	4.52%	-20.54%	-3.57%	-10.80%
	(0.63)	(-1.53)	(-4.08)	(-1.94)
Light Buy × 2021	-0.08%	-3.43%	-0.81%	-4.93%
	(-0.09)	(-1.41)	(-0.85)	(-2.27)
Sell × 2021	4.42%	3.88%	-1.58%	-6.59%
	(0.98)	(0.52)	(-0.80)	(-1.85)

Table 5: SA Reports and Future Returns

This table repeats the analysis in Table 3 after adding two additional variables: *Net SA* and *Net SA × 2021*. *Net SA* is the number of SA reports issuing a buy recommendation for stock *i* on day *t* less the number of *SA* reports issuing a sell recommendation for stock *i* on day *t*, and *Net SA × 2021*, interacts *Net SA* with an indicator equal to one for the post-GME period (January-June 2021) and zero otherwise. All other variables are defined in Table 3 (with more detailed variable definitions in Appendix D). Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether the *Net DD − Net SA* and *Net DD 21 - Net SA 21* are significantly different from zero.

	Ret [1,5]	Ret [1,21]	Ret [1,5]	Ret [1,21]
	[1]	[2]	[3]	[4]
Net DD	0.86%	6.04%	0.92%	2.32%
	(1.80)	(2.19)	(1.92)	(2.21)
Net DD × 2021	0.58%	-5.21%	-0.84%	-3.83%
	(0.55)	(-2.61)	(-1.39)	(-2.54)
Net SA	0.34%	0.64%	0.33%	0.64%
	(4.15)	(2.68)	(4.19)	(2.68)
Net SA × 2021	0.04%	0.25%	-0.05%	0.37%
	(0.31)	(0.62)	(-0.32)	(0.97)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.58)	(-1.26)	(-1.58)	(-1.27)
Log (BM)	-0.07%	-0.25%	-0.08%	-0.26%
	(-0.95)	(-0.84)	(-0.97)	(-0.86)
Ret [0]	-7.22%	-9.35%	-7.28%	-9.30%
	(-5.15)	(-4.67)	(-5.23)	(-4.54)
Ret [-5, -1]	-2.50%	-3.32%	-2.50%	-3.38%
	(-2.27)	(-2.30)	(-2.27)	(-2.36)
Ret [-26, -6]	-0.38%	-0.81%	-0.38%	-0.82%
	(-1.48)	(-0.85)	(-1.48)	(-0.85)
News Sentiment [0]	0.08%	0.08%	0.06%	0.06%
	(2.74)	(0.82)	(1.69)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.22)	(0.14)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.07%
	(0.47)	(0.84)	(0.73)	(1.09)
Net DD - Net SA	0.52%	5.45%	0.58%	1.66%
	(1.06)	(1.96)	(1.20)	(1.53)
Net DD 21 - Net SA 21	0.55%	-5.50%	-0.79%	-4.20%
	(0.50)	(-2.50)	(-1.21)	(-2.89)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table 6: WSB Reports and Future Returns - Information Processing vs. Information

This table repeats the analysis in Table 3 after partitioning all DD reports into *Confounded* and *Non-Confounded Reports*. *Confounded Report* is an indicator equal to one if the report is issued around a confounding information event, defined as a DD report issued on the previous day (-1), an earnings announcement issued on the previous or current day (-1, 0), or abnormal media coverage on the previous or current day (-1, 0). *Net DD Processing* is the *Net DD* measure computed for the subset of reports where *Confounded Report = 1,* and Net *DD Production* is the *Net DD* measure computed for the subset of reports where *Confounded Report = 0*. All other variables are defined in Table 3 (with more detailed variable definitions in Appendix D). Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether *Net DD Processing - Net DD Production* and *Net DD Processing × D2021 - Net DD Production × D2021* are significantly different from zero.

| | Ret [1,5] | Ret [1,21] | Ret [1,5] | Ret [1,21] |
	[1]	[2]	[3]	[4]
Net DD Processing	0.48%	7.91%	0.39%	3.22%
	(0.90)	(2.11)	(0.79)	(2.17)
Net DD Processing × 2021	-2.11%	-11.77%	-1.85%	-6.80%
	(-2.43)	(-2.71)	(-2.39)	(-2.73)
Net DD Production	1.06%	5.17%	1.18%	1.88%
	(1.88)	(2.24)	(2.09)	(2.08)
Net DD Production ×2021	0.53%	-4.12%	-0.90%	-3.11%
	(0.46)	(-2.59)	(-1.28)	(-2.26)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.60)	(-1.27)	(-1.60)	(-1.27)
Log (BM)	-0.08%	-0.25%	-0.08%	-0.26%
	(-0.97)	(-0.84)	(-0.99)	(-0.86)
Ret [0]	-7.21%	-9.35%	-7.27%	-9.30%
	(-5.13)	(-4.63)	(-5.21)	(-4.54)
Ret [-5, -1]	-2.49%	-3.34%	-2.49%	-3.38%
	(-2.26)	(-2.31)	(-2.26)	(-2.36)
Ret [-26, -6]	-0.38%	-0.82%	-0.38%	-0.82%
	(-1.48)	(-0.85)	(-1.48)	(-0.85)
News Sentiment [0]	0.08%	0.08%	0.06%	0.06%
	(2.73)	(0.82)	(1.67)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.22)	(0.15)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.07%
	(0.46)	(0.84)	(0.72)	(1.09)
Net DD Processing - Net DD Production	-0.58%	2.74%	-0.80%	1.34%
	(-0.97)	(1.86)	(-1.44)	(1.57)
Net DD Process 21 - Net DD Prod. 21	-2.64%	-7.65%	-0.96%	-3.69%
	(-1.03)	(-2.96%)	(-0.92%)	(-1.91)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table 7: WSB Reports and Future Returns - DD Reports vs. Non-Research Postings

This table repeats the analysis in Table 3 after adding two additional variables: *Non-Research Posts* and *Non-Research Posts × 2021*. We define a WSB posts as non-research related if it belongs to the one of the following categories: *News*, *Losses*, *Gains*, *Charts*, and *Shi$posts*. *Non-Research Posts* is the number of stock *i* was listed in the title of a non-research posts on day *t* and *Non-Research Posts × 2021*, interacts *Non-Research Posts* with an indicator equal to one for the post-GME period (January-June 2021) and zero otherwise. All other variables are defined in Table 3 (with more detailed variable definitions in Appendix D). Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether *Net DD – Non-Research Posts* and *Net DD 21 - Non-Research Posts 21* are significantly different from zero.

| | Ret [1,5] | Ret [1,21] | Ret [1,5] | Ret [1,21] |
	[1]	[2]	[3]	[4]
Net DD	0.85%	5.82%	0.90%	2.34%
	(1.79)	(2.30)	(1.92)	(2.26)
Net DD × 2021	-0.35%	-4.41%	-0.77%	-3.46%
	(-0.42)	(-2.87)	(-1.25)	(-2.39)
Non-Research Posts	0.18%	2.58%	0.22%	-0.16%
	(0.61)	(0.87)	(0.74)	(-0.19)
Non-Research Posts × 2021	2.00%	-3.90%	-0.61%	-2.78%
	(1.59)	(-1.24)	(-1.27)	(-2.77)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.60)	(-1.27)	(-1.60)	(-1.26)
Log (BM)	-0.08%	-0.25%	-0.08%	-0.26%
	(-0.97)	(-0.84)	(-0.99)	(-0.86)
Ret [0]	-7.26%	-9.34%	-7.27%	-9.28%
	(-5.18)	(-4.61)	(-5.21)	(-4.52)
Ret [-5, -1]	-2.49%	-3.34%	-2.49%	-3.38%
	(-2.26)	(-2.32)	(-2.26)	(-2.36)
Ret [-26, -6]	-0.38%	-0.82%	-0.38%	-0.82%
	(-1.48)	(-0.86)	(-1.48)	(-0.85)
News Sentiment [0]	0.09%	0.07%	0.06%	0.06%
	(2.76)	(0.81)	(1.67)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.21)	(0.15)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.07%
	(0.45)	(0.85)	(0.72)	(1.09)
Net DD - Non-Research	0.66%	3.23%	0.68%	2.50%
	(1.20)	(1.96)	(1.25)	(2.10)
Net DD 21 - Non-Research 21	-2.35%	-0.51%	-0.15%	-0.68%
	(-1.43)	(-0.25)	(-0.18)	(-0.43)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table 8: WSB Reports and Cash Flow News

This table reports results from the estimation of Equation (3):

$$Y_{it+1,t+x} = \beta_1 Net\ DD_{it} + \beta_1 Net\ DD \times 2021_{it} + Controls_{it} + Day_t + \varepsilon_{it.}$$

The dependent variable, *Y*, is a measure of cash flow news over the subsequent week (i.e., x = 5 trading days) or the subsequent month (x = 21 trading days). Cash flow news is measured as either *Media Sentiment (Media)*, computed as the sum of the daily Bloomberg sentiment score; *Positive Forecast Error* (*Pos FE*), an indicator equal to one if the realized earnings exceed the median quarterly forecast across all I/B/E/S analysts as of day *t,* and *Positive Forecast Revision* (*Pos FR*) computed as the number of upward revisions by I/B/E/S analysts scaled by the total number of revisions. All other variables are defined as in Table 3. More detailed variable definitions are available in Appendix D. Specifications 1 and 2 report the results for the full sample for five-day and 21-day measures of *Media Sentiment*. Specifications 3 and 4 and 5 and 6 report analogous results for *Positive Forecast Error* and *Positive Forecast Revision,* respectively. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether the *Net DD + Net DD × 2021* is significantly different from zero.

	Media [1,5]	Media [1,21]	Pos FE [1,5]	Pos FE [1,21]	Pos FR [1,5]	Pos FR [1,21]
	[1]	[2]	[3]	[4]	[5]	[6]
Net DD	4.75%	16.33%	5.41%	3.88%	2.83%	2.42%
	(1.94)	(1.84)	(3.40)	(1.91)	(2.23)	(1.90)
Net DD × 2021	-6.12%	-26.17%	-15.31%	-10.00%	-2.97%	-2.57%
	(-2.40)	(-2.84)	(-6.36)	(-4.55)	(-2.34)	(-2.06)
Log (Size)	0.64%	3.18%	3.00%	3.51%	-0.10%	0.13%
	(5.44)	(5.62)	(10.05)	(10.79)	(-0.46)	(0.46)
Log (BM)	-0.85%	-4.24%	-1.37%	-0.89%	-0.52%	-1.16%
	(-6.60)	(-6.71)	(-2.05)	(-1.13)	(-3.27)	(-3.69)
Ret [0]	18.71%	30.47%	14.87%	12.77%	9.06%	12.85%
	(8.03)	(7.71)	(3.11)	(5.84)	(7.56)	(8.62)
Ret [-5, -1]	2.87%	6.51%	7.32%	7.96%	5.29%	8.82%
	(4.59)	(2.87)	(2.10)	(3.51)	(6.77)	(7.33)
Ret [-26, -6]	0.61%	3.21%	6.36%	4.58%	3.73%	6.39%
	(1.76)	(1.94)	(3.71)	(2.71)	(4.91)	(5.21)
News Sentiment [0]	33.24%	86.97%	2.22%	2.14%	2.45%	2.55%
	(34.16)	(24.96)	(3.24)	(4.07)	(8.05)	(10.31)
News Sentiment [-5, -1]	15.50%	56.17%	1.03%	1.66%	1.05%	1.38%
	(22.75)	(20.88)	(1.94)	(4.10)	(10.75)	(10.56)
News Sentiment [-26, -6]	7.89%	30.74%	0.97%	0.76%	0.33%	0.62%
	(18.59)	(16.15)	(3.97)	(3.17)	(5.60)	(6.36)
Net DD + Net DD 21	-1.37%	-9.84%	-9.90%	-6.12%	-0.14%	-0.15%
	(-2.14)	(-3.95)	(-4.97)	(-4.33)	(-1.23)	(-1.05)
Obs. (Firm-Days)	2,782,100	2,782,100	164,081	643,752	1,967,098	1,965,704
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All Firms		All Firms		All Firms	

Table 9: WSB Reports and Future Returns - Price Pressure Reports

This table repeats the analysis in Table 3 after partitioning all DD reports into *Price Pressure (PP) Reports* and *Non-PP Reports*. *PP Report* is an indicator equal to one if the number of "price pressure" words in the report exceeds the number of "fundamental" words in the report (see Appendix C for the list of "price pressure" and "fundamental" words). *Net DD PP* is the *Net DD* measure computed for the subset of reports where *PP Report = 1,* and Net *DD Non-PP* is the *Net DD* measure computed for the subset of reports where *PP Report = 0.* All other variables are defined in Table 3 (with more detailed variable definitions in Appendix D). Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report formal tests of whether *Net DD PP - Net DD Non-PP* and *Net DD PP × D2021 - Net DD Non-PP PP × D2021* are significantly different from zero.

	Ret [1,5]	Ret [1,21]	Ret [1,5]	Ret [1,21]
	[1]	[2]	[3]	[4]
Net DD PP	0.89%	36.70%	0.96%	3.84%
	(1.24)	(1.93)	(1.14)	(2.73)
Net DD PP × D2021	-3.33%	-42.02%	-3.09%	-8.62%
	(-4.49)	(-2.18)	(-3.52)	(-3.91)
Net DD Non-PP	0.86%	2.76%	0.92%	2.19%
	(1.80)	(2.91)	(1.98)	(1.98)
Net DD Non-PP × D2021	0.73%	-1.70%	-0.60%	-3.34%
	(0.70)	(-1.83)	(-0.97)	(-2.21)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.60)	(-1.27)	(-1.60)	(-1.27)
Log (BM)	-0.08%	-0.25%	-0.08%	-0.26%
	(-0.97)	(-0.84)	(-0.99)	(-0.86)
Ret [0]	-7.21%	-9.34%	-7.27%	-9.30%
	(-5.13)	(-4.62)	(-5.21)	(-4.54)
Ret [-5, -1]	-2.49%	-3.34%	-2.49%	-3.38%
	(-2.26)	(-2.31)	(-2.26)	(-2.36)
Ret [-26, -6]	-0.38%	-0.82%	-0.38%	-0.82%
	(-1.48)	(-0.85)	(-1.48)	(-0.85)
News Sentiment [0]	0.08%	0.08%	0.06%	0.06%
	(2.73)	(0.82)	(1.68)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.22)	(0.10)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.07%
	(0.46)	(0.81)	(0.72)	(1.09)
Net DD PP - Net DD Non-PP	0.03%	33.94%	0.04%	1.65%
	(0.05)	(1.77)	(0.07)	(1.04)
Net DD PP × 21 - Net DD Non-PP × 21	-4.07%	-40.32%	-2.49%	-5.28%
	(-3.28)	(-2.01)	(-3.02)	(-2.84)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table 10: WSB Reports and Future Returns - Comment Agreement

This table reports the estimates of future one-month ahead returns on Net *DD Comment Agree, Net DD Comment Disagree,* and controls. *Net DD Comment Agree* is the *Net DD* measure computed for the subset of reports where *Comment Agreement = 1,* and Net *DD Comment Disagree* is the *Net DD* measure computed for the subset of reports where *Comment Agreement = 0.* All other variables are as defined in Table 3 (with more detailed variable definitions in Appendix D). Specification 1 report the results for all firms for the full time series, and Specification 2 allows the estimates on *Net DD Agree* and *Net DD Disagree* to vary in the pre-GME and post-GME period by interacting *Net DD* with *2021,* an indicator equal to one for the 2021 sample period and zero otherwise. Specifications 3 and 4 present analogous results after excluding GME and AMC from the sample. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether the *Net DD Agree - Net DD Disagree* and *Net DD Agree × 2021 - Net DD Disagree × 2021* is significantly different from zero.

	Ret [1,21]	Ret [1,21]	Ret [1,21]	Ret [1,21]
	[1]	[2]	[4]	[5]
Net DD Comments Agree	6.59%	8.41%	2.10%	2.93%
	(1.76)	(2.16)	(1.79)	(2.55)
Net DD Comments Disagree	1.10%	2.86%	-0.81%	1.48%
	(1.45)	(2.54)	(-0.81)	(1.42)
Net DD Comments Agree ×2021		-10.67%		-4.60%
		(-2.20)		(-1.66)
Net DD Comments Disagree × 2021		-1.90%		-2.97%
		(-2.61)		(-2.04)
Log (Size)	-0.27%	-0.27%	-0.27%	-0.27%
	(-1.27)	(-1.27)	(-1.26)	(-1.27)
Log (BM)	-0.25%	-0.25%	-0.26%	-0.26%
	(-0.84)	(-0.84)	(-0.86)	(-0.86)
Ret [0]	-9.35%	-9.35%	-9.30%	-9.30%
	(-4.63)	(-4.63)	(-4.54)	(-4.54)
Ret [-5, -1]	-3.33%	-3.34%	-3.38%	-3.38%
	(-2.30)	(-2.31)	(-2.36)	(-2.36)
Ret [-26, -6]	-0.81%	-0.82%	-0.82%	-0.82%
	(-0.85)	(-0.85)	(-0.85)	(-0.85)
News Sentiment [0]	0.08%	0.08%	0.06%	0.06%
	(0.82)	(0.82)	(0.72)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.01%	0.01%
	(0.16)	(0.15)	(0.12)	(0.11)
News Sentiment [-26, -6]	0.05%	0.05%	0.07%	0.07%
	(0.85)	(0.84)	(1.10)	(1.09)
Agree - Disagree	5.49%	5.55%	2.91%	1.44%
	(1.87)	(1.74)	(3.32)	(2.04)
Agree × 2021- Disagree ×2021		-8.76%		-1.63%
		(-1.97)		(-0.80)
Obs. (Firm-Days)	2,782,100	2,274,064	2,780,590	2,272,802
Day FE	Yes	Yes	Yes	Yes
Firm Sample		All Firms	Exclude GME & AMC	

Table 11: WSB Reports and Future Returns - Trader Agreement

This table reports the estimates of future one-month ahead returns on Net *DD Trader Agree, Net DD Trader Disagree*, and controls. *Net DD Trader Agree* is the *Net DD* measure computed when trader order imbalances on the day the report is issued are in the same direction as the report recommendation (i.e., net buying following a buy recommendation)*,* and Net *DD Trader Disagree* is the *Net DD* measure computed when trader order imbalances are not in the same direction of the report recommendation All other variables are as defined in Table 3 (with more detailed variable definitions in Appendix D). We report the results for three separate groups of investors: institutional trading (Specifications 1 and 2), large retail trading (Specifications 3 and 4), and small retail trading (Specifications 5 and 6). More detailed variable definitions are in Appendix D. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether the *Net DD Trader Agree - Net DD Trader Disagree* and *Net DD Trader Agree × 2021 - Net DD Trader Disagree × 2021 are* significantly different from zero.

	Inst. Traders		Large Retail Traders		Small Retail Traders	
	Ret [1,21]	*Ret [1,21]*	*Ret [1,21]*	*Ret [1,21]*	*Ret [1,21]*	*Ret [1,21]*
	[1]	[2]	[4]	[5]	[4]	[5]
Net DD Traders Agree	1.36%	4.15%	2.51%	5.61%	1.65%	7.42%
	(2.03)	(2.56)	(2.07)	(2.35)	(1.61)	(2.54)
Net DD Trades Disagree	2.59%	7.90%	1.30%	6.61%	3.05%	3.09%
	(1.62)	(2.01)	(1.27)	(1.90)	(1.56)	(1.39)
Net DD Trader Agree × 2021		-3.51%		-3.89%		-7.11%
		(-2.74)		(-2.20)		(-2.80)
Net DD Trader Disagree × 2021		-6.83%		-6.88%		-0.06%
		(-2.48)		(-2.50)		(-0.21)
Log (Size)	-0.27%	-0.27%	-0.27%	-0.27%	-0.27%	-0.28%
	(-1.26)	(-1.26)	(-1.26)	(-1.26)	(-1.26)	(-1.26)
Log (BM)	-0.25%	-0.25%	-0.25%	-0.25%	-0.25%	-0.25%
	(-0.83)	(-0.83)	(-0.82)	(-0.82)	(-0.82)	(-0.81)
Ret [0]	-9.17%	-9.18%	-9.21%	-9.22%	-9.24%	-9.26%
	(-4.26)	(-4.27)	(-4.38)	(-4.39)	(-4.39)	(-4.41)
Ret [-5, -1]	-3.31%	-3.32%	-3.31%	-3.32%	-3.30%	-3.32%
	(-2.26)	(-2.28)	(-2.26)	(-2.28)	(-2.26)	(-2.28)
Ret [-26, -6]	-0.73%	-0.74%	-0.73%	-0.74%	-0.73%	-0.74%
	(-0.76)	(-0.77)	(-0.76)	(-0.77)	(-0.76)	(-0.77)
News Sentiment [0]	0.07%	0.07%	0.07%	0.07%	0.07%	0.06%
	(0.71)	(0.70)	(0.71)	(0.70)	(0.69)	(0.68)
News Sentiment [-5, -1]	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
	(0.11)	(0.10)	(0.11)	(0.10)	(0.10)	(0.09)
News Sentiment [-26, -6]	0.05%	0.05%	0.05%	0.05%	0.05%	0.05%
	(0.87)	(0.86)	(0.87)	(0.86)	(0.86)	(0.85)
Buy	0.00%	0.00%	0.21%	0.21%	0.37%	0.36%
	(-0.03)	(-0.00)	(2.60)	(2.60)	(2.36)	(2.33)
Buy * 2021	0.07%	0.06%	-0.08%	-0.08%	-0.42%	-0.40%
	(0.31)	(0.29)	(-0.64)	(-0.66)	(-0.94)	(-0.91)
Agree - Disagree	-1.23%	-3.75%	1.20%	-1.00%	-1.40%	4.32%
	(-1.10)	(-1.36)	(1.69)	(-0.55)	(-0.78)	(4.69)
Agree × 2021– Disagree × 2021		3.32%		2.99%		-7.05%
		(1.51)		(1.54)		(2.50)
Obs. (Firm-Days)	2,720,394	2,720,394	2,720,394	2,720,394	2,720,394	2,720,394
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm Sample	All Firms		All Firms		All Firms	

Internet Appendix for:

Place your bets? The market consequences of investment research on Reddit's Wallstreetbets

Daniel Bradley, Jan Hanousek Jr.*, Russell Jame, and Zicheng Xiao

In this appendix, we tabulate results from select robustness tests referenced in the paper. The set of tables are as follows:

Table IA.1 Determinants of WSB Coverage -Exclude GME & AMC

This table repeats the analysis in Table 2 after excluding GME and AMC.

	WSB Coverage [1]	SA Coverage [2]	WSB - SA Coverage [3]
Inst Ownership	-0.03	-0.06	0.03
	(-3.93)	(-5.87)	(3.71)
Inst Ownership × 2021	-0.04	0.04	-0.08
	(-1.47)	(3.68)	(-2.92)
Log (Breadth of Ownership)	0.01	0.03	-0.03
	(0.77)	(2.10)	(-2.06)
Log (Breadth of Ownership) × 2021	-0.01	-0.01	0.00
	(-0.65)	(-0.86)	(0.03)
Log (Size)	0.25	0.80	-0.55
	(4.51)	(12.30)	(-8.24)
Log (Size) × 2021	0.47	-0.11	0.57
	(4.44)	(-1.59)	(5.20)
Log (BM) × 2021	-0.02	-0.03	0.01
	(-2.45)	(-2.46)	(1.29)
Log (BM)	0.03	0.00	0.03
	(1.78)	(-0.13)	(1.34)
Negative BM	0.02	0.03	0.00
	(1.20)	(0.62)	(-0.07)
Negative BM × 2021	-0.06	-0.01	-0.05
	(-0.60)	(-0.22)	(-0.59)
Log (Vol)	0.10	0.45	-0.35
	(4.68)	(9.49)	(-8.07)
Log (Vol) × 2021	0.16	-0.24	0.40
	(2.72)	(-4.46)	(6.27)
Log (Turn)	0.00	-0.03	0.04
	(0.27)	(-1.96)	(2.73)
Log (Turn) × 2021	0.11	0.08	0.03
	(3.10)	(5.01)	(1.03)
Return (m-1)	0.01	0.01	0.00
	(2.03)	(2.12)	(0.09)
Return (m-1) × 2021	0.00	-0.01	0.01
	(-0.16)	(-0.63)	(0.40)
Return (m-2, m-12)	0.28	0.22	0.07
	(3.00)	(4.95)	(0.92)
Return (m-2, m-12) × 2021	-0.25	-0.20	-0.05
	(-2.83)	(-4.79)	(-0.72)
Log (Media Coverage)	0.17	0.38	-0.21
	(3.69)	(6.26)	(-4.06)
Log (Media Coverage) × 2021	0.03	-0.20	0.24
	(0.57)	(-3.93)	(2.88)
Log (IBES Coverage)	-0.01	0.04	-0.05
	(-0.76)	(2.88)	(-4.52)
Log (IBES Coverage) × 2021	-0.08	-0.04	-0.04
	(-2.60)	(-2.87)	(-1.39)
High Max	0.03	(0.06)	-0.03
	(3.05)	(3.69)	(-1.75)
High Max × 2021	0.17	(0.03)	0.15
	(7.60)	(1.21)	(7.19)
Heavy Short	0.05	0.15	-0.10
	(2.67)	(5.52)	(-3.43)
Heavy Short × 2021	0.28	-0.08	0.36
	(3.91)	(-1.94)	(4.44)
Recent IPO	0.57	0.97	-0.40

	(3.81)	(6.83)	(-2.50)
Recent IPO × 2021	1.22	0.40	0.83
	(4.09)	(3.08)	(2.75)
Observations (Firm-Months)	117,453	117,453	117,453
Fixed Effects	Month	Month	Month
R-square	4.44%	15.43%	6.09%
Sample		Exclude GME & AMC	

Table IA.2: WSB Reports and Future Daily Returns – Robustness

This table reports results from the estimation of the following equation:

$$R_{it+1} = \beta_1 Net\ DD_{it-1,t-x} + \beta_1 Net\ DD_{it-1,t-x} \times 2021 + Controls_{it} + Day_t + \varepsilon_{it.}$$

The dependent variable is the one-day ahead stock return. *Net DD* is the number of buy DD recommendations for stock *i* computed over days *t-1* through *t-x* less the number of sell DD recommendations for stock *i* over the same horizon. We measure *Net DD* over the prior week (i.e., x =5) or the prior month (i.e. x = 21). All other variables are defined as in Table 3. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether *Net DD + Net DD × 2021* is significantly different from zero.

	Ret [1]	Ret [1]	Ret [1]	Ret [1]
	[1]	[2]	[3]	[4]
Net DD [1,5]	0.11%		0.12%	
	(2.17)		(2.48)	
Net DD [1,5] × 2021	-0.04%		-0.12%	
	(-0.45)		(-1.87)	
Net DD [1,21]		0.06%		0.06%
		(2.75)		(2.78)
Net DD [1,21] × 2021		-0.03%		-0.06%
		(-1.23)		(-2.17)
Log (Size)	-0.02%	-0.02%	-0.02%	-0.02%
	(-2.38)	(-2.40)	(-2.38)	(-2.39)
Log (BM)	-0.01%	-0.01%	-0.01%	-0.01%
	(-0.95)	(-0.94)	(-0.96)	(-0.96)
Ret [0]	-3.29%	-3.29%	-3.29%	-3.29%
	(-5.23)	(-5.23)	(-5.22)	(-5.22)
Ret [-5, -1]	-0.90%	-0.90%	-0.90%	-0.90%
	(-2.90)	(-2.90)	(-2.89)	(-2.89)
Ret [-26, -6]	-0.08%	-0.08%	-0.08%	-0.08%
	(-1.64)	(-1.67)	(-1.64)	(-1.66)
News Sentiment [0]	0.07%	0.07%	0.07%	0.07%
	(3.71)	(3.73)	(3.50)	(3.50)
News Sentiment [-5, -1]	0.00%	0.00%	0.00%	0.00%
	(0.08)	(0.10)	(-0.28)	(-0.29)
News Sentiment [-26, -6]	0.00%	0.00%	0.00%	0.00%
	(0.34)	(0.30)	(0.57)	(0.54)
Net DD + Net DD × 2021	0.07%	0.03%	0.00%	0.00%
	(1.02)	(1.77)	(0.12)	(-0.03)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table IA.3: WSB Reports and Future Returns - DD Reports vs Non-Research Postings (Robustness)

This table repeats the analysis in Table 7 of the paper using a broader definition of non-research posts. *Non-Research* is the number of non-research posts that include the firm's ticker in the title and the post belongs to one of the following *WSB* categories: *News, Losses, Gains, Charts, Shi$posts,* **Yolo**, and **Discussion**. All other details are identical to Table 7.

	Ret [1,5]	Ret [1,21]	Ret [1,5]	Ret [1,21]
	[1]	[2]	[3]	[4]
NET_DD	0.83%	5.62%	0.87%	2.28%
	(1.79)	(2.34)	(1.92)	(2.29)
NET_DD × 2021	-0.29%	-4.40%	-0.75%	-3.18%
	(-0.35)	(-2.54)	(-1.22)	(-2.28)
Non- Research Posts	0.20%	2.53%	0.27%	0.22%
	(0.77)	(1.15)	(1.18)	(0.35)
Non-Research Posts × 2021	1.44%	-3.23%	-0.39%	-1.88%
	(1.42)	(-1.61)	(-1.53)	(-2.77)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.60)	(-1.27)	(-1.60)	(-1.26)
Log (BM)	-0.08%	-0.25%	-0.08%	-0.26%
	(-0.97)	(-0.84)	(-0.99)	(-0.86)
Ret [0]	-7.27%	-9.36%	-7.27%	-9.27%
	(-5.20)	(-4.64)	(-5.21)	(-4.51)
Ret [-5, -1]	-2.49%	-3.35%	-2.49%	-3.38%
	(-2.26)	(-2.32)	(-2.26)	(-2.36)
Ret [-26, -6]	-0.38%	-0.82%	-0.38%	-0.82%
	(-1.48)	(-0.86)	(-1.48)	(-0.86)
News Sentiment [0]	0.08%	0.07%	0.06%	0.06%
	(2.76)	(0.80)	(1.67)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.21)	(0.15)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.07%
	(0.45)	(0.85)	(0.72)	(1.10)
Net DD - Non-Research	0.63%	3.09%	0.60%	2.06%
	(1.26)	(2.45)	(1.25)	(2.14)
Net DD 21 - Non-Research 21	-1.72%	-1.17%	-0.39%	-1.30%
	(-1.29)	(-0.89)	(0.50)	(-0.91)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample		All Firms		Exclude GME & AMC

Table IA.4: WSB Reports and Cash Flow News – Exclude GME & AMC

This table repeats the analysis in Table 8 after excluding GME and AMC.

	Media [1,5]	Media [1,21]	Pos FE [1,5]	Pos FE [1,21]	Pos FR [1,5]	Pos FR [1,21]
	[1]	[2]	[3]	[4]	[5]	[6]
Net DD	4.79%	16.30%	4.93%	3.66%	3.07%	2.72%
	(2.08)	(1.74)	(2.88)	(1.63)	(2.38)	(2.08)
NET_DD × 2021	-5.00%	-23.90%	-16.02%	-11.80%	-3.49%	-3.05%
	(-2.18)	(-2.51)	(-3.68)	(-4.05)	(-2.74)	(-2.19)
Log (Size)	0.64%	3.18%	3.00%	3.51%	-0.10%	0.13%
	(5.44)	(5.62)	(10.05)	(10.79)	(-0.47)	(0.46)
Log (BM)	-0.84%	-4.21%	-1.35%	-0.87%	-0.52%	-1.15%
	(-6.56)	(-6.67)	(-2.01)	(-1.10)	(-3.24)	(-3.66)
Ret [0]	18.47%	30.53%	14.88%	12.86%	9.06%	12.92%
	(8.09)	(8.08)	(3.13)	(5.87)	(7.53)	(8.67)
Ret [-5, -1]	2.84%	6.44%	7.36%	7.93%	5.29%	8.83%
	(4.54)	(2.85)	(2.11)	(3.50)	(6.79)	(7.34)
Ret [-26, -6]	0.60%	3.16%	6.33%	4.58%	3.73%	6.42%
	(1.74)	(1.92)	(3.69)	(2.68)	(4.91)	(5.21)
News Sentiment [0]	33.21%	86.97%	2.24%	2.15%	2.46%	2.54%
	(33.80)	(24.92)	(3.28)	(4.06)	(8.01)	(10.22)
News Sentiment [-5, -1]	15.51%	56.23%	1.07%	1.67%	1.05%	1.38%
	(22.78)	(20.96)	(1.99)	(4.12)	(10.76)	(10.65)
News Sentiment [-26, -6]	7.90%	30.77%	0.96%	0.75%	0.33%	0.62%
	(18.68)	(16.17)	(3.96)	(3.14)	(5.64)	(6.34)
NET DD + Net DD × 2021	-0.21%	-7.60%	-11.09%	-8.14%	-0.42%	-0.33%
	(-0.40)	(-3.17)	(-2.81)	(-3.73)	(-0.97)	(-0.49)
Obs. (Firm-Days)	2,780,590	2,780,590	163,969	643,126	1,965,704	1,965,704
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Exclude GME& AMC		Exclude GME& AMC		Exclude GME& AMC	

Table IA.5: WSB Reports and Cash Flow News – Non-Research Postings

This table repeats the analysis in Table 8 after including *Non-Research Posts* (as defined in Table 7) and *Non-Research Posts* × 2021.

| | Media [1,5] | Media [1,21] | Pos FE [1,5] | Pos FE [1,21] | Pos FR [1,5] | Pos FR [1,21] |
	[1]	[2]	[3]	[4]	[5]	[6]
Net DD	4.96%	17.05%	5.38%	3.92%	2.72%	2.35%
	(2.06)	(2.03)	(3.32)	(1.94)	(2.19)	(1.90)
NET_DD × 2021	-5.00%	-20.37%	-15.29%	-9.85%	-2.87%	-2.41%
	(-2.13)	(-2.44)	(-6.32)	(-4.61)	(-2.19)	(-1.93)
Non- Research Posts	-1.89%	-6.49%	0.59%	-0.73%	1.68%	1.03%
	(-1.06)	(-0.79)	(0.11)	(-0.23)	(1.43)	(0.61)
Non-Research Posts × 2021	-1.16%	-8.44%	-25.37%	-3.74%	-1.65%	-1.19%
	(-0.61)	(-0.93)	(-3.92)	(-1.31)	(-1.46)	(-0.73)
Log (Size)	0.64%	3.18%	3.00%	3.51%	-0.10%	0.13%
	(5.45)	(5.62)	(10.09)	(10.80)	(-0.47)	(0.46)
Log (BM)	-0.85%	-4.24%	-1.37%	-0.89%	-0.52%	-1.16%
	(-6.60)	(-6.71)	(-2.05)	(-1.13)	(-3.27)	(-3.69)
Ret [0]	18.78%	30.84%	14.89%	12.82%	9.05%	12.85%
	(8.07)	(8.06)	(3.12)	(5.85)	(7.57)	(8.61)
Ret [-5, -1]	2.88%	6.52%	7.34%	7.96%	5.29%	8.82%
	(4.59)	(2.87)	(2.11)	(3.51)	(6.77)	(7.33)
Ret [-26, -6]	0.61%	3.21%	6.36%	4.58%	3.73%	6.39%
	(1.77)	(1.95)	(3.71)	(2.71)	(4.91)	(5.21)
News Sentiment [0]	33.24%	86.97%	2.23%	2.14%	2.45%	2.55%
	(34.16)	(24.97)	(3.24)	(4.07)	(8.05)	(10.31)
News Sentiment [-5, -1]	15.50%	56.17%	1.03%	1.66%	1.05%	1.38%
	(22.76)	(20.90)	(1.93)	(4.10)	(10.74)	(10.55)
News Sentiment [-26, -6]	7.89%	30.75%	0.97%	0.76%	0.33%	0.62%
	(18.59)	(16.15)	(3.98)	(3.17)	(5.60)	(6.36)
NET DD + Net DD × 2021	6.85%	23.54%	4.79%	4.65%	1.04%	1.32%
	(2.35)	(2.49)	(0.81)	(1.25)	(0.94)	(0.78)
Net DD 21 - Non-Research 21	-3.84%	-11.93%	10.08%	-6.11%	-1.22%	-1.22%
	(-1.36)	(-1.18)	(1.58)	(-1.62)	(-0.97)	(-0.74)
Obs. (Firm-Days)	2,780,590	2,780,590	163,969	643,126	1,965,704	1,965,704
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All Firms		All Firms		All Firms	

IA.7

Table IA6: WSB Reports and Future Returns - Price Pressure Reports Robustness

This table repeats the analysis in Table 9 using an alternative definition of *price pressure (PP)* report. In this table, *PP Report* is an indicator equal to one if report contains at least one "price pressure" word, and zero otherwise. See Appendix C for the list of "price pressure" words.

	Ret [1,5]	Ret [1,21]	Ret [1,5]	Ret [1,21]
	[1]	[2]	[3]	[4]
Net DD PP	1.23%	22.10%	1.41%	3.91%
	(2.31)	(2.00)	(3.15)	(12.27)
Net DD PP × D2021	-2.73%	-25.65%	-2.73%	-7.16%
	(-4.44)	(-2.25)	(-4.73)	(-3.07)
Net DD Non-PP	0.77%	2.01%	0.81%	1.98%
	(1.45)	(1.60)	(1.58)	(1.55)
Net DD Non-PP × D2021	0.83%	-0.95%	-0.52%	-3.21%
	(0.77)	(-0.69)	(-0.80)	(-2.03)
Log (Size)	-0.08%	-0.27%	-0.08%	-0.27%
	(-1.60)	(-1.27)	(-1.60)	(-1.27)
Log (BM)	-0.08%	-0.25%	-0.08%	-0.26%
	(-0.97)	(-0.84)	(-0.99)	(-0.86)
Ret [0]	-7.21%	-9.34%	-7.27%	-9.30%
	(-5.13)	(-4.62)	(-5.21)	(-4.54)
Ret [-5, -1]	-2.49%	-3.33%	-2.49%	-3.38%
	(-2.26)	(-2.31)	(-2.26)	(-2.36)
Ret [-26, -6]	-0.38%	-0.81%	-0.38%	-0.82%
	(-1.48)	(-0.85)	(-1.48)	(-0.85)
News Sentiment [0]	0.08%	0.08%	0.06%	0.06%
	(2.73)	(0.83)	(1.67)	(0.72)
News Sentiment [-5, -1]	0.01%	0.01%	0.00%	0.01%
	(0.22)	(0.13)	(-0.09)	(0.11)
News Sentiment [-26, -6]	0.01%	0.05%	0.01%	0.06%
	(0.46)	(0.82)	(0.72)	(1.09)
Net DD PP - Net DD Non-PP	0.46%	20.09%	0.59%	1.93%
	(0.79)	(1.86)	(1.34)	(1.50)
Net DD PP × 21 - Net DD Non-PP × 21	-3.56%	-24.70%	-2.21%	-3.95%
	(-3.09)	(-1.96)	(-3.31)	(-1.91)
Obs. (Firm-Days)	2,782,100	2,782,100	2,780,590	2,780,590
Day FE	Yes	Yes	Yes	Yes
Sample	All Firms		Exclude GME & AMC	

Table IA.7: Investor Order Imbalances Following WSB Reports

This table reports results from the following equation:

$$OIB_{it} = \beta_1 Net\ DD_{it-1,t} + \beta_2 Net\ DD_{it-1,t} \times 2021 + Controls_{it} + Day_t + \varepsilon_{it.}$$

The dependent variable, *OIB*, is one of three measures of directional trading for firm *i* on day *t*: *Inst. Vol OIB, Retail Vol OIB*, or *Retail Trade OIB*. *Inst Vol OIB* is defined as institutional buy share volume less institutional sell share volume scaled by total institutional share volume. *Retail Vol OIB* and *Retail Trade OIB* are defined analogously after replacing *Institutional Share Volume* with *Retail Share Volume* and *Retail Number of Trades*, respectively. *Net DD* is the number of buy DD recommendations for stock *i* across days *t* and *t-1* less the number of sell DD recommendations for stock *i* on days *t* and *t-1*. Detailed definitions of all control variables are available in Appendix D. Standard errors are clustered by firm and month, and t-statistics are reported below each estimate. Below the regression estimates, we also report a formal test of whether *Net DD + Net DD × 2021* is significantly different from zero.

	Inst. Vol OIB [1]	Ret Vol OIB [2]	Ret Trade OIB [3]	Inst. Vol OIB [5]	Ret Vol OIB [5]	Ret Trade OIB [6]
Net DD	-0.32%	1.37%	5.04%	-0.38%	1.24%	4.93%
	(-1.10)	(3.36)	(6.72)	(-1.39)	(2.93)	(6.74)
Net DD × 2021	0.50%	-1.42%	-0.48%	0.51%	-1.32%	0.45%
	(1.36)	(-3.36)	(-0.45)	(1.32)	(-2.92)	(0.60)
Log (Size)	0.70%	0.27%	0.14%	0.70%	0.27%	0.14%
	(17.22)	(10.17)	(3.20)	(17.22)	(10.16)	(3.19)
Log (BM)	0.02%	-0.03%	-0.17%	0.02%	-0.03%	-0.18%
	(0.50)	(-0.65)	(-3.72)	(0.47)	(-0.70)	(-3.82)
Ret [-5, -1]	0.71%	-2.36%	-2.40%	0.71%	-2.36%	-2.39%
	(2.18)	(-5.53)	(-5.38)	(2.18)	(-5.53)	(-5.35)
Ret [-26, -6]	0.32%	-0.85%	-0.96%	0.32%	-0.85%	-0.96%
	(2.26)	(-4.80)	(-5.01)	(2.26)	(-4.80)	(-5.01)
News Sentiment [-5, -1]	0.10%	0.11%	0.06%	0.10%	0.11%	0.06%
	(3.88)	(2.00)	(1.34)	(3.83)	(2.01)	(1.41)
News Sentiment [-26, -6]	-0.01%	0.03%	0.01%	-0.01%	0.03%	0.01%
	(-0.91)	(1.34)	(0.51)	(-0.88)	(1.40)	(0.59)
Retail Trade OIB [-5, - 1]	1.00%	10.26%	26.64%	1.00%	10.25%	26.57%
	(3.82)	(31.75)	(26.89)	(3.81)	(31.67)	(26.98)
Retail Vol OIB [-5, - 1]	-1.27%	4.46%	-5.72%	-1.27%	4.46%	-5.69%
	(-7.82)	(12.34)	(-14.27)	(-7.81)	(12.32)	(-14.28)
Inst Vol OIB [-5, - 1]	30.40%	-1.88%	-1.45%	30.39%	-1.88%	-1.45%
	(60.95)	(-4.64)	(-4.57)	(60.96)	(-4.64)	(-4.58)
Net DD + Net DD × 2021	0.18%	-0.05%	4.56%	0.13%	-0.08%	5.38%
	(0.75)	(-0.67)	(5.85)	(0.45)	(-0.66)	(15.64)
Obs. (Firm-Days)	2,705,492	2,705,492	2,705,492	2,703,986	2,703,986	2,703,986
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample		Full Sample			Exclude GME & AMC	

Table IA.8: WSB Reports and Future Returns - Trader Agreement (Exclude GME & AMC)

This table reports the analysis in Table 11 after excluding GME and AMC.

	Inst. Traders		Large Retail Traders		Small Retail Traders	
	Ret [1,21]	*Ret [1,21]*	*Ret [1,21]*	*Ret [1,21]*	*Ret [1,21]*	*Ret [1,21]*
	[1]	[2]	[4]	[5]	[4]	[5]
Net DD Traders Agree	0.04%	2.13%	0.12%	2.46%	0.21%	3.09%
	(0.04)	(1.68)	(0.10)	(1.65)	(0.18)	(2.16)
Net DD Trades Disagree	0.13%	2.46%	0.04%	2.11%	-0.32%	0.56%
	(0.11)	(2.33)	(0.04)	(2.73)	(-0.37)	(1.52)
Net DD Trader Agree × 2021		-3.83%		-3.93%		-4.57%
		(-2.33)		(-2.17)		(-2.67)
Net DD Trader Disagree × 2021		-3.78%		-3.60%		-1.93%
		(-2.53)		(-2.40)		(-1.19)
Log (Size)	-0.27%	-0.27%	-0.27%	-0.27%	-0.27%	-0.27%
	(-1.25)	(-1.26)	(-1.25)	(-1.26)	(-1.26)	(-1.26)
Log (BM)	-0.26%	-0.26%	-0.26%	-0.26%	-0.25%	-0.25%
	(-0.85)	(-0.85)	(-0.85)	(-0.85)	(-0.84)	(-0.84)
Ret [0]	-9.12%	-9.13%	-9.17%	-9.17%	-9.20%	-9.20%
	(-4.19)	(-4.19)	(-4.31)	(-4.31)	(-4.32)	(-4.32)
Ret [-5, -1]	-3.36%	-3.37%	-3.36%	-3.36%	-3.36%	-3.37%
	(-2.33)	(-2.33)	(-2.33)	(-2.34)	(-2.33)	(-2.34)
Ret [-26, -6]	-0.74%	-0.74%	-0.74%	-0.74%	-0.73%	-0.74%
	(-0.77)	(-0.77)	(-0.77)	(-0.77)	(-0.77)	(-0.77)
News Sentiment [0]	0.05%	0.06%	0.05%	0.05%	0.05%	0.05%
	(0.61)	(0.61)	(0.59)	(0.60)	(0.57)	(0.57)
News Sentiment [-5, -1]	0.01%	0.01%	0.01%	0.01%	0.01%	0.00%
	(0.07)	(0.06)	(0.07)	(0.06)	(0.06)	(0.05)
News Sentiment [-26, -6]	0.07%	0.07%	0.07%	0.07%	0.07%	0.07%
	(1.12)	(1.12)	(1.12)	(1.12)	(1.11)	(1.11)
Buy	0.00%	(0.00)	0.21%	0.21%	(0.00)	0.36%
	(-0.00)	(-0.00)	(2.60)	(2.59)	(2.35)	(2.33)
Buy * 2021	(0.00)	(0.00)	-0.11%	-0.11%	(-0.00)	-0.43%
	(0.21)	(0.21)	(-1.05)	(-1.02)	(-1.05)	(-1.02)
Agree - Disagree	-0.08%	-0.33%	0.08%	0.34%	0.53%	2.53%
	(-0.15)	(-0.34)	(0.11)	(0.28)	(0.93)	(2.05)
Agree × 2021– Disagree × 2021		-0.05%		-0.33%		-2.63%
		(-0.06)		(0.22)		(-1.67)
Obs. (Firm-Days)	2,720,394	2,720,394	2,720,394	2,720,394	2,720,394	2,720,394
Day FE	YES	YES	YES	YES	YES	YES
Firm Sample	Exclude GME & AMC		Exclude GME & AMC		Exclude GME & AMC	